

PLUMED

2.3.8

Generated by Doxygen 1.8.14



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	About this manual . . . . .	1
1.2	Codes interfaced with PLUMED . . . . .	1
<b>2</b>	<b>Change Log</b>	<b>3</b>
2.1	Version 2.0 . . . . .	4
2.2	Version 2.1 . . . . .	7
2.3	Version 2.2 . . . . .	13
2.4	Version 2.3 . . . . .	18
<b>3</b>	<b>Installation</b>	<b>25</b>
3.1	Configuring PLUMED . . . . .	25
3.1.1	BLAS and LAPACK . . . . .	28
3.1.2	VMD trajectory plugins . . . . .	29
3.2	Compiling PLUMED . . . . .	29
3.3	Installing PLUMED . . . . .	31
3.4	Patching your MD code . . . . .	32
3.5	Cross compiling . . . . .	33
3.6	Installing PLUMED with MacPorts . . . . .	34
3.7	Installing PLUMED on a cluster . . . . .	35
3.8	Other hints . . . . .	36
3.9	Code specific notes . . . . .	36
3.9.1	amber14 . . . . .	36
3.9.2	gromacs-2016.5 . . . . .	37
3.9.3	gromacs-4.5.7 . . . . .	37
3.9.4	gromacs-5.0.7 . . . . .	37
3.9.5	gromacs-5.1.4 . . . . .	37
3.9.6	lammps-6Apr13 . . . . .	38
3.9.7	namd-2.8 . . . . .	38
3.9.8	namd-2.9 . . . . .	38
3.9.9	qespresso-5.0.2 . . . . .	38

<b>4</b>	<b>Getting started</b>	<b>39</b>
4.1	Plumed units . . . . .	40
4.2	UNITS . . . . .	40
<b>5</b>	<b>Collective variables</b>	<b>43</b>
5.1	Groups and Virtual Atoms . . . . .	44
5.1.1	Specifying Atoms . . . . .	44
5.1.1.1	Molecules . . . . .	44
5.1.1.2	Broken Molecules and PBC . . . . .	44
5.1.2	Virtual Atoms . . . . .	45
5.1.3	GROUP . . . . .	45
5.1.4	MOLINFO . . . . .	47
5.1.5	WHOLEMOLECULES . . . . .	50
5.1.6	FIT_TO_TEMPLATE . . . . .	51
5.1.7	WRAPAROUND . . . . .	53
5.1.8	RESET_CELL . . . . .	55
5.1.9	CENTER_OF_MULTICOLVAR . . . . .	56
5.1.10	CENTER . . . . .	57
5.1.11	COM . . . . .	58
5.1.12	FIXEDATOM . . . . .	59
5.1.13	GHOST . . . . .	60
5.2	CV Documentation . . . . .	61
5.2.1	ALPHABETA . . . . .	62
5.2.2	ALPHARMSD . . . . .	63
5.2.3	ANGLE . . . . .	66
5.2.4	ANTIBETARMSD . . . . .	67
5.2.5	CELL . . . . .	70
5.2.6	CONSTANT . . . . .	70
5.2.7	CONTACTMAP . . . . .	71
5.2.8	COORDINATION . . . . .	74
5.2.9	CS2BACKBONE . . . . .	75

5.2.10	DHENERGY	78
5.2.11	DIHCOR	79
5.2.12	DIPOLE	80
5.2.13	DISTANCE_FROM_CONTOUR	81
5.2.14	DISTANCE	83
5.2.15	ENERGY	85
5.2.16	ERMSD	85
5.2.17	FAKE	87
5.2.18	FRET	88
5.2.19	GPROPERTYMAP	89
5.2.20	GYRATION	91
5.2.21	JCOUPLING	92
5.2.22	NOE	94
5.2.23	PARABETARMSD	95
5.2.24	PATHMSD	98
5.2.25	PATH	99
5.2.26	PCAVARS	101
5.2.27	POSITION	104
5.2.28	PRE	105
5.2.29	PROPERTYMAP	107
5.2.30	PUCKERING	108
5.2.31	RDC	110
5.2.32	TEMPLATE	112
5.2.33	TORSION	113
5.2.34	VOLUME	114
5.3	Distances from reference configurations	115
5.3.1	DRMSD	116
5.3.2	MULTI-RMSD	117
5.3.3	PCARMSD	119
5.3.4	RMSD	120

5.3.5	TARGET	122
5.4	Functions	124
5.4.1	COMBINE	125
5.4.2	ENSEMBLE	126
5.4.3	FUNCPATHMSD	127
5.4.4	FUNCSUMHILLS	129
5.4.5	LOCALENSEMBLE	130
5.4.6	MATHEVAL	131
5.4.6.1	TIME	134
5.4.7	PIECEWISE	134
5.4.8	SORT	135
5.4.9	STATS	136
5.5	MultiColvar	138
5.5.1	MultiColvar functions	139
5.5.2	MultiColvar bias	141
5.5.3	ANGLES	141
5.5.4	BRIDGE	144
5.5.5	COORDINATIONNUMBER	145
5.5.6	DENSITY	148
5.5.7	DISTANCES	149
5.5.8	FCCUBIC	153
5.5.9	INPLANEDISTANCES	157
5.5.10	MOLECULES	160
5.5.11	PLANES	161
5.5.12	Q3	163
5.5.13	Q4	166
5.5.14	Q6	170
5.5.15	SIMPLECUBIC	174
5.5.16	TETRAHEDRAL	178
5.5.17	TORSIONS	181

5.5.18	XDISTANCES	183
5.5.19	XYDISTANCES	186
5.5.20	XZDISTANCES	189
5.5.21	YDISTANCES	192
5.5.22	YZDISTANCES	196
5.5.23	ZDISTANCES	199
5.5.24	MFILTER_BETWEEN	202
5.5.25	MFILTER_LESS	205
5.5.26	MFILTER_MORE	207
5.5.27	AROUND	210
5.5.28	CAVITY	213
5.5.29	INCYLINDER	216
5.5.30	INSPHERE	218
5.5.31	TETRAHEDRALPORE	220
5.5.32	GRADIENT	223
5.5.33	INTERMOLECULARTORSIONS	225
5.5.34	LOCAL_AVERAGE	227
5.5.35	LOCAL_Q3	230
5.5.36	LOCAL_Q4	234
5.5.37	LOCAL_Q6	238
5.5.38	NLINKS	242
5.5.39	SMAC	243
5.5.40	MTRANSFORM_BETWEEN	247
5.5.41	MTRANSFORM_LESS	250
5.5.42	MTRANSFORM_MORE	253
5.5.43	UWALLS	256
5.6	Exploiting contact matrices	258
5.6.1	ALIGNED_MATRIX	259
5.6.2	CONTACT_MATRIX	261
5.6.3	HBOND_MATRIX	262
5.6.4	SMAC_MATRIX	265
5.6.5	CLUSTER_WITHSURFACE	267
5.6.6	COLUMNSUMS	268
5.6.7	DFSCCLUSTERING	270
5.6.8	ROWSUMS	272
5.6.9	SPRINT	274
5.6.10	CLUSTER_DIAMETER	276
5.6.11	CLUSTER_DISTRIBUTION	277
5.6.12	CLUSTER_NATOMS	280
5.6.13	CLUSTER_PROPERTIES	281
5.6.14	OUTPUT_CLUSTER	284

<b>6 Analysis</b>	<b>287</b>
6.1 Dimensionality Reduction	289
6.2 COMMITTOR	289
6.3 DUMPATOMS	290
6.4 DUMPDERIVATIVES	291
6.5 DUMPFORCES	293
6.6 DUMPMASSCHARGE	294
6.7 DUMPMULTICOLVAR	295
6.8 DUMPPROJECTIONS	296
6.9 PRINT	297
6.9.1 FLUSH	298
6.10 UPDATE_IF	298
6.11 REWEIGHT_BIAS	300
6.12 REWEIGHT_METAD	301
6.13 REWEIGHT_TEMP	302
6.14 AVERAGE	303
6.15 HISTOGRAM	304
6.16 MULTICOLVARDENS	307
6.17 CONVERT_TO_FES	309
6.18 DUMPCUBE	310
6.19 DUMPGRID	311
6.20 FIND_CONTOUR_SURFACE	313
6.21 FIND_CONTOUR	314
6.22 FIND_SPHERICAL_CONTOUR	316
6.23 FOURIER_TRANSFORM	318
6.24 INTERPOLATE_GRID	319
6.25 CLASSICAL_MDS	320
6.25.1 Method of optimisation	322
6.26 PCA	324



<b>7</b>	<b>Bias</b>	<b>327</b>
7.1	ABMD	327
7.2	BIASVALUE	329
7.3	EXTENDED_LAGRANGIAN	331
7.4	EXTERNAL	333
7.5	LOWER_WALLS	334
7.6	METAD	336
7.7	METAINFERENCE	342
7.8	MOVINGRESTRAINT	344
7.9	PBMETAD	346
7.10	RESTRAINT	349
7.11	UPPER_WALLS	351
7.12	RESTART	352
<b>8</b>	<b>Command Line Tools</b>	<b>355</b>
8.1	driver-float	355
8.2	driver	357
8.2.1	READ	359
8.3	gentemplate	360
8.4	info	361
8.5	kt	362
8.6	manual	362
8.7	simplemd	363
8.8	sum_hills	364

<b>9</b>	<b>Miscellaneous</b>	<b>367</b>
9.1	Comments . . . . .	367
9.2	Continuation lines . . . . .	368
9.3	Using VIM syntax file . . . . .	369
9.4	Including other files . . . . .	374
9.4.1	INCLUDE . . . . .	374
9.5	Loading shared libraries . . . . .	375
9.5.1	LOAD . . . . .	375
9.6	Debugging the code . . . . .	376
9.6.1	DEBUG . . . . .	376
9.7	Changing exchange patterns in replica exchange . . . . .	377
9.7.1	RANDOM_EXCHANGES . . . . .	377
9.8	List of modules . . . . .	378
9.9	Frequently used tools . . . . .	379
9.9.1	histogrambead . . . . .	380
9.9.2	kernelfunctions . . . . .	381
9.9.3	landmarkselection . . . . .	382
9.9.4	pdbreader . . . . .	382
9.9.5	switchingfunction . . . . .	383
9.9.6	Regular Expressions . . . . .	384
9.9.7	Files . . . . .	385
9.9.7.1	Restart . . . . .	385
9.9.7.2	Backup . . . . .	386
9.9.7.3	Replica suffix . . . . .	386
<b>10</b>	<b>Performances</b>	<b>387</b>
10.1	GROMACS and PLUMED with GPU . . . . .	388
10.2	Metadynamics . . . . .	388
10.3	Multiple time stepping . . . . .	389
10.3.1	EFFECTIVE_ENERGY_DRIFT . . . . .	389
10.4	Multicolvar . . . . .	390
10.5	Neighbour Lists . . . . .	390
10.6	OpenMP . . . . .	391
10.7	Secondary Structure . . . . .	391
10.8	Time your Input . . . . .	391

<b>11 Tutorials</b>	<b>393</b>
11.1 Belfast tutorial: Analyzing CVs	394
11.1.1 Aims	394
11.1.2 Learning Outcomes	394
11.1.3 Resources	394
11.1.4 Instructions	394
11.1.4.1 A note on units	395
11.1.4.2 Introduction to the PLUMED input file	395
11.1.4.3 MULTICOLVAR	397
11.1.4.4 Analysis of Collective Variables	398
11.2 Belfast tutorial: Adaptive variables I	399
11.2.1 Aim	399
11.2.2 Resources	399
11.2.3 What happens when in a complex reaction?	399
11.2.4 Path collective variables	400
11.2.5 A note on the path topology	402
11.2.6 How many frames do I need?	403
11.2.7 Some tricks of the trade: the neighbors list.	403
11.2.8 The molecule of the day: alanine dipeptide	403
11.2.9 Examples	404
11.2.10 How to format my input?	406
11.2.11 Fast forward: metadynamics on the path	406
11.3 Belfast tutorial: Adaptive variables II	408
11.3.1 Aims	408
11.3.2 Learning Outcomes	408
11.3.3 Resources	409
11.3.4 Instructions	409
11.3.4.1 Visualising the trajectory	409
11.3.4.2 Finding collective variables	409
11.3.4.3 Dimensionality reduction	410

11.3.5	Extensions	411
11.3.6	Further Reading	411
11.4	Belfast tutorial: Umbrella sampling	411
11.4.1	Aims	411
11.4.2	Summary of theory	412
11.4.2.1	Biased sampling	412
11.4.2.2	Umbrella sampling	413
11.4.2.3	Weighted histogram analysis method	413
11.4.3	Learning Outcomes	414
11.4.4	Resources	414
11.4.5	Instructions	415
11.4.5.1	The model system	415
11.4.5.2	Restrained simulations	415
11.4.5.3	Reweighting the results	416
11.4.5.4	A free-energy landscape	416
11.4.5.5	Combining multiple restraints	417
11.4.6	Comments	418
11.4.6.1	How does PLUMED work	418
11.4.7	Further Reading	419
11.5	Belfast tutorial: Out of equilibrium dynamics	419
11.5.1	Resources	419
11.5.2	Steered MD	419
11.5.3	Moving on a more complex path	422
11.5.4	Why work is important?	422
11.5.5	Targeted MD	424
11.6	Belfast tutorial: Metadynamics	424
11.6.1	Aims	424
11.6.2	Summary of theory	425
11.6.3	Learning Outcomes	426
11.6.4	Resources	426

11.6.5	Instructions	426
11.6.5.1	The model system	426
11.6.5.2	Exercise 1. Setup and run a metadynamics simulation	426
11.6.5.3	Exercise 2. Restart a metadynamics simulation	428
11.6.5.4	Exercise 3. Calculate free-energies and monitor convergence	428
11.6.5.5	Exercise 4. Setup and run a well-tempered metadynamics simulation, part I	430
11.6.5.6	Exercise 5. Setup and run a well-tempered metadynamics simulation, part II	431
11.7	Belfast tutorial: Replica exchange I	431
11.7.1	Aims	431
11.7.2	Summary of theory	432
11.7.3	Learning Outcomes	433
11.7.4	Resources	433
11.7.5	Instructions	433
11.7.5.1	The model system	433
11.7.5.2	Exercise 1. Setup and run a PT simulation, part I	433
11.7.5.3	Exercise 2. Setup and run a PT simulation, part II	435
11.7.5.4	Exercise 3. Setup and run a PTMetaD simulation	436
11.7.5.5	Exercise 4. The Well-Tempered Ensemble	437
11.8	Belfast tutorial: Replica exchange II and Multiple walkers	439
11.8.1	Aims	439
11.8.1.1	Learning Outcomes	439
11.8.2	Resources	439
11.8.3	Instructions	439
11.8.3.1	Bias-Exchange Metadynamics	439
11.8.3.2	Convergence of the Simulations	441
11.8.3.3	Bias-Exchange Analysis with METAGUI	441
11.8.3.4	Multiple Walker Metadynamics	444
11.8.4	Reference	444
11.9	Belfast tutorial: NMR restraints	445
11.9.1	Aims	445

11.9.1.1 Learning Outcomes . . . . .	445
11.9.2 Resources . . . . .	445
11.9.3 Instructions . . . . .	445
11.9.3.1 Experimental data as Collective Variables . . . . .	445
11.9.3.2 Replica-Averaged Restrained Simulations . . . . .	446
11.9.4 Reference . . . . .	447
11.10Belfast tutorial: Steinhardt Parameters . . . . .	448
11.10.1 Aims . . . . .	448
11.10.1.1 Learning Outcomes . . . . .	448
11.10.2 Resources . . . . .	448
11.10.3 Instructions . . . . .	448
11.10.3.1 Simplemd . . . . .	448
11.10.3.2 Coordination Numbers . . . . .	449
11.10.3.3 Steinhard parameter . . . . .	449
11.10.3.4 Local versus Global . . . . .	449
11.10.3.5 Local Steinhardt parameters . . . . .	450
11.10.4 Further Reading . . . . .	450
11.11Cambridge tutorial . . . . .	451
11.11.1 Alanine dipeptide: our toy model . . . . .	451
11.11.2 Exercise 1. Metadynamics . . . . .	451
11.11.2.1 Resources . . . . .	451
11.11.2.2 Summary of theory . . . . .	451
11.11.2.3 Setup, run, and analyse a well-tempered metadynamics simulation . . . . .	452
11.11.2.4 Calculate free-energies and monitor convergence . . . . .	453
11.11.3 Exercise 2. Bias-Exchange Metadynamics . . . . .	454
11.11.3.1 Resources . . . . .	454
11.11.3.2 Summary of theory . . . . .	455
11.11.3.3 Setup, run, and analyse a well-tempered bias-exchange metadynamics simulation . . . . .	455
11.11.3.4 Calculate free-energies and monitor convergence . . . . .	456
11.11.4 Exercise 3. Replica-Average Metadynamics . . . . .	457

11.11.4.1 Resources . . . . .	457
11.11.4.2 Summary of theory . . . . .	458
11.11.4.3 The system: Chignolin . . . . .	458
11.11.4.4 Setup, run and analysis . . . . .	459
11.12Cineca tutorial . . . . .	459
11.12.1 Resources . . . . .	460
11.12.2 Alanine dipeptide: our toy model . . . . .	460
11.12.3 Monitoring collective variables . . . . .	460
11.12.3.1 Exercise 1: on-the-fly analysis . . . . .	461
11.12.3.2 Exercise 2: analysis with the driver tool . . . . .	462
11.12.4 Biasing collective variables . . . . .	463
11.12.4.1 Metadynamics . . . . .	463
11.12.4.2 Restraints . . . . .	467
11.12.4.3 Using multiple replicas . . . . .	469
11.12.4.4 Using multiple restraints with replica exchange . . . . .	470
11.13Using Hamiltonian replica exchange with GROMACS . . . . .	473
11.13.1 Generate scaled topologies . . . . .	473
11.13.2 Run GROMACS . . . . .	474
11.14Julich tutorial: Developing CVs in plumed . . . . .	476
11.14.1 Aims . . . . .	476
11.14.2 Learning Outcomes . . . . .	476
11.14.3 Resources . . . . .	476
11.14.4 Introduction . . . . .	477
11.14.5 Instructions . . . . .	477
11.14.5.1 Calculating a reasonably complex collective variable . . . . .	477
11.14.5.2 Implementing a new collective variable . . . . .	478
11.14.6 Final thoughts . . . . .	479
11.15Moving from PLUMED 1 to PLUMED 2 . . . . .	479
11.15.1 New syntax . . . . .	480
11.15.2 Groups . . . . .	480

11.15.3 Names in output files . . . . .	481
11.15.4 Units . . . . .	482
11.15.5 Directives . . . . .	482
11.16 Munster tutorial . . . . .	483
11.16.1 Alanine dipeptide: our toy model . . . . .	484
11.16.2 Monitoring collective variables . . . . .	484
11.16.2.1 Analyze on the fly . . . . .	485
11.16.2.2 Analyze using the driver . . . . .	486
11.16.2.3 Periodic boundaries and explicit water . . . . .	487
11.16.2.4 Other analysis tools . . . . .	488
11.16.3 Biasing collective variables . . . . .	488
11.16.3.1 Metadynamics . . . . .	488
11.16.3.2 Restraints . . . . .	493
11.16.3.3 Moving restraints . . . . .	496
11.16.3.4 Using multiple replicas . . . . .	496
11.16.3.5 Using multiple restraints with replica exchange . . . . .	497
<b>12 Index of Actions</b>	<b>501</b>
12.1 Full list of actions . . . . .	501
<b>13 Bug List</b>	<b>509</b>
<b>Bibliography</b>	<b>514</b>



# Chapter 1

## Introduction

PLUMED is a plugin that works with a large number of molecular dynamics codes ([Codes interfaced with PLUMED](#)). It can be used to analyse features of the dynamics on-the-fly or to perform a wide variety of free energy methods. PLUMED can also work as a [Command Line Tools](#) to perform analysis on trajectories saved in most of the existing formats. If PLUMED is useful for your work please read and cite [\[1\]](#), if you are interested in the PLUMED 1 original publication please read and cite [\[2\]](#).

To follow the development of PLUMED 2, you can look at the detailed [Change Log](#).

To install PLUMED, see this page: [Installation](#), while in [Getting started](#) you can find a brief introduction on how to write your first PLUMED input file.

[Tutorials](#) are available to introduce basic as well as more advanced features of PLUMED.

### 1.1 About this manual

This manual has been compiled from PLUMED version **2.3.8** (git version: **90a5f73**). Manual built on Travis CI for branch v2.3.

Regtest results for this version can be found [here](#).

This is the user manual - if you want to modify PLUMED or to understand how it works internally, have a look at the [developer manual](#).

### 1.2 Codes interfaced with PLUMED

PLUMED can be incorporated into an MD code and used to analyse or bias a molecular dynamics run on the fly. Some MD code could already include calls to the PLUMED library and be PLUMED-ready in its original distribution. As far as we know, the following MD codes can be used with PLUMED out of the box:

- [AmberTools](#), sander module, since version 15.
- [CP2K](#), since Feb 2015.
- [ESPResSo](#), in a Plumedized version that can be found [here](#).
- [PINY-MD](#), in its plumed branch.

- [IPHIGENIE](#).
- [AceMD](#), see [this link](#).
- [OpenMM](#), using the [openmmp-plumed plugin](#).
- [DL\\_POLY4](#).
- [VNL-ATK](#), see [this link](#).
- [ABIN](#).

Please refer to the documentation of the MD code to know how to use it with the latest PLUMED release. If you maintain another MD code that is PLUMED-ready let us know and we will add it to this list.

Additionally, we provide patching procedures for the following codes:

- [amber14](#)
- [gromacs-2016-5](#)
- [gromacs-4-5-7](#)
- [gromacs-5-0-7](#)
- [gromacs-5-1-4](#)
- [lammmps-6Apr13](#)
- [namd-2-8](#)
- [namd-2-9](#)
- [qespresso-5-0-2](#)

Alternatively, one can use PLUMED as a [Command Line Tools](#) for postprocessing the results from molecular dynamics or enhanced sampling calculations. Notice that PLUMED can be used as an analysis tool also from the following packages:

- [PLUMED-GUI](#) is a [VMD](#) plugin that computes PLUMED collective variables.
- [HTMD](#) can use PLUMED collective variables for analysis.

## Chapter 2

# Change Log

Here you can find a history of changes across different PLUMED versions. The future releases are expected to follow more or less the pace of the old release. This means:

- Approximately once per year, after summer, a new release (2.X). These releases typically group together all the features that were contributed during the year.
- Approximately every three month, we announce a patch (e.g. 2.2.X). This typically contains bug fixes, and could occasionally contain a new feature.

A few months before each new release we provide a beta release. We typically maintain release branches until the fifth patch release (2.X.5), which should come out approximately 15 month after the original release (2.X). After that, branches are not supported anymore.

Notice that occasionally we publish patches on the mailing list. These patches are always included in the following release, but we encourage users that want to be up to date to follow the mailing list.

Below you can find change logs for all the published releases. We mostly add new features without breaking existing ones. However, some of the changes lead to incompatible behavior. In the Change Log we try to give as much visibility as possible to these changes to avoid surprises.

We also log changes that are relevant if you are developing the code. These change lists are however not complete, and if you want to put your hands in the code and maintain your own collective variables we suggest you to follow the development on github.

- Changes for [Version 2.0](#)
- Changes for [Version 2.1](#)
- Changes for [Version 2.2](#)
- Changes for [Version 2.3](#)

## 2.1 Version 2.0

### Version 2.0.0 (Sep 27, 2013)

Version 2.0 is a complete rewrite, so there is no way to write a complete set of difference with respect to plumed 1.3. Here is a possibly incomplete summary of the difference:

- The input is simpler, more flexible, and more error proof. Many checks are now performed and in this way common errors are avoided.
- The units are now the same for all MD codes. If you want to use a different unit than the default you set it in the input file.
- The analysis tools are now much more flexible. As an example of this it is now possible to write different collective variables with different frequencies.
- Many complex collective variables are considerably faster than they were in plumed1. In particular, all variables based on RMSD distances.
- Centers of mass can be used as if they were atoms. Hence, unlike plumed 1.3, you can use center of mass positions in ALL collective variables.
- The virial contribution is now computed and passed to the MD code. Plumed can thus now be used to perform biased NPT simulations.
- Variables can be dumped on different files, and are computed only when this is necessary.
- PLUMED is now compiled as a separate library. This simplifies the patching procedure, but might require some extra work to configure PLUMED properly. Since PLUMED can be loaded as a shared library, it is possible to setup everything such that PLUMED and MD codes can be updated independently from each other.

In addition, it is now much easier to contribute new functionality to the code because:

- There is a much simpler interface between plumed and the base MD codes. This makes it much easier to add plumed to a new MD code. Hopefully, in the future, interfaces with MD codes will be maintained by the developers of the MD codes independently from PLUMED developers. This will allow more MD codes to be compatible with PLUMED.
- There is C++ object oriented programming and full compatibility with the C++ standard library
- A modular structure.
- New collective variables and methods can be released independently.
- There is an extensive developer documentation.
- User documentation is provided together inside the implementation files.

Caveats:

- PLUMED 2 input file (plumed.dat) has a syntax which is not compatible with PLUMED 1. Transition should be easy, but cannot be done just using the new version with the old input file.
- PLUMED 2 is written in C++, thus requires a C++ compiler
- PLUMED 2 may not include all the features that were available in PLUMED 1.

A tutorial explaining how to move from PLUMED 1 to PLUMED 2 is available (see [Moving from PLUMED 1 to PLUMED 2](#)).

### Version 2.0.1 (Nov 14, 2013)

For users:

- Fixed a bug in [HISTOGRAM](#) with REWEIGHT\_BIAS. Reweighting was only done when also temperature-reweighting was enabled.
- Fixed a bug that was sometime crashing code with domain decomposition and non-dense simulation boxes (e.g. implicit solvent).
- Performance improvements for [GYRATION](#).
- Flush all files every 10000 steps by default, without need to use [FLUSH](#)
- Errors when writing input for [switchingfunction](#) are now properly recognized.
- Added message when [simplemd](#) is used on a non-existing file.
- Fixed `plumed mklib` such that it deletes the target shared library in case of compilation error.
- Several small fixes in documentation and log file.

For developers:

- Added possibility to setup replica exchange from MD codes in fortran (commands "GREX setMPIIntercomm" and "GREX setMPIIntracomm").
- `cmd("setStopFlag")` should now be called after PLUMED initialization.
- Several small fixes in documentation.

### Version 2.0.2 (Feb 11, 2014)

For users:

- Fixed bug with [METAD](#) with INTERVAL and replica exchange, including bias exchange. Now the bias is correctly computed outside the boundaries. Notice that this is different from what was done in PLUMED 1.3. Also notice that INTERVAL now works correctly with grids and splines.
- Fixed bug with [READ](#) and periodic variables.
- Fixed bug with [HISTOGRAM](#) (option USE\_ALL\_DATA was not working properly).
- Gromacs patch updated to 4.6.5.
- Gromacs patch for 4.6 has been modified to allow for better load balancing when using GPUs.
- Added option 'plumed info --long-version' and 'plumed info --git-version'.
- Added full reference (page/number) to published paper in doc and log.
- Fixed a bug in file backups (only affecting Windows version - thanks to T. Giorgino).
- Added possibility to search in the documentation.
- Several small fixes in documentation and log file.

For developers:

- Fixed makefile dependencies in some auxiliary files in src/lib (\*cmake and \*inc).
- Changed way modules are linked in src/. E.g. src/colvar/tools/ is not anymore a symlink to src/colvar but a real directory. (Notice that this introduces a regression: when using plumed as an external library some include files could not work - this only applies when plumed is installed; also notice that this is fixed in 2.0.3)
- Patch for gromacs 4.6 now also include original code so as to simplify its modification.
- Added option 'plumed patch --save-originals'.
- Fixed regtest regtest/secondarystructure/rt32 to avoid problems with NUMERICAL\_DERIVATIVES.
- Removed include graphs in the documentation (too large).
- Several small fixes in documentation.

### Version 2.0.3 (June 30, 2014)

For users:

- Now compiles on Blue Gene Q with IBM compilers.
- Fixed bug in [CENTER](#) where default WEIGHTS were missing.
- Fixed broken [CONTACTMAP](#) with SUM
- Fixed [DUMPATOMS](#) with gro file and more than 100k atoms.
- Added CMDIST in [CONTACTMAP](#) to emulate plumed1 CMAP.
- Several small fixes in documentation and log file.

For developers:

- Fixed cmd("getBias") to retrieve bias. It was not working with single precision codes and it was not converting units properly.
- Fixed a regression in 2.0.2 concerning include files from installed plumed (see commit 562d5ea9dfc3).
- Small fix in tools/Random.cpp that allows Random objects to be declared as static.
- Small fix in user-doc compilation, so that if plumed is not found the sourceme.sh file is sourced
- Fixed non-ansi syntax in a few points and a non-important memory leakage.
- Split cltools/Driver.cpp to make parallel compilation faster.

### Version 2.0.4 (Sep 15, 2014)

For users:

- Fixed a bug in [BIASVALUE](#) that could produce wrong acceptance with replica exchange simulations.
- Fixed a few innocuous memory leaks.
- Fixed reader for xyz files, that now correctly detects missing columns. Also a related regtest has been changed.
- Several small fixes in documentation and log file.

For developers:

- Renamed Value.cpp to BiasValue.cpp

## Version 2.0.5 (Dec 15, 2014)

### Warning

This branch is not maintained. Users are invited to upgrade to a newer version

For users:

- Fixed a bug in replica exchange with different Hamiltonians (either lamdba-dynamics or plumed XX-hrex branch) possibly occuring when using charge or mass dependent variables.
- Fixed a bug in analysis (e.g. [HISTOGRAM](#)) leading to wrong accumulation of statistics when running a replica exchange simulation.
- Fixed a bug in the calculation of derivatives in histograms. This should be harmless since people usually only consider the value in histograms and not the derivatives.
- Fixed an issue in Makefile that could results in problems when patching an MD code with `--shared` option (pointed out by Abhi Acharya). This fixes a regression introduced in 2.0.2.
- Small fixes in documentation.

For developers:

- Added warning when performing regtests using an instance of plumed from a different directory

## 2.2 Version 2.1

### Version 2.1.0 (Sep 15, 2014)

Version 2.1 contains several improvements with respect to 2.0. Users currently working with 2.0 should have a look at the section "Changes leading to incompatible behavior" below and might need tiny adjustments in their input files. In 2.1 we restored more features of 1.3 that were missing in 2.0, so users still working with 1.↔3 could opt for an upgrade. A tutorial explaining how to move from PLUMED 1 to PLUMED 2 is available (see [Moving from PLUMED 1 to PLUMED 2](#)).

Below you find a list of all the changes with respect to version 2.0. Notice that version 2.1 includes already all the fixes in branch 2.0 up to 2.0.4.

Changes from version 2.0 which are relevant for users:

- Changes leading to incompatible behavior:
  - [COORDINATION](#) now skips pairs of one atom with itself.
  - Labels of quantities calculated by [BIASVALUE](#) have changed from `label.bias.argname` to `label.argname_bias`, which is more consistent with steered MD
  - Labels of quantities calculated by [ABMD](#) have change from `label.min_argname` to `label.argname_min`, which is more consistent with steered MD
  - Labels of quantities calculated by [PIECEWISE](#) have change from `label.argnumber` to `label.argname↔_pfunc`, which is more consistent with steered MD
  - For multicolvars components calculated with `LESS_THAN` and `MORE_THAN` keywords are now labelled `lessthan` and `morethan`. This change is necessary as the underscore character now has a special usage in component names.

- In **CONTACTMAP** components are now labelled *label.contact- n*.
- The command **SPHERE** has been replaced by **UWALLS**.
- New configuration system based on autoconf (use `./configure` from root directory). Optional packages are detected at compile time and correctly enabled or disabled. An internal version of lapack and blas will be used if these libraries are not installed.
- New actions:
  - **SPRINT** topological collective variables.
  - **CH3SHIFTS** collective variable.
  - **POSITION** collective variable.
  - **FIT\_TO\_TEMPLATE**.
  - **COMMITTOR** analysis.
  - **LOCAL\_AVERAGE**.
  - **NLINKS**.
  - **DIHCOR**.
  - **NOE**.
  - **RDC**.
  - **CLASSICAL\_MDS**.
  - **XDISTANCES**.
  - **YDISTANCES**.
  - **ZDISTANCES**.
  - **DUMPMULTICOLVAR**.
  - Crystallization module, including **Q3**, **LOCAL\_Q3**, **Q4**, **Q6**, **LOCAL\_Q4**, **LOCAL\_Q6**, **MOLECULES**, **SIMPLECUBIC**, **TETRAHEDRAL** and **FCCUBIC**.
  - **ENSEMBLE** to perform Replica-Averaging on any collective variable.
- New features for existing actions:
  - **METAD** : **WALKERS\_MPI** flag (multiple walkers in a mpi-based multi-replica framework), **ACCELERATION** flag (calculate on the fly the Metadynamics acceleration factor), **TAU** option (alternative way to set Gaussian height in well-tempered metadynamics), **GRID\_SPACING** (alternative to **GRID\_BIN** to set grid spacing). Notice that now one can also omit **GRID\_BIN** and **GRID\_SPACING** when using fixed size Gaussian, and the grid spacing will be automatically set.
  - **DISTANCE** : added **SCALED\_COMPONENTS**
  - **COORDINATION** : if a single group is provided, it avoids permuted atom indexes and runs at twice the speed.
  - **DUMPATOMS** : **PRECISION** option to set number of digits in output file.
  - **GROUP** : **NDX\_FILE** and **NDX\_GROUP** options to import atom lists from ndx (gromacs) files.
  - In many multicolvars, **MIN** and **MAX** options can be used.
  - **HISTOGRAM** : **GRID\_SPACING** (alternative to **GRID\_BIN** to set grid spacing), **FREE-ENERGY** flags in addition to standard probability density, additional option for **KERNEL=DISCRETE** to accumulate standard histograms.
  - **sum\_hills** : added options `–spacing` (alternative to `–bin` to set grid spacing) and `–setmintozero` to translate the minimum of the output files to zero.
  - **CONTACTMAP** : parallelised and added weights.
- New features in MD patches (require repatch):
  - New patch for Gromacs 5.0
  - Gromacs 4.6.X patch updated to 4.6.7



- Gromacs 4.6.7 supports [COMMITTOR](#) analysis; can be now be used to perform energy minimization; now passes temperature to PLUMED (this allows temperature to be omitted in some actions, namely [METAD](#) and analysis actions).

Notice that if you use runtime binding it is not compulsory to repatch, and that all combinations should work correctly (new/old PLUMED with repatched/non-repatched MD code).

- Other new features:
  - [driver](#) can now read trajectories in many formats using VMD molfile plugin (requires VMD plugins to be compiled and installed). In case VMD plugins are not installed, the configuration system falls back to an internal version which implements a minimal list of plugins (gromacs and dcd) (kindly provided by T. Giorgino).
  - [switchingfunction](#) : added STRETCH flag.
  - Negative strides in atom ranges (e.g. ATOMS=10-1:-3 is expanded to ATOMS=10,7,4,1).
  - [COORDINATION](#) and [DHENERGY](#) with NLIST now work correctly in replica exchange simulations.
  - Multicolvars with neighbor lists now work correctly in replica exchange simulations.
  - Improved multicolvar neighbor lists.
- Optimizations:
  - Root-mean-square deviations with align weights different from displace weights are now considerably faster. This will affect [RMSD](#) calculations plus other variables based on RMSD.
  - [WHOLEMOLECULES](#) is slightly faster.
  - [COORDINATION](#) is slightly faster when NN and MM are even and D\_0=0.
  - Atom scattering with domain decomposition is slightly faster.
  - Link cells are now exploited in some multicolvars.
  - Derivatives are not calculated unless they are specifically required, because for instance you are adding a bias.
- Documentation:
  - All tutorial material from the recent plumed meeting in Belfast is now in the manual
  - Improvements to documentation, including lists of referenceable quantities outputted by each action
  - Manual has been re-organized following suggestions received at the plumed meeting.
  - An experimental PDF version of the manual is now provided (a link can be found in the documentation homepage).

Changes from version 2.0 which are relevant for developers:

- Added regtests for plumed as a library (e.g. basic/rt-make-0). plumed command has an additional flag (`--is-installed`) to probe if running from a compilation directory or from a fully installed copy (this is needed for regtests to work properly).
- Improved class Communicator. Many operations can now be done directly on Vectors, Tensors, `std::vector` and `PLMD::Matrix`.
- Modified class RMSD.
- Patches for GPL codes (QuantumEspresso and Gromacs) now also include original code so as to simplify their modification.
- Fixed dependencies among actions such that it is now possible (and reliable) to use MPI calls inside `Action::prepare()`
- `colvar/CoordinationBase.cpp` has been changed to make it faster. If you devised a class which inherits from here, consider that `CoordinationBase::pairing` now needs *squared* distance instead of distance

- It is possible to run "make install" from subdirectories (e.g. from src/colvar)
- There is a small script which disables/enables all optional modules (make mod-light/mod-heavy/mod-reset)
- Added "-q" option to plumed patch
- You can now create new metrics to measure distances from a reference configurations. If you do so such metrics can then be used in paths straightforwardly
- You can now use multicolvars in tandem with manyrestraints in order to add a large numbers of restraints.
- Can now do multicolvar like things in which each colvar is a vector rather than a scalar.
- Updated script that generated header files so that they properly show years. Notice that the script should now be run from within a git repository

This list is likely incompleted, if you are developing in PLUMED you are encouraged to follow changes on github.

### Version 2.1.1 (Dec 15, 2014)

This release includes all the fixes available in branch 2.0 until 2.0.5.

For users:

- New patch for AMBER 14 (sander module only). This patch should be compatible with any PLUMED 2 version (including 2.0). It includes most PLUMED features with the notable exception of multi-replica framework.
- Changed definition in arbitrary phase of eigenvectors. This will change the result of some analysis method where the phase does matter (e.g. [CLASSICAL\\_MDS](#)) and make some regression test better reproducible.
- Fixed a portability issue in BG/P where gettimeofday is not implemented. Notice that this fix implies that one should execute again ./configure to have plumed timing working correctly.
- CS2Backbone: fixed a bug that resulted in only a fraction of the chemical shifts being printed with WRITE\_CS and parallel simulations (requires to get the last almost updated from SVN)
- NOE: fixed a bug in the replica-averaging
- Fixed a linking issue with ALMOST, where bz2 was always used to link ALMOST to PLUMED even if it is not compulsory to build ALMOST.
- Fixed a wrong include in the GMX5 patch.
- [FUNCPATHMSD](#) can now be used together with [CONTACTMAP](#) to define pathways in contactmaps space
- Configuration is more verbose, a warning is given if a default option cannot be enabled and an error is given if an option explicitly enabled cannot be enabled.
- Compilation is less verbose (use "make VERBOSE=1" to have old behavior)
- Small fixes in documentation.

For developers:

- Tests are now performed at every single push on travis-ci.org
- Manual is built and pushed to the online server from travis-ci.org (see developer doc)
- Fixes in developer doc.

### Version 2.1.2 (Mar 16, 2015)

For users:

- Added two new short tutorials to the manual ([Cambridge tutorial](#) and [Munster tutorial](#)).
- Fixed a severe bug on [DRMSD](#) - cutoff values were ignored by PLUMED. Notice that this bug was introduced in 2.1.0, so that it should not affect the 2.0.x series.
- Fixed a bug affecting LAMMPS patch used with a single processor. Notice that the fix is inside PLUMED, thus it does not necessarily requires repatching.
- Sander patch now works with multiple replica (no replica exchange yet). It also contains some fix from J. Swails.
- GMX5 patch was not working for bias-exchange like cases
- Patching system now checks for the availability of shared/static/runtime version of plumed before patching
- Configure now check better if compiler flag are accepted by the compiler. This makes configure on bluegene more robust.
- Sourceme.sh now sets proper library path in linux also.

### Version 2.1.3 (June 30, 2015)

For users:

- Fixed bug in [ENSEMBLE](#) derivatives when more than 1 argument was provided
- Fixed bug in [GHOST](#) : virial is now computed correctly.
- Fixed a serious bug in virial communicated from plumed to gromacs, for both gromacs versions 4.6 and 5.↔ 0. See [#132](#). This fix requires gromacs to be repatched and could be very important if you run biased simulations in the NPT ensemble.
- Fixed a bug in the virial computed with [FIT\\_TO\\_TEMPLATE](#) when the reference pdb had center non located at the origin.
- Fixed a bug in the the forces computed with [FIT\\_TO\\_TEMPLATE](#) when used in combination with [COM](#), [CENTER](#), or [GHOST](#)
- Fixed a bug that could lead plumed to be stuck with domain decomposition in some extreme case (one domain with all atoms, other domains empty).
- Fixed a bug when [COMBINE](#) or [MATHEVAL](#) are used with PERIODIC keyword. Now when PERIODIC keyword is used the result of the calculation is brought within the periodicity domain. See [#139](#).
- Fixed a bug related to [RANDOM\\_EXCHANGES](#) followed by [INCLUDE](#)
- Fixed bug in derivatives of histogram bead with triangular kernels
- Updated gromacs patch 4.5.5 to 4.5.7
- Updated internal molfile plugins to VMD 1.9.2.
- Included crd and crdbox formats to internal molfile.
- Added `-natoms` to [driver](#) . This is required to read coordinate files with VMD plugins when number of atoms is not present (e.g. amber crd files)
- Added the checks in the driver to detect cases where molinfo does not provide box information (e.g. pdb).

- Added support for `readdir_r` when available, which makes opening files thread safe.
- CFLAGS now include `-fPIC` by default
- Added a warning when using [METAD](#) without grids with a large number of hills.
- Fixes in user documentation.

For developers:

- Allow external VMD plugins to be detected with `--has-external-molfile`. This is required to enable some regtest with amber files.
- Added `--dump-full-virial` to [driver](#)
- Allow definition of variables where some of the components have derivatives and some haven't ([#131](#)).
- Improved travis tests with more debug options.
- Improved some regtest to check out-of-diagonal virial components
- Improved make `cppcheck` options.
- Fixes in developer documentation.

### Version 2.1.4 (Oct 13, 2015)

For users:

- Fixed NAMD patch. Masses and charges were not passed correctly, thus resulting in wrong [COM](#) or [CENTER](#) with MASS. This fix required repatching NAMD. Notice that this bug was present also in v2.0 but in a different form. More information here ([#162](#)), including a workaround that allows masses to be fixed without repatching.
- When installing with `PLUMED_LIBSUFFIX` an underscore is used as separator instead of a dash. E.g. `make install PLUMED_LIBSUFFIX=2.1` will result in an executable named `plumed_v2.1`. This fix a potential problem (see [Installation](#)).
- Fixed erroneously reported message about MPI at the end of `./configure`.
- Changed warning message about undocumented components.
- PLUMED now says in the log file if it was compiled from a dirty git repository.
- Fixed a problem leading to rare random crashes when using [METAD](#) with `WALKERS_MPI` and multiple processors per replica.
- Small change in numerical accuracy of lattice reduction. Should be more robust when running with highly optimizing compilers.
- Fixed a bug in normalisation of kernel functions. This affects [HISTOGRAM](#) if these actions were used with previous versions of the code care should be taken when analysing the results.
- Fixed a bug in derivatives of kernel functions with non-diagonal covariances. This affects the derivatives output by [sum\\_hills](#)

### Version 2.1.5 (Jan 18, 2016)

#### Warning

This branch is not maintained. Users are invited to upgrade to a newer version

For users:

- PLUMED now reports an error when using [HISTOGRAM](#) with FREE-ENERGY without USE\_ALL\_DATA. See [#175](#)
- Fixed a bug in configure together with `--enable-almost`. The check for libz2 library was not working properly.

## 2.3 Version 2.2

### Version 2.2 (Oct 13, 2015)

Version 2.2 contains several improvements with respect to 2.1. Users currently working with 2.1 should have a look at the section "Changes leading to incompatible behavior" below and might need tiny adjustments in their input files. In 2.2 we restored more features of 1.3 that were missing in 2.1, so users still working with 1.↔3 could opt for an upgrade. A tutorial explaining how to move from PLUMED 1 to PLUMED 2 is available (see [Moving from PLUMED 1 to PLUMED 2](#)).

Below you find a list of all the changes with respect to version 2.1. Notice that version 2.2 includes already all the fixes in branch 2.1 up to 2.1.4 indicated in [Version 2.1](#).

Changes from version 2.1 which are relevant for users:

- Changes leading to incompatible behavior:
  - Labels of quantities calculates by [SPRINT](#) have changed from *label.coord\_num* to *label.coord-num*
  - [METAD](#) with WALKERS\_MPI now writes a single hills file, without suffixes
  - removed the `./configure.sh` script of v2.0.x, now plumed can only be configured using autotools (`./configure`)
  - [COM](#), [CENTER](#), and [GYRATION](#) now automatically make molecules whole. In case you do not want them to do it, use NOPBC flag, which recovers plumed 2.1 behavior
  - Some MD code could now automatically trigger restart (e.g. gromacs when starting from cpt files). This can be overwritten using [RESTART](#) NO.
  - Replica suffixes are now added by PLUMED *before* extension (e.g. use `plumed.0.dat` instead of `plumed.dat.0`)
  - When using [switchingfunction](#) the STRETCH keyword is now implicit. NOSTRETCH is available to enforce the old behavior.
- Module activation can now be controlled during configure with `--enable-modules` option.
- Almost complete refactoring of installation procedure. Now DESTDIR and other standard autoconf directories (e.g. bindir) are completely supported. Additionally, everything should work properly also when directory names include spaces ([#157](#)). Finally, compiler is not invoked on install unless path are explicitly changed ([#107](#)).
- Related to installation refactoring, upon install a previously installed plumed is not removed. This is to avoid data loss if prefix variable is not properly set

- Several changes have been made in the Makefile.conf that makes it not compatible with those packaged with plumed 2.0/2.1. Please use ./configure to generate a new configuration file.
- Added partial OpenMP parallelization, see [OpenMP](#)
- Added multiple time step integration for bias potentials, see [Multiple time stepping](#)
- Link cells are now used in all multicolvars that involve [switchingfunction](#). The link cell cutoff is set equal to  $2 * d_{\text{max}}$ . Where  $d_{\text{max}}$  is the (user-specified) point at which the switching function goes to zero. Users should always set this parameter when using a switching function in order to achieve optimal performance.
- DHENERGY option is no longer possible within [DISTANCES](#). You can still calculate the DHENERGY colvar by using [DHENERGY](#)
- Reweighting in the manner described in [3] is now possible using a combination of the [METAD](#) and [HISTOGRAM](#) actions. The relevant keywords in [METAD](#) are REWEIGHTING\_NGRID and REWEIGHTING\_NHILLS. The  $c(t)$  and the appropriate weight to apply to the configurations are given by the values labeled rct and rbias.
- News in configure and install:
  - ./configure now allows external blas to be used with internal lapack. This is done automatically if only blas are available, and can be enforced with `--disable-external-lapack`.
  - ./configure supports `--program-prefix`, `--program-suffix`, and `--program-transform-name`.
  - make install supports DESTDIR and prefix.
  - Environment variables PLUMED\_LIBSUFFIX and PLUMED\_PREFIX are deprecated and will be removed in a later version.
- New actions
  - [DUMPMASSCHARGE](#) to dump a file with mass and charges during MD.
  - [EFFECTIVE\\_ENERGY\\_DRIFT](#) to check that plumed forces are not screwing the MD integrator.
  - [EXTENDED\\_LAGRANGIAN](#) : in combination with [METAD](#) it implements metadynamics with Extended Lagrangian; standalone it implements TAMD/dAFED.
  - [DFSCLUSTERING](#) calculate the size of clusters
  - [DUMPMULTICOLVAR](#) print out a multicolvar
  - [MFILTER\\_LESS](#) filter multicolvar by the value of the colvar
  - [MFILTER\\_MORE](#)
  - [MFILTER\\_BETWEEN](#)
  - [PCARMSD](#) pca collective variables using OPTIMAL rmsd measure
  - [PCAVARS](#) pca collective variables using any one of the measures in reference
  - [GRADIENT](#) can be used to calculate the gradient of a quantity. Used to drive nucleation
  - [CAVITY](#)
  - [PUCKERING](#) implemented for 5-membered rings (thanks to Alejandro Gil-Ley).
  - [WRAPAROUND](#) to fix periodic boundary conditions.
- New features for existing actions:
  - Keywords UPDATE\_FROM and UPDATE\_UNTIL to limit update step in a defined time window, available only for actions where it would be useful.
  - Keyword UNNORMALIZED for [HISTOGRAM](#).
  - Possibility to use Tiwary-Parrinello reweighting for [METAD](#)
  - Keywords for [GROUP](#) (REMOVE, SORT, UNIQUE) to allow more flexible editing of groups.
  - [DUMPATOMS](#) now supports dumping xtc and trr files (requires xdrfile library).
  - [driver](#) can now read xtc and trr files also with xdrfile library.

- **driver** accepts a `-mc` flag to read charges and masses from a file produced during molecular dynamics with **DUMPMASSCHARGE**
- Possibility to enable or disable **RESTART** on a per action basis, available only for actions where it would be useful.
- **MOLINFO** now supports many more special names for rna and dna (thanks to Alejandro Gil-Ley).
- **VMEAN** and **VSUM** allow one to calculate the sum of a set of vectors calculated by **VectorMultiColvar**. Note these can also be used in tandem with **AROUND** or **MFILTER\_MORE** to calculate the average vector within a particular part of the cell or the average vector amongst those that have a magnitude greater than some tolerance
- New way of calculating the minimum value in multicolvars (**ALT\_MIN**). This is less susceptible to overflow for certain values of  $\beta$ .
  - New keywords for calculating the **LOWEST** and **HIGHEST** colvar calculated by a multicolvar
- Added components to **DIPOLE** ([#160](#)).
- Other changes:
  - File reader now supports dos newlines as well as files with no newline at the end.

For developers:

- In order to be able to use openmp parallelism within multicolvar, secondarystructure, manyrestraints and crystallisation we had to make some substantial changes to the code that underlies these routines that is contained within vesselbase. In particular we needed to get rid of the derivatives and buffer private variables in the class **ActionWithVessel**. As a consequence the derivatives calculated in the various **performTask** methods are stored in an object of type **MultiValue**. Within multicolvar this is contained within an object of type **Atom↔ValuePack**, which stores information on the atom indices. If you have implemented a new multicolvar it should be relatively straightforward to translate them so they can exploit this new version of the code. Look at what has been done to the other multicolvars in there for guidance. Sorry for any inconvenience caused.
- Changed the logic of several PLUMED `ifdef` macros so as to make them consistent. Now every feature based on external libraries is identified by a `__PLUMED_HAS_*` macro.

### Version 2.2.1 (Jan 18, 2016)

For users:

- **PBMETAD** implement the new Parallel Bias Metadynamics flavor of the Metadynamics sampling method.
- PLUMED now reports an error when using **HISTOGRAM** with **UNNORMALIZED** without **USE\_ALL\_DATA**. See [#175](#)
- Fixed a bug in `configure` together with `-enable-almost`. The check for `libz2` library was not working properly.
- Fixed a bug in install procedure that was introducing an error in linking with `CP2K`.
- Fixed a bug that sometimes was preventing the printing of a usefull error message.

For developers:

- **Vector** and **Tensor** now support direct output with `<<`.
- Added some missing `matmul` operation **Vector** and **Tensor**.
- `./configure` is automatically relaunched when changing `./configure` or `Makefile.conf`. This makes it more robust to switch between branches.

### Version 2.2.2 (Apr 13, 2016)

For users:

- [MOLINFO](#) for RNA accepts more residue names, see [#180](#).
- added two mpi barriers (one was missing in PBMetaD for multiple walkers) to help synchronise initialisation
- Fixed a bug in internal stopwatches that was making [DEBUG](#) logRequestedAtoms not working
- Some multicolvars (including [BRIDGE](#), [ANGLES](#), and [INPLANEDISTANCES](#)) now crashes if one asks for too many atoms, see [#185](#).
- Optimisations (activation of the dependencies, secondary structures, DRMSD)
- Fixed a performance regression with RMSD=OPTIMAL-FAST
- Fixed a bug in the normalization of kernel functions (relevant for [HISTOGRAM](#)).
- Fixed a regression introduced in v2.2 that was making [METAD](#) with non-MPI multiple walkers crash if reading frequently. See [#190](#)
- Updated patch for gromacs 5.x. Patches for gromacs 5.0 and 5.1 have been fixed so as to allow patching in runtime mode.
- Possibility to control manual generation (including pdf) from `./configure`. Pdf manual is now off by default. Notice that on travis CI it is still generated.

For developers:

- Fixed a bug in the interpretation of cmd strings. Namely, an erroneous string was not triggering an error. This is harmless for MD codes properly patched, but could have introduced problems in MD codes with typos in cmd strings.
- `./configure` is not automatically relaunched anymore when doing `make clean`.

### Version 2.2.3 (Jun 30, 2016)

For users:

- Updated patches for gromacs 5.1.x and 5.0.x to fix a problem when plumed was trying to write to an already closed gromacs log file.
- When looking for a value outside the GRID now the error include the name of the responsible collective variable
- Numerical check in LatticeReduction made less picky. This should solve some of the internal errors reported by `LatticeReduction.cpp` when using aggressive compilers.
- Files are now flushed at the correct step. Before this fix, they were flushed at the step before the requested one (e.g. with [FLUSH](#) STRIDE=100 at step 99, 199, etc).
- In [METAD](#), INTERVAL with periodic variables now report an error.
- [LOAD](#) now works also when plumed is installed with a suffix.
- Added `--md-root` option to `plumed patch` which allows it to be run from a directory different from the one where the md code is located.
- Wham script in [Munster tutorial](#) tutorial now writes weights in scientific notation.



For developers:

- `./configure` checks if dependencies can be generated. If not, they are disabled.
- Added `--disable-dependency-tracking` to `./configure`
- Added a make target `all_plus_doc` that builds both code and docs.
- Added possibility to set a default location for plumed library in runtime binding. If the plumed wrapped is compiled with `-D__PLUMED_DEFAULT_KERNEL=/path/libplumedKernel.so`, then if the env var `PLUMED_KERNEL` is undefined or empty PLUMED will look in the path at compile time.
- Tentative port files are now available at [this link](#). They can be used to install PLUMED using MacPorts.

### Version 2.2.4 (Dec 12, 2016)

For users:

- Fix a bug in [PBMETAD](#) when biasing periodic and not periodic collective variables at the same time
- GSL library is now treated by `./configure` in the same way as other libraries, that is `-lgsl` `-lgslcblas` are only added if necessary.
- Fix a bug in [METAD](#) when using INTERVAL and ADAPTIVE gaussians at the same time
- Updated gromacs patch for 5.1.x to 5.1.4
- Fix a performance regression in the calculate loop where derivatives and forces were set to zero even if an action was not active, this is relevant for postprocessing and for the on-the-fly analysis
- Torsion calculation has been made slightly faster and improved so as to provide correct derivatives even for special angles (e.g.  $+\pi/2$  and  $-\pi/2$ ).

For developers:

- Macports portfile is now tested on travis at every plumed push.

### Version 2.2.5 (Mar 31, 2017)

#### Warning

This branch is not maintained. Users are invited to upgrade to a newer version

For users:

- Fixed a problem with large step numbers in driver (see [#209](#)).
- Fixed a problem leading to crashes when using switching functions without cutoff with some compiler (see [#210](#)).
- Fixed a bug when using [FIT\\_TO\\_TEMPLATE](#) and domain decomposition (see [#214](#)).
- Added an automatic flush of HILLS files when using [METAD](#) with file-based multiple walkers.
- Root dir is logged to allow easier debugging of problems.

## 2.4 Version 2.3

### Version 2.3 (Dec 12, 2016)

Version 2.3 contains several improvements with respect to 2.2. Users currently working with 2.2 should have a look at the section "Changes leading to incompatible behavior" below and might need tiny adjustments in their input files.

Below you find a list of all the changes with respect to version 2.2. Notice that version 2.3 includes already all the fixes in branch 2.2 up to 2.2.3 indicated in [Version 2.2](#).

Changes from version 2.2 which are relevant for users:

- Changes leading to incompatible behavior:
  - [COMMITTOR](#) can now be used to define multiple basins, but the syntax has been changed
  - Syntax for [SPRINT](#) and [DFSCLUSTERING](#) has changed. We have separated the Actions that calculate the contact matrix from these actions. These actions thus now take a contact matrix as input. This means that we these actions can be used with contact matrices that measures whether or not a pair of atoms are hydrogen bonded. For more details on this see [Exploiting contact matrices](#). For clustering the output can now be passed to the actions [CLUSTER\\_PROPERTIES](#), [CLUSTER\\_DIAMETER](#), [CLUSTER\\_NATOMS](#), [OUTPUT\\_CLUSTER](#) and [CLUSTER\\_DISTRIBUTION](#). These provide various different kinds of information about the connected components found by clustering
  - In [driver](#) masses and charges are set by default to NaN. This makes it less likely to do mistakes trying to compute centers of mass or electrostatic-dependent variables when masses or charges were not set. To compute these variables from the driver you are now forced to use `--pdb` or `--mc`.
  - In rational switching functions, by default MM is twice NN. This is valid both in [switchingfunction](#) with expanded syntax and when specifying MM on e.g. [COORDINATION](#)
  - Patch script `plumed patch` now patches by default with `--shared`. This should make the procedure more robust (see [#186](#)).
  - Faster [GYRATION](#) but new default behavior is not mass weighted
  - When using [HISTOGRAM](#) you now output the accumulated grid using [DUMPGRID](#) or [DUMPCUBE](#) to get the free energy you use the method [CONVERT\\_TO\\_FES](#). These changes allow one to use grids calculated within PLUMED in a work flow of tasks similarly to the way that you can currently use Values.
  - The way that reweighting is performed is now different. There are three separate actions [REWEIGHT\\_BIAS](#), [REWEIGHT\\_TEMP](#) and [REWEIGHT\\_METAD](#). These actions calculate the quantities that were calculated using the keywords [REWEIGHT\\_BIAS](#) and [REWEIGHT\\_TEMP](#) that used to appear in the old [HISTOGRAM](#) method. Now those these methods can be used in any methods that calculate ensemble averages for example [HISTOGRAM](#) and [AVERAGE](#)
  - Manual is now build with locally compiled plumed
  - Removed CH3SHIFT
  - [CS2BACKBONE](#) is now native in PLUMED removing the need to link ALMOST, small syntax differences
  - [CS2BACKBONE](#), [NOE](#), [RDC](#), removed the keyword ENSEMBLE: now ensemble averages can only be calculated using [ENSEMBLE](#)
  - [RDC](#), syntax changes
  - It is not possible anymore to select modules using `modulename.on` and `modulename.off` files. Use `./configure --enable-modules` instead.
  - Removed IMD modules. In case someone is interested in restoring it, please contact the PLUMED developers.
- New actions:
  - [FIXEDATOM](#)
  - [HBOND\\_MATRIX](#)

- [CLUSTER\\_PROPERTIES](#)
- [CLUSTER\\_DIAMETER](#)
- [CLUSTER\\_NATOMS](#)
- [OUTPUT\\_CLUSTER](#)
- [CLUSTER\\_DISTRIBUTION](#)
- [ROWSUMS](#)
- [COLUMNSUMS](#)
- [UPDATE\\_IF](#)
- [DUMPGRID](#)
- [DUMPCUBE](#)
- [CONVERT\\_TO\\_FES](#)
- [INTERPOLATE\\_GRID](#)
- [FIND\\_CONTOUR](#)
- [FIND\\_SPHERICAL\\_CONTOUR](#)
- [FIND\\_CONTOUR\\_SURFACE](#)
- [AVERAGE](#)
- [REWEIGHT\\_BIAS](#)
- [REWEIGHT\\_TEMP](#)
- [REWEIGHT\\_METAD](#)
- [PCA](#)
- [PRE](#)
- [STATS](#)
- [METAINTERFERENCE](#)
- [LOCALENSEMBLE](#)
- [FRET](#)
- [RESET\\_CELL](#)
- [JCOUPLING](#)
- [ERMSD](#)

- New features in MD patches (require repatch):

- Patch for amber 14 now passes charges with appropriate units (fixes [#165](#)). Notice that the patch is still backward compatible with older PLUMED version, but the charges will only be passed when using PLUMED 2.3 or later.
- Patch for GROMACS 5.1 incorporates Hamiltonian replica exchange, see [Using Hamiltonian replica exchange with GROMACS](#)
- Gromacs 2016, 5.1.x, 5.0.x, flush the plumed output files upon checkpointing
- Added patch for Gromacs 2016.1
- gromacs 5.1.x patch updated to 5.1.4
- Removed the patch for Gromacs 4.6.x
- LAMMPS patch updated to support multiple walkers and report plumed bias to LAMMPS (thanks to Pablo Piaggi).

- New features for existing actions:

- The SPECIES and SPECIESA keyword in MultiColvars can now take a multicolvar as input. This allows one to calculate quantities such as the Q4 parameters for those atoms that have a coordination number greater than x.
- Added MATHEVAL type in [switchingfunction](#)
- Added Q type native contacts in [switchingfunction](#) (thanks to Jan Domanski).

- [COMMITTOR](#) can now be used to define multiple basins
- The number of atoms admitted in [BRIDGE](#) has been significantly increased, see [#185](#).
- [driver](#) now allows `–trajectory-stride` to be set to zero when reading with `–ixtc/–itrr`. In this case, step number is read from the trajectory file.
- [METAD](#) and [PBMETAD](#) can now be restarted from a GRID
- Added keywords TARGET and DAMPFACTOR in [METAD](#)
- When using [METAD](#) with file-based multiple walkers and parallel jobs (i.e. mpirun) extra suffix is not added (thanks to Marco De La Pierre).
- [ENSEMBLE](#) added keywords for weighted averages, and calculation of higher momenta
- [MOLINFO](#) now allows single atoms to be picked by name.
- [FIT\\_TO\\_TEMPLATE](#) now supports optimal alignment.
- [CONSTANT](#) added the possibility of storing more values as components with or without derivatives
- [PUCKERING](#) now supports 6 membered rings.
- Extended checkpoint infrastructure, now [METAD](#) and [PBMETAD](#) will write GRIDS also on checkpoint step (only the GROMACS patch is currently using the checkpointing interface)
- Other features:
  - Added a plumed-config command line tool. Can be used to inspect configuration also when cross compiling.
  - Added a `--mpi` option to `plumed`, symmetric to `--no-mpi`. Currently, it has no effect (MPI is initialized by default when available).
  - PLUMED now generate a VIM syntax file, see [Using VIM syntax file](#)
  - The backward cycle is now parallelised in MPI/OpenMP in case many collective variables are used.
  - GSL library is now searched by default during `./configure`.
  - Tutorials have been (partially) updated to reflect some of the changes in the syntax
  - Parser now reports errors when passing numbers that cannot be parsed instead of silently replacing their default value. See [#104](#).
  - More and more documentation
- Bug fixes:
- Fixed a bug in [PBMETAD](#) that was preventing the writing of GRIDS if a hill was not added in that same step

#### For developers:

- IMPORTANT: BIAS can now be BIASED as well, this changes can lead to some incompatibility: now the "bias" component is always defined automatically by the constructure as a `componentWithDerivatives`, derivatives are automaticcaly obtained by forces. The main change is that you don't have to define the bias component anymore in your constructor and that you can use `setBias(value)` to set the value of the bias component in calculate.
- Added new strings for plumed cmd: `setMDMassUnits`, `setMDChargeUnits`, `readInputLine`, `performCalcNo↵Update`, `update` and `doCheckPoint`.
- Easier to add actions with multiple arguments
- New functions to access local quantities in domain decomposition
- Active modules to enable regtests are chosen using `plumed config`.
- A script is available to check if source code complies plumed standard. Notice that this script is run together with `cppcheck` on `travis-ci`.
- Cppcheck on `travis-ci` has been updated to 1.75. Several small issues triggering errors on 1.75 were fixed (e.g. structures passed by value are now passed by const ref) and false positives marked as such.
- Added coverage scan.

### Version 2.3.1 (Mar 31, 2017)

- Fix to FIT\_TO\_TEMPLATE as in 2.2.5. Notice that in 2.3.0 also the case with TYPE=OPTIMAL was affected. This is fixed now.
- small change in CS2BACKBONE to symmetrise the ring current contribution with respect to ring rotations (also faster)
- fixed plumed-config that was not working.
- log file points to the config.txt files to allow users to check which features were available in that compiled version.
- make clean in root dir now also cleans vim subdirectory.
- Updated gromacs patch to version 2016.3

For developers:

- Cppcheck on travis-ci has been updated to 1.77.
- Doxygen on travis-ci has been updated to 1.8.13

### Version 2.3.2 (Jun 12, 2017)

See branch [v2.3](#) on git repository.

- Resolved problem with nan in SMAC with SPECIESA and SPECIESB involving molecules that are the same
- PDB reader is now able to read files with dos newlines (see [#223](#)).
- Fixed bug in CS2BACKBONE (v2.3.1) related to ring currents of HIS and TRP
- Fixed bug in if condition in PCAVARS so that you can run with only one eigenvector defined in input
- Fixed bug with timers in sum\_hills [#194](#).
- Fixed bug when using MOVINGRESTRAINT with periodic variables such as TORSION [#225](#).
- Fixed bug in HBOND\_MATRIX that used to appear when you used DONORS and ACCEPTORS with same numbers of atoms
- Fixed bug in DISTANCES that appears when using BETWEEN and link cells.
- Prevented users from causing segfaults by storing derivatives without LOWMEM flag. In these cases PLUMED crashes with meaningful errors.
- Fixed bug in HISTOGRAM that causes nans when using KERNEL=DISCRETE option
- Fixed a bug in the parser related to braces, see [#229](#)
- Fixed a bug that appeared when using Q3, Q4 and Q6 with LOWEST or HIGHEST flag
- Fixed a bug that appears when you use MFILTER\_LESS as input to COORDINATIONNUMBER with SPECIESA and SPECIESB flags
- Fixed a bug that was making flushing when gromacs checkpoints not functional (thanks to Summer Snow).
- Fixed a bug affecting EXTENDED\_LAGRANGIAN and METAD with ADAPT=DIFF when using an argument with periodicity (min,max) such that min is different from -max. This does not affect normal TORSION, but would affect PUCKERING component phi with 6-membered rings. In addition, it would affect any variable that is created by the user with a periodicity domain not symmetric around zero. See [#235](#) (thanks to Summer Snow for reporting this bug).

- Fixed numerical issue leading to simulations stuck (LatticeReduction problem) with intel compiler and large simulation cells.
- Fixed a bug affecting `LOCAL_AVERAGE` and outputting all multicolvars calculated by `Q6` with `DUMPMULTICOLVAR`
- `plumed info --user-doc` and `plumed info --developer-doc` now fall back to online manual when local doc is not installed, see [#240](#).

For developers:

- **IMPORTANT:** we started to enforce code formatting using `astyle`. Check the developer documentation to learn how to take care of not-yet-formatted branches.
- `plumedcheck` validation has been made stricter. All the checks are now described in the developer manual.
- New flag `--disable-libsearch` for `configure`, allowing an easier control of linked libraries when installing PLUMED with a package manager such as MacPorts.
- Added `--disable-static-patch` to `./configure` to disable tests related to static patching. It can be used when static patching is not needed to make sure a wrong c++ library is not linked by mistake.
- Using `install_name_tool` to fix the name of the installed library on OSX. Allows linking the PLUMED shared library without explicitly setting `DYLD_LIBRARY_PATH`.
- Added environment variable `PLUMED_ASYNC_SHARE` to enforce synchronous/asynchronous atom sharing (mostly for debug purpose).
- On travis-ci, using `ccache` to speedup builds.
- On travis-ci, added a regtest using Docker with gcc6 and MPI.
- On travis-ci, docs for unofficial or unsupported branches are set not to be indexed by search engines (see [#239](#))
- Cppcheck on travis-ci has been updated to 1.79.

### Version 2.3.3 (Oct 3, 2017)

For users:

- Fixed a bug in `switchingfunction` `MATHEVAL`, leading to inconsistent results when using OpenMP with multiple threads (see [#249](#)).
- `FIT_TO_TEMPLATE` now reports when it is used with a reference file with zero weights.
- Fixed logging of `UNITS` (thanks to Omar Valsson).
- Fixed a possible bug with `EFFECTIVE_ENERGY_DRIFT` and domain decomposition with a domain containing zero atoms.

For developers:

- Fixed a bug in `./configure --disable-libsearch` when searching for molfile plugins.
- Cppcheck on travis-ci has been updated to 1.80.
- Configure script now has a list of better alternatives to find a working `ld -r -o` tool to merge object files. This solves linking issues on some peculiar systems (see [#291](#), thanks to Massimiliano Culpo).
- Using `install_name_tool` also on non-installed libraries. This makes it possible to link them and later find them without explicitly setting `DYLD_LIBRARY_PATH`. This should also make the `DYLD_LIBRARY_PATH` irrelevant. Notice that `DYLD_LIBRARY_PATH` is not well behaved in OSX El Capitan.

### Version 2.3.4 (Dec 15, 2017)

For users:

- GROMACS patch updated to gromacs-2016.4. This patch was also fixed in order to properly work with [ENERGY](#) (see [#316](#)) and to implement `-hrex` option (see [#197](#)).
- Patch for GROMACS 5.1.4 updated to fix an error with [ENERGY](#) (see [#316](#)).
- Solved a bug in [ERMSD](#) leading to incorrect results when using non-default length units (e.g. with `UNITS LENGTH=A`).

For developers:

- Regtest script also reports when exitcode different from zero is returned.
- Patch script reports errors returning a nonzero exit code.
- cppcheck update to 1.81
- Solved small bug in stored `PLUMED_ROOT` directory as obtained from statically patched MD codes. Namely, the compilation directory was stored rather than the installation one.

### Version 2.3.5 (Mar 2, 2018)

For users:

- Fixed `plumed partial_tempering` to agree with GROMACS conventions for the choice of dihedrals (see [#337](#)). Should be irrelevant for the vast majority of cases.
- Fixed small bug in regexp parser - the part outside the parentheses was just ignored.

For developers:

- Doxygen on travis-ci has been updated to 1.8.14.
- Embedded astyle updated to 3.1.
- `make clean` now correctly removes the `src/lib/plumed` executable.

### Version 2.3.6 (Jul 2, 2018)

For users:

- Fixed a problem leading to NaN derivatives of [switchingfunction](#)  $\phi$  when distance between two atoms is large.
- GROMACS patch updated to gromacs-2016.5.
- `./configure` crashes if prefix is set to present working directory (notice that this choice was already leading to issues).
- [DUMPATOMS](#) reports an error when trying to write xtc/xdr files without the xdrfile library installed.

- Fixed a bug appearing when using `PATH` or `GPROPERTYPATH` with virtual atoms without simultaneously using the same atoms in a different action.
- Fixed incorrect format of the pdb file written by `PCA` (see #363).
- Fixed behavior of natural units. When an MD code asks for natural units, it is not necessary to also set units within PLUMED using `UNITS` (see #364).

For developers:

- Fixed small issue in debug options of `driver` (see #245).
- `plumed patch -e` now accepts a name closely matching the patch name (e.g. `plumed patch -e gromacs2016.5` will try to patch even if the stored patch is for `gromacs-2016.4`). This simplifies managing Portfiles. Nothing changes when picking the patch from the interactive menu.
- Install newer ccache on travis-ci, build faster.
- Small fix in provided env modules (`PLUMED_VIMPATH` is set also when shared libraries are disabled).

### Version 2.3.7 (Oct 5, 2018)

For users:

- Fixed flag `DETAILED_TIMERS` in `DEBUG` (flag was ignored and detailed timers always written).
- Small fix in `DUMPMASSCHARGE` (atoms are now correctly requested only at first step).

### Version 2.3.8 (Dec 19, 2018)

#### Warning

This branch is not maintained. Users are invited to upgrade to a newer version

For users:

- Fixed some openMP regression (some related to the whole codes and some specifics for Coordination and Multicolvar), this were compiler dependent so not all users may have experienced them
- Fixed an issue with `CS2BACKBONE` when more than 2 chains were used
- Fixed memory leak in `RDC`.
- Fixed segmentation fault with more than two CVs in reweighting `METAD` (see #399, thanks to fiskissimo).

For developers:

- Small fix in `LD_FLAGS` when enabling coverage.
- Fixed order of flags in tests for static linking done by configure (see #407).
- Fixed the way paths are hardcoded so as to facilitate conda packaging (see #416).



## Chapter 3

# Installation

In this page you can learn how to [configure](#), [compile](#), and [install](#) PLUMED. For those of you who are impatient, the following might do the job:

```
> ./configure --prefix=/usr/local
> make -j 4
> make doc # this is optional and requires proper doxygen version installed
> make install
```

Notice that `make install` is not strictly necessary as `plumed` can be used from the compilation directory. This is very useful so as to quickly test the implementation of new features. However, we strongly recommend to perform a full install.

Once the above is completed the `plumed` executable should be in your execution path and you will be able to use PLUMED to analyze existing trajectories or play with the Lennard-Jones code that is included. However, because PLUMED is mostly used to bias on the fly simulations performed with serious molecular dynamics packages, you can find instructions about how to [patch](#) your favorite MD code so that it can be combined with PLUMED below. Again, if you are impatient, something like this will do the job:

```
> cd /md/root/dir
> plumed patch -p
```

Then compile your MD code. For some MD codes these instructions are insufficient. It is thus recommended that you read the instructions at the end of this page. Notice that MD codes could in principle be "PLUMED ready" in their official distribution. If your favorite MD code is available "PLUMED ready" you will have to compile PLUMED first, then (optionally) install it, then check the MD codes' manual to discover how to link it.

### 3.1 Configuring PLUMED

The `./configure` command just generates a `Makefile.conf` file and a `sourceme.sh` file. In PLUMED 2.0 these files were pre-prepared and stored in the directory `configurations/`. The new ones generated by `./configure` are similar to the old ones but are not completely compatible. In particular, some of the `-D` options have been changed in version 2.2, and several new variables so as to specify the installation directories have been added. For this reason, you now should run `./configure` again. Anyway, it should be easy to enforce a similar setup with `autoconf` by passing the proper arguments on the command line. If you have problems on your architecture, please report them to the mailing list.

Useful command line options for `./configure` can be found by typing

```
> ./configure --help
```

PLUMED is made up of modules. Some of them are on by default, some others aren't. Since version 2.3, the activation of modules should be made during configuration using the `--enable-modules` option (see [List of modules](#)).

Notice that some functionalities of PLUMED depend on external libraries which are looked for by configure. You can typically avoid looking for a library using the "disable" syntax, e.g.

```
> ./configure --disable-mpi --disable-matheval
```

Notice that when mpi search is enabled (by default) compilers such as "mpic++" and "mpicxx" are searched for first. On the other hand, if mpi search is disabled ("`./configure --disable-mpi`") non-mpi compilers are searched for. Notice that only a few of the possible compiler name are searched. Thus, compilers such as "g++-mp-4.8" should be explicitly requested with the CXX option.

You can better control which compiler is used by setting the variables CXX and CC. E.g., to use Intel compilers use the following command:

```
> ./configure CXX=icpc CC=icc
```

Notice that we are using icpc in this example, which is not an mpi compiler as a result mpi will not be enabled. Also consider that this is different with respect to what some other configure script does in that variables such as MPICXX are completely ignored here. In case you work on a machine where CXX is set to a serial compiler and MPICXX to a MPI compiler, to compile with MPI you should use

```
> ./configure CXX="$MPICXX"
```

### Warning

This procedure could be somehow confusing since many other programs behave in a different way. The flag `--enable-mpi` is perfectly valid but is not needed here. Autoconf will check if a code containing MPI calls can be compiled, and if so it will enable it. `--disable-mpi` could be used if you are using a compiler that supports MPI but you don't want PLUMED to be compiled with MPI support. Thus the correct way to enable MPI is to pass to `./configure` the name of a C++ compiler that implements MPI using the CXX option. In this way, MPI library is treated similarly to all the other libraries that PLUMED tries to link by default.

To tune the compilation options you can use the CXXFLAGS variable:

```
> ./configure CXXFLAGS=-O3
```

If you are implementing new functionality and want to build with debug flags in place so as to do some checking you can use

```
> ./configure --enable-debug
```

This will perform some extra check during execution (possibly slowing down PLUMED) and write full symbol tables in the executable (making the final executable much larger).

The main goal of the automatic configure is to find the libraries. When they are stored in unconventional places it is thus sensible to tell autoconf where to look! To do this there are some environment variable that can be used to instruct the linker which directories it should search for libraries inside. These variables are compiler dependent, but could have been set by the system administrator so that libraries are found without any extra flag. Our suggested procedure is to first try to configure without any additional flags and to then check the log so as to see whether or not the libraries were properly detected.

If a library is not found during configuration, you can try to use options to modify the search path. For example if your matheval libraries is in `/opt/local` (this is where MacPorts put it) and configure is not able to find it you can try

```
> ./configure LDFLAGS=-L/opt/local/lib CPPFLAGS=-I/opt/local/include
```

Notice that PLUMED will first try to link a routine from say `matheval` without any additional flag, and then in case of failure will retry adding `-lmatheval` to the `LIBS` options. If also this does not work, the `matheval` library will be disabled and some features will not be available. This procedure allows you to use libraries with custom names. So, if your `matheval` library is called `/opt/local/lib/libmymatheval.so` you can link it with

```
> ./configure LDFLAGS=-L/opt/local/lib CPPFLAGS=-I/opt/local/include LIBS=-lmymatheval
```

In this example, the linker will directly try to link `/opt/local/lib/libmymatheval.so`. This rule is true for all the libraries, so that you will always be able to link a specific version of a library by specifying it using the `LIBS` variable.

Since version 2.3.2, the search for the library functions passing to the linker a flag with the standard library name (in the `matheval` example, it would be `-lmatheval`) can be skipped by using the option `--disable-libsearch`. Notice that in this manner only libraries that are explicitly passed using the `LIBS` option will be linked. For instance

```
> ./configure --disable-libsearch LIBS=-lmatheval
```

will make sure that only `matheval` is linked and, for instance, `blas` and `lapack` libraries are not. This might be useful when installing PLUMED within package managers such as MacPorts to make sure that only desired libraries are linked and thus to avoid to introduce spurious dependencies. The only exception to this rule is `-ldl`, which is anyway a system library on Linux.

#### Warning

On Linux you might have problems using the `LDFLAGS` option. In particular, if `makefile` complains that it cannot link the file `'src/lib/plumed-shared'`, try to set correctly the runtime path by using

```
> ./configure LDFLAGS="-L/opt/local/lib -Wl,-rpath,/opt/local/lib" \
  CPPFLAGS=-I/opt/local/include LIBS=-lmymatheval
```

Notice that although the file `'src/lib/plumed-shared'` is not necessary, being able to produce it means that it will be possible to link PLUMED dynamically with MD codes later.

PLUMED needs `blas` and `lapack`. These are treated slightly different from other libraries. The search is done in the usual way (i.e., first look for them without any link flag, then add `-lblas` and `-llapack`, respectively). As such if you want to use a specific version of `blas` or `lapack` you can make them available to configure by using

```
> ./configure LDFLAGS=-L/path/to/blas/lib LIBS=-lnameoflib
```

If the functions of these libraries are not found, the compiler looks for a version with a final underscore added. Finally, since `blas` and `lapack` are compulsory in PLUMED, you can use a internal version of these libraries that comes as part of PLUMED. If all else fails the internal version of BLAS and LAPACK are the ones that will be used by PLUMED. If you wish to disable any search for external libraries (e.g. because the system libraries have problems) this can be done with

```
> ./configure --disable-external-blas
```

Notice that you can also disable external `lapack` only, that is use internal `lapack` with external `blas` using

```
> ./configure --disable-external-lapack
```

Since typically it is the `blas` library that can be heavily optimized, this configuration should not provide significant slowing down and could be used on systems where native `lapack` libraries have problems.

As a final resort, you can also edit the resulting `Makefile.conf` file. Notable variables in this file include:

- `DYNAMIC_LIB` : these are the libraries needed to compile the PLUMED library (e.g. `-L/path/to/matheval -lmatheval` etc). Notice that for the PLUMED shared library to be compiled properly these should be dynamic libraries. Also notice that PLUMED preferentially requires BLAS and LAPACK library; see [BLAS and LAPACK](#) for further info. Notice that the variables that you supply with `configure LIBS=something` will end up in this variable. This is a bit misleading but is required to keep the configuration files compatible with PLUMED 2.0.
- `LIBS` : these are the libraries needed when patching an MD code; typically only `-ldl` (needed to have functions for dynamic loading).
- `CPPFLAGS` : add here definition needed to enable specific optional functions; e.g. use `-D__PLUMED_HAS_MATHEVAL` to enable the matheval library
- `SOEXT` : this gives the extension for shared libraries in your system, typically `"so"` on unix, `"dylib"` on mac; If your system does not support dynamic libraries or, for some other reason, you would like only static executables you can just set this variable to a blank (`"SOEXT="`).

### 3.1.1 BLAS and LAPACK

We tried to keep PLUMED as independent as possible from external libraries and as such those features that require external libraries (e.g. Matheval) are optional. However, to have a properly working version of plumed PLUMED you need BLAS and LAPACK libraries. We would strongly recommend you download these libraries and install them separately so as to have the most efficient possible implementations of the functions contained within them. However, if you cannot install blas and lapack, you can use the internal ones. Since version 2.1, PLUMED uses a configure script to detect libraries. In case system LAPACK or BLAS are not found on your system, PLUMED will use the internal replacement.

We have had a number of emails (and have struggled ourselves) with ensuring that PLUMED can link BLAS and LAPACK. The following describes some of the pitfalls that you can fall into and a set of sensible steps by which you can check whether or not you have set up the configuration correctly.

Notice first of all that the `DYNAMIC_LIB` variable in the `Makefile.conf` should contain the flag necessary to load the BLAS and LAPACK libraries. Typically this will be `-llapack -lblas`, in some case followed by `-lgfortran`. Full path specification with `-L` may be necessary and on some machines the blas and lapack libraries may not be called `-llapack` and `-lblas`. Everything will depend on your system configuration.

Some simple to fix further problems include:

- If the linker complains and suggests recompiling lapack with `-fPIC`, it means that you have static lapack libraries. Either install dynamic lapack libraries or switch to static compilation of PLUMED by unsetting the `SOEXT` variable in the configuration file.
- If the linker complains about other missing functions (typically starting with `"for_"` prefix) then you should also link some Fortran libraries. PLUMED is written in C++ and often C++ linkers do not include Fortran libraries by default. These libraries are required for lapack and blas to work. Please check the documentation of your compiler.
- If the linker complains that `dsyevr_` cannot be found, try adding `-DF77_NO_UNDERSCORE` to `CPPFLAGS`. Notice that `"./configure"` should automatically try this solution.

### 3.1.2 VMD trajectory plugins

If you configure PLUMED with VMD's plugins you will be able to read many more trajectory formats. To this aim, you need to download the SOURCE of VMD, which contains a plugins directory. Adapt build.sh and compile it. At the end, you should get the molfile plugins compiled as a static library `libmolfile_plugin.a`. Locate said file and `libmolfile_plugin.h`, they should be in a directory called `/pathtovmdplugins/ARCH/molfile` (e.g. `/pathtovmdplugins/MACOSXX86_64/molfile`). Also locate file `molfile_plugin.h`, which should be in `/pathtovmdplugins/include`. Then customize the configure command with something along the lines of:

```
./configure LDFLAGS="-L/pathtovmdplugins/ARCH/molfile" CPPFLAGS="-I/pathtovmdplugins/include -I/pathtovmdplugins/ARCH/molfile"
```

Notice that it might be necessary to add to `LDFLAGS` the path to your TCL interpreter, e.g.

```
./configure LDFLAGS="-ltcl8.5 -L/mypathtotcl -L/pathtovmdplugins/ARCH/molfile" \
CPPFLAGS="-I/pathtovmdplugins/include -I/pathtovmdplugins/ARCH/molfile"
```

Then, rebuild plumed.

## 3.2 Compiling PLUMED

Once configured, PLUMED can be compiled using the following command:

```
> make -j 4
```

This will compile the entire code and produce a number of files in the 'src/lib' directory, including the executable 'src/lib/plumed'. When shared libraries are enabled, a shared libraries called 'src/lib/libKernel.so' should also be present. Notice that the extension could be '.dylib' on a Mac.

In case you want to run PLUMED *without installing it* (i.e. from the compilation directory), you can use the file 'sourceme.sh' that has been created by the configure script in the main PLUMED directory. This file can be "sourced" (presently only working for bash shell) if you want to use PLUMED *before installing it* (i.e. from the compilation directory). It is a good idea to source it now, so that you can play with the just compiled PLUMED:

```
> source sourceme.sh
```

Now a "plumed" executable should be in your path. Try to type

```
> plumed -h
```

### Warning

If you are cross compiling, the `plumed` executable will not work. As a consequence, you won't be able to run regtests or compile the manual. This is not a problem.

You can also check if PLUMED is correctly compiled by performing our regression tests. Be warned that some of them fail because of the different numerical accuracy on different machines.

```
> cd regtest
> make
```

Notice that regtests are performed using the "plumed" executable that is currently in the path. You can check the exact version they will use by using the command

```
> which plumed
```

This means that if you do not source "sourceme.sh", the tests will fail. This does not mean that plumed is not working it just means that you haven't told them shell where to find plumed! As a related note, you can easily run the test suite using a different version of PLUMED (maybe an earlier version that you already installed), just making sure that it can be found in the path. Clearly, if you test a given version of PLUMED with a test suite from a different version you can expect two possible kinds of innocuous errors:

- If `plumed` executable is older than the test suite, the tests might fail since they rely on some feature introduced in PLUMED in a newer version.
- If `plumed` executable is newer than the test suite, the tests might fail since some non-backward compatible change was made in PLUMED. We try to keep the number of non-backward compatible changes small, but as you can see in the ChangeLog there are typically a few of them at every new major release.

### Attention

Even though we regularly perform tests on [Travis-CI](#), it is possible that aggressive optimizations or even architecture dependent features trigger bugs that did not show up on travis. So please always perform regtests when you install PLUMED.

Notice that the compiled executable, which now sits in 'src/lib/plumed', relies on other resource files present in the compilation directory. This directory should thus stay in the correct place. One should thus not rename or delete it. In fact the path to the PLUMED root directory is hardcoded in the `plumed` executable as can be verified using

```
> plumed info --root
```

In case you try to use the `plumed` executable without the compilation directory in place (e.g. you move away the `src/lib/plumed` static executable and delete or rename the compilation directory) PLUMED will not work correctly and will give you an error message

```
> plumed help
ERROR: I cannot find /xxx/yyy/patches directory
```

You can force `plumed` to run anyway by using the option `--standalone-executable`:

```
> plumed --standalone-executable help
```

Many features will not be available if you run in this way. However, this is currently the only way to use the PLUMED static executable on Windows.

## 3.3 Installing PLUMED

It is strongly suggested to install PLUMED in a predefined location. This is done using

```
> make install
```

This will allow you to remove the original compilation directory, or to recompile a different PLUMED version in the same place.

To install PLUMED one should first decide the location. The standard way to do it is during the configure step:

```
> ./configure --prefix=$HOME/opt
> make
> make install
```

However, you can also change it after compilation setting the variable `prefix`.

```
> ./configure
> make
> make install prefix=$HOME/opt
```

If you didn't specify the `--prefix` option during configure, and you did not set the `prefix` variable when installing, PLUMED will be installed in `/usr/local`. The install command should be executed with root permissions (e.g. "sudo make install") if you want to install PLUMED on a system directory.

### Warning

Please **do not** set prefix to the current directory (`./configure --prefix=$PWD`). PLUMED expects the installation directory to be a different one! You might want to use something like `./configure --prefix=$PWD/install` instead.

Notice that upon installation PLUMED might need to relink a library. This was always true until version 2.1, but in version 2.2 libraries should only be relinked if one changes the install prefix during when typing `make install`. If root user does not have access to compilers, "sudo -E make install" might solve the issue.

Upon install, executables are copied to `$prefix/bin`, libraries to `$prefix/lib`, include files to `$prefix/include`, and documentation to `$prefix/shared/doc/plumed`. Additionally, a directory `$prefix/lib/plumed` is created containing several other files, including patch files, object files (for static patches), etc. Notice also that these path can be further customized using standard autoconf directories (e.g. `./configure --bindir=/usr/bin64`).

One should then set the environment properly. We suggest to do it using the module framework (<http://modules.sourceforge.net>). An ad hoc generated module file for PLUMED can be found in `$prefix/lib/plumed/src/lib/modulefile`. Just edit it as you wish and put it in your modulefile directory. This will also allow you to install multiple PLUMED versions on your machine and to switch amongst them. If you do not want to use modules, you can still have a look at the modulefile we did so as to know which environment variables should be set for PLUMED to work correctly.

If the environment is properly configured one should be able to do the following things:

- use the "plumed" executable from the command line. This is also possible before installing.
- link against the PLUMED library using the "-lplumed" flag for the linker. This allows one to use PLUMED library in general purpose programs

- use the PLUMED internal functionalities (C++ classes) including header files such as `"#include <plumed/tools/Vector.h>".` This is useful as it may be expedient to exploit the PLUMED library in general purpose programs

As a final note, if you want to install several PLUMED versions without using modules then you should provide a different suffix and/or prefix at configure time:

```
> ./configure prefix=$HOME/opt --program-suffix=_2.2 --program-prefix=mpi-  
> make install
```

This will install a plumed executable named "mpi-plumed\_2.2". All the other files will be renamed similarly, e.g. the PLUMED library will be loaded with `"-lmpi-plumed_2.2"` and the PLUMED header files will be included with `"#include <mpi-plumed_2.2/tools/Vector.h>".` Notice that you can also use arbitrary scripts to edit the name of the executable with the option `--program-transform-name=PROGRAM` (see [autoconf documentation](#) for more info). These options are useful if you do not want to set up modules, but we believe that using modules as described above is more flexible.

### 3.4 Patching your MD code

A growing number of MD codes can use PLUMED without any modification. If you are using one of these codes, refer to its manual to know how to activate PLUMED. In case your MD code is not supporting PLUMED already, you should modify it. We provide scripts to adjust some of the most popular MD codes so as to provide PLUMED support. At the present times we support patching the following list of codes:

- amber14
- gromacs-2016-5
- gromacs-4-5-7
- gromacs-5-0-7
- gromacs-5-1-4
- lammmps-6Apr13
- namd-2-8
- namd-2-9
- qespresso-5-0-2

In the section [Code specific notes](#) you can find information specific for each MD code.

To patch your MD code, you should have already installed PLUMED properly. This is necessary as you need to have the command "plumed" in your execution path. As described above this executable will be in your paths if plumed was installed or if you have run `sourceme.sh`

Once you have a compiled and working version of plumed, follow these steps to add it to an MD code

- Configure and compile your MD engine (look for the instructions in its documentation).
- Test if the MD code is working properly.
- Go to the root directory for the source code of the MD engine.
- Patch with PLUMED using:



```
> plumed patch -p
```

The script will interactively ask which MD engine you are patching.

- Once you have patched recompile the MD code (if dependencies are set up properly in the MD engine, only modified files will be recompiled)

There are different options available when patching. You can check all of them using

```
> plumed patch --help
```

Particularly interesting options include:

- `--static` just link PLUMED as a collection of object files. This is only suggested if for external reasons you absolutely need a static executable. Notice that with this setting it is often more complicated to configure properly the MD code, since all the libraries that PLUMED depends on should be properly specified. The `./configure` script does its best in this sense, but sometime it cannot solve the problem. Additionally, this patching mode has been reported not to work properly on OSX.
- `--shared` (default) allows you to link PLUMED as a shared library. As a result when PLUMED is updated, there will be no need to recompile the MD code. This is way better than `--static` since the libraries that PLUMED depends on should be automatically linked. Notice that if you later remove the directory where PLUMED is installed also the MD code will not run anymore.
- `--runtime` allows you to choose the location of the PLUMED library at runtime by setting the variable `PLUMED_KERNEL`. This is probably the most flexible option, and we encourage system administrators to use this option when installing PLUMED on shared facilities. Indeed, using this setting it will be possible to update separately the PLUMED library and the MD code, leaving to the user the possibility to combine different versions at will. We also recommend to use the provided modulefile (see above) to properly set the runtime environment.

Notice that it is not currently possible to link PLUMED as a static library (something like 'libplumed.a'). The reason for this is that PLUMED heavily relies on C++ static constructors that do not behave well in static libraries. For this reason, to produce a static executable with an MD code + PLUMED we link PLUMED as a collection of object files.

If your MD code is not supported, you may want to implement an interface for it. Refer to the [developer manual](#).

## 3.5 Cross compiling

If you are compiling an executable from a different machine, then `plumed` executable will not be available in the compilation environment. This means that you won't be able to perform regtests on the machine nor to compile the manual. You can try to run the regtests on the computing nodes, but this might require some tweak since often machines where people do cross compiling have architectures with limited capabilities on the compute nodes. Also notice that many of the `plumed` options (e.g. `patch`) are implemented as shell scripts launched from within the `plumed` executable. If the compute nodes have some limitation (e.g. they do not allow to fork new processes) these options will not work. Anyway, the PLUMED library in combination with an MD software should work if both PLUMED and the MD software have been properly compiled.

Also notice that it will not be possible to use the command `plumed patch` on the machine where you are compiling. You should thus use `plumed-patch` instead of `plumed patch` (notice that it should be written as a single word).

Try e.g.:

```
> plumed-patch --help
```

This script provides a "shell only" implementation of `plumed patch` that will skip the launch of the `plumed` executable.

Notice that other command line tools will be available in the directory `prefix/lib/progname/`. If configuring with default values this would be `/usr/local/lib/plumed/plumed-*`. These files are not included in the execution path (`prefix/bin`) to avoid clashes, but can be executed also when `plumed` is cross compiled and the main `plumed` executable cannot be launched.

## 3.6 Installing PLUMED with MacPorts

If you are using a Mac, notice that you can take advantage of a MacPorts package. Installing a working plumed should be as easy as:

- Install [MacPorts](#)
- Type `sudo port install plumed`

Notice that plumed comes with many variants that can be inspected with the command

```
> sudo port info plumed
```

Plumed uses variants to support different compilers. For instance, you can install plumed with openmpi using

```
> sudo port install plumed +openmpi
```

Using gcc instead of native compilers is recommended so as to take advantage of openMP

```
> sudo port install plumed +openmpi +gcc7
```

Variants can be also used to compile with debug flags (+debug), to pick a linear algebra library (e.g. +openblas) and to enable all optional modules (+allmodules). Notice that the default variant installed with `sudo port install plumed` is shipped as a precompiled binary, which is significantly faster to install.

In addition, we provide a developer version (typically: a later version not yet considered as stable) under the subport `plumed-devel` that can be installed with

```
> sudo port install plumed-devel
```

`plumed-devel` also supports the same variants as `plumed` in order to customize the compilation. `plumed-devel` and `plumed` cannot be installed at the same time.

It is also possible to install a plumed-patched version of gromacs. For instance, you can use the following command to install gromacs patched with plumed with gcc compiler and openmpi:

```
> sudo port install plumed +openmpi +gcc7
> sudo port install gromacs-plumed +openmpi +gcc7
```

In case you want to combine gromacs with the unstable version of plumed, use this instead:

```
> sudo port install plumed-devel +openmpi +gcc7
> sudo port install gromacs-plumed +openmpi +gcc7
```

Notice that gromacs should be compiled using the same compiler variant as plumed (in this example +openmpi +gcc7). In case this is not true, compilation will fail.

Also notice that gromacs is patched with plumed in runtime mode but that the path of `libplumedKernel.dylib` in the MacPorts tree is hardcoded. As a consequence:

- If gromacs is run with `PLUMED_KERNEL` environment variable unset (or set to empty), then the MacPorts plumed is used.
- If gromacs is run with `PLUMED_KERNEL` environment variable pointing to another instance of the plumed library, the other instance is used.

This is especially useful if you are developing PLUMED since you will be able to install gromacs once for all and combine it with your working version of PLUMED.

## 3.7 Installing PLUMED on a cluster

If you are installing PLUMED on a cluster and you want several users to take advantage of it consider the following suggestions.

First of all, we highly recommend using the module file that PLUMED provides to set up the environment. Just edit it as necessary to make it suitable for your environment.

Notice that PLUMED can take advantage of many additional features if specific libraries are available upon compiling it. Install `libmatheval` first and check if PLUMED `./configure` is detecting it. `libmatheval` is a must have with PLUMED. If someone uses `gromacs`, install `libxdrfile` first and check if PLUMED `./configure` is detecting it. PLUMED will be able to write `trr/xtc` file, simplifying analysis.

Try to patch all MD codes with the `--runtime` option. This will allow independent update of PLUMED and MD codes. Users will be able to combine any of the installed `gromacs/amber/etc` versions with any of the installed PLUMED versions. Notice that it is sometime claimed that statically linked codes are faster. In our experience, this is not true. In case you absolutely need a static executable, be ready to face non trivial linking issues. PLUMED is written in C++, thus required the appropriate C++ library to be linked, and might require additional libraries (e.g. `libmatheval`).

Sometime we make small fixes on the patches. For this reason, keep track of which version of PLUMED you used to patch each of the MD code. Perhaps you can call the MD code modules with names such as `gromacs/4.6.7p1`, `gromacs/4.6.7p2` and write somewhere in the module file which version of PLUMED you used. Alternatively, call them something like `gromacs/4.6.7p2.2.0`. In this way, when we report a bug on the mailing list, users will know if the version they are using is affected by it.

Usually it is not necessary to install both a MPI and a non-MPI PLUMED version. PLUMED library only calls MPI functions when the MD code is compiled with MPI. PLUMED executable calls MPI functions only when it is invoked without `--no-mpi`. In many machines it is thus sufficient to run the `plumed` executable on the login node as

```
> plumed --no-mpi
```

even though PLUMED was compiled with MPI and the login node does not support MPI. The only case where you might need two different PLUMED installation for compute and login node is when you are cross compiling.

PLUMED needs to be well optimized to run efficiently. If you need a single PLUMED binary to run efficiency on machines with different levels of hardware (e.g.: some of your workstations support AVX and some do not), with intel compiler you can use something like

```
./configure CXX=mpicxx CXXFLAGS="-O3 -axSSE2,AVX"
```

It will take more time to compile but it will allow you to use a single module. Otherwise, you should install two PLUMED version with different optimization levels.

Using modules, it is not necessary to make the PLUMED module explicitly dependent on the used library. Imagine a scenario where you first installed a module `libmatheval`, then load it while you compile PLUMED. If you provide the following option to configure `LDFLAGS="-Wl,-rpath,$LD_LIBRARY_PATH"`, the PLUMED executable and library will remember where `libmatheval` is, without the need to load `libmatheval` module at runtime. Notice that this trick often does not work for fundamental libraries such as C++ and MPI library. As a consequence, usually the PLUMED module should load the compiler and MPI modules.

### Attention

In case you found out how to compile PLUMED on some fancy architecture please share your tricks! You can either post it in your blog, send it to the mailing list, or ask as to update this paragraph in the manual, we will be happy to do so.

## 3.8 Other hints

We here collect a list of suggestions that might be useful on particular machines.

- On Blue Gene Q (likely on AIX) the prelinking made with `ld -r` is not working properly. There is no easy way to detect this at configure time. If during `make` you receive an error in the form

```
ld: TOC section size exceeds 64k
```

please configure plumed again with the following flag

```
./configure --disable-ld-r
```

- On Cray machines, you might have to set the following environment variable before configuring and building both PLUMED and the MD code that you want to patch with PLUMED (kindly reported by Marco De La Pierre):

```
export CRAYPE_LINK_TYPE=dynamic
```

- Intel MPI seems to require the flags `-lmpi_mt -mt_mpi` for compiling and linking and the flag `-DMPICH_IGNORE_CXX_SEEK` for compiling (kindly reported by Abhishek Acharya). You might want to try to configure using

```
./configure LDFLAGS=-lmpi_mt CXXFLAGS="-DMPICH_IGNORE_CXX_SEEK -mt_mpi" STATIC_LIBS=-mt_mpi
```

Adding libraries to `STATIC_LIBS` uses them for all the linking steps, whereas those in `LIBS` are only used when linking the PLUMED kernel library. See more at [this thread](#).

## 3.9 Code specific notes

Here you can find instructions that are specific for patching each of the supported MD codes. Notice that MD codes with native PLUMED support are not listed here.

- [amber14](#)
- [gromacs-2016.5](#)
- [gromacs-4.5.7](#)
- [gromacs-5.0.7](#)
- [gromacs-5.1.4](#)
- [lammmps-6Apr13](#)
- [namd-2.8](#)
- [namd-2.9](#)
- [gespresso-5.0.2](#)

### 3.9.1 amber14

PLUMED can be incorporated into amber (sander module) using the standard patching procedure. Patching must be done in the root directory of amber *before* compilation.

To enable PLUMED in a sander simulation one should use add to the `cntrl` input namelist these two fields:

```
plumed=1 , plumedfile='plumed.dat'
```

The first is switching plumed on, the second is specifying the name of the plumed input file.

This patch is compatible with the MPI version of sander and support multisander. However, replica exchange is not supported. Multisander can thus only be used for multiple walkers metadynamics or for ensemble restraints.

For more information on amber you should visit <http://ambermd.org>

### 3.9.2 gromacs-2016.5

PLUMED can be incorporated into gromacs using the standard patching procedure. Patching must be done in the gromacs root directory *before* the cmake command is invoked.

On clusters you may want to patch gromacs using the static version of plumed, in this case building gromacs can result in multiple errors. One possible solution is to configure gromacs with these additional options:

```
cmake -DBUILD_SHARED_LIBS=OFF -DGMX_PREFER_STATIC_LIBS=ON
```

To enable PLUMED in a gromacs simulation one should use mdrun with an extra -plumed flag. The flag can be used to specify the name of the PLUMED input file, e.g.:

```
gmh mdrun -plumed plumed.dat
```

For more information on gromacs you should visit <http://www.gromacs.org>

### 3.9.3 gromacs-4.5.7

PLUMED can be incorporated into gromacs using the standard patching procedure. Patching must be done in the gromacs source directory *after* gromacs has been configured but *before* gromacs is compiled. Gromcas should be configured with ./configure (not cmake).

To enable PLUMED in a gromacs simulation one should use mdrun with an extra -plumed flag. The flag can be used to specify the name of the PLUMED input file, e.g.:

```
mdrun -plumed plumed.dat
```

For more information on gromacs you should visit <http://www.gromacs.org>

### 3.9.4 gromacs-5.0.7

PLUMED can be incorporated into gromacs using the standard patching procedure. Patching must be done in the gromacs root directory *before* the cmake command is invoked.

To enable PLUMED in a gromacs simulation one should use mdrun with an extra -plumed flag. The flag can be used to specify the name of the PLUMED input file, e.g.:

```
gmh mdrun -plumed plumed.dat
```

For more information on gromacs you should visit <http://www.gromacs.org>

### 3.9.5 gromacs-5.1.4

PLUMED can be incorporated into gromacs using the standard patching procedure. Patching must be done in the gromacs root directory *before* the cmake command is invoked.

On clusters you may want to patch gromacs using the static version of plumed, in this case building gromacs can result in multiple errors. One possible solution is to configure gromacs with these additional options:

```
cmake -DBUILD_SHARED_LIBS=OFF -DGMX_PREFER_STATIC_LIBS=ON
```

To enable PLUMED in a gromacs simulation one should use mdrun with an extra -plumed flag. The flag can be used to specify the name of the PLUMED input file, e.g.:

```
gmh mdrun -plumed plumed.dat
```

This patch also implements the -hrex keyword for gromacs. See [Using Hamiltonian replica exchange with GROMACS](#)

For more information on gromacs you should visit <http://www.gromacs.org>

### 3.9.6 lammps-6Apr13

PLUMED can be incorporated into LAMMPS using a simple patching procedure. Patching must be done *before* LAMMPS is configured. After patching, one should enable PLUMED using the command `make yes-user-plumed`. In the same way, before reverting one should disable PLUMED using the command `make no-user-plumed`.

Also notice that command "fix plumed" should be used in lammps input file *after* the relevant input parameters have been set (e.g. after "timestep" command)

**Bug** The provided patch does not pass the virial correctly to PLUMED (see <https://github.com/plumed/plumed2/issues/377>). If you need to perform constant pressure simulations you should download a LAMMPS version newer than November 2018, where the mentioned bug has been fixed, that natively supports PLUMED 2.4. If you need to use LAMMPS with fixed virial with PLUMED 2.3, you can compile LAMMPS with PLUMED 2.4 using runtime binding and then set the variable `PLUMED_KERNEL` to load PLUMED 2.3.

See also [http://lammps.sandia.gov/doc/Section\\_commands.html](http://lammps.sandia.gov/doc/Section_commands.html) for further info on processing LAMMPS input, as well as this discussion on github: <http://github.com/plumed/plumed2/issues/67>.

For more information on LAMMPS you should visit <http://lammps.sandia.gov/>

### 3.9.7 namd-2.8

To enable PLUMED in a NAMD simulation one should add the following lines in the NAMD configuration file (often named as `xxx.namd` or `xxx.conf`) and run the plumed-patched version of NAMD with it:

```
plumed on
plumedfile plumed.dat
```

The first line tells NAMD to run with PLUMED and the second line specifies the PLUMED input file.

**Bug** NAMD does not currently take into account virial contributions from PLUMED. Please use constant volume simulations only.

For more information on NAMD you should visit <http://www.ks.uiuc.edu/Research/namd/>

### 3.9.8 namd-2.9

To enable PLUMED in a NAMD simulation one should add the following lines in the NAMD configuration file (often named as `xxx.namd` or `xxx.conf`) and run the plumed-patched version of NAMD with it:

```
plumed on
plumedfile plumed.dat
```

The first line tells NAMD to run with PLUMED and the second line specifies the PLUMED input file.

**Bug** NAMD does not currently take into account virial contributions from PLUMED. Please use constant volume simulations only.

For more information on NAMD you should visit <http://www.ks.uiuc.edu/Research/namd/>

### 3.9.9 qespresso-5.0.2

For more information on Quantum Espresso you should visit <http://www.quantum-espresso.org>

## Chapter 4

# Getting started

To run PLUMED you need to provide one input file. In this file you specify what it is that PLUMED should do during the course of the run. Typically this will involve calculating one or more collective variables, perhaps calculating a function of these CVs and then doing some analysis of values of your collective variables/functions or running some free energy method. A very brief introduction to the syntax used in the PLUMED input file is provided in this [10-minute video](#).

Within this input file every line is an instruction for PLUMED to perform some particular action. This could be the calculation of a colvar, an occasional analysis of the trajectory or a biasing of the dynamics. The first word in these lines specify what particular action is to be performed. This is then followed by a number of keywords which provide PLUMED with more details as to how the action is to be performed. These keywords are either single words (in which they tell PLUMED to do the calculation in a particular way - for example NOPBC tells PLUMED to not use the periodic boundary conditions when calculating a particular colvar) or they can be words followed by an equals sign and a comma separated list *with no spaces* of numbers or characters (so for example ATOMS=1,2,3,4 tells PLUMED to use atom numbers 1,2,3 and 4 in the calculation of a particular colvar). The reason why spaces are not admitted is that PLUMED should be able to understand when the list of atoms ended and a new keyword should be expected. Space separated lists can be used instead of comma separated list if the entire list is enclosed in curly braces (e.g. ATOMS={1 2 3 4}). Please note that you can split commands over multiple lines by using [Continuation lines](#).

The most important of these keywords is the label keyword as it is only by using these labels that we can pass data from one action to another. As an example if you do:

```
DISTANCE ATOMS=1,2
```

(see [DISTANCE](#))

Then PLUMED will do nothing other than read in your input file. In contrast if you do:

```
DISTANCE ATOMS=1,2 LABEL=d1  
PRINT ARG=d1 FILE=colvar STRIDE=10
```

(see [PRINT](#))

then PLUMED will print out the value of the distance between atoms 1 and 2 every 10 steps to the file colvar as you have told PLUMED to take the value calculated by the action d1 and to print it. You can use any character string to label your actions as long as it does not begin with the symbol @. Strings beginning with @ are used by within PLUMED to reference special, code-generated groups of atoms and to give labels to any Actions for which the user does not provide a label in the input.

Notice that if a word followed by a column is added at the beginning of the line (e.g. pippo:), PLUMED automatically removes it and adds an equivalent label (LABEL=pippo). Thus, a completely equivalent result can be obtained with the following shortcut:

```
d1: DISTANCE ATOMS=1,2
PRINT ARG=d1 FILE=colvar STRIDE=10
```

Also notice that all the actions can be labeled, and that many actions besides normal collective variables can define one or more value, which can be then referred using the corresponding label.

Actions can be referred also with POSIX regular expressions (see [Regular Expressions](#)) if regex library is available on your system and detected at configure time. You can also add [Comments](#) to the input or set up your input over multiple files and then create a composite input by [Including other files](#).

More information on the input syntax as well as details on the the various trajectory analysis tools that come with PLUMED are given in:

- [Collective variables](#) tells you about the ways that you can calculate functions of the positions of the atoms.
- [Analysis](#) tells you about the various forms of analysis you can run on trajectories using PLUMED.
- [Bias](#) tells you about the methods that you can use to bias molecular dynamics simulations with PLUMED.

## 4.1 Plumed units

By default the PLUMED inputs and outputs quantities in the following units:

- Energy - kJ/mol
- Length - nanometers
- Time - picoseconds

Unlike PLUMED 1 the units used are independent of the MD engine you are using. If you want to change these units you can do this using the [UNITS](#) keyword.

## 4.2 UNITS

<b>This is part of the setup <a href="#">module</a></b>
---------------------------------------------------------

This command sets the internal units for the code. A new unit can be set by either specifying how to convert from the plumed default unit into that new unit or by using the shortcuts described below. This directive **MUST** appear at the **BEGINNING** of the plumed.dat file. The same units must be used throughout the plumed.dat file.

Notice that all input/output will then be made using the specified units. That is: all the input parameters, all the output files, etc. The only exceptions are file formats for which there is a specific convention concerning the units. For example, trajectories written in .gro format (with [DUMPATOMS](#)) are going to be always in nm.

### Options



<b>NATURAL</b>	( default=off ) use natural units
<b>LENGTH</b>	the units of lengths. Either specify a conversion factor from the default, nm, or A (for angstroms) or um
<b>ENERGY</b>	the units of energy. Either specify a conversion factor from the default, kj/mol, or use j/mol or kcal/mol
<b>TIME</b>	the units of time. Either specify a conversion factor from the default, ps, or use ns or fs
<b>MASS</b>	the units of masses. Specify a conversion factor from the default, amu
<b>CHARGE</b>	the units of charges. Specify a conversion factor from the default, e

### Examples

```
# this is using nm - kj/mol - fs
UNITS LENGTH=nm TIME=fs
```

If a number, x, is found, the new unit is equal to x (default units)

```
# this is using nm - kj/mol - fs
UNITS LENGTH=nm TIME=0.001
```



## Chapter 5

# Collective variables

Chemical systems contain an enormous number of atoms, which, in most cases, makes it simply impossible for us to understand anything by monitoring the atom positions directly. Consequently, we introduce Collective variables (CVs) that describe the chemical processes we are interested in and monitor these simpler quantities instead. These CVs are used in many of the methods implemented in PLUMED - their values can be monitored using [PRINT](#), [Functions](#) of them can be calculated or they can be analyzed or biased using the [Analysis](#) and [Biasing](#) methods implemented in PLUMED. Before doing any of these things however we first have to tell PLUMED how to calculate them.

The simplest collective variables that are implemented in PLUMED take in a set of atomic positions and output one or multiple scalar CV values. Information on these variables is given on the page entitled [CV Documentation](#) while information as to how sets of atoms can be selected can be found in the pages on [Groups and Virtual Atoms](#). Please be aware that PLUMED contains implementations of many other collective variables but that the input for these variables may be less transparent when it is first encountered. In particular, the page on [Distances from reference configurations](#) describes the various ways that you can calculate the distance from a particular reference configuration. So you will find instructions on how to calculate the RMSD distance from the folded state of a protein here. Meanwhile, the page on [Functions](#) describes the various functions of collective variables that can be used in the code. This is a very powerful feature of PLUMED as you can use the [Functions](#) commands to calculate any function or combination of the simple collective variables listed on the page [CV Documentation](#). Lastly the page on [MultiColvar](#) describes MultiColvars.

MultiColvars allow you to use many different colvars and allow us to implement all these collective variables without implementing having an unmanageably large amount of code. For some things (e.g. [DISTANCES GROUPA=1 GROUPB=2-100 LESS\\_THAN={RATIONAL R\\_0=3}](#)) there are more computationally efficient options available in plumed (e.g. [COORDINATION](#)). However, MultiColvars are worth investigating as they provide a flexible syntax for many quite-complex CVs.

- [Groups and Virtual Atoms](#)
- [CV Documentation](#)
- [Distances from reference configurations](#)
- [Functions](#)
- [MultiColvar](#)
- [Exploiting contact matrices](#)

## 5.1 Groups and Virtual Atoms

### 5.1.1 Specifying Atoms

The vast majority of the CVs implemented in PLUMED are calculated from a list of atom positions. Within PLUMED atoms are specified using their numerical indices in the molecular dynamics input file.

In PLUMED lists of atoms can be either provided directly inside the definition of each collective variable, or predefined as a [GROUP](#) that can be reused multiple times. Lists of atoms can be written as:

- comma separated lists of numbers (GROUP ATOMS=10,11,15,20 LABEL=g1)
- numerical ranges. So GROUP ATOMS=10-20 LABEL=g2 is equivalent to GROUP ATOMS=10,11,12,13,14,15,16,17,18,19,20 LABEL=g2
- numerical ranges with a stride. So GROUP ATOMS=10-100:10 LABEL=g3 is equivalent to GROUP ATOMS=10,20,30,40,50,60,70,80,90,100 LABEL=g3
- atoms ranges with a negative stride. So GROUP ATOMS=100-10:-10 LABEL=g4 is equivalent to GROUP ATOMS=100,90,80,70,60,50,40,30,20,10 LABEL=g4
- all the above methods together. For example GROUP ATOMS=1,2,10-20,40-60:5,100-70:-2 LABEL=g5.

Some collective variable must accept a fixed number of atoms, for example a [DISTANCE](#) is calculated using two atoms only, an [ANGLE](#) is calculated using either 3 or 4 atoms and [TORSION](#) is calculated using 4 atoms.

Additional material and examples can be also found in the tutorial [Belfast tutorial: Analyzing CVs](#).

#### 5.1.1.1 Molecules

In addition, for certain colvars, pdb files can be read in using the following keywords and used to select ATOMS:

<a href="#">MOLINFO</a>	This command is used to provide information on the molecules that are present in your system.
-------------------------	-----------------------------------------------------------------------------------------------

#### 5.1.1.2 Broken Molecules and PBC

PLUMED is designed so that for the majority of the CVs implemented the periodic boundary conditions are treated in the same manner as they would be treated in the host code. In some codes this can be problematic when the colvars you are using involve some property of a molecule. These codes allow the atoms in the molecules to become separated by periodic boundaries, a fact which PLUMED could only deal with were the topology passed from the MD code to PLUMED. Making this work would involve a lot laborious programming and goes against our original aim of having a general patch that can be implemented in a wide variety of MD codes. Consequentially, we have implemented a more pragmatic solution to this problem - the user specifies in input any molecules (or parts of molecules) that must be kept in tact throughout the simulation run. In PLUMED 1 this was done using the `ALIGN_ATOMS` keyword. In PLUMED 2 the same effect can be achieved using the [WHOLEMOLECULES](#) command.

The following input computes the end-to-end distance for a polymer of 100 atoms and keeps it at a value around 5.

```
WHOLEMOLECULES ENTITY0=1-100
e2e: DISTANCE ATOMS=1,100 NOPBC
RESTRAINT ARG=e2e KAPPA=1 AT=5
```

Notice that NOPBC is used to be sure in [DISTANCE](#) that if the end-to-end distance is larger than half the simulation box the distance is compute properly. Also notice that, since many MD codes break molecules across cell boundary, it might be necessary to use the [WHOLEMOLECULES](#) keyword (also notice that it should be before distance).

Notice that most expressions are invariant with respect to a change in the order of the atoms, but some of them depend on that order. E.g., with [WHOLEMOLECULES](#) it could be useful to specify atom lists in a reversed order.

```
# to see the effect, one could dump the atoms as they were before molecule reconstruction:
# DUMPATOMS FILE=dump-broken.xyz ATOMS=1-20
WHOLEMOLECULES STRIDE=1 ENTITY0=1-20
DUMPATOMS FILE=dump.xyz ATOMS=1-20
```

Notice that there are other ways to manipulate the coordinates stored within PLUMED:

- Using the [FIT\\_TO\\_TEMPLATE](#) they can be aligned to a template structure.
- Using [WRAPAROUND](#) you can bring a set of atom as close as possible to another set of atoms.
- Using [RESET\\_CELL](#) you can rotate the periodic cell.

### 5.1.2 Virtual Atoms

Sometimes, when calculating a colvar, you may not want to use the positions of a number of atoms directly. Instead you may wish to use the position of a virtual atom whose position is generated based on the positions of a collection of other atoms. For example you might want to use the center of mass of a group of atoms. Plumed has a number of routines for calculating the positions of these virtual atoms from lists of atoms:

<a href="#">CENTER_OF_MULTICOLVAR</a>	Calculate a a weighted average position based on the value of some multi-colvar.
<a href="#">CENTER</a>	Calculate the center for a group of atoms, with arbitrary weights.
<a href="#">COM</a>	Calculate the center of mass for a group of atoms.
<a href="#">FIXEDATOM</a>	Add a virtual atom in a fixed position.
<a href="#">GHOST</a>	Calculate the absolute position of a ghost atom with fixed coordinates in the local reference frame formed by three atoms. The computed ghost atom is stored as a virtual atom that can be accessed in an atom list through the the label for the GHOST action that creates it.

To specify to a colvar that you want to use the position of a virtual atom to calculate a colvar rather than one of the atoms in your system you simply use the label for your virtual atom in place of the usual numerical index. Virtual atoms and normal atoms can be mixed together in the input to colvars as shown below:

```
COM ATOMS=1,10 LABEL=com1
DISTANCE ATOMS=11,com1
```

If you don't want to calculate CVs from the virtual atom. That is to say you just want to monitor the position of a virtual atom (or any set of atoms) over the course of your trajectory you can do this using [DUMPATOMS](#).

### 5.1.3 GROUP

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

Define a group of atoms so that a particular list of atoms can be referenced with a single label in definitions of CVs or virtual atoms.

Atoms can be listed as comma separated numbers (i.e. 1,2,3,10,45,7,9,...) , simple positive ranges (i.e. 20-40), ranges with a stride either positive or negative (i.e. 20-40:2 or 80-50:-2) or as combinations of all the former methods (1,2,4,5,10-20,21-40:2,80-50:-2).

Moreover, lists can be imported from ndx files (GROMACS format). Use `NDX_FILE` to set the name of the index file and `NDX_GROUP` to set the name of the group to be imported (default is first one).

It is also possible to remove atoms from a list and or sort them using keywords `REMOVE`, `SORT`, and `UNIQUE`. The flow is the following:

- If `ATOMS` is present take the ordered list of atoms from the `ATOMS` keyword.
- If `NDX_FILE` is present append the list from the the gromacs group.
- If `REMOVE` is present remove the first occurrence of each of these atoms from the list. If one tries to remove an atom that was not listed plumed adds a notice in the output.
- If `SORT` is present resulting list is sorted.
- If `UNIQUE` is present the resulting list is sorted and duplicate elements are removed.

Notice that this command just creates a shortcut, and does not imply any real calculation. So, having a huge group defined does not slow down your calculation in any way. It is just convenient to better organize input files. Might be used in combination with the `INCLUDE` command so as to store long group definitions in a separate file.

The atoms involved can be specified using

<b>ATOMS</b>	the numerical indexes for the set of atoms in the group. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
<b>REMOVE</b>	remove these atoms from the list. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>

Options

<b>SORT</b>	( default=off ) sort the resulting list
<b>UNIQUE</b>	( default=off ) sort atoms and remove duplicated ones
<b>NDX_FILE</b>	the name of index file (gromacs syntax)
<b>NDX_GROUP</b>	the name of the group to be imported (gromacs syntax) - first group found is used by default

Examples

This command create a group of atoms containing atoms 1,4,7,11 and 14 (labeled 'o'), and another containing atoms 2,3,5,6,8,9,12,13 (labeled 'h'):

```
o: GROUP ATOMS=1,4,7,11,14
h: GROUP ATOMS=2,3,5,6,8,9,12,13
# compute the coordination among the two groups
c: COORDINATION GROUPA=o GROUPB=h R_0=0.3
# same could have been obtained without GROUP, just writing:
# c: COORDINATION GROUPA=1,4,7,11,14 GROUPB=2,3,5,6,8,9,12,13

# print the coordination on file 'colvar'
PRINT ARG=c FILE=colvar
```

(see also [COORDINATION](#) and [PRINT](#))

Groups can be conveniently stored in a separate file. E.g. one could create a file named 'groups.dat' which reads

```
o: GROUP ATOMS=1,4,7,11,14
h: GROUP ATOMS=2,3,5,6,8,9,12,13
```

and then include it in the main 'plumed.dat' file

```
INCLUDE FILE=groups.dat
# compute the coordination among the two groups
c: COORDINATION GROUPA=o GROUPB=h R_0=0.3
# print the coordination on file 'colvar'
PRINT ARG=c FILE=colvar
```

(see also [INCLUDE](#), [COORDINATION](#), and [PRINT](#)). The groups.dat file could be very long and include lists of thousand atoms without cluttering the main plumed.dat file.

A GROMACS index file can also be imported

```
# import group named 'protein' from file index.ndx
pro: GROUP NDX_FILE=index.ndx NDX_GROUP=protein
# dump all the atoms of the protein on a trajectory file
DUMPATOMS ATOMS=pro FILE=traj.gro
```

(see also [DUMPATOMS](#))

A list can be edited with REMOVE

```
# take one atom every three
ox: GROUP ATOMS=1-90:3
# take the remaining atoms
hy: GROUP ATOMS=1-90 REMOVE=ox
DUMPATOMS ATOMS=ox FILE=ox.gro
DUMPATOMS ATOMS=hy FILE=hy.gro
```

(see also [DUMPATOMS](#))

### 5.1.4 MOLINFO

This is part of the setup <a href="#">module</a>
--------------------------------------------------

This command is used to provide information on the molecules that are present in your system.

The information on the molecules in your system can either be provided in the form of a pdb file or as a set of lists of atoms that describe the various chains in your system. If a pdb file is used plumed the MOLINFO command will

endeavor to recognize the various chains and residues that make up the molecules in your system using the chain↔ IDs and resnumbers from the pdb file. You can then use this information in later commands to specify atom lists in terms residues. For example using this command you can find the backbone atoms in your structure automatically.

### Warning

Please be aware that the PDB parser in plumed is far from perfect. You should thus check the log file and examine what plumed is actually doing whenever you use the MOLINFO action. Also make sure that the atoms are listed in the pdb with the correct order. If you are using gromacs, the safest way is to use reference pdb file generated with `gmx editconf -f topol.tpr -o reference.pdb`.

More information of the PDB parser implemented in PLUMED can be found [at this page](#).

Providing `MOLTYPE=protein`, `MOLTYPE=rna`, or `MOLTYPE=dna` will instruct plumed to look for known residues from these three types of molecule. In other words, this is available for historical reasons and to allow future extensions where alternative lists will be provided. As of now, you can just ignore this keyword.

Using MOLINFO with a protein's or nucleic acid's pdb extends the possibility of atoms selection using the @ special symbol in the form

```
@ "definition"-chainresiduenum
@ "definition"-residuenum
```

### So for example

```
@psi-1 will select the atoms defining the psi torsion of residue 1
@psi-C1 will define the same torsion for residue 1 of chain C.
```

In the following are listed the current available definitions:

For protein residues, the following groups are available:

```
@phi-#
@psi-#
@omega-#
@chi1-#
```

that select the appropriate atoms that define each dihedral angle for residue #.

For DNA or RNA residues, the following groups are available:

```
# quadruplets for backbone dihedral angles
@alpha-#
@beta-#
@gamma-#
@delta-#
@epsilon-#
@zeta-#

# quadruplets for sugar dihedral angles
@v0-#
@v1-#
@v2-#
@v3-#
@v4-#

# quadruplet corresponding to the chi torsional angle
@chi-#
```



```
# backbone, sugar, and base heavy atoms
@back-#
@sugar-#
@base-#

# ordered triplets of atoms on the 6-membered ring of nucleobases
# namely:
#  C2/C4/C6 for pyrimidines
#  C2/C6/C4 for purines
@lcs-#
```

Notice that `zeta` and `epsilon` groups should not be used on 3' end residue and `alpha` and `beta` should not be used on 5' end residue.

Furthermore it is also possible to pick single atoms using the syntax `@atom-chainresiduenum` or `@atom-residuenum`.

#### Warning

If a residue-chain is repeated twice in the reference pdb only the first entry will be selected.

**Bug** At the moment the HA1 atoms in a GLY residues are treated as if they are the CB atoms. This may or may not be true - GLY is problematic for secondary structure residues as it is achiral.

**Bug** If you use `WHOLEMOLECULES RESIDUES=1-10` for a 18 amino acid protein ( 18 amino acids + 2 terminal groups = 20 residues ) the code will fail as it will not be able to interpret terminal residue 1.

The atoms involved can be specified using

<b>CHAIN</b>	(for masochists ( mostly Davide Branduardi ) ) The atoms involved in each of the chains of interest in the structure.. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Compulsory keywords

<b>STRUCTURE</b>	a file in pdb format containing a reference structure. This is used to defines the atoms in the various residues, chains, etc . For more details on the PDB file format visit <a href="http://www.wwpdb.org/docs.html">http://www.wwpdb.org/docs.html</a>
<b>MOLTYPE</b>	( default=protein ) what kind of molecule is contained in the pdb file - usually not needed since protein/RNA/DNA are compatible

#### Examples

In the following example the MOLINFO command is used to provide the information on which atoms are in the backbone of a protein to the ALPHARMSD CV.

```
MOLINFO STRUCTURE=reference.pdb
ALPHARMSD RESIDUES=all TYPE=DRMSD LESS_THAN={RATIONAL R_0=0.08 NN=8 MM=12} LABEL=a
```

(see also [ALPHARMSD](#))

The following example prints the distance corresponding to the hydrogen bonds in a GC Watson-Crick pair.

```
MOLINFO STRUCTURE=reference.pdb
hb1: DISTANCE ATOMS=@N2-1,@O2-14
hb2: DISTANCE ATOMS=@N1-1,@N3-14
hb3: DISTANCE ATOMS=@O6-1,@N4-14
PRINT ARG=hb1,hb2,hb3
```

(see also [DISTANCE](#)).

This example use MOLINFO to calculate torsions angles

```
MOLINFO MOLTYPE=protein STRUCTURE=myprotein.pdb
t1: TORSION ATOMS=@phi-3
t2: TORSION ATOMS=@psi-4
PRINT ARG=t1,t2 FILE=colvar STRIDE=10
```

## 5.1.5 WHOLEMOLECULES

This is part of the generic <a href="#">module</a>
----------------------------------------------------

This action is used to rebuild molecules that can become split by the periodic boundary conditions.

It is similar to the `ALIGN_ATOMS` keyword of `plumed1`, and is needed since some MD dynamics code (e.g. `GR↔OMACS`) can break molecules during the calculation.

Running some CVs without this command can cause there to be discontinuities changes in the CV value and artifacts in the calculations. This command can be applied more than once. To see what effect it has use a variable without `pbs` or use the [DUMPATOMS](#) directive to output the atomic positions.

### Attention

This directive modifies the stored position at the precise moment it is executed. This means that only collective variables which are below it in the input script will see the corrected positions. As a general rule, put it at the top of the input file. Also, unless you know exactly what you are doing, leave the default stride (1), so that this action is performed at every MD step.

The way `WHOLEMOLECULES` modifies each of the listed entities is this:

- First atom of the list is left in place
- Each atom of the list is shifted by a lattice vectors so that it becomes as close as possible to the previous one, iteratively.

In this way, if an entity consists of a list of atoms such that consecutive atoms in the list are always closer than half a box side the entity will become whole. This can be usually achieved selecting consecutive atoms (1-100), but it is also possible to skip some atoms, provided consecutive chosen atoms are close enough.

The atoms involved can be specified using

<b>ENTITY</b>	the atoms that make up a molecule that you wish to align. To specify multiple molecules use a list of ENTITY keywords: ENTITY0, ENTITY1,... You can use multiple instances of this keyword i.e. ENTITY1, ENTITY2, ENTITY3...
---------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>RESIDUES</b>	this command specifies that the backbone atoms in a set of residues all must be aligned. It must be used in tandem with the <a href="#">MOLINFO</a> action and the MOLTYPE keyword. If you wish to use all the residues from all the chains in your system you can do so by specifying all. Alternatively, if you wish to use a subset of the residues you can specify the particular residues you are interested in as a list of numbers
-----------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which molecules are reassembled. Unless you are completely certain about what you are doing leave this set equal to 1!
---------------	---------------------------------------------------------------------------------------------------------------------------------------------------------

Options

<b>MOLTYPE</b>	the type of molecule that is under study. This is used to define the backbone atoms
----------------	-------------------------------------------------------------------------------------

Examples

This command instructs plumed to reconstruct the molecule containing atoms 1-20 at every step of the calculation and dump them on a file.

```
# to see the effect, one could dump the atoms as they were before molecule reconstruction:
# DUMPATOMS FILE=dump-broken.xyz ATOMS=1-20
WHOLEMOLECULES ENTITY0=1-20
DUMPATOMS FILE=dump.xyz ATOMS=1-20
```

(see also [DUMPATOMS](#))

This command instructs plumed to reconstruct two molecules containing atoms 1-20 and 30-40

```
WHOLEMOLECULES ENTITY0=1-20 ENTITY1=30-40
DUMPATOMS FILE=dump.xyz ATOMS=1-20,30-40
```

(see also [DUMPATOMS](#))

This command instructs plumed to reconstruct the chain of backbone atoms in a protein

```
MOLINFO STRUCTURE=helix.pdb
WHOLEMOLECULES RESIDUES=all MOLTYPE=protein
```

(See also [MOLINFO](#))

### 5.1.6 FIT\_TO\_TEMPLATE

	<b>This is part of the generic <a href="#">module</a></b>
--	-----------------------------------------------------------

This action is used to align a molecule to a template.

This can be used to move the coordinates stored in plumed so as to be aligned with a provided template in PDB format. Pdb should contain also weights for alignment (see the format of PDB files used e.g. for [RMSD](#)). Make sure your PDB file is correctly formatted as explained [in this page](#). Weights for displacement are ignored, since no displacement is computed here. Notice that all atoms (not only those in the template) are aligned. To see what effect try the [DUMPATOMS](#) directive to output the atomic positions.

Also notice that PLUMED propagate forces correctly so that you can add a bias on a CV computed after alignment. For many CVs this has no effect, but in some case the alignment can change the result. Examples are:

- [POSITION](#) CV since it is affected by a rigid shift of the system.
- [DISTANCE](#) CV with COMPONENTS. Since the alignment could involve a rotation (with TYPE=OPTIMAL) the actual components could be different from the original ones.
- [CELL](#) components for a similar reason.

#### Attention

The implementation of TYPE=OPTIMAL is available but should be considered in testing phase. Please report any strange behavior.

This directive modifies the stored position at the precise moment it is executed. This means that only collective variables which are below it in the input script will see the corrected positions. As a general rule, put it at the top of the input file. Also, unless you know exactly what you are doing, leave the default stride (1), so that this action is performed at every MD step.

#### Warning

The molecule used for alignment should be whole. In case it is broken by the host MD code, please use [WHOLEMOLECULES](#) to reconstruct it before [FIT\\_TO\\_TEMPLATE](#).

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which molecules are reassembled. Unless you are completely certain about what you are doing leave this set equal to 1!
<b>REFERENCE</b>	a file in pdb format containing the reference structure and the atoms involved in the CV.
<b>TYPE</b>	( default=SIMPLE ) the manner in which RMSD alignment is performed. Should be OPTIMAL or SIMPLE.

#### Examples

Align the atomic position to a template then print them

```
# to see the effect, one could dump the atoms before alignment
DUMPATOMS FILE=dump-before.xyz ATOMS=1-20
FIT_TO_TEMPLATE STRIDE=1 REFERENCE=ref.pdb TYPE=SIMPLE
DUMPATOMS FILE=dump-after.xyz ATOMS=1-20
```

(see also [DUMPATOMS](#))

### 5.1.7 WRAPAROUND

This is part of the generic <a href="#">module</a>
----------------------------------------------------

Rebuild periodic boundary conditions around chosen atoms.

Modify position of atoms indicated by ATOMS by shifting them by lattice vectors so that they are as close as possible to the atoms indicated by AROUND. More precisely, for every atom *i* in the ATOMS list the following procedure is performed:

- The atom *j* among those in the AROUND list is searched that is closest to atom *i*.
- The atom *i* is replaced with its periodic image that is closest to atom *j*.

This action works similarly to [WHOLEMOLECULES](#) in that it replaces atoms coordinate. Notice that only atoms specified with ATOMS are replaced, and that, at variance with [WHOLEMOLECULES](#), the order in which atoms are specified is irrelevant.

This is often convenient at a post processing stage (using the [driver](#)), but sometime it is required during the simulation if collective variables need atoms to be in a specific periodic image.

#### Attention

This directive modifies the stored position at the precise moment it is executed. This means that only collective variables which are below it in the input script will see the corrected positions. As a general rule, put it at the top of the input file. Also, unless you know exactly what you are doing, leave the default stride (1), so that this action is performed at every MD step.

Consider that the computational cost grows with the product of the size of the two lists (ATOMS and AROUND), so that this action can become very expensive. If you are using it to analyse a trajectory this is usually not a big problem. If you use it to analyze a simulation on the fly, e.g. with [DUMPATOMS](#) to store a properly wrapped trajectory, consider the possibility of using the STRIDE keyword here (with great care).

The atoms involved can be specified using

<b>AROUND</b>	reference atoms. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
<b>ATOMS</b>	wrapped atoms. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which molecules are reassembled. Unless you are completely certain about what you are doing leave this set equal to 1!
<b>GROUPBY</b>	( default=1 ) group atoms so as not to break molecules

## Examples

This command instructs plumed to move all the ions to their periodic image that is as close as possible to the rna group.

```
rna: GROUP ATOMS=1-100
ions: GROUP ATOMS=101-110
# first make the rna molecule whole
WHOLEMOLECULES ENTITY0=rna
WRAPAROUND ATOMS=ions AROUND=rna
DUMPATOMS FILE=dump.xyz ATOMS=rna,ions
```

(see also [WHOLEMOLECULES](#), [GROUP](#) and [DUMPATOMS](#))

In case you want to do it during a simulation and you only care about wrapping the ions in the `dump.xyz` file, you can use the following:

```
# add some restraint that do not require molecules to be whole:
a: TORSION ATOMS=1,2,10,11
RESTRAINT ARG=a AT=0.0 KAPPA=5

# then do the things that are required for dumping the trajectory
# notice that they are all done every 100 steps, so as not to
# unnecessarily overload the calculation

rna: GROUP ATOMS=1-100
ions: GROUP ATOMS=101-110
# first make the rna molecule whole
WHOLEMOLECULES ENTITY0=rna STRIDE=100
WRAPAROUND ATOMS=ions AROUND=rna STRIDE=100
DUMPATOMS FILE=dump.xyz ATOMS=rna,ions STRIDE=100
```

(see also [TORSION](#), [GROUP](#), [WHOLEMOLECULES](#) and [DUMPATOMS](#))

Notice that if the biased variable requires a molecule to be whole, you might have to put just the [WHOLEMOLECULES](#) command before computing that variable and leave the default `STRIDE=1`.

This command instructs plumed to center all atoms around the center of mass of a solute molecule.

```
solute: GROUP ATOMS=1-100
all: GROUP ATOMS=1-1000
# center of the solute:
# notice that since plumed 2.2 this also works if the
# solute molecule is broken
com: COM ATOMS=solute
# notice that we wrap around a single atom. this should be fast
WRAPAROUND ATOMS=all AROUND=com
DUMPATOMS FILE=dump.xyz ATOMS=all
```

(see also [GROUP COM](#) [DUMPATOMS](#))

Notice that whereas [WHOLEMOLECULES](#) is designed to make molecules whole, [WRAPAROUND](#) can easily break molecules. In the last example, if solvent (atoms 101-1000) is made e.g. of water, then water molecules could be broken by [WRAPAROUND](#) (hydrogen could end up in an image and oxygen in another one). One solution is to use [WHOLEMOLECULES](#) on *all* the water molecules after [WRAPAROUND](#). This is tedious. A better solution is to use the `GROUPBY` option which is going to consider the atoms listed in `ATOMS` as a list of groups each of size `GROUPBY`. The first atom of the group will be brought close to the `AROUND` atoms. The following atoms of the group will be just brought close to the first atom of the group. Assuming that oxygen is the first atom of each water molecules, in the following examples all the water oxygens will be brought close to the solute, and all the hydrogens will be kept close to their related oxygen.

```

solute: GROUP ATOMS=1-100
water: GROUP ATOMS=101-1000
com: COM ATOMS=solute
# notice that we wrap around a single atom. this should be fast
WRAPAROUND ATOMS=solute AROUND=com
# notice that we wrap around a single atom. this should be fast
WRAPAROUND ATOMS=water AROUND=com GROUPBY=3
DUMPATOMS FILE=dump.xyz ATOMS=solute,water

```

(see also [GROUP COM DUMPATOMS](#))

### 5.1.8 RESET\_CELL

This is part of the generic <a href="#">module</a>
----------------------------------------------------

This action is used to rotate the full cell

This can be used to modify the periodic box. Notice that this is done at fixed scaled coordinates, so that also atomic coordinates for the entire system are affected. To see what effect try the [DUMPATOMS](#) directive to output the atomic positions.

Also notice that PLUMED propagate forces correctly so that you can add a bias on a CV computed after rotation. See also [FIT\\_TO\\_TEMPLATE](#)

Currently, only TYPE=TRIANGULAR is implemented, which allows one to reset the cell to a lower triangular one. Namely, a proper rotation is found that allows rotating the box so that the first lattice vector is in the form (ax,0,0), the second lattice vector is in the form (bx,by,0), and the third lattice vector is arbitrary.

#### Attention

The implementation of this action is available but should be considered in testing phase. Please report any strange behavior.

This directive modifies the stored position at the precise moment it is executed. This means that only collective variables which are below it in the input script will see the corrected positions. Unless you know exactly what you are doing, leave the default stride (1), so that this action is performed at every MD step.

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which molecules are reassembled. Unless you are completely certain about what you are doing leave this set equal to 1!
<b>TYPE</b>	( default=TRIANGULAR ) the manner in which the cell is reset

#### Examples

Reset cell to be triangular after a rototranslational fit

```

DUMPATOMS FILE=dump-original.xyz ATOMS=1-20
FIT_TO_TEMPLATE STRIDE=1 REFERENCE=ref.pdb TYPE=OPTIMAL
DUMPATOMS FILE=dump-fit.xyz ATOMS=1-20
RESET_CELL TYPE=TRIANGULAR
DUMPATOMS FILE=dump-reset.xyz ATOMS=1-20

```

(see also [DUMPATOMS](#))

### 5.1.9 CENTER\_OF\_MULTICOLVAR

	<b>This is part of the multicolvar <a href="#">module</a></b>
--	---------------------------------------------------------------

Calculate a a weighted average position based on the value of some multicolvar.

This action calculates the position of a new virtual atom using the following formula:

$$x_{\alpha} = \frac{1}{2\pi} \arctan \left[ \frac{\sum_i w_i f_i \sin(2\pi x_{i,\alpha})}{\sum_i w_i f_i \cos(2\pi x_{i,\alpha})} \right]$$

Where in this expression the  $w_i$  values are a set of weights calculated within a multicolvar action and the  $f_i$  are the values of the multicolvar functions. The  $x_{i,\alpha}$  values are the positions (in scaled coordinates) associated with each of the multicolvars calculated.

**Bug** The virial contribution for this type of virtual atom is not currently evaluated so do not use in bias functions unless the volume of the cell is fixed

The atoms involved can be specified using

<b>ATOMS</b>	the list of atoms which are involved the virtual atom's definition. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>DATA</b>	find the average value for a multicolvar
-------------	------------------------------------------

Options

<b>COMPONENT</b>	if your input multicolvar is a vector then specify which component you would like to use in calculating the weight
------------------	--------------------------------------------------------------------------------------------------------------------

Examples

Lets suppose that you are examining the formation of liquid droplets from gas. You may want to determine the center of mass of any of the droplets formed. In doing this calculation you recognise that the atoms in the liquid droplets will have a higher coordination number than those in the surrounding gas. As you want to calculate the position of



the droplets you thus recognise that these atoms with high coordination numbers should have a high weight in the weighted average you are using to calculate the position of the droplet. You can thus calculate the position of the droplet using an input like the one shown below:

```
c1: COORDINATIONNUMBER SPECIES=1-512 SWITCH={EXP D_0=4.0 R_0=0.5}
cc: CENTER_OF_MULTICOLVAR DATA=c1
```

The first line here calculates the coordination numbers of all the atoms in the system. The virtual atom then uses the values of the coordination numbers calculated by the action labelled c1 when it calculates the Berry Phase average described above. (N.B. the  $w_i$  in the above expression are all set equal to 1 in this case)

The above input is fine we can, however, refine this somewhat by making use of a multicolvar transform action as shown below:

```
c1: COORDINATIONNUMBER SPECIES=1-512 SWITCH={EXP D_0=4.0 R_0=0.5}
cf: MTRANSFORM_MORE DATA=c1 SWITCH={RATIONAL D_0=2.0 R_0=0.1} LOWMEM
cc: CENTER_OF_MULTICOLVAR DATA=cf
```

This input once again calculates the coordination numbers of all the atoms in the system. The middle line then transforms these coordinations numbers to numbers between 0 and 1. Essentially any atom with a coordination number larger than 2.0 is given a weight of one and below this value the transformed value decays to zero. It is these transformed coordination numbers that are used to calculate the Berry phase average described in the previous section.

### 5.1.10 CENTER

This is part of the <a href="#">vatom module</a>
--------------------------------------------------

Calculate the center for a group of atoms, with arbitrary weights.

The computed center is stored as a virtual atom that can be accessed in an atom list through the label for the CENTER action that creates it. Notice that the generated virtual atom has charge equal to the sum of the charges and mass equal to the sum of the masses. If used with the MASS flag, then it provides a result identical to [COM](#).

When running with periodic boundary conditions, the atoms should be in the proper periodic image. This is done automatically since PLUMED 2.2, by considering the ordered list of atoms and rebuilding PBCs with a procedure that is equivalent to that done in [WHOLEMOLECULES](#). Notice that rebuilding is local to this action. This is different from [WHOLEMOLECULES](#) which actually modifies the coordinates stored in PLUMED.

In case you want to recover the old behavior you should use the NOPBC flag. In that case you need to take care that atoms are in the correct periodic image.

The atoms involved can be specified using

<b>ATOMS</b>	the list of atoms which are involved the virtual atom's definition. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>MASS</b>	( default=off ) If set center is mass weighted
<b>WEIGHTS</b>	Center is computed as a weighted average.

## Examples

```
# a point which is on the line connecting atoms 1 and 10, so that its distance
# from 10 is twice its distance from 1:
c1: CENTER ATOMS=1,1,10
# this is another way of stating the same:
c1bis: CENTER ATOMS=1,10 WEIGHTS=2,1

# center of mass among these atoms:
c2: CENTER ATOMS=2,3,4,5 MASS

d1: DISTANCE ATOMS=c1,c2

PRINT ARG=d1
```

(See also [DISTANCE](#), [COM](#) and [PRINT](#)).

## 5.1.11 COM

This is part of the <a href="#">vatom module</a>
--------------------------------------------------

Calculate the center of mass for a group of atoms.

The computed center of mass is stored as a virtual atom that can be accessed in an atom list through the label for the COM action that creates it.

For arbitrary weights (e.g. geometric center) see [CENTER](#).

When running with periodic boundary conditions, the atoms should be in the proper periodic image. This is done automatically since PLUMED 2.2, by considering the ordered list of atoms and rebuilding PBCs with a procedure that is equivalent to that done in [WHOLEMOLECULES](#). Notice that rebuilding is local to this action. This is different from [WHOLEMOLECULES](#) which actually modifies the coordinates stored in PLUMED.

In case you want to recover the old behavior you should use the NOPBC flag. In that case you need to take care that atoms are in the correct periodic image.

The atoms involved can be specified using

<b>ATOMS</b>	the list of atoms which are involved the virtual atom's definition. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
--------------	------------------------------------------------------------------------------------

## Examples

The following input instructs plumed to print the distance between the center of mass for atoms 1,2,3,4,5,6,7 and that for atoms 15,20:

```
COM ATOMS=1-7          LABEL=c1
COM ATOMS=15,20        LABEL=c2
DISTANCE ATOMS=c1,c2   LABEL=d1
PRINT ARG=d1
```

(See also [DISTANCE](#) and [PRINT](#)).

## 5.1.12 FIXEDATOM

This is part of the <a href="#">vatom module</a>
--------------------------------------------------

Add a virtual atom in a fixed position.

This action creates a virtual atom at a fixed position. The coordinates can be specified in cartesian components (by default) or in scaled coordinates (SCALED\_COMPONENTS). It is also possible to assign a predefined charge or mass to the atom.

## Attention

Similar to [POSITION](#) this variable is not invariant for translation of the system. Adding a force on it can create serious troubles.

Notice that the distance between to atoms created using FIXEDATOM is invariant for translation. Additionally, if one first align atoms to a reference using [FIT\\_TO\\_TEMPLATE](#), then it is safe to add further fixed atoms without breaking translational invariance.

The atoms involved can be specified using

<b>ATOMS</b>	the list of atoms which are involved the virtual atom's definition. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Compulsory keywords

<b>AT</b>	coordinates of the virtual atom
<b>SET_MASS</b>	( default=1 ) mass of the virtual atom
<b>SET_CHARGE</b>	( default=0 ) charge of the virtual atom

### Options

<b>SCALED_COMPONENTS</b>	( default=off ) use scaled components
--------------------------	---------------------------------------

### Examples

The following input instructs plumed to compute the angle between distance of atoms 15 and 20 and the z axis and keeping it close to zero.

```
a: FIXEDATOM AT=0,0,0
b: FIXEDATOM AT=0,0,1
an: ANGLE ATOMS=a,b,15,20
RESTRAINT ARG=an AT=0.0 KAPPA=100.0
```

(See also [ANGLE](#) and [RESTRAINT](#)).

The following input instructs plumed to align a protein on a template and then compute the distance of one of its atom from the point (10,20,30).

```
FIT_TO_TEMPLATE STRIDE=1 REFERENCE=ref.pdb TYPE=SIMPLE
a: FIXEDATOM AT=10,20,30
d: DISTANCE ATOMS=a,20
PRINT ARG=d FILE=colvar
```

(See also [FIT\\_TO\\_TEMPLATE](#) and [DISTANCE](#)).

### 5.1.13 GHOST

	This is part of the <a href="#">vatom module</a>
--	--------------------------------------------------

Calculate the absolute position of a ghost atom with fixed coordinates in the local reference frame formed by three atoms. The computed ghost atom is stored as a virtual atom that can be accessed in an atom list through the the label for the GHOST action that creates it.

The atoms involved can be specified using

<b>ATOMS</b>	the list of atoms which are involved the virtual atom's definition. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
<b>COORDINATES</b>	coordinates of the ghost atom in the local reference frame. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>

## Examples

The following input instructs plumed to print the distance between the ghost atom and the center of mass for atoms 15,20:

```
GHOST ATOMS=1,5,10 COORDINATES=10.0,10.0,10.0 LABEL=c1
COM ATOMS=15,20 LABEL=c2
DISTANCE ATOMS=c1,c2 LABEL=d1
PRINT ARG=d1
```

(See also [DISTANCE](#) and [PRINT](#)).

## 5.2 CV Documentation

The following list contains descriptions of a number of the colvars that are currently implemented in PLUMED.

<a href="#">ALPHABETA</a>	Measures a distance including pbc between the instantaneous values of a set of torsional angles and set of reference values.
<a href="#">ALPHARMSD</a>	Probe the alpha helical content of a protein structure.
<a href="#">ANGLE</a>	Calculate an angle.
<a href="#">ANTIBETARMSD</a>	Probe the antiparallel beta sheet content of your protein structure.
<a href="#">CELL</a>	Calculate the components of the simulation cell
<a href="#">CONSTANT</a>	Return one or more constant quantities with or without derivatives.
<a href="#">CONTACTMAP</a>	Calculate the distances between a number of pairs of atoms and transform each distance by a switching function. The transformed distance can be compared with a reference value in order to calculate the squared distance between two contact maps. Each distance can also be weighted for a given value. CONTACTMAP can be used together with <a href="#">FUNCPATHMSD</a> to define a path in the contactmap space.
<a href="#">COORDINATION</a>	Calculate coordination numbers.
<a href="#">CS2BACKBONE</a>	This collective variable calculates the backbone chemical shifts for a protein.
<a href="#">DHENERGY</a>	Calculate Debye-Huckel interaction energy among GROUPA and GROUPB.
<a href="#">DIHCOR</a>	Measures the degree of similarity between dihedral angles.
<a href="#">DIPOLE</a>	Calculate the dipole moment for a group of atoms.
<a href="#">DISTANCE_FROM_CONTOUR</a>	Calculate the perpendicular distance from a Willard-Chandler dividing surface.
<a href="#">DISTANCE</a>	Calculate the distance between a pair of atoms.
<a href="#">ENERGY</a>	Calculate the total energy of the simulation box.
<a href="#">ERMSD</a>	Calculate eRMSD with respect to a reference structure.
<a href="#">FAKE</a>	This is a fake colvar container used by cltools or various other actions and just support input and period definition
<a href="#">FRET</a>	Calculate the FRET efficiency between a pair of atoms. The efficiency is calculated using the Forster relation:
<a href="#">GPROPERTYMAP</a>	Property maps but with a more flexible framework for the distance metric being used.
<a href="#">GYRATION</a>	Calculate the radius of gyration, or other properties related to it.
<a href="#">JCOUPLING</a>	Calculates $^3J$ coupling constants for a dihedral angle.
<a href="#">NOE</a>	Calculates NOE intensities as sums of $1/r^6$ , also averaging over multiple equivalent atoms or ambiguous NOE.
<a href="#">PARABETARMSD</a>	Probe the parallel beta sheet content of your protein structure.

<a href="#">PATHMSD</a>	This Colvar calculates path collective variables.
<a href="#">PATH</a>	Path collective variables with a more flexible framework for the distance metric being used.
<a href="#">PCAVARS</a>	Projection on principal component eigenvectors or other high dimensional linear subspace
<a href="#">POSITION</a>	Calculate the components of the position of an atom.
<a href="#">PRE</a>	Calculates the Paramagnetic Resonance Enhancement intensity ratio between two atoms. The reference atom for the spin label is added with SPI↔NLABEL, the affected atom(s) are given as numbered GROUPA1, GROUPA2, ... The additional parameters needed for the calculation are given as INEPT, the inepttime, TAUC the correlation time, OMEGA, the larmor frequency and RTWO for the relaxation time.
<a href="#">PROPERTYMAP</a>	Calculate generic property maps.
<a href="#">PUCKERING</a>	Calculate sugar pseudorotation coordinates.
<a href="#">RDC</a>	Calculates the (Residual) Dipolar Coupling between two atoms.
<a href="#">TEMPLATE</a>	This file provides a template for if you want to introduce a new CV.
<a href="#">TORSION</a>	Calculate a torsional angle.
<a href="#">VOLUME</a>	Calculate the volume of the simulation box.

### 5.2.1 ALPHABETA

<b>This is part of the multicolvar <a href="#">module</a></b>
---------------------------------------------------------------

Measures a distance including pbc between the instantaneous values of a set of torsional angles and set of reference values.

This colvar calculates the following quantity.

$$s = \frac{1}{2} \sum_i \left[ 1 + \cos(\phi_i - \phi_i^{\text{Ref}}) \right]$$

where the  $\phi_i$  values are the instantaneous values for the [TORSION](#) angles of interest. The  $\phi_i^{\text{Ref}}$  values are the user-specified reference values for the torsional angles.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM↔S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>REFERENCE</b>	the reference values for each of the torsional angles. If you use a single REFERENCE value the same reference value is used for all torsions You can use multiple instances of this keyword i.e. REFERENCE1, REFERENCE2, REFERENCE3...
------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

### Examples

The following provides an example of the input for an alpha beta similarity.

```
ALPHABETA ...
ATOMS1=168,170,172,188 REFERENCE1=3.14
ATOMS2=170,172,188,190 REFERENCE2=3.14
ATOMS3=188,190,192,230 REFERENCE3=3.14
LABEL=ab
... ALPHABETA
PRINT ARG=ab FILE=colvar STRIDE=10
```

Because all the reference values are the same we can calculate the same quantity using

```
ALPHABETA ...
ATOMS1=168,170,172,188 REFERENCE=3.14
ATOMS2=170,172,188,190
ATOMS3=188,190,192,230
LABEL=ab
... ALPHABETA
PRINT ARG=ab FILE=colvar STRIDE=10
```

Writing out the atoms involved in all the torsions in this way can be rather tedious. Thankfully if you are working with protein you can avoid this by using the [MOLINFO](#) command. PLUMED uses the pdb file that you provide to this command to learn about the topology of the protein molecule. This means that you can specify torsion angles using the following syntax:

```
MOLINFO MOLTYPE=protein STRUCTURE=myprotein.pdb
ALPHABETA ...
ATOMS1=@phi-3 REFERENCE=3.14
ATOMS2=@psi-3
ATOMS3=@phi-4
LABEL=ab
... ALPHABETA
PRINT ARG=ab FILE=colvar STRIDE=10
```

Here, @phi-3 tells plumed that you would like to calculate the  $\phi$  angle in the third residue of the protein. Similarly @psi-4 tells plumed that you want to calculate the  $\psi$  angle of the 4th residue of the protein.

## 5.2.2 ALPHARMSD

This is part of the secondary structure module
------------------------------------------------

Probe the alpha helical content of a protein structure.

Any chain of six contiguous residues in a protein chain can form an alpha helix. This colvar thus generates the set of all possible six residue sections and calculates the RMSD distance between the configuration in which the residues find themselves and an idealized alpha helical structure. These distances can be calculated by either aligning the instantaneous structure with the reference structure and measuring each atomic displacement or by calculating differences between the set of interatomic distances in the reference and instantaneous structures.

This colvar is based on the following reference [4]. The authors of this paper use the set of distances from the alpha helix configurations to measure the number of segments that have an alpha helical configuration. This is done by calculating the following sum of functions of the rmsd distances:

$$s = \sum_i \frac{1 - \left( \frac{r_i - d_0}{r_0} \right)^n}{1 - \left( \frac{r_i - d_0}{r_0} \right)^m}$$

where the sum runs over all possible segments of alpha helix. By default the NN, MM and D\_0 parameters are set equal to those used in [4]. The R\_0 parameter must be set by the user - the value used in [4] was 0.08 nm.

If you change the function in the above sum you can calculate quantities such as the average distance from a purely the alpha helical configuration or the distance between the set of residues that is closest to an alpha helix and the reference configuration. To do these sorts of calculations you can use the AVERAGE and MIN keywords. In addition you can use the LESS\_THAN keyword if you would like to change the form of the switching function. If you use any of these options you no longer need to specify NN, R\_0, MM and D\_0.

Please be aware that for codes like gromacs you must ensure that plumed reconstructs the chains involved in your CV when you calculate this CV using anything other than TYPE=DRMSD. For more details as to how to do this see [WHOLEMOLECULES](#).

#### Description of components

By default this Action calculates the number of structural units that are within a certain distance of a idealised secondary structure element. This quantity can then be referenced elsewhere in the input by using the label of the action. However, this Action can also be used to calculate the following quantities by using the keywords as described below. The quantities then calculated can be referenced using the label of the action followed by a dot and then the name from the table below. Please note that you can use the LESS\_THAN keyword more than once. The resulting components will be labelled *label.lessthan-1*, *label.lessthan-2* and so on unless you exploit the fact that these labels are customizable. In particular, by using the LABEL keyword in the description of you LESS\_THAN function you can set name of the component that you are calculating

Quantity	Keyword	Description
altmin	ALT_MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
highest	HIGHEST	the lowest of the quantities calculated by this action
lessthan	LESS_THAN	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
lowest	LOWEST	the lowest of the quantities calculated by this action
min	MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.

Generated by Doxygen



The atoms involved can be specified using

<b>RESIDUES</b>	this command is used to specify the set of residues that could conceivably form part of the secondary structure. It is possible to use residues numbers as the various chains and residues should have been identified else using an instance of the <a href="#">MOLINFO</a> action. If you wish to use all the residues from all the chains in your system you can do so by specifying all. Alternatively, if you wish to use a subset of the residues you can specify the particular residues you are interested in as a list of numbers. Please be aware that to form secondary structure elements your chain must contain at least N residues, where N is dependent on the particular secondary structure you are interested in. As such if you define portions of the chain with fewer than N residues the code will crash.
-----------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Compulsory keywords

<b>TYPE</b>	( default=DRMSD ) the manner in which RMSD alignment is performed. Should be OPTIMAL, SIMPLE or DRMSD. For more details on the OPTIMAL and SIMPLE methods see <a href="#">RMSD</a> . For more details on the DRMSD method see <a href="#">DRMSD</a> .
<b>R_0</b>	The r_0 parameter of the switching function.
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>NN</b>	( default=8 ) The n parameter of the switching function
<b>MM</b>	( default=12 ) The m parameter of the switching function

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>VERBOSE</b>	( default=off ) write a more detailed output
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...

<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

### Examples

The following input calculates the number of six residue segments of protein that are in an alpha helical configuration.

```
MOLINFO STRUCTURE=helix.pdb
ALPHARMSD RESIDUES=all TYPE=DRMSD LESS_THAN={RATIONAL R_0=0.08 NN=8 MM=12} LABEL=a
```

(see also [MOLINFO](#))

### 5.2.3 ANGLE

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

Calculate an angle.

This command can be used to compute the angle between three atoms. Alternatively if four atoms appear in the atom specification it calculates the angle between two vectors identified by two pairs of atoms.

If *three* atoms are given, the angle is defined as:

$$\theta = \arccos \left( \frac{\mathbf{r}_{21} \cdot \mathbf{r}_{23}}{|\mathbf{r}_{21}| |\mathbf{r}_{23}|} \right)$$

Here  $\mathbf{r}_{ij}$  is the distance vector among the i-th and the j-th listed atom.

If *four* atoms are given, the angle is defined as:

$$\theta = \arccos \left( \frac{\mathbf{r}_{21} \cdot \mathbf{r}_{34}}{|\mathbf{r}_{21}| |\mathbf{r}_{34}|} \right)$$

Notice that angles defined in this way are non-periodic variables and their value is limited by definition between 0 and  $\pi$ .

The vectors  $\mathbf{r}_{ij}$  are by default evaluated taking periodic boundary conditions into account. This behavior can be changed with the NOPBC flag.

The atoms involved can be specified using

<b>ATOMS</b>	the list of atoms involved in this collective variable (either 3 or 4 atoms). For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances

## Examples

This command tells plumed to calculate the angle between the vector connecting atom 1 to atom 2 and the vector connecting atom 2 to atom 3 and to print it on file COLVAR1. At the same time, the angle between vector connecting atom 1 to atom 2 and the vector connecting atom 3 to atom 4 is printed on file COLVAR2.

```
a: ANGLE ATOMS=1,2,3
# equivalently one could state:
# a: ANGLE ATOMS=1,2,2,3

b: ANGLE ATOMS=1,2,3,4

PRINT ARG=a FILE=COLVAR1
PRINT ARG=b FILE=COLVAR2
```

(see also [PRINT](#))

## 5.2.4 ANTIBETARMSD

This is part of the <a href="#">secondarystructure module</a>
---------------------------------------------------------------

Probe the antiparallel beta sheet content of your protein structure.

Two protein segments containing three contiguous residues can form an antiparallel beta sheet. Although if the two segments are part of the same protein chain they must be separated by a minimum of 2 residues to make room for the turn. This colvar thus generates the set of all possible six residue sections that could conceivably form an antiparallel beta sheet and calculates the RMSD distance between the configuration in which the residues find themselves and an idealized antiparallel beta sheet structure. These distances can be calculated by either aligning the instantaneous structure with the reference structure and measuring each atomic displacement or by calculating differences between the set of interatomic distances in the reference and instantaneous structures.

This colvar is based on the following reference [4]. The authors of this paper use the set of distances from the antiparallel beta sheet configurations to measure the number of segments that have an configuration that resembles an antiparallel beta sheet. This is done by calculating the following sum of functions of the rmsd distances:

$$s = \sum_i \frac{1 - \left( \frac{r_i - d_0}{r_0} \right)^n}{1 - \left( \frac{r_i - d_0}{r_0} \right)^m}$$

where the sum runs over all possible segments of antiparallel beta sheet. By default the NN, MM and D\_0 parameters are set equal to those used in [4]. The R\_0 parameter must be set by the user - the value used in [4] was 0.08 nm.

If you change the function in the above sum you can calculate quantities such as the average distance from a purely configuration composed of pure anti-parallel beta sheets or the distance between the set of residues that is closest to an anti-parallel beta sheet and the reference configuration. To do these sorts of calculations you can use the AVERAGE and MIN keywords. In addition you can use the LESS\_THAN keyword if you would like to change the form of the switching function. If you use any of these options you no longer need to specify NN, R\_0, MM and D\_0.

Please be aware that for codes like gromacs you must ensure that plumed reconstructs the chains involved in your CV when you calculate this CV using anything other than TYPE=DRMSD. For more details as to how to do this see [WHOLEMOLECULES](#).

### Description of components

By default this Action calculates the number of structural units that are within a certain distance of a idealised secondary structure element. This quantity can then be referenced elsewhere in the input by using the label of the action. However, this Action can also be used to calculate the following quantities by using the keywords as described below. The quantities then calculated can be referenced using the label of the action followed by a dot and then the name from the table below. Please note that you can use the LESS\_THAN keyword more than once. The resulting components will be labelled *label.less-than-1*, *label.less-than-2* and so on unless you exploit the fact that these labels are customizable. In particular, by using the LABEL keyword in the description of you LESS\_THAN function you can set name of the component that you are calculating

Quantity	Keyword	Description
altmin	ALT_MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
highest	HIGHEST	the lowest of the quantities calculated by this action
lessthan	LESS_THAN	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
lowest	LOWEST	the lowest of the quantities calculated by this action
min	MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.

The atoms involved can be specified using

<b>RESIDUES</b>	this command is used to specify the set of residues that could conceivably form part of the secondary structure. It is possible to use residues numbers as the various chains and residues should have been identified else using an instance of the <a href="#">MOLINFO</a> action. If you wish to use all the residues from all the chains in your system you can do so by specifying all. Alternatively, if you wish to use a subset of the residues you can specify the particular residues you are interested in as a list of numbers. Please be aware that to form secondary structure elements your chain must contain at least N residues, where N is dependent on the particular secondary structure you are interested in. As such if you define portions of the chain with fewer than N residues the code will crash.
-----------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Compulsory keywords

<b>TYPE</b>	( default=DRMSD ) the manner in which RMSD alignment is performed. Should be OPTIMAL, S↔IMPLE or DRMSD. For more details on the OPTIMAL and SIMPLE methods see <a href="#">RMSD</a> . For more details on the DRMSD method see <a href="#">DRMSD</a> .
-------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<b>R_0</b>	The r_0 parameter of the switching function.
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>NN</b>	( default=8 ) The n parameter of the switching function
<b>MM</b>	( default=12 ) The m parameter of the switching function
<b>STYLE</b>	( default=all ) Antiparallel beta sheets can either form in a single chain or from a pair of chains. If STYLE=all all chain configuration with the appropriate geometry are counted. If STYLE=inter only sheet-like configurations involving two chains are counted, while if STYLE=intra only sheet-like configurations involving a single chain are counted

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>VERBOSE</b>	( default=off ) write a more detailed output
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp(\frac{\beta}{s_i})}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>STRANDS_CUTOFF</b>	If in a segment of protein the two strands are further apart then the calculation of the actual RMSD is skipped as the structure is very far from being beta-sheet like. This keyword speeds up the calculation enormously when you are using the LESS_THAN option. However, if you are using some other option, then this cannot be used

## Examples

The following input calculates the number of six residue segments of protein that are in an antiparallel beta sheet configuration.

```
MOLINFO STRUCTURE=helix.pdb
ANTIBETARMSD RESIDUES=all TYPE=DRMSD LESS_THAN={RATIONAL R_0=0.08 NN=8 MM=12} LABEL=a
```

(see also [MOLINFO](#))

### 5.2.5 CELL

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

Calculate the components of the simulation cell

#### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>ax</b>	the ax component of the cell matrix
<b>ay</b>	the ay component of the cell matrix
<b>az</b>	the az component of the cell matrix
<b>bx</b>	the bx component of the cell matrix
<b>by</b>	the by component of the cell matrix
<b>bz</b>	the bz component of the cell matrix
<b>cx</b>	the cx component of the cell matrix
<b>cy</b>	the cy component of the cell matrix
<b>cz</b>	the cz component of the cell matrix

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
------------------------------	----------------------------------------------------------------------------

#### Examples

The following input tells plumed to print the squared modulo of each of the three lattice vectors

```
cell: CELL
aaa: COMBINE ARG=cell.ax,cell.ay,cell.az POWERS=2,2,2 PERIODIC=NO
bbb: COMBINE ARG=cell.bx,cell.by,cell.bz POWERS=2,2,2 PERIODIC=NO
ccc: COMBINE ARG=cell.cx,cell.cy,cell.cz POWERS=2,2,2 PERIODIC=NO
PRINT ARG=aaa,bbb,ccc
```

(See also [COMBINE](#) and [PRINT](#)).

### 5.2.6 CONSTANT

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

Return one or more constant quantities with or without derivatives.

Useful in combination with functions that takes in input constants or parameters.

### Description of components

The names of the components in this action can be customized by the user in the actions input file. However, in addition to these customizable components the following quantities will always be output

Quantity	Description
<b>v</b>	the # value

### Compulsory keywords

<b>VALUES</b>	( default=NAN ) The values of the constants
<b>VALUE</b>	( default=NAN ) The value of the constant

### Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>NODE↔ RIV</b>	( default=off ) Set to TRUE if you want values without derivatives.

### Examples

The following input instructs plumed to compute the distance between atoms 1 and 2. If this distance is between 1.0 and 2.0, it is printed. If it is lower than 1.0 (larger than 2.0), 1.0 (2.0) is printed

```
cn: CONSTANT VALUES=1.0,2.0
dis: DISTANCE ATOMS=1,2
sss: SORT ARG=cn.v_0,dis,cn.v_1
PRINT ARG=sss.2
```

(See also [DISTANCE](#), [SORT](#), and [PRINT](#)).

In case you want to pass a single value you can use VALUE:

```
cn: CONSTANT VALUE=1.0
dis: DISTANCE ATOMS=1
sss: SORT ARG=cn,dis
PRINT ARG=sss.1
```

## 5.2.7 CONTACTMAP

This is part of the <b>colvar</b> <a href="#">module</a>
----------------------------------------------------------

Calculate the distances between a number of pairs of atoms and transform each distance by a switching function. The transformed distance can be compared with a reference value in order to calculate the squared distance between two contact maps. Each distance can also be weighted for a given value. CONTACTMAP can be used together with [FUNCPATHMSD](#) to define a path in the contactmap space.

The individual contact map distances related to each contact can be accessed as components named `cm.contact-1`, `cm.contact-2`, etc, assuming that the label of the CONTACTMAP is `cm`.

### Description of components

By default the value of the calculated quantity can be referenced elsewhere in the input file by using the label of the action. Alternatively this Action can be used to calculate the following quantities by employing the keywords listed below. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>contact</b>	By not using SUM or CMDIST each contact will be stored in a component

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the contacts you wish to calculate. Keywords like ATOMS1, ATOMS2, ATOMS3,... should be listed and one contact will be calculated for each ATOM keyword you specify. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Compulsory keywords

<b>SWITCH</b>	The switching functions to use for each of the contacts in your map. You can either specify a global switching function using SWITCH or one switching function for each contact. Details of the various switching functions you can use are provided on <a href="#">switchingfunction</a> . You can use multiple instances of this keyword i.e. SWITCH1, SWITCH2, SWITCH3...
---------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SUM</b>	( default=off ) calculate the sum of all the contacts in the input
<b>CMDIST</b>	( default=off ) calculate the distance with respect to the provided reference contact map
<b>SERIAL</b>	( default=off ) Perform the calculation in serial - for debug purpose



<b>REFERENCE</b>	A reference value for a given contact, by default is 0.0 You can either specify a global reference value using REFERENCE or one reference value for each contact. You can use multiple instances of this keyword i.e. REFERENCE1, REFERENCE2, REFERENCE3...
<b>WEIGHT</b>	A weight value for a given contact, by default is 1.0 You can either specify a global weight value using WEIGHT or one weight value for each contact. You can use multiple instances of this keyword i.e. WEIGHT1, WEIGHT2, WEIGHT3...

### Examples

The following example calculates switching functions based on the distances between atoms 1 and 2, 3 and 4 and 4 and 5. The values of these three switching functions are then output to a file named colvar.

```
CONTACTMAP ATOMS1=1,2 ATOMS2=3,4 ATOMS3=4,5 ATOMS4=5,6 SWITCH={RATIONAL R_0=1.5} LABEL=f1
PRINT ARG=f1.* FILE=colvar
```

The following example calculates the difference of the current contact map with respect to a reference provided. In this case REFERENCE is the fraction of contact that is formed (i.e. the distance between two atoms transformed with the SWITCH), while R\_0 is the contact distance. WEIGHT gives the relative weight of each contact to the final distance measure.

```
CONTACTMAP ...
ATOMS1=1,2 REFERENCE1=0.1 WEIGHT1=0.5
ATOMS2=3,4 REFERENCE2=0.5 WEIGHT2=1.0
ATOMS3=4,5 REFERENCE3=0.25 WEIGHT3=1.0
ATOMS4=5,6 REFERENCE4=0.0 WEIGHT4=0.5
SWITCH={RATIONAL R_0=1.5}
LABEL=cmap
CMDIST
... CONTACTMAP

PRINT ARG=cmap FILE=colvar
```

The next example calculates fraction of native contacts (Q) for Trp-cage mini-protein. R\_0 is the distance at which the switch function is guaranteed to be 1.0 – it doesn't really matter for Q and should be something very small, like 1 Å. REF is the reference distance for the contact, e.g. the distance from a crystal structure. LAMBDA is the tolerance for the distance – if set to 1.0, the contact would have to have exactly the reference value to be formed; instead for lambda values of 1.5–1.8 are usually used to allow some slack. BETA is the softness of the switch function, default is 50nm. WEIGHT is the 1/(number of contacts) giving equal weight to each contact.

When using native contact Q switch function, please cite [5]

```
# Full example available in regtest/basic/rt72/

CONTACTMAP ...
ATOMS1=1,67 SWITCH1={Q R_0=0.01 BETA=50.0 LAMBDA=1.5 REF=0.4059} WEIGHT1=0.003597
ATOMS2=1,68 SWITCH2={Q R_0=0.01 BETA=50.0 LAMBDA=1.5 REF=0.4039} WEIGHT2=0.003597
ATOMS3=1,69 SWITCH3={Q R_0=0.01 BETA=50.0 LAMBDA=1.5 REF=0.3215} WEIGHT3=0.003597
[snip]
ATOMS275=183,213 SWITCH275={Q R_0=0.01 BETA=50.0 LAMBDA=1.5 REF=0.355} WEIGHT275=0.003597
ATOMS276=183,234 SWITCH276={Q R_0=0.01 BETA=50.0 LAMBDA=1.5 REF=0.428} WEIGHT276=0.003597
ATOMS277=183,250 SWITCH277={Q R_0=0.01 BETA=50.0 LAMBDA=1.5 REF=0.3832} WEIGHT277=0.003597
ATOMS278=197,220 SWITCH278={Q R_0=0.01 BETA=50.0 LAMBDA=1.5 REF=0.3827} WEIGHT278=0.003597
LABEL=cmap
SUM
... CONTACTMAP

PRINT ARG=cmap FILE=colvar
```

(See also [PRINT](#) and [switchingfunction](#))

## 5.2.8 COORDINATION

This is part of the **colvar** module

Calculate coordination numbers.

This keyword can be used to calculate the number of contacts between two groups of atoms and is defined as

$$\sum_{i \in A} \sum_{j \in B} s_{ij}$$

where  $s_{ij}$  is 1 if the contact between atoms  $i$  and  $j$  is formed, zero otherwise. In practise,  $s_{ij}$  is replaced with a switching function to make it differentiable. The default switching function is:

$$s_{ij} = \frac{1 - \left( \frac{\mathbf{r}_{ij} - d_0}{r_0} \right)^n}{1 - \left( \frac{\mathbf{r}_{ij} - d_0}{r_0} \right)^m}$$

but it can be changed using the optional SWITCH option.

To make your calculation faster you can use a neighbor list, which makes it that only a relevant subset of the pairwise distance are calculated at every step.

If GROUPB is empty, it will sum the  $\frac{N(N-1)}{2}$  pairs in GROUPA. This avoids computing twice permuted indexes (e.g. pair (i,j) and (j,i)) thus running at twice the speed.

Notice that if there are common atoms between GROUPA and GROUPB the switching function should be equal to one. These "self contacts" are discarded by plumed (since version 2.1), so that they actually count as "zero".

The atoms involved can be specified using

<b>GROUPA</b>	First list of atoms. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
<b>GROUPB</b>	Second list of atoms (if empty, $N*(N-1)/2$ pairs in GROUPA are counted). For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>

Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D<sub>↔</sub> _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R<sub>↔</sub> _0</b>	The r_0 parameter of the switching function

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) Perform the calculation in serial - for debug purpose
<b>PAIR</b>	( default=off ) Pair only 1st element of the 1st group with 1st element in the second, etc
<b>NLIST</b>	( default=off ) Use a neighbour list to speed up the calculation
<b>NL_CUTOFF</b>	The cutoff for the neighbour list
<b>NL_STRIDE</b>	The frequency with which we are updating the atoms in the neighbour list
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.

### Examples

The following example instructs plumed to calculate the total coordination number of the atoms in group 1-10 with the atoms in group 20-100. For atoms 1-10 coordination numbers are calculated that count the number of atoms from the second group that are within 0.3 nm of the central atom. A neighbour list is used to make this calculation faster, this neighbour list is updated every 100 steps.

```
COORDINATION GROUPA=1-10 GROUPB=20-100 R_0=0.3 NLIST NL_CUTOFF=0.5 NL_STRIDE=100
```

The following is a dummy example which should compute the value 0 because the self interaction of atom 1 is skipped. Notice that in plumed 2.0 "self interactions" were not skipped, and the same calculation should return 1.

```
c: COORDINATION GROUPA=1 GROUPB=1 R_0=0.3
PRINT ARG=c STRIDE=10
```

```
c1: COORDINATION GROUPA=1-10 GROUPB=1-10 R_0=0.3
x: COORDINATION GROUPA=1-10 R_0=0.3
c2: COMBINE ARG=x COEFFICIENTS=2
# the two variables c1 and c2 should be identical, but the calculation of c2 is twice faster
# since it runs on half of the pairs. Notice that to get the same result you
# should double it
PRINT ARG=c1,c2 STRIDE=10
```

See also [PRINT](#) and [COMBINE](#)

### 5.2.9 CS2BACKBONE

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

This collective variable calculates the backbone chemical shifts for a protein.

The functional form is that of CamShift [6]. The chemical shifts of the selected nuclei/residues are saved as components. Reference experimental values can also be stored as components. The two components can then be used

to calculate either a scoring function as in [7] [8], using the keyword CAMSHIFT or to calculate ensemble averages chemical shift per chemical shift as in [9] [10] (see [STATS](#) and [ENSEMBLE](#)).

CamShift calculation is relatively heavy because it often uses a large number of atoms, in order to make it faster it is currently parallelised with [OpenMP](#).

As a general rule, when using [CS2BACKBONE](#) or other experimental restraints it is better to increase the accuracy of the constraint algorithm due to the increased strain on the bonded structure. In the case of GROMACS it is safer to use lincs-iter=2 and lincs-order=6.

In general the system for which chemical shifts are calculated must be completely included in ATOMS and a TEMP↔LATE pdb file for the same atoms should be provided as well in the folder DATA. The atoms are made automatically whole unless NOPBC is used, in particular if the system is made of by multiple chains it is usually better to use NOPBC and make the molecule whole [WHOLEMOLECULES](#) selecting an appropriate order.

In addition to a pdb file one needs to provide a list of chemical shifts to be calculated using one file per nucleus type (CAshifts.dat, CBshifts.dat, Cshifts.dat, Hshifts.dat, HAshts.dat, Nshifts.dat), all the six files should always be present. A chemical shift for a nucleus is calculated if a value greater than 0 is provided. For practical purposes the value can correspond to the experimental value. Residues numbers should go from 1 to N irrespectively of the numbers used in the pdb file. The first and last residue of each chain should be preceded by a # character. Termini groups like ACE or NME should be removed from the PDB.

```
CAshifts.dat:
#1 0.0
2 55.5
3 58.4
.
.
#last 0.0
#last+1 (first) of second chain
.
#last of second chain
```

The default behaviour is to store the values for the active nuclei in components (ca\_#, cb\_#, co\_#, ha\_#, hn\_#, nh\_# and expca\_#, expcb\_#, expco\_#, expha\_#, exphn\_#, exp\_nh#) with NOEXP it is possible to only store the backcalculated values.

A pdb file is needed to generate a simple topology of the protein. For histidines in protonation states different from D the HIE/HSE HIP/HSP name should be used. GLH and ASH can be used for the alternative protonation of GLU and ASP. Non-standard amino acids and other molecules are not yet supported, but in principle they can be named UNK. If multiple chains are present the chain identifier must be in the standard PDB format, together with the TER keyword at the end of each chain.

One more standard file is also needed in the folder DATA: camshift.db. This file includes all the CamShift parameters and can be found in regtest/basic/rt45/data/.

All the above files must be in a single folder that must be specified with the keyword DATA.

Additional material and examples can be also found in the tutorial [Belfast tutorial: NMR restraints](#)

## Description of components

The names of the components in this action can be customized by the user in the actions input file. However, in addition to these customizable components the following quantities will always be output

Quantity	Description
<b>ha</b>	the calculated Ha hydrogen chemical shifts
<b>hn</b>	the calculated H hydrogen chemical shifts
<b>nh</b>	the calculated N nitrogen chemical shifts
<b>ca</b>	the calculated Ca carbon chemical shifts
<b>cb</b>	the calculated Cb carbon chemical shifts
<b>co</b>	the calculated C' carbon chemical shifts
<b>expha</b>	the experimental Ha hydrogen chemical shifts
<b>exphn</b>	the experimental H hydrogen chemical shifts
<b>expnh</b>	the experimental N nitrogen chemical shifts
<b>expca</b>	the experimental Ca carbon chemical shifts
<b>expcb</b>	the experimental Cb carbon chemical shifts
<b>expco</b>	the experimental C' carbon chemical shifts

The atoms involved can be specified using

<b>ATOMS</b>	The atoms to be included in the calculation, e.g. the whole protein.. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>DATA</b>	( default=data/ ) The folder with the experimental chemical shifts.
<b>TEMPLATE</b>	( default=template.pdb ) A PDB file of the protein system to initialise ALMOST.
<b>NEIGH_FREQ</b>	( default=20 ) Period in step for neighbour list update.
<b>NRES</b>	Number of residues, corresponding to the number of chemical shifts.

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>CAMSHIFT</b>	( default=off ) Set to TRUE if you to calculate a single CamShift score.
<b>NOEXP</b>	( default=off ) Set to TRUE if you don't want to have fixed components with the experimental values.

Examples

In this first example the chemical shifts are used to calculate a scoring function to be used in NMR driven Metadynamics [8] :

```
whole: GROUP ATOMS=2612-2514:-1,961-1:-1,2466-962:-1,2513-2467:-1
```

```
WHOLEMOLECULES ENTITY0=whole
cs: CS2BACKBONE ATOMS=1-2612 NRES=176 DATA=./data/ TEMPLATE=template.pdb CAMSHIFT NOPBC
metad: METAD ARG=cs HEIGHT=0.5 SIGMA=0.1 PACE=200 BIASFACTOR=10
PRINT ARG=cs,metad.bias FILE=COLVAR STRIDE=100
```

In this second example the chemical shifts are used as replica-averaged restrained as in [9] [10].

```
cs: CS2BACKBONE ATOMS=1-174 DATA=data/ NRES=13
encs: ENSEMBLE ARG=(cs\hn_*), (cs\nh_*)
stcs: STATS ARG=encs.* SQDEVSUM PARARG=(cs\exphn_*), (cs\expnh_*)
RESTRAINT ARG=stcs.sqdevsum AT=0 KAPPA=0 SLOPE=24

PRINT ARG=(cs\hn_*), (cs\nh_*) FILE=RESTRAINT STRIDE=100
```

(See also [WHOLEMOLECULES](#), [STATS](#), [METAD](#), [RESTRAINT](#) and [PRINT](#))

## 5.2.10 DHENERGY

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

Calculate Debye-Huckel interaction energy among GROUPA and GROUPB.

This variable calculates the electrostatic interaction among GROUPA and GROUPB using a Debye-Huckel approximation defined as

$$\frac{1}{4\pi\epsilon_r\epsilon_0} \sum_{i \in A} \sum_{j \in B} q_i q_j \frac{e^{-\kappa|\mathbf{r}_{ij}|}}{|\mathbf{r}_{ij}|}$$

This collective variable can be used to analyze or induce electrostatically driven reactions [11]. Notice that the value of the DHENERGY is returned in plumed units (see [UNITS](#)).

If GROUPB is empty, it will sum the  $N*(N-1)/2$  pairs in GROUPA. This avoids computing twice permuted indexes (e.g. pair (i,j) and (j,i)) thus running at twice the speed.

Notice that if there are common atoms between GROUPA and GROUPB their interaction is discarded.

The atoms involved can be specified using

<b>GROUPA</b>	First list of atoms. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
<b>GROUPB</b>	Second list of atoms (if empty, $N*(N-1)/2$ pairs in GROUPA are counted). For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>

Compulsory keywords

<b>I</b>	( default=1.0 ) Ionic strength (M)
<b>TEMP</b>	( default=300.0 ) Simulation temperature (K)
<b>EPSILON</b>	( default=80.0 ) Dielectric constant of solvent

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) Perform the calculation in serial - for debug purpose
<b>PAIR</b>	( default=off ) Pair only 1st element of the 1st group with 1st element in the second, etc
<b>NLIST</b>	( default=off ) Use a neighbour list to speed up the calculation
<b>NL_CUTOFF</b>	The cutoff for the neighbour list
<b>NL_STRIDE</b>	The frequency with which we are updating the atoms in the neighbour list

## Examples

```
# this is printing the electrostatic interaction between two groups of atoms
dh: DHENERGY GROUPA=1-10 GROUPB=11-20 EPSILON=80.0 I=0.1 TEMP=300.0
PRINT ARG=dh
```

(see also [PRINT](#))

## 5.2.11 DIHCOR

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Measures the degree of similarity between dihedral angles.

This colvar calculates the following quantity.

$$s = \frac{1}{2} \sum_i [1 + \cos(\phi_i - \psi_i)]$$

where the  $\phi_i$  and  $\psi_i$  values and the instantaneous values for the [TORSION](#) angles of interest.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM↵S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

## Examples

The following provides an example input for the dihcov action

```
DIHCOR ...
  ATOMS1=1,2,3,4,5,6,7,8
  ATOMS2=5,6,7,8,9,10,11,12
  LABEL=dih
... DIHCOR
PRINT ARG=dih FILE=colvar STRIDE=10
```

In the above input we are calculating the correlation between the torsion angle involving atoms 1, 2, 3 and 4 and the torsion angle involving atoms 5, 6, 7 and 8. This is then added to the correlation between the torsion angle involving atoms 5, 6, 7 and 8 and the correlation angle involving atoms 9, 10, 11 and 12.

Writing out the atoms involved in all the torsions in this way can be rather tedious. Thankfully if you are working with protein you can avoid this by using the [MOLINFO](#) command. PLUMED uses the pdb file that you provide to this command to learn about the topology of the protein molecule. This means that you can specify torsion angles using the following syntax:

```
MOLINFO MOLTYPE=protein STRUCTURE=myprotein.pdb
DIHCOR ...
  ATOMS1=@phi-3,@psi-3
  ATOMS2=@psi-3,@phi-4
  ATOMS4=@phi-4,@psi-4
... DIHCOR
PRINT ARG=dih FILE=colvar STRIDE=10
```

Here, @phi-3 tells plumed that you would like to calculate the  $\phi$  angle in the third residue of the protein. Similarly @psi-4 tells plumed that you want to calculate the  $\psi$  angle of the 4th residue of the protein.

## 5.2.12 DIPOLE

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

Calculate the dipole moment for a group of atoms.



**Warning**

The atoms used for [DIPOLE](#) calculation should be from a whole molecule. In case the molecule is broken by the host MD code, please use [WHOLEMOLECULES](#) to reconstruct it before [DIPOLE](#) calculation.

**Description of components**

By default the value of the calculated quantity can be referenced elsewhere in the input file by using the label of the action. Alternatively this Action can be used to calculate the following quantities by employing the keywords listed below. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Keyword	Description
<b>x</b>	<b>COMPONENTS</b>	the x-component of the dipole
<b>y</b>	<b>COMPONENTS</b>	the y-component of the dipole
<b>z</b>	<b>COMPONENTS</b>	the z-component of the dipole

The atoms involved can be specified using

<b>GROUP</b>	the group of atoms we are calculating the dipole moment for. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------

**Options**

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>COMPONENTS</b>	( default=off ) calculate the x, y and z components of the dipole separately and store them as label.x, label.y and label.z

**Examples**

The following tells plumed to calculate the dipole of the group of atoms containing the atoms from 1-10 and print it every 5 steps

```
d: DIPOLE GROUP=1-10
PRINT FILE=output STRIDE=5 ARG=5
```

(see also [PRINT](#))

**Attention**

If the total charge  $Q$  of the group is non zero, then a charge  $Q/N$  will be subtracted to every atom, where  $N$  is the number of atoms. This implies that the dipole (which for a charged system depends on the position) is computed on the geometric center of the group.

**5.2.13 DISTANCE\_FROM\_CONTOUR**

<b>This is part of the multicolvar <a href="#">module</a></b>
---------------------------------------------------------------

Calculate the perpendicular distance from a Willard-Chandler dividing surface.

Suppose that you have calculated a multicolvar. By doing so you have calculated a set of colvars,  $s_i$ , and each of these colvars has a well defined position in space  $(x_i, y_i, z_i)$ . You can use this information to calculate a phase-field model of the colvar density using:

$$p(x, y, z) = \sum_i s_i K \left[ \frac{x - x_i}{\sigma_x}, \frac{y - y_i}{\sigma_y}, \frac{z - z_i}{\sigma_z} \right]$$

In this expression  $\sigma_x, \sigma_y$  and  $\sigma_z$  are bandwidth parameters and  $K$  is one of the [kernel functions](#). This is what is done within [MULTICOLVARDENS](#)

The Willard-Chandler surface is a surface of constant density in the above phase field  $p(x, y, z)$ . In other words, it is a set of points,  $(x', y', z')$ , in your box which have:

$$p(x', y', z') = \rho$$

where  $\rho$  is some target density. This action calculates the distance projected on the  $x, y$  or  $z$  axis between the position of some test particle and this surface of constant field density.

The atoms involved can be specified using

<b>ATOM</b>	The atom whose perpendicular distance we are calculating from the contour. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
-------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>DATA</b>	The input base multicolvar which is being used to calculate the contour
<b>BANDWIDTH</b>	the bandwidths for kernel density estimation
<b>KERNEL</b>	( default=gaussian ) the kernel function you are using. More details on the kernels available in plumed plumed can be found in <a href="#">kernel functions</a> .
<b>DIR</b>	the direction perpendicular to the contour that you are looking for
<b>CONTOUR</b>	the value we would like for the contour

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize

<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

### Examples

In this example atoms 2-100 are assumed to be concentrated along some part of the  $z$  axis so that you have an interface between a liquid/solid and the vapour. The quantity `dc` measures the distance between the surface at which the density of 2-100 atoms is equal to 0.2 and the position of the test particle atom 1.

```
dens: DENSITY SPECIES=2-100
dc: DISTANCE_FROM_CONTOUR DATA=dens ATOM=1 BANDWIDTH=0.5,0.5,0.5 DIR=z CONTOUR=0.2
```

#### 5.2.14 DISTANCE

This is part of the [colvar module](#)

Calculate the distance between a pair of atoms.

By default the distance is computed taking into account periodic boundary conditions. This behavior can be changed with the `NOPBC` flag. Moreover, single components in cartesian space ( $x, y$ , and  $z$ , with `COMPONENTS`) or single components projected to the three lattice vectors ( $a, b$ , and  $c$ , with `SCALED_COMPONENTS`) can be also computed.

Notice that Cartesian components will not have the proper periodicity! If you have to study e.g. the permeation of a molecule across a membrane, better to use `SCALED_COMPONENTS`.

#### Description of components

By default the value of the calculated quantity can be referenced elsewhere in the input file by using the label of the action. Alternatively this Action can be used to calculate the following quantities by employing the keywords listed below. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Keyword	Description
<b>x</b>	<b>COMPONENTS</b>	the x-component of the vector connecting the two atoms
<b>y</b>	<b>COMPONENTS</b>	the y-component of the vector connecting the two atoms
<b>z</b>	<b>COMPONENTS</b>	the z-component of the vector connecting the two atoms
<b>a</b>	<b>SCALED_COMPONENTS</b>	the normalized projection on the first lattice vector of the vector connecting the two atoms
<b>b</b>	<b>SCALED_COMPONENTS</b>	the normalized projection on the second lattice vector of the vector connecting the two atoms
<b>c</b>	<b>SCALED_COMPONENTS</b>	the normalized projection on the third lattice vector of the vector connecting the two atoms

The atoms involved can be specified using

<b>ATOMS</b>	the pair of atom that we are calculating the distance between. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>COMPONENTS</b>	( default=off ) calculate the x, y and z components of the distance separately and store them as label.x, label.y and label.z
<b>SCALED_COMPONENTS</b>	( default=off ) calculate the a, b and c scaled components of the distance separately and store them as label.a, label.b and label.c

#### Examples

The following input tells plumed to print the distance between atoms 3 and 5, the distance between atoms 2 and 4 and the x component of the distance between atoms 2 and 4.

```
d1:  DISTANCE ATOMS=3,5
d2:  DISTANCE ATOMS=2,4
d2c: DISTANCE ATOMS=2,4 COMPONENTS
PRINT ARG=d1,d2,d2c.x
```

(See also [PRINT](#)).

The following input computes the end-to-end distance for a polymer of 100 atoms and keeps it at a value around 5.

```
WHOLEMOLECULES ENTITY0=1-100
e2e: DISTANCE ATOMS=1,100 NOPBC
RESTRAINT ARG=e2e KAPPA=1 AT=5
```

(See also [WHOLEMOLECULES](#) and [RESTRAINT](#)).

Notice that NOPBC is used to be sure that if the end-to-end distance is larger than half the simulation box the distance is compute properly. Also notice that, since many MD codes break molecules across cell boundary, it might be necessary to use the [WHOLEMOLECULES](#) keyword (also notice that it should be *before* distance). The list of atoms provided to WHOLEMOLECULES here contains all the atoms between 1 and 100. Strictly speaking, this is not necessary. If you know for sure that atoms with difference in the index say equal to 10 are *not* going to be farther than half cell you can e.g. use

```
WHOLEMOLECULES ENTITY0=1,10,20,30,40,50,60,70,80,90,100
e2e: DISTANCE ATOMS=1,100 NOPBC
RESTRAINT ARG=e2e KAPPA=1 AT=5
```

Just be sure that the ordered list provide to WHOLEMOLECULES has the following properties:

- Consecutive atoms should be closer than half-cell throughout the entire simulation.

- Atoms required later for the distance (e.g. 1 and 100) should be included in the list

The following example shows how to take into account periodicity e.g. in z-component of a distance

```
# this is a center of mass of a large group
c: COM ATOMS=1-100
# this is the distance between atom 101 and the group
d: DISTANCE ATOMS=c,101 COMPONENTS
# this makes a new variable, dd, equal to d and periodic, with domain -10,10
# this is the right choice if e.g. the cell is orthorombic and its size in
# z direction is 20.
dz: COMBINE ARG=d,z PERIODIC=-10,10
# metadynamics on dd
METAD ARG=dz SIGMA=0.1 HEIGHT=0.1 PACE=200
```

(see also [COM](#), [COMBINE](#), and [METAD](#))

Using `SCALED_COMPONENTS` this problem should not arise because they are always periodic with domain  $(-0.5, +0.5)$ .

### 5.2.15 ENERGY

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

Calculate the total energy of the simulation box.

Total energy can be biased with umbrella sampling [\[12\]](#) or with well tempered metadynamics [\[13\]](#).

Notice that this CV could be unavailable with some MD code. When it is available, and when also replica exchange is available, metadynamics applied to `ENERGY` can be used to decrease the number of required replicas.

**Bug** Acceptance for replica exchange when [ENERGY](#) is biased is computed correctly only if all the replicas has the same potential energy function. This is for instance not true when using GROMACS with lambda replica exchange or with plumed-hrex branch.

#### Examples

The following input instructs plumed to print the energy of the system

```
ENERGY LABEL=ene
PRINT ARG=ene
```

(See also [PRINT](#)).

### 5.2.16 ERMSD

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

Calculate eRMSD with respect to a reference structure.

eRMSD is a metric developed for measuring distances between three-dimensional RNA structures. The standard RMSD measure is highly inaccurate when measuring distances among three-dimensional structures of nucleic acids. It is not unusual, for example, that two RNA structures with low RMSD (i.e. less than 0.4nm) display a completely different network of base-base interactions.

eRMSD measures the distance between structures by considering only the relative positions and orientations of nucleobases. The eRMSD can be considered as a vectorial version of contact maps and it is calculated as follows:

1. Set up a local reference system in the center of the six-membered ring of each nucleobase in a molecule. The xy plane lies on the plane of the nucleobase, and it is oriented such that the Watson-Crick interaction is always at  $\theta \approx 60^\circ$ .
2. Calculate all pairwise distance vectors  $\vec{r}_{i,j}$  among base centers.
3. Rescale distance vectors as  $\tilde{\vec{r}}_{i,j} = (r_x/a, r_y/a, r_z/b)$ , where  $a=b=5$ ,  $c=3$ . This rescaling has the effect of weighting more deviations on the z-axis with respect to the x/y directions.
4. Calculate the G vectors

$$\vec{G}(\tilde{\vec{r}}) = (\sin(\gamma\tilde{r})\tilde{r}_x/\tilde{r}, \sin(\gamma\tilde{r})\tilde{r}_y/\tilde{r}, \sin(\gamma\tilde{r})\tilde{r}_z/\tilde{r}, 1 + \cos(\gamma\tilde{r})) \times \frac{\Theta(\tilde{r}_{cutoff} - \tilde{r})}{\gamma}$$

Here,  $\gamma = \pi/\tilde{r}_{cutoff}$  and  $\Theta$  is the Heaviside step function. The default cutoff is set to 2.4.

1. The eRMSD between two structures  $\alpha$  and  $\beta$  reads

$$eRMSD = \sqrt{\frac{1}{N} \sum_{j,k} |\vec{G}(\tilde{\vec{r}}_{jk}^\alpha) - \vec{G}(\tilde{\vec{r}}_{jk}^\beta)|^2}$$

Using the default cutoff, two structures with eRMSD of 0.7 or lower can be considered as significantly similar. A full description of the eRMSD can be found in [14]

ERMSD is computed using the position of three atoms on the 6-membered ring of each involved nucleobase. The atoms should be:

- C2,C4,C6 for pyrimidines
- C2,C6,C4 for purines

The different order for purines and pyrimidines is fundamental and allows you to compute ERMSD between structures with different sequences as well! Notice that the simplest way to avoid mistakes in choosing these atoms is to use the `@lcs-#` strings as shown in the examples (see also [MOLINFO](#)).

#### Warning

Notice that the ERMSD implemented here is not integrated with the other metrics in plumed. As a consequence, it is not (yet) possible to e.g. build path collective variables using ERMSD

Notice that ERMSD expect a single molecule and makes coordinate whole before anything else. As such, results might be unexpected for a multi molecular system.

The atoms involved can be specified using

<b>ATOMS</b>	the list of atoms (use lcs). For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	---------------------------------------------------------------------------------------------------------------------------------

## Compulsory keywords

<b>REFERENCE</b>	a file in pdb format containing the reference structure and the atoms involved in the CV.
<b>CUTOFF</b>	( default=2.4 ) only pairs of atoms closer than CUTOFF are considered in the calculation.

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>PAIRS</b>	List of pairs considered. All pairs are considered if this value is not specified.

## Examples

Calculate the eRMSD from reference structure reference.pdb using the default cutoff (2.4). The list of residues involved in the calculation has to be specified. In this example, the eRMSD is calculated considering residues 1,2,3,4,5,6.

```
MOLINFO STRUCTURE=reference.pdb
eRMSD1: ERMSD REFERENCE=reference.pdb ATOMS=@lcs-1,@lcs-2,@lcs-3,@lcs-4,@lcs-5,@lcs-6
```

## 5.2.17 FAKE

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

This is a fake colvar container used by cltools or various other actions and just support input and period definition

The atoms involved can be specified using

<b>ATOMS</b>	the fake atom index, a number is enough. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	---------------------------------------------------------------------------------------------------------------------------------------------

## Compulsory keywords

<b>PERIODIC</b>	if the output of your function is periodic then you should specify the periodicity of the function. If the output is not periodic you must state this using PERIODIC=NO,NO (one for the lower and the other for the upper boundary). For multicomponents then it is PERIODIC=IC=mincomp1,maxcomp1,mincomp2,maxcomp2 etc
-----------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>COMPONENTS</b>	additional components that this variable is supposed to have. Periodicity is ruled by PERIODIC keyword

## Examples

```
FAKE ATOMS=1 PERIODIC=-3.14,3.14 LABEL=d2
```

(See also [PRINT](#)).

## 5.2.18 FRET

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

Calculate the FRET efficiency between a pair of atoms. The efficiency is calculated using the Forster relation:

$$E = \frac{1}{1 + (R/R_0)^6}$$

where  $R$  is the distance and  $R_0$  is the Forster radius.

By default the distance is computed taking into account periodic boundary conditions. This behavior can be changed with the NOPBC flag.

The atoms involved can be specified using

<b>ATOMS</b>	the pair of atom that we are calculating the distance between. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Compulsory keywords



<b>R0</b>	The value of the Forster radius.
-----------	----------------------------------

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances

### Examples

The following input tells plumed to print the FRET efficiencies calculated as a function of the distance between atoms 3 and 5 and the distance between atoms 2 and 4.

```
fe1: FRET ATOMS=3,5 R0=5.5
fe2: FRET ATOMS=2,4 R0=5.5
PRINT ARG=fe1,fe2
```

(See also [PRINT](#)).

The following input computes the FRET efficiency calculated on the terminal atoms of a polymer of 100 atoms and keeps it at a value around 0.5.

```
WHOLEMOLECULES ENTITY0=1-100
fe: FRET ATOMS=1,100 R0=5.5 NOPBC
RESTRAINT ARG=fe KAPPA=100 AT=0.5
```

(See also [WHOLEMOLECULES](#) and [RESTRAINT](#)).

Notice that NOPBC is used to be sure that if the distance is larger than half the simulation box the distance is compute properly. Also notice that, since many MD codes break molecules across cell boundary, it might be necessary to use the [WHOLEMOLECULES](#) keyword (also notice that it should be *before* FRET). Just be sure that the ordered list provide to WHOLEMOLECULES has the following properties:

- Consecutive atoms should be closer than half-cell throughout the entire simulation.
- Atoms required later for the distance (e.g. 1 and 100) should be included in the list

#### 5.2.19 GPROPERTYMAP

	<b>This is part of the mapping <a href="#">module</a></b>
--	-----------------------------------------------------------

Property maps but with a more flexible framework for the distance metric being used.

This colvar calculates a property map using the formalism developed by Spiwok [15]. In essence if you have the value of some property,  $X_i$ , that it takes at a set of high-dimensional positions then you calculate the value of the property at some arbitrary point in the high-dimensional space using:

$$X = \frac{\sum_i X_i * \exp(-\lambda D_i(x))}{\sum_i \exp(-\lambda D_i(x))}$$

Within PLUMED there are multiple ways to define the distance from a high-dimensional configuration,  $D_i$ . You could calculate the RMSD distance or you could calculate the amount by which a set of collective variables change. As such this implementation of the `propertymap` allows one to use all the different distance metric that are discussed in [Distances from reference configurations](#). This is as opposed to the alternative implementation `PROPERTYMAP` which is a bit faster but which only allows one to use the RMSD distance.

#### Compulsory keywords

<b>REFERENCE</b>	a pdb file containing the set of reference configurations
<b>PROPERTY</b>	the property to be used in the index. This should be in the REMARK of the reference
<b>TYPE</b>	( default=OPTIMAL-FAST ) the manner in which distances are calculated. More information on the different metrics that are available in PLUMED can be found in the section of the manual on <a href="#">Distances from reference configurations</a>
<b>LAMBDA</b>	the value of the lambda parameter for paths

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>DISABLE_CHECKS</b>	( default=off ) disable checks on reference input structures.
<b>NOZPATH</b>	( default=off ) do not calculate the zpath position
<b>NOMAPPING</b>	( default=off ) do not calculate the position on the manifold

#### Examples

The input shown below can be used to calculate the interpolated values of two properties called X and Y based on the values that these properties take at a set of reference configurations and using the formula above. For this input the distances between the reference configurations and the instantaneous configurations are calculated using the OPTIMAL metric that is discussed at length in the manual pages on [RMSD](#).

```
p2: GPROPERTYMAP REFERENCE=allv.pdb PROPERTY=X,Y LAMBDA=69087
PRINT ARG=p2.X,p2.Y,p2.zpath STRIDE=1 FILE=colvar
```

The additional input file for this calculation, which contains the reference frames and the values of X and Y at these reference points has the following format.

```

REMARK X=1 Y=2
ATOM      1  CL  ALA      1      -3.171    0.295    2.045    1.00    1.00
ATOM      5  CLP ALA      1      -1.819   -0.143    1.679    1.00    1.00
ATOM      6  OL  ALA      1      -1.177   -0.889    2.401    1.00    1.00
ATOM      7  NL  ALA      1      -1.313    0.341    0.529    1.00    1.00
ATOM      8  HL  ALA      1      -1.845    0.961   -0.011    1.00    1.00
ATOM      9  CA  ALA      1      -0.003   -0.019    0.021    1.00    1.00
ATOM     10  HA  ALA      1       0.205   -1.051    0.259    1.00    1.00
ATOM     11  CB  ALA      1       0.009    0.135   -1.509    1.00    1.00
ATOM     15  CRP ALA      1       1.121    0.799    0.663    1.00    1.00
ATOM     16  OR  ALA      1       1.723    1.669    0.043    1.00    1.00
ATOM     17  NR  ALA      1       1.423    0.519    1.941    1.00    1.00
ATOM     18  HR  ALA      1       0.873   -0.161    2.413    1.00    1.00
ATOM     19  CR  ALA      1       2.477    1.187    2.675    1.00    1.00
END
FIXED
REMARK X=2 Y=3
ATOM      1  CL  ALA      1      -3.175    0.365    2.024    1.00    1.00
ATOM      5  CLP ALA      1      -1.814   -0.106    1.685    1.00    1.00
ATOM      6  OL  ALA      1      -1.201   -0.849    2.425    1.00    1.00
ATOM      7  NL  ALA      1      -1.296    0.337    0.534    1.00    1.00
ATOM      8  HL  ALA      1      -1.807    0.951   -0.044    1.00    1.00
ATOM      9  CA  ALA      1       0.009   -0.067    0.033    1.00    1.00
ATOM     10  HA  ALA      1       0.175   -1.105    0.283    1.00    1.00
ATOM     11  CB  ALA      1       0.027    0.046   -1.501    1.00    1.00
ATOM     15  CRP ALA      1       1.149    0.725    0.654    1.00    1.00
ATOM     16  OR  ALA      1       1.835    1.491   -0.011    1.00    1.00
ATOM     17  NR  ALA      1       1.380    0.537    1.968    1.00    1.00
ATOM     18  HR  ALA      1       0.764   -0.060    2.461    1.00    1.00
ATOM     19  CR  ALA      1       2.431    1.195    2.683    1.00    1.00
END

```

### 5.2.20 GYRATION

This is part of the [colvar module](#)

Calculate the radius of gyration, or other properties related to it.

The different properties can be calculated and selected by the TYPE keyword: the Radius of Gyration (RADIUS); the Trace of the Gyration Tensor (TRACE); the Largest Principal Moment of the Gyration Tensor (GTPC\_1); the middle Principal Moment of the Gyration Tensor (GTPC\_2); the Smallest Principal Moment of the Gyration Tensor (GTPC\_3); the Asphericity (ASPHERICITY); the Acylindricity (ACYLINDRICITY); the Relative Shape Anisotropy (KAPPA2); the Smallest Principal Radius Of Gyration (GYRATION\_3); the Middle Principal Radius of Gyration (GYRATION\_2); the Largest Principal Radius of Gyration (GYRATION\_1). A derivation of all these different variants can be found in [16]

The radius of gyration is calculated using:

$$s_{\text{Gyr}} = \left( \frac{\sum_i^n m_i |r_i - r_{\text{COM}}|^2}{\sum_i^n m_i} \right)^{1/2}$$

with the position of the center of mass  $r_{\text{COM}}$  given by:

$$r_{\text{COM}} = \frac{\sum_i^n r_i m_i}{\sum_i^n m_i}$$

The radius of gyration usually makes sense when atoms used for the calculation are all part of the same molecule. When running with periodic boundary conditions, the atoms should be in the proper periodic image. This is done automatically since PLUMED 2.2, by considering the ordered list of atoms and rebuilding PBCs with a procedure that is equivalent to that done in [WHOLEMOLECULES](#). Notice that rebuilding is local to this action. This is different from [WHOLEMOLECULES](#) which actually modifies the coordinates stored in PLUMED.

In case you want to recover the old behavior you should use the NOPBC flag. In that case you need to take care that atoms are in the correct periodic image.

The atoms involved can be specified using

<b>ATOMS</b>	the group of atoms that you are calculating the Gyration Tensor for. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>TYPE</b>	( default=RADIUS ) The type of calculation relative to the Gyration Tensor you want to perform
-------------	------------------------------------------------------------------------------------------------

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>MASS_WEIGHTED</b>	( default=off ) set the masses of all the atoms equal to one

Examples

The following input tells plumed to print the radius of gyration of the chain containing atoms 10 to 20.

```
GYRATION TYPE=RADIUS ATOMS=10-20 LABEL=rg
PRINT ARG=rg STRIDE=1 FILE=colvar
```

(See also [PRINT](#))

### 5.2.21 JCOUPLING

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

Calculates  $^3J$  coupling constants for a dihedral angle.

The J-coupling between two atoms is given by the Karplus relation:

$$^3J(\theta) = A \cos^2(\theta + \Delta\theta) + B \cos(\theta + \Delta\theta) + C$$

where  $A$ ,  $B$  and  $C$  are the Karplus parameters and  $\Delta\theta$  is an additional constant added on to the dihedral angle  $\theta$ . The Karplus parameters are determined empirically and are dependent on the type of J-coupling.

This collective variable computes the J-couplings for a set of atoms defining a dihedral angle. You can specify the atoms involved using the [MOLINFO](#) notation. You can also specify the experimental couplings using the `ADD↔COUPLINGS` flag and `COUPLING` keywords. These will be included in the output. You must choose the type of coupling using the type keyword, you can also supply custom Karplus parameters using `TYPE=CUSTOM` and the `A`, `B`, `C` and `SHIFT` keywords. You will need to make sure you are using the correct dihedral angle:

- Ha-N:  $\psi$
- Ha-HN:  $\phi$
- N-C  $\gamma$ :  $\chi_1$
- CO-C  $\gamma$ :  $\chi_1$

#### Description of components

The names of the components in this action can be customized by the user in the actions input file. However, in addition to these customizable components the following quantities will always be output

Quantity	Description
j	the calculated J-coupling

In addition the following quantities can be calculated by employing the keywords listed below

Quantity	Keyword	Description
exp	ADDCOUPINGS	the experimental J-coupling

The atoms involved can be specified using

<b>ATOMS</b>	the 4 atoms involved in each of the bonds for which you wish to calculate the J-coupling. Keywords like ATOMS1, ATOMS2, ATOMS3,... should be listed and one J-coupling will be calculated for each ATOMS keyword you specify. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Compulsory keywords

<b>TYPE</b>	Type of J-coupling to compute (HAN,HAHN,CCG,NGC,CUSTOM)
-------------	---------------------------------------------------------

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>ADDCOUPINGS</b>	( default=off ) Set this flag if you want to have fixed components with the experimental values.
<b>A</b>	Karplus parameter A
<b>B</b>	Karplus parameter B
<b>C</b>	Karplus parameter C
<b>SHIFT</b>	Angle shift in radians

<b>COUPLING</b>	Add an experimental value for each coupling You can use multiple instances of this keyword i.e. COUPLING1, COUPLING2, COUPLING3...
-----------------	------------------------------------------------------------------------------------------------------------------------------------

### Examples

In the following example we calculate the Ha-N J-coupling from a set of atoms involved in dihedral  $\psi$  angles in the peptide backbone. We also add the experimental datapoints and compute the correlation and other measures and finally print the results.

```
MOLINFO MOLTYPE=protein STRUCTURE=peptide.pdb
WHOLEMOLECULES ENTITY0=1-111

JCOUPLING ...
  ADDCOUPLINGS
  TYPE=HAN
  ATOMS1=@psi-2 COUPLING1=-0.49
  ATOMS2=@psi-4 COUPLING2=-0.54
  ATOMS3=@psi-5 COUPLING3=-0.53
  ATOMS4=@psi-7 COUPLING4=-0.39
  ATOMS5=@psi-8 COUPLING5=-0.39
  LABEL=jhan
... JCOUPLING

jhanst: STATS ARG=(jhan\j_*) PARARG=(jhan\exp_*)

PRINT ARG=jhanst.*,jhan.* FILE=COLVAR STRIDE=100

ENDPLUMED
```

(See also [PRINT](#), [STATS](#))

### 5.2.22 NOE

This is part of the [colvar module](#)

Calculates NOE intensities as sums of  $1/r^6$ , also averaging over multiple equivalent atoms or ambiguous NOE.

Each NOE is defined by two groups containing the same number of atoms, distances are calculated in pairs, transformed in  $1/r^6$ , summed and saved as components.

$$NOE() = \left( \frac{1}{N_{eq}} \sum_j^{N_{eq}} \left( \frac{1}{r_j^6} \right) \right)^{\frac{-1}{6}}$$

Intensities can then in principle ensemble averaged using [ENSEMBLE](#) and used to calculate a scoring function for example with [METAINTERFERENCE](#).

### Description of components

The names of the components in this action can be customized by the user in the actions input file. However, in addition to these customizable components the following quantities will always be output

Quantity	Description
<b>noe</b>	the # NOE

In addition the following quantities can be calculated by employing the keywords listed below

Quantity	Keyword	Description
<b>exp</b>	<b>ADDEXP</b>	the # NOE experimental distance

The atoms involved can be specified using

<b>GROUPA</b>	the atoms involved in each of the contacts you wish to calculate. Keywords like GROUPA1, GROUPA2, GROUPA3,... should be listed and one contact will be calculated for each ATOM keyword you specify. You can use multiple instances of this keyword i.e. GROUPA1, GROUPA2, GROUPA3...
<b>GROUPB</b>	the atoms involved in each of the contacts you wish to calculate. Keywords like GROUPB1, GROUPB2, GROUPB3,... should be listed and one contact will be calculated for each ATOM keyword you specify. You can use multiple instances of this keyword i.e. GROUPB1, GROUPB2, GROUPB3...

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>ADDEXP</b>	( default=off ) Set to TRUE if you want to have fixed components with the experimental reference values.
<b>NOEDIST</b>	Add an experimental value for each NOE. You can use multiple instances of this keyword i.e. NOEDIST1, NOEDIST2, NOEDIST3...

#### Examples

In the following examples three noes are defined, the first is calculated based on the distances of atom 1-2 and 3-2; the second is defined by the distance 5-7 and the third by the distances 4-15,4-16,8-15,8-16.

```
NOE ...
GROUPA1=1, 3 GROUPB1=2, 2
GROUPA2=5 GROUPB2=7
GROUPA3=4, 4, 8, 8 GROUPB3=15, 16, 15, 16
LABEL=noes
... NOE

PRINT ARG=noes.* FILE=colvar
```

(See also [PRINT](#))

#### 5.2.23 PARABETARMSD

<b>This is part of the secondarystructure module</b>
------------------------------------------------------

Probe the parallel beta sheet content of your protein structure.

Two protein segments containing three contiguous residues can form a parallel beta sheet. Although if the two segments are part of the same protein chain they must be separated by a minimum of 3 residues to make room for the turn. This colvar thus generates the set of all possible six residue sections that could conceivably form a parallel beta sheet and calculates the RMSD distance between the configuration in which the residues find themselves and an idealized parallel beta sheet structure. These distances can be calculated by either aligning the instantaneous structure with the reference structure and measuring each atomic displacement or by calculating differences between the set of interatomic distances in the reference and instantaneous structures.

This colvar is based on the following reference [4]. The authors of this paper use the set of distances from the parallel beta sheet configurations to measure the number of segments whose configuration resembles a parallel beta sheet. This is done by calculating the following sum of functions of the rmsd distances:

$$s = \sum_i \frac{1 - \left( \frac{r_i - d_0}{r_0} \right)^n}{1 - \left( \frac{r_i - d_0}{r_0} \right)^m}$$

where the sum runs over all possible segments of parallel beta sheet. By default the NN, MM and D\_0 parameters are set equal to those used in [4]. The R\_0 parameter must be set by the user - the value used in [4] was 0.08 nm.

If you change the function in the above sum you can calculate quantities such as the average distance from a structure composed of only parallel beta sheets or the distance between the set of residues that is closest to a parallel beta sheet and the reference configuration. To do these sorts of calculations you can use the AVERAGE and MIN keywords. In addition you can use the LESS\_THAN keyword if you would like to change the form of the switching function. If you use any of these options you no longer need to specify NN, R\_0, MM and D\_0.

Please be aware that for codes like gromacs you must ensure that plumed reconstructs the chains involved in your CV when you calculate this CV using anything other than TYPE=DRMSD. For more details as to how to do this see [WHOLEMOLECULES](#).

#### Description of components

By default this Action calculates the number of structural units that are within a certain distance of a idealised secondary structure element. This quantity can then be referenced elsewhere in the input by using the label of the action. However, this Action can also be used to calculate the following quantities by using the keywords as described below. The quantities then calculated can be referenced using the label of the action followed by a dot and then the name from the table below. Please note that you can use the LESS\_THAN keyword more than once. The resulting components will be labelled *label.lessthan-1*, *label.lessthan-2* and so on unless you exploit the fact that these labels are customizable. In particular, by using the LABEL keyword in the description of you LESS\_THAN function you can set name of the component that you are calculating

Quantity	Keyword	Description
altmin	ALT_MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
highest	HIGHEST	the lowest of the quantities calculated by this action
lessthan	LESS_THAN	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
lowest	LOWEST	the lowest of the quantities calculated by this action
min	MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.



The atoms involved can be specified using

<b>RESIDUES</b>	this command is used to specify the set of residues that could conceivably form part of the secondary structure. It is possible to use residues numbers as the various chains and residues should have been identified else using an instance of the <a href="#">MOLINFO</a> action. If you wish to use all the residues from all the chains in your system you can do so by specifying all. Alternatively, if you wish to use a subset of the residues you can specify the particular residues you are interested in as a list of numbers. Please be aware that to form secondary structure elements your chain must contain at least N residues, where N is dependent on the particular secondary structure you are interested in. As such if you define portions of the chain with fewer than N residues the code will crash.
-----------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Compulsory keywords

<b>TYPE</b>	( default=DRMSD ) the manner in which RMSD alignment is performed. Should be OPTIMAL, $S_{\leftrightarrow}$ IMPL or DRMSD. For more details on the OPTIMAL and SIMPLE methods see <a href="#">RMSD</a> . For more details on the DRMSD method see <a href="#">DRMSD</a> .
<b>R_0</b>	The r_0 parameter of the switching function.
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>NN</b>	( default=8 ) The n parameter of the switching function
<b>MM</b>	( default=12 ) The m parameter of the switching function
<b>STYLE</b>	( default=all ) Parallel beta sheets can either form in a single chain or from a pair of chains. If $S_{\leftrightarrow}$ STYLE=all all chain configuration with the appropriate geometry are counted. If STYLE=inter only sheet-like configurations involving two chains are counted, while if STYLE=intra only sheet-like configurations involving a single chain are counted

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>VERBOSE</b>	( default=off ) write a more detailed output
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...

<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>STRANDS_CUTOFF</b>	If in a segment of protein the two strands are further apart then the calculation of the actual RMSD is skipped as the structure is very far from being beta-sheet like. This keyword speeds up the calculation enormously when you are using the LESS_THAN option. However, if you are using some other option, then this cannot be used

### Examples

The following input calculates the number of six residue segments of protein that are in an parallel beta sheet configuration.

```
MOLINFO STRUCTURE=helix.pdb
PARABETARMSD RESIDUES=all TYPE=DRMSD LESS_THAN={RATIONAL R_0=0.08 NN=8 MM=12} LABEL=a
```

(see also [MOLINFO](#))

### 5.2.24 PATHMSD

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

This Colvar calculates path collective variables.

This is the Path Collective Variables implementation ( see [17] ). This variable computes the progress along a given set of frames that is provided in input ("sss" component) and the distance from them ("zzz" component). (see below).

#### Warning

The molecule used for [PATHMSD](#) calculation should be whole (both atoms used in alignment and in displacement calculation). In case it is broken by the host MD code, please use [WHOLEMOLECULES](#) to reconstruct it before [PATHMSD](#) calculation.

#### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>sss</b>	the position on the path
<b>zzz</b>	the distance from the path

#### Compulsory keywords

<b>LAMBDA</b>	the lambda parameter is needed for smoothing, is in the units of plumed
<b>REFERENCE</b>	the pdb is needed to provide the various milestones

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NEIGH_SIZE</b>	size of the neighbor list
<b>NEIGH_STRIDE</b>	how often the neighbor list needs to be calculated in time units

#### Examples

Here below is a case where you have defined three frames and you want to calculate the progress along the path and the distance from it in p1

```
p1: PATHMSD REFERENCE=file.pdb LAMBDA=500.0 NEIGH_STRIDE=4 NEIGH_SIZE=8
PRINT ARG=p1.sss,p1.zzz STRIDE=1 FILE=colvar FMT=%8.4f
```

note that NEIGH\_STRIDE=4 NEIGH\_SIZE=8 control the neighborlist parameter (optional but recommended for performance) and states that the neighbor list will be calculated every 4 timesteps and consider only the closest 8 member to the actual md snapshots.

In the REFERENCE PDB file the frames must be separated either using END or ENDMDL.

#### Note

The implementation of this collective variable and of [PROPERTYMAP](#) is shared, as well as most input options.

### 5.2.25 PATH

<b>This is part of the mapping <a href="#">module</a></b>
-----------------------------------------------------------

Path collective variables with a more flexible framework for the distance metric being used.

The Path Collective Variables developed by Branduardi and co-workers [17] allow one to compute the progress along a high-dimensional path and the distance from the high-dimensional path. The progress along the path (s) is computed using:

$$s = \frac{\sum_{i=1}^N i \exp(-\lambda R[X - X_i])}{\sum_{i=1}^N \exp(-\lambda R[X - X_i])}$$

while the distance from the path ( $z$ ) is measured using:

$$z = -\frac{1}{\lambda} \ln \left[ \sum_{i=1}^N \exp(-\lambda R[X - X_i]) \right]$$

In these expressions  $N$  high-dimensional frames ( $X_i$ ) are used to describe the path in the high-dimensional space. The two expressions above are then functions of the distances from each of the high-dimensional frames  $R[X - X_i]$ . Within PLUMED there are multiple ways to define the distance from a high-dimensional configuration. You could calculate the RMSD distance or you could calculate the amount by which a set of collective variables change. As such this implementation of the path cv allows one to use all the difference distance metrics that are discussed in [Distances from reference configurations](#). This is as opposed to the alternative implementation of path ([PATHMSD](#)) which is a bit faster but which only allows one to use the RMSD distance.

#### Compulsory keywords

<b>REFERENCE</b>	a pdb file containing the set of reference configurations
<b>TYPE</b>	( default=OPTIMAL-FAST ) the manner in which distances are calculated. More information on the different metrics that are available in PLUMED can be found in the section of the manual on <a href="#">Distances from reference configurations</a>
<b>LAMBDA</b>	the value of the lambda parameter for paths

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>DISABLE_CHECKS</b>	( default=off ) disable checks on reference input structures.
<b>NOZPATH</b>	( default=off ) do not calculate the zpath position
<b>NOSPATH</b>	( default=off ) do not calculate the spath position

#### Examples

The following input instructs PLUMED to calculate the values of the path collective variables. The frames that make up this path are defined in the file all.pdb and all distances are measured using the OPTIMAL metric that is discussed in the manual page on [RMSD](#).

```
p2: PATH REFERENCE=all.pdb LAMBDA=69087
PRINT ARG=p2.spath,p2.zpath STRIDE=1 FILE=colvar
```

If you wish to use collective variable values in the definition of your path you would use an input file with something like this:

```
d1: DISTANCE ATOMS=1,2
d2: DISTANCE ATOMS=3,4a
p2: PATH REFERENCE=myspath.pdb LAMBDA=2 TYPE=EUCLIDEAN
PRINT ARG=p2.spath,p2.zpath STRIDE=1 FILE=colvar
```

The corresponding pdb file containing the definitions of the frames in the path would then look like this:

```
DESCRIPTION: a defintiion of a PATH
REMARK TYPE=EUCLIDEAN
REMARK ARG=d1,d2
REMARK d1=1.0 d2=1.0
END
REMARK TYPE=EUCLIDEAN
REMARK ARG=d1,d2
REMARK d1=2.0 d2=2.0
END
```

For each frame in the path you must specify the arguments that should be used to calculate the distance between the instantaneous configuration of the system and the reference configurations together with the values that these arguments take in each of the reference configurations.

### 5.2.26 PCAVARS

This is part of the mapping <a href="#">module</a>
----------------------------------------------------

Projection on principal component eigenvectors or other high dimensional linear subspace

The collective variables described in [Distances from reference configurations](#) allow one to calculate the distance between the instantaneous structure adopted by the system and some high-dimensional, reference configuration. The problem with doing this is that, as one gets further and further from the reference configuration, the distance from it becomes a progressively poorer and poorer collective variable. This happens because the "number" of structures at a distance  $d$  from a reference configuration is proportional to  $d^N$  in an  $N$  dimensional space. Consequently, when  $d$  is small the distance from the reference configuration may well be a good collective variable. However, when  $d$  is large it is unlikely that the distance from the reference structure is a good CV. When the distance is large there will almost certainly be markedly different configurations that have the same CV value and hence barriers in transverse degrees of freedom.

For these reasons dimensionality reduction is often employed so a projection  $s$  of a high-dimensional configuration  $\mathbf{X}$  in a lower dimensionality space using a function:

$$s = F(\mathbf{X} - \mathbf{X}^{ref})$$

where here we have introduced some high-dimensional reference configuration  $\mathbf{X}^{ref}$ . By far the simplest way to do this is to use some linear operator for  $F$ . That is to say we find a low-dimensional projection by rotating the basis vectors using some linear algebra:

$$\mathbf{s}_i = \sum_k A_{ik} (X_k - X_k^{ref})$$

Here  $A$  is a  $d$  by  $D$  matrix where  $D$  is the dimensionality of the high dimensional space and  $d$  is the dimensionality of the lower dimensional subspace. In plumed when this kind of projection you can use the majority of the metrics detailed on [Distances from reference configurations](#) to calculate the displacement,  $\mathbf{X} - \mathbf{X}^{ref}$ , from the reference configuration. The matrix  $A$  can be found by various means including principal component analysis and normal mode analysis. In both these methods the rows of  $A$  would be the principle eigenvectors of a square matrix. For PCA the covariance while for normal modes the Hessian.

**Bug** It is not possible to use the [DRMSD](#) metric with this variable. You can get around this by listing the set of distances you wish to calculate for your DRMSD in the plumed file explicitly and using the EUCLIDEAN metric. MAHALONOBIS and NORM-EUCLIDEAN also do not work with this variable but using these options makes little sense when projecting on a linear subspace.

### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>eig</b>	the projections on each eigenvalue are stored on values labeled eig-1, eig-2, ...
<b>residual</b>	the distance of the configuration from the linear subspace defined by the vectors, $e_i$ , that are contained in the rows of $A$ . In other words this is $\sqrt{r^2 - \sum_i [\mathbf{r} \cdot \mathbf{e}_i]^2}$ where $r$ is the distance between the instantaneous position and the reference point.

### Compulsory keywords

<b>REFERENCE</b>	a pdb file containing the reference configuration and configurations that define the directions for each eigenvector
<b>TYPE</b>	( default=OPTIMAL ) The method we are using for alignment to the reference structure

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NORMALIZE</b>	( default=off ) calculate the length of the eigenvector input and divide the components by it so as to have a normalised vector

### Examples

The following input calculates a projection on a linear subspace where the displacements from the reference configuration are calculated using the OPTIMAL metric. Consequently, both translation of the center of mass of the atoms and rotation of the reference frame are removed from these displacements. The matrix  $A$  and the reference configuration  $R^{ref}$  are specified in the pdb input file reference.pdb and the value of all projections (and the residual) are output to a file called colvar2.

```
PCAVARS REFERENCE=reference.pdb TYPE=OPTIMAL LABEL=pca2
PRINT ARG=pca2.* FILE=colvar2
```

The reference configurations can be specified using a pdb file. The first configuration that you provide is the reference configuration, which is referred to in the above as  $X^{ref}$  subsequent configurations give the directions of row vectors that are contained in the matrix  $A$  above. These directions can be specified by specifying a second configuration - in this case a vector will be constructed by calculating the displacement of this second configuration from the reference configuration. A pdb input prepared in this way would look as follows:

```
ATOM      2  CH3  ACE      1      12.932 -14.718  -6.016  1.00  1.00
ATOM      5  C    ACE      1      21.312  -9.928  -5.946  1.00  1.00
ATOM      9  CA   ALA      2      19.462 -11.088  -8.986  1.00  1.00
ATOM     13  HB2  ALA      2      21.112 -10.688 -12.476  1.00  1.00
ATOM     15  C    ALA      2      19.422   7.978 -14.536  1.00  1.00
ATOM     20  HH31 NME      3      20.122  -9.928 -17.746  1.00  1.00
ATOM     21  HH32 NME      3      18.572 -13.148 -16.346  1.00  1.00
END
ATOM      2  CH3  ACE      1      13.932 -14.718  -6.016  1.00  1.00
ATOM      5  C    ACE      1      20.312  -9.928  -5.946  1.00  1.00
ATOM      9  CA   ALA      2      18.462 -11.088  -8.986  1.00  1.00
ATOM     13  HB2  ALA      2      20.112 -11.688 -12.476  1.00  1.00
ATOM     15  C    ALA      2      19.422   7.978 -12.536  1.00  1.00
ATOM     20  HH31 NME      3      20.122  -9.928 -17.746  1.00  1.00
ATOM     21  HH32 NME      3      18.572 -13.148 -16.346  1.00  1.00
END
```

Alternatively, the second configuration can specify the components of  $A$  explicitly. In this case you need to include the keyword TYPE=DIRECTION in the remarks to the pdb as shown below.

```
ATOM      2  CH3  ACE      1      12.932 -14.718  -6.016  1.00  1.00
ATOM      5  C    ACE      1      21.312  -9.928  -5.946  1.00  1.00
ATOM      9  CA   ALA      2      19.462 -11.088  -8.986  1.00  1.00
ATOM     13  HB2  ALA      2      21.112 -10.688 -12.476  1.00  1.00
ATOM     15  C    ALA      2      19.422   7.978 -14.536  1.00  1.00
ATOM     20  HH31 NME      3      20.122  -9.928 -17.746  1.00  1.00
ATOM     21  HH32 NME      3      18.572 -13.148 -16.346  1.00  1.00
END
REMARK TYPE=DIRECTION
ATOM      2  CH3  ACE      1      0.1414  0.3334 -0.0302  1.00  0.00
ATOM      5  C    ACE      1      0.0893 -0.1095 -0.1434  1.00  0.00
ATOM      9  CA   ALA      2      0.0207 -0.321   0.0321  1.00  0.00
ATOM     13  HB2  ALA      2      0.0317 -0.6085  0.0783  1.00  0.00
ATOM     15  C    ALA      2      0.1282 -0.4792  0.0797  1.00  0.00
ATOM     20  HH31 NME      3      0.0053 -0.465   0.0309  1.00  0.00
ATOM     21  HH32 NME      3     -0.1019 -0.4261 -0.0082  1.00  0.00
END
```

If your metric involves arguments the labels of these arguments in your plumed input file should be specified in the REMARKS for each of the frames of your path. An input file in this case might look like this:

```
DESCRIPTION: a pca eigenvector specified using the start point and direction in the HD space.
REMARK WEIGHT=1.0
REMARK ARG=d1,d2
REMARK d1=1.0 d2=1.0
END
REMARK TYPE=DIRECTION
REMARK ARG=d1,d2
REMARK d1=0.1 d2=0.25
END
```

Here we are working with the EUCLIDEAN metric and notice that we have specified the components of  $A$  using DIRECTION. Consequently, the values of d1 and d2 in the second frame above do not specify a particular coordinate in the high-dimensional space as in they do in the first frame. Instead these values are the coefficients that can be used to construct a linear combination of d1 and d2. If we wanted to specify the direction in this metric using the start and end point of the vector we would write:

```

DESCRIPTION: a pca eigenvector specified using the start and end point of a vector in the HD space.
REMARK WEIGHT=1.0
REMARK ARG=d1,d2
REMARK d1=1.0 d2=1.0
END
REMARK ARG=d1,d2
REMARK d1=1.1 d2=1.25
END

```

### 5.2.27 POSITION

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

Calculate the components of the position of an atom.

Notice that single components will not have the proper periodicity! If you need the values to be consistent through PBC you should use `SCALED_COMPONENTS`, which defines values that by construction are in the  $-0.5, 0.5$  domain. This is similar to the equivalent flag for [DISTANCE](#). Also notice that by default the minimal image distance from the origin is considered (can be changed with `NOPBC`).

#### Attention

This variable should be used with extreme care since it allows to easily go into troubles. See comments below.

This variable can be safely used only if Hamiltonian is not invariant for translation (i.e. there are other absolute positions which are biased, e.g. by position restraints) and cell size and shapes are fixed through the simulation.

If you are not in this situation and still want to use the absolute position of an atom you should first fix the reference frame. This can be done e.g. using [FIT\\_TO\\_TEMPLATE](#).

#### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>x</b>	the x-component of the atom position
<b>y</b>	the y-component of the atom position
<b>z</b>	the z-component of the atom position

In addition the following quantities can be calculated by employing the keywords listed below

Quantity	Keyword	Description
<b>a</b>	<b>SCALED_COMPONENTS</b>	the normalized projection on the first lattice vector of the atom position
<b>b</b>	<b>SCALED_COMPONENTS</b>	the normalized projection on the second lattice vector of the atom position
<b>c</b>	<b>SCALED_COMPONENTS</b>	the normalized projection on the third lattice vector of the atom position



The atoms involved can be specified using

<b>ATOM</b>	the atom number. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
-------------	---------------------------------------------------------------------------------------------------------------------

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SCALED_COMPONENTS</b>	( default=off ) calculate the a, b and c scaled components of the position separately and store them as label.a, label.b and label.c

#### Examples

```
# align to a template
FIT_TO_TEMPLATE REFERENCE=ref.pdb
p: POSITION ATOM=3
PRINT ARG=p.x,p.y,p.z
```

(see also [FIT\\_TO\\_TEMPLATE](#))

### 5.2.28 PRE

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

Calculates the Paramagnetic Resonance Enhancement intensity ratio between two atoms. The reference atom for the spin label is added with SPINLABEL, the affected atom(s) are given as numbered GROUPA1, GROUPA2, ... The additional parameters needed for the calculation are given as INEPT, the inept time, TAUC the correlation time, OMEGA, the larmor frequency and RTWO for the relaxation time.

#### Description of components

The names of the components in this action can be customized by the user in the actions input file. However, in addition to these customizable components the following quantities will always be output

Quantity	Description
<b>pre</b>	the # PRE

In addition the following quantities can be calculated by employing the keywords listed below

Quantity	Keyword	Description
exp	ADDEXP	the # PRE experimental intensity

The atoms involved can be specified using

<b>SPINLABEL</b>	The atom to be used as the paramagnetic center.. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
<b>GROUPA</b>	the atoms involved in each of the contacts you wish to calculate. Keywords like GROUPA1, GROUPA2, GROUPA3,... should be listed and one contact will be calculated for each ATOM keyword you specify. You can use multiple instances of this keyword i.e. GROUPA1, GROUPA2, GROUPA3...

Compulsory keywords

<b>INEPT</b>	is the INEPT time (in ms).
<b>TAUC</b>	is the correlation time (in ns) for this electron-nuclear interaction.
<b>OMEGA</b>	is the Larmor frequency of the nuclear spin (in MHz).

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>ADDEXP</b>	( default=off ) Set to TRUE if you want to have fixed components with the experimental values.
<b>RTWO</b>	The relaxation of the atom/atoms in the corresponding GROUPA of atoms. Keywords like RTWO1, RTWO2, RTWO3,... should be listed. You can use multiple instances of this keyword i.e. RTWO1, RTWO2, RTWO3...
<b>PREINT</b>	Add an experimental value for each PRE. You can use multiple instances of this keyword i.e. PREINT1, PREINT2, PREINT3...

Examples

In the following example five PRE intensities are calculated using the distance between the oxygen of the spin label and the backbone hydrogens. Omega is the NMR frequency, RTWO the R2 for the hydrogens, INEPT of 8 ms for the experiment and a TAUC of 1.21 ns

```
PRE ...
LABEL=HN_pre
INEPT=8
TAUC=1.21
OMEGA=900
SPINLABEL=1818
```

```

GROUPA1=86 RTWO1=0.0120272827
GROUPA2=177 RTWO2=0.0263953158
GROUPA3=285 RTWO3=0.0058899829
GROUPA4=335 RTWO4=0.0102072646
GROUPA5=451 RTWO5=0.0086341843
... PRE

PRINT ARG=HN_pre.* FILE=PRE.dat STRIDE=1

```

### 5.2.29 PROPERTYMAP

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

Calculate generic property maps.

This Colvar calculates the property maps according to the work of Spiwok [15].

Basically it calculates

$$X = \frac{\sum_i X_i * \exp(-\lambda D_i(x))}{\sum_i \exp(-\lambda D_i(x))} \quad (5.1)$$

$$Y = \frac{\sum_i Y_i * \exp(-\lambda D_i(x))}{\sum_i \exp(-\lambda D_i(x))} \quad (5.2)$$

$$\dots \quad (5.3)$$

$$zzz = -\frac{1}{\lambda} \log\left(\sum_i \exp(-\lambda D_i(x))\right) \quad (5.4)$$

where the parameters  $X_i$  and  $Y_i$  are provided in the input pdb (allv.pdb in this case) and  $D_i(x)$  is the MSD after optimal alignment calculated on the pdb frames you input (see Kearsley).

#### Warning

The molecule used for [PROPERTYMAP](#) calculation should be whole (both atoms used in alignment and in displacement calculation). In case it is broken by the host MD code, please use [WHOLEMOLECULES](#) to reconstruct it before [PROPERTYMAP](#) calculation.

#### Description of components

The names of the components in this action can be customized by the user in the actions input file. However, in addition to these customizable components the following quantities will always be output

Quantity	Description
<b>zzz</b>	the minimum distance from the reference points

#### Compulsory keywords

<b>LAMBDA</b>	the lambda parameter is needed for smoothing, is in the units of plumed
<b>REFERENCE</b>	the pdb is needed to provide the various milestones
<b>PROPERTY</b>	the property to be used in the indexing: this goes in the REMARK field of the reference

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NEIGH_SIZE</b>	size of the neighbor list
<b>NEIGH_STRIDE</b>	how often the neighbor list needs to be calculated in time units

## Examples

```
p3: PROPERTYMAP REFERENCE=../../trajectories/path_msd/allv.pdb PROPERTY=X,Y LAMBDA=69087 NEIGH_SIZE=8 NEIGH_ST
PRINT ARG=p3.X,p3.Y,p3.zzz STRIDE=1 FILE=colvar FMT=%8.4f
```

note that NEIGH\_STRIDE=4 NEIGH\_SIZE=8 control the neighborlist parameter (optional but recommended for performance) and states that the neighbor list will be calculated every 4 timesteps and consider only the closest 8 member to the actual md snapshots.

In this case the input line instructs plumed to look for two properties X and Y with attached values in the REMARK line of the reference pdb (Note: No spaces from X and = and 1 !!!!). e.g.

```
REMARK X=1 Y=2
ATOM      1  CL  ALA      1      -3.171    0.295    2.045    1.00    1.00
ATOM      5  CLP ALA      1      -1.819   -0.143    1.679    1.00    1.00
.....
END
REMARK X=2 Y=3
ATOM      1  CL  ALA      1      -3.175    0.365    2.024    1.00    1.00
ATOM      5  CLP ALA      1      -1.814   -0.106    1.685    1.00    1.00
....
END
```

## Note

The implementation of this collective variable and of [PATHMSD](#) is shared, as well as most input options.

## 5.2.30 PUCKERING

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

Calculate sugar pseudorotation coordinates.

This command can be used to calculate ring's pseudorotations in sugars (puckers). It works for both 5-membered and 6-membered rings. Notice that there are two different implementations depending if one passes 5 or 6 atoms in the ATOMS keyword.

For 5-membered rings the implementation is the one discussed in [18] . This implementation is simple and can be used in RNA to distinguish C2'-endo and C3'-endo conformations. Both the polar coordinates (phs and amp) and the cartesian coordinates (Zx and Zy) are provided. C2'-endo conformations have negative Zx, whereas C3'-endo conformations have positive Zy. Notation is consistent with [18] . The five atoms should be provided as

C4',O4',C1',C2',C3'. Notice that this is the same order that can be obtained using the [MOLINFO](#) syntax (see example below).

For 6-membered rings the implementation is the general Cremer-Pople one [19] as also discussed in [20]. This implementation provides both a triplet with Cartesian components (qx, qy, and qz) and a triplet of polar components (amplitude, phi, and theta). Applications of this particular implementation are to be published (paper in preparation).

**Bug** The 6-membered ring implementation distributed with this version of PLUMED leads to qx and qy values that have an opposite sign with respect to those originally defined in [19]. The bug will be fixed in a later version but is here kept to preserve reproducibility.

Components of this action are:

#### Description of components

By default the value of the calculated quantity can be referenced elsewhere in the input file by using the label of the action. Alternatively this Action can be used to calculate the following quantities by employing the keywords listed below. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>phs</b>	Pseudorotation phase (5 membered rings)
<b>amp</b>	Pseudorotation amplitude (5 membered rings)
<b>Zx</b>	Pseudorotation x cartesian component (5 membered rings)
<b>Zy</b>	Pseudorotation y cartesian component (5 membered rings)
<b>phi</b>	Pseudorotation phase (6 membered rings)
<b>theta</b>	Theta angle (6 membered rings)
<b>amplitude</b>	Pseudorotation amplitude (6 membered rings)
<b>qx</b>	Cartesian component x (6 membered rings)
<b>qy</b>	Cartesian component y (6 membered rings)
<b>qz</b>	Cartesian component z (6 membered rings)

The atoms involved can be specified using

<b>ATOMS</b>	the five or six atoms of the sugar ring in the proper order. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
------------------------------	----------------------------------------------------------------------------

#### Examples

This input tells plumed to print the puckering phase angle of the 3rd nucleotide of a RNA molecule on file COLVAR.

```
MOLINFO STRUCTURE=rna.pdb MOLTYPE=rna
PUCKERING ATOMS=@sugar-3 LABEL=puck
PRINT ARG=puck.phs FILE=COLVAR
```

### 5.2.31 RDC

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

Calculates the (Residual) Dipolar Coupling between two atoms.

The RDC between two atomic nuclei depends on the  $\theta$  angle between the inter-nuclear vector and the external magnetic field. In isotropic media RDCs average to zero because of the orientational averaging, but when the rotational symmetry is broken, either through the introduction of an alignment medium or for molecules with highly anisotropic paramagnetic susceptibility, RDCs become measurable.

$$D = D_{max} 0.5(3 \cos^2(\theta) - 1)$$

where

$$D_{max} = -\mu_0 \gamma_1 \gamma_2 \hbar / (8\pi^3 r^3)$$

that is the maximal value of the dipolar coupling for the two nuclear spins with gyromagnetic ratio  $\gamma$ .  $\mu$  is the magnetic constant and  $h$  is the Planck constant.

Common Gyromagnetic Ratios (C.G.S)

- H(1) 26.7513
- C(13) 6.7261
- N(15) -2.7116
- NH -72.5388
- CH 179.9319
- CN -18.2385
- CC 45.2404

This collective variable calculates the Residual Dipolar Coupling for a set of couple of atoms using the above definition. From the calculated RDCs and a set of experimental values it calculates either their correlation or the squared quality factor

$$Q^2 = \frac{\sum_i (D_i - D_i^{exp})^2}{\sum_i (D_i^{exp})^2}$$

RDCs report only on the fraction of molecules that is aligned, this means that comparing the RDCs from a single structure in a MD simulation to the experimental values is not particularly meaningful, from this point of view it is better to compare their correlation. The fraction of aligned molecules can be obtained by maximising the correlation

between the calculated and the experimental RDCs. This fraction can be used as a scaling factor in the calculation of the RDCs in order to compare their values. The averaging of the RDCs calculated with the above definition from a standard MD should result to 0 because of the rotational diffusion, but this variable can be used to break the rotational symmetry.

RDCs can also be calculated using a Single Value Decomposition approach, in this case the code rely on the a set of function from the GNU Scientific Library (GSL). (With SVD forces are not currently implemented).

Replica-Averaged restrained simulations can be performed with this CV and the function [ENSEMBLE](#).

Additional material and examples can be also found in the tutorial [Belfast tutorial: NMR restraints](#)

#### Description of components

The names of the components in this action can be customized by the user in the actions input file. However, in addition to these customizable components the following quantities will always be output

Quantity	Description
<b>rdc</b>	the calculated # RDC

In addition the following quantities can be calculated by employing the keywords listed below

Quantity	Keyword	Description
<b>exp</b>	<b>SVD/ADDCOUPplings</b>	the experimental # RDC

The atoms involved can be specified using

<b>ATOMS</b>	the couple of atoms involved in each of the bonds for which you wish to calculate the RDC. Keywords like ATOMS1, ATOMS2, ATOMS3,... should be listed and one dipolar coupling will be calculated for each ATOMS keyword you specify. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Compulsory keywords

<b>GYROM</b>	Add the product of the gyromagnetic constants for the bond.
<b>SCALE</b>	Add the scaling factor to take into account concentration and other effects.

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SVD</b>	( default=off ) Set to TRUE if you want to backcalculate using Single Value Decomposition (need GSL at compilation time).

<b>ADDCOUPPLINGS</b>	( default=off ) Set to TRUE if you want to have fixed components with the experimental values.
<b>COUPLING</b>	Add an experimental value for each coupling (needed by SVD and useful for ef STATS). You can use multiple instances of this keyword i.e. COUPLING1, COUPLING2, COUPLING3...

### Examples

In the following example five N-H RDCs are defined and their correlation with respect to a set of experimental data is calculated and restrained. In addition, and only for analysis purposes, the same RDCs are calculated using a Single Value Decomposition algorithm.

```
RDC ...
GYROM=-72.5388
SCALE=1.0
ATOMS1=20,21
ATOMS2=37,38
ATOMS3=56,57
ATOMS4=76,77
ATOMS5=92,93
LABEL=nh
... RDC

STATS ARG=nh.* PARAMETERS=8.17,-8.271,-10.489,-9.871,-9.152

rdce: RESTRAINT ARG=nh.corr KAPPA=0. SLOPE=-25000.0 AT=1.

RDC ...
GYROM=-72.5388
SCALE=1.0
SVD
ATOMS1=20,21 COUPLING1=8.17
ATOMS2=37,38 COUPLING2=-8.271
ATOMS3=56,57 COUPLING3=-10.489
ATOMS4=76,77 COUPLING4=-9.871
ATOMS5=92,93 COUPLING5=-9.152
LABEL=svd
... RDC

PRINT ARG=nh.corr,rdce.bias FILE=colvar
PRINT ARG=svd.* FILE=svd
```

(See also [PRINT](#), [RESTRAINT](#))

### 5.2.32 TEMPLATE

	This is part of the colvar <a href="#">module</a>
--	---------------------------------------------------

This file provides a template for if you want to introduce a new CV.

The atoms involved can be specified using



<b>ATOMS</b>	the keyword with which you specify what atoms to use should be added like this. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Compulsory keywords

<b>TEMPLATE_COMPULSORY</b>	all compulsory keywords should be added like this with a description here
----------------------------	---------------------------------------------------------------------------

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>TEMPLATE_DEFAULT_OFF_FLAG</b>	( default=off ) flags that are by default not performed should be specified like this
<b>TEMPLATE_DEFAULT_ON_FLAG</b>	( default=on ) flags that are by default performed should be specified like this
<b>TEMPLATE_OPTIONAL</b>	all optional keywords that have input should be added like a description here

#### Examples

```
# This should be a sample input.
t: TEMPLATE ATOMS=1,2
PRINT ARG=t STRIDE=100 FILE=COLVAR
```

(see also [PRINT](#))

### 5.2.33 TORSION

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

Calculate a torsional angle.

This command can be used to compute the torsion between four atoms or alternatively to calculate the angle between two vectors projected on the plane orthogonal to an axis.

The atoms involved can be specified using

<b>ATOMS</b>	the four atoms involved in the torsional angle
--------------	------------------------------------------------

Or alternatively by using

<b>AXIS</b>	two atoms that define an axis. You can use this to find the angle in the plane perpendicular to the axis between the vectors specified using the VECTOR1 and VECTOR2 keywords.
<b>VECTOR1</b>	two atoms that define a vector. You can use this in combination with VECTOR2 and AXIS
<b>VECTOR2</b>	two atoms that define a vector. You can use this in combination with VECTOR1 and AXIS

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>COSINE</b>	( default=off ) calculate cosine instead of dihedral

#### Examples

This input tells plumed to print the torsional angle between atoms 1, 2, 3 and 4 on file COLVAR.

```
t: TORSION ATOMS=1,2,3,4
# this is an alternative, equivalent, definition:
# t: TORSION VECTOR1=2,1 AXIS=2,3 VECTOR2=3,4
PRINT ARG=t FILE=COLVAR
```

If you are working with a protein you can specify the special named torsion angles  $\phi$ ,  $\psi$ ,  $\omega$  and  $\chi_1$  by using TORSION in combination with the **MOLINFO** command. This can be done by using the following syntax.

```
MOLINFO MOLTYPE=protein STRUCTURE=myprotein.pdb
t1: TORSION ATOMS=@phi-3
t2: TORSION ATOMS=@psi-4
PRINT ARG=t1,t2 FILE=colvar STRIDE=10
```

Here, @phi-3 tells plumed that you would like to calculate the  $\phi$  angle in the third residue of the protein. Similarly @psi-4 tells plumed that you want to calculate the  $\psi$  angle of the 4th residue of the protein.

### 5.2.34 VOLUME

This is part of the colvar <b>module</b>
------------------------------------------

Calculate the volume of the simulation box.

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
------------------------------	----------------------------------------------------------------------------

### Examples

The following input tells plumed to print the volume of the system

```
VOLUME LABEL=vol
PRINT ARG=vol
```

(See also [PRINT](#)).

## 5.3 Distances from reference configurations

One colvar that has been shown to be very successful in studying protein folding is the distance between the instantaneous configuration and a reference configuration - often the structure of the folded state. When the free energy of a protein is shown as a function of this collective variable there is a minima for low values of the CV, which is due to the folded state of the protein. There is then a second minima at higher values of the CV, which is the minima corresponding to the unfolded state.

A slight problem with this sort of collective variable is that there are many different ways of calculating the distance from a particular reference structure. The simplest - adding together the distances by which each of the atoms has been translated in going from the reference configuration to the instantaneous configuration - is not particularly sensible. A distance calculated in this way does not neglect translation of the center of mass of the molecule and rotation of the frame of reference. A common practise is thus to remove these components by calculating the [RMSD](#) distance between the reference and instantaneous configurations. This is not the only way to calculate the distance, however. One could also calculate the total amount by which a large number of collective variables change in moving from the reference to the instantaneous configurations. One could even combine RMSD distances with the amount the collective variables change. A full list of the ways distances can be measured in PLUMED is given below:

<a href="#">DRMSD</a>	Calculate the distance RMSD with respect to a reference structure.
<a href="#">MULTI-RMSD</a>	Calculate the RMSD distance moved by a number of separated domains from their positions in a reference structure.
<a href="#">PCARMSD</a>	Calculate the PCA components ( see <a href="#">[21]</a> and <a href="#">[22]</a> ) for a number of provided eigenvectors and an average structure. Performs optimal alignment at every step and reports the rmsd so you know if you are far or close from the average structure. It takes the average structure and eigenvectors in form of a pdb. Note that beta and occupancy values in the pdb are neglected and all the weights are placed to 1 (differently from the RMSD colvar for example)
<a href="#">RMSD</a>	Calculate the RMSD with respect to a reference structure.
<a href="#">TARGET</a>	This function measures the pythagorean distance from a particular structure measured in the space defined by some set of collective variables.

These options for calculating distances are re-used in a number of places in the code. For instance they are used in some of the analysis algorithms that are implemented in PLUMED and in [PATH](#) collective variables. Notice that most of these actions read the reference configuration from a PDB file. Be sure you understand how to format properly a PDB file to use in PLUMED (see [pdbreader](#)).

## 5.3.1 DRMSD

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

Calculate the distance RMSD with respect to a reference structure.

To calculate the root-mean-square deviation between the atoms in two configurations you must first superimpose the two structures in some ways. Obviously, it is the internal vibrational motions of the structure - i.e. not the translations and rotations - that are interesting. However, aligning two structures by removing the translational and rotational motions is not easy. Furthermore, in some cases there can be alignment issues caused by so-called frame-fitting problems. It is thus often cheaper and easier to calculate the distances between all the pairs of atoms. The distance between the two structures,  $\mathbf{X}^a$  and  $\mathbf{X}^b$  can then be measured as:

$$d(\mathbf{X}^A, \mathbf{X}^B) = \sqrt{\frac{1}{N(N-1)} \sum_{i \neq j} [d(\mathbf{x}_i^a, \mathbf{x}_j^a) - d(\mathbf{x}_i^b, \mathbf{x}_j^b)]^2}$$

where  $N$  is the number of atoms and  $d(\mathbf{x}_i, \mathbf{x}_j)$  represents the distance between atoms  $i$  and  $j$ . Clearly, this representation of the configuration is invariant to translation and rotation. However, it can become expensive to calculate when the number of atoms is large. This can be resolved within the DRMSD colvar by setting LOWER\_CUTOFF and UPPER\_CUTOFF. These keywords ensure that only pairs of atoms that are within a certain range are incorporated into the above sum.

In PDB files the atomic coordinates and box lengths should be in Angstroms unless you are working with natural units. If you are working with natural units then the coordinates should be in your natural length unit. For more details on the PDB file format visit <http://www.wwpdb.org/docs.html>

## Compulsory keywords

<b>REFERENCE</b>	a file in pdb format containing the reference structure and the atoms involved in the CV.
<b>LOWER_CUTOFF</b>	only pairs of atoms further than LOWER_CUTOFF are considered in the calculation.
<b>UPPER_CUTOFF</b>	only pairs of atoms closer than UPPER_CUTOFF are considered in the calculation.
<b>TYPE</b>	( default=DRMSD ) what kind of DRMSD would you like to calculate. You can use either the normal DRMSD involving all the distances between the atoms in your molecule. Alternatively, if you have multiple molecules you can use the type INTER-DRMSD to compute DRMSD values involving only those distances between the atoms at least two molecules or the type INTRA-DRMSD to compute DRMSD values involving only those distances between atoms in the same molecule

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances

## Examples

The following tells plumed to calculate the distance RMSD between the positions of the atoms in the reference file and their instantaneous position. Only pairs of atoms whose distance in the reference structure is within 0.1 and 0.8 nm are considered.

```
DRMSD REFERENCE=file.pdb LOWER_CUTOFF=0.1 UPPER_CUTOFF=0.8
```

The following tells plumed to calculate a DRMSD value for a pair of molecules.

```
DRMSD REFERENCE=file.pdb LOWER_CUTOFF=0.1 UPPER_CUTOFF=0.8 TYPE=INTER-DRMSD
```

In the input reference file (file.pdb) the atoms in each of the two molecules are separated by a TER command as shown below.

```
ATOM      8  HT3  ALA    2      -1.480  -1.560   1.212   1.00   1.00      DIA  H
ATOM      9  CAY  ALA    2      -0.096   2.144  -0.669   1.00   1.00      DIA  C
ATOM     10  HY1  ALA    2       0.871   2.385  -0.588   1.00   1.00      DIA  H
TER
ATOM     12  HY3  ALA    2      -0.520   2.679  -1.400   1.00   1.00      DIA  H
ATOM     14  OY   ALA    2      -1.139   0.931  -0.973   1.00   1.00      DIA  O
END
```

In this example the INTER-DRMSD type ensures that the set of distances from which the final quantity is computed involve one atom from each of the two molecules. If this is replaced by INTRA-DRMSD then only those distances involving pairs of atoms that are both in the same molecule are computed.

### 5.3.2 MULTI-RMSD

This is part of the <a href="#">colvar module</a>
---------------------------------------------------

Calculate the RMSD distance moved by a number of separated domains from their positions in a reference structure.

When you have large proteins the calculation of the root mean squared deviation between all the atoms in a reference structure and the instantaneous configuration becomes prohibitively expensive. You may thus instead want to calculate the RMSD between the atoms in a set of domains of your protein and your reference structure. That is to say:

$$d(X, X_r) = \sqrt{\sum_i w_i |X_i - X'_i|^2}$$

where here the sum is over the domains of the protein,  $X_i$  represents the positions of the atoms in domain  $i$  in the instantaneous configuration and  $X'_i$  is the positions of the atoms in domain  $i$  in the reference configuration.  $w_i$  is an optional weight.

The distances for each of the domains in the above sum can be calculated using the [DRMSD](#) or [RMSD](#) measures or using a combination of these distance. The reference configuration is specified in a pdb file like the one below:

```
ATOM      2  O   ALA    2      -0.926  -2.447  -0.497   1.00   1.00      DIA  O
ATOM      4  HNT ALA    2       0.533  -0.396   1.184   1.00   1.00      DIA  H
ATOM      6  HT1 ALA    2      -0.216  -2.590   1.371   1.00   1.00      DIA  H
ATOM      7  HT2 ALA    2      -0.309  -1.255   2.315   1.00   1.00      DIA  H
```

```

ATOM      8  HT3  ALA      2      -1.480  -1.560   1.212   1.00   1.00      DIA  H
ATOM      9  CAY  ALA      2      -0.096   2.144  -0.669   1.00   1.00      DIA  C
ATOM     10  HY1  ALA      2       0.871   2.385  -0.588   1.00   1.00      DIA  H
TER
ATOM     12  HY3  ALA      2      -0.520   2.679  -1.400   1.00   1.00      DIA  H
ATOM     14  OY   ALA      2      -1.139   0.931  -0.973   1.00   1.00      DIA  O
ATOM     16  HN   ALA      2       1.713   1.021  -0.873   1.00   1.00      DIA  H
ATOM     18  HA   ALA      2       0.099  -0.774  -2.218   1.00   1.00      DIA  H
ATOM     19  CB   ALA      2       2.063  -1.223  -1.276   1.00   1.00      DIA  C
ATOM     20  HB1  ALA      2       2.670  -0.716  -2.057   1.00   1.00      DIA  H
ATOM     21  HB2  ALA      2       2.556  -1.051  -0.295   1.00   1.00      DIA  H
ATOM     22  HB3  ALA      2       2.070  -2.314  -1.490   1.00   1.00      DIA  H
END

```

with the TER keyword being used to separate the various domains in you protein.

### Compulsory keywords

<b>REFERENCE</b>	a file in pdb format containing the reference structure and the atoms involved in the CV.
<b>TYPE</b>	( default=MULTI-SIMPLE ) the manner in which RMSD alignment is performed. Should be MULTI-OPTIMAL, MULTI-OPTIMAL-FAST, MULTI-SIMPLE or MULTI-DRMSD.

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>SQUARED</b>	( default=off ) This should be setted if you want MSD instead of RMSD

### Examples

The following tells plumed to calculate the RMSD distance between the positions of the atoms in the reference file and their instantaneous position. The Kearseley algorithm for each of the domains.

```
MULTI-RMSD REFERENCE=file.pdb TYPE=MULTI-OPTIMAL
```

The following tells plumed to calculate the RMSD distance between the positions of the atoms in the domains of reference the reference structure and their instantaneous positions. Here distances are calculated using the [DRMSD](#) measure.

```
MULTI-RMSD REFERENCE=file.pdb TYPE=MULTI-DRMSD
```

in this case it is possible to use the following DRMSD options in the pdb file using the REMARK syntax:

```

NOPBC to calculate distances without PBC
LOWER_CUTOFF=# only pairs of atoms further than LOWER_CUTOFF are considered in the calculation
UPPER_CUTOFF=# only pairs of atoms further than UPPER_CUTOFF are considered in the calculation

```

as shown in the following example

```

REMARK NOPBC
REMARK LOWER_CUTOFF=0.1
REMARK UPPER_CUTOFF=0.8
ATOM      2  O  ALA      2      -0.926  -2.447  -0.497  1.00  1.00      DIA  O
ATOM      4  HNT ALA      2       0.533  -0.396   1.184  1.00  1.00      DIA  H
ATOM      6  HT1 ALA      2      -0.216  -2.590   1.371  1.00  1.00      DIA  H
ATOM      7  HT2 ALA      2      -0.309  -1.255   2.315  1.00  1.00      DIA  H
ATOM      8  HT3 ALA      2     -1.480  -1.560   1.212  1.00  1.00      DIA  H
ATOM      9  CAY ALA      2     -0.096   2.144  -0.669  1.00  1.00      DIA  C
ATOM     10  HY1 ALA      2       0.871   2.385  -0.588  1.00  1.00      DIA  H
TER
ATOM     12  HY3 ALA      2     -0.520   2.679  -1.400  1.00  1.00      DIA  H
ATOM     14  OY  ALA      2     -1.139   0.931  -0.973  1.00  1.00      DIA  O
ATOM     16  HN  ALA      2       1.713   1.021  -0.873  1.00  1.00      DIA  H
ATOM     18  HA  ALA      2       0.099  -0.774  -2.218  1.00  1.00      DIA  H
ATOM     19  CB  ALA      2       2.063  -1.223  -1.276  1.00  1.00      DIA  C
ATOM     20  HB1 ALA      2       2.670  -0.716  -2.057  1.00  1.00      DIA  H
ATOM     21  HB2 ALA      2       2.556  -1.051  -0.295  1.00  1.00      DIA  H
ATOM     22  HB3 ALA      2       2.070  -2.314  -1.490  1.00  1.00      DIA  H
END

```

### 5.3.3 PCARMSD

This is part of the colvar <a href="#">module</a>
---------------------------------------------------

Calculate the PCA components ( see [21] and [22] ) for a number of provided eigenvectors and an average structure. Performs optimal alignment at every step and reports the rmsd so you know if you are far or close from the average structure. It takes the average structure and eigenvectors in form of a pdb. Note that beta and occupancy values in the pdb are neglected and all the weights are placed to 1 (differently from the RMSD colvar for example)

#### Description of components

By default the value of the calculated quantity can be referenced elsewhere in the input file by using the label of the action. Alternatively this Action can be used to calculate the following quantities by employing the keywords listed below. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>eig</b>	the projections on each eigenvalue are stored on values labeled eig-1, eig-2, ...
<b>residual</b>	the distance of the present configuration from the configuration supplied as AVERAGE in terms of MSD after optimal alignment

#### Compulsory keywords

<b>AVERAGE</b>	a file in pdb format containing the reference structure and the atoms involved in the CV.
<b>EIGENVECTORS</b>	a file in pdb format containing the reference structure and the atoms involved in the CV.

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SQUARED-ROOT</b>	( default=off ) This should be setted if you want RMSD instead of MSD

### Examples

```
PCARMSD AVERAGE=file.pdb EIGENVECTORS=eigenvectors.pdb
```

The input is taken so to be compatible with the output you get from `g_covar` utility of gromacs (suitably adapted to have a pdb input format).

### 5.3.4 RMSD

This is part of the **colvar module**

Calculate the RMSD with respect to a reference structure.

The aim with this colvar is to calculate something like:

$$d(X, X') = |X - X'|$$

where  $X$  is the instantaneous position of all the atoms in the system and  $X'$  is the positions of the atoms in some reference structure provided as input.  $d(X, X')$  thus measures the distance all the atoms have moved away from this reference configuration. Oftentimes, it is only the internal motions of the structure - i.e. not the translations of the center of mass or the rotations of the reference frame - that are interesting. Hence, when calculating the the root-mean-square deviation between the atoms in two configurations you must first superimpose the two structures in some way. At present PLUMED provides two distinct ways of performing this superposition. The first method is applied when you use `TYPE=SIMPLE` in the input line. This instruction tells PLUMED that the root mean square deviation is to be calculated after the positions of the geometric centers in the reference and instantaneous configurations are aligned. In other words  $d(X, x')$  is to be calculated using:

$$d(X, X') = \sqrt{\sum_i \sum_{\alpha}^{x,y,z} \frac{w_i}{\sum_j w_j} (X_{i,\alpha} - com_{\alpha}(X) - X'_{i,\alpha} + com_{\alpha}(X'))^2}$$

with

$$com_{\alpha}(X) = \sum_i \frac{w'_i}{\sum_j w'_j} X_{i,\alpha}$$

and

$$com_{\alpha}(X') = \sum_i \frac{w_i}{\sum_j w_j} X'_{i,\alpha}$$

Obviously,  $com_{\alpha}(X)$  and  $com_{\alpha}(X')$  represent the positions of the center of mass in the reference and instantaneous configurations if the weights  $w$  are set equal to the atomic masses. If the weights are all set equal to one, however,  $com_{\alpha}(X)$  and  $com_{\alpha}(X')$  are the positions of the geometric centers. Notice that there are sets of weights:  $w'$  and  $w$ . The first is used to calculate the position of the center of mass (so it determines how the atoms are *aligned*). Meanwhile, the second is used when calculating how far the atoms have actually been *displaced*.



These weights are assigned in the reference configuration that you provide as input (i.e. they appear in the input file to this action that you set using `REFERENCE=whatever.pdb`). This input reference configuration consists of a simple pdb file containing the set of atoms for which you want to calculate the RMSD displacement and their positions in the reference configuration. It is important to note that the indices in this pdb need to be set correctly. The indices in this file determine the indices of the instantaneous atomic positions that are used by PLUMED when calculating this colvar. As such if you want to calculate the RMSD distance moved by the 1st, 4th, 6th and 28th atoms in the MD codes input file then the indices of the corresponding reference positions in this pdb file should be set equal to 1, 4, 6 and 28.

The pdb input file should also contain the values of  $w$  and  $w'$ . In particular, the OCCUPANCY column (the first column after the coordinates) is used to provide the values of  $w'$  that are used to calculate the position of the centre of mass. The BETA column (the second column after the Cartesian coordinates) is used to provide the  $w$  values which are used in the calculation of the displacement. Please note that it is possible to use fractional values for beta and for the occupancy. However, we recommend you only do this when you really know what you are doing however as the results can be rather strange.

In PDB files the atomic coordinates and box lengths should be in Angstroms unless you are working with natural units. If you are working with natural units then the coordinates should be in your natural length unit. For more details on the PDB file format visit <http://www.wwpdb.org/docs.html>. Make sure your PDB file is correctly formatted as explained [in this page](#).

A different method is used to calculate the RMSD distance when you use `TYPE=OPTIMAL` on the input line. In this case the root mean square deviation is calculated after the positions of geometric centers in the reference and instantaneous configurations are aligned AND after an optimal alignment of the two frames is performed so that motion due to rotation of the reference frame between the two structures is removed. The equation for  $d(X, X')$  in this case reads:

$$d(X, X') = \sqrt{\sum_i \sum_{\alpha}^{x,y,z} \frac{w_i}{\sum_j w_j} [X_{i,\alpha} - com_{\alpha}(X) - \sum_{\beta} M(X, X', w')_{\alpha,\beta} (X'_{i,\beta} - com_{\beta}(X'))]^2}$$

where  $M(X, X', w')$  is the optimal alignment matrix which is calculated using the Kearsley [23] algorithm. Again different sets of weights are used for the alignment ( $w'$ ) and for the displacement calculations ( $w$ ). This gives a great deal of flexibility as it allows you to use a different sets of atoms (which may or may not overlap) for the alignment and displacement parts of the calculation. This may be very useful when you want to calculate how a ligand moves about in a protein cavity as you can use the protein as a reference system and do no alignment of the ligand.

(Note: when this form of RMSD is used to calculate the secondary structure variables ([ALPHARMSD](#), [ANTIBETARMSD](#) and [PARABETARMSD](#) all the atoms in the segment are assumed to be part of both the alignment and displacement sets and all weights are set equal to one)

Please note that there are a number of other methods for calculating the distance between the instantaneous configuration and a reference configuration that are available in plumed. More information on these various methods can be found in the section of the manual on [Distances from reference configurations](#).

#### Warning

The molecule used for [RMSD](#) calculation should be whole (both atoms used in alignment and in displacement calculation). In case it is broken by the host MD code, please use [WHOLEMOLECULES](#) to reconstruct it before [RMSD](#) calculation.

#### Compulsory keywords

<b>REFERENCE</b>	a file in pdb format containing the reference structure and the atoms involved in the CV.
<b>TYPE</b>	( default=SIMPLE ) the manner in which RMSD alignment is performed. Should be OPTIMAL or SIMPLE.

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>SQUARED</b>	( default=off ) This should be setted if you want MSD instead of RMSD

### Examples

The following tells plumed to calculate the RMSD distance between the positions of the atoms in the reference file and their instantaneous position. The Kearsley algorithm is used so this is done optimally.

```
RMSD REFERENCE=file.pdb TYPE=OPTIMAL
```

...

### 5.3.5 TARGET

This is part of the function <a href="#">module</a>
-----------------------------------------------------

This function measures the pythagorean distance from a particular structure measured in the space defined by some set of collective variables.

This collective variable can be used to calculate something akin to:

$$d(X, X') = |X - X'|$$

where  $X$  is the instaneous values for a set of collective variables for the system and  $X'$  is the values that these self-same set of collective variables take in some reference structure provided as input. If we call our set of collective variables  $\{s_i\}$  then this CV computes:

$$d = \sqrt{\sum_{i=1}^N (s_i - s_i^{(ref)})^2}$$

where  $s_i^{(ref)}$  are the values of the CVs in the reference structure and  $N$  is the number of input CVs.

We can also calculate normalized euclidean differences using this action and the METRIC=NORM-EUCLIDEAN flag. In other words, we can compute:

$$d = \sqrt{\sum_{i=1}^N \sigma_i (s_i - s_i^{(ref)})^2}$$

where  $\sigma_i$  is a vector of weights. Lastly, by using the METRIC=MAHALONOBIS we can compute mahalonobis distances using:

$$d = \left( \mathbf{s} - \mathbf{s}^{(ref)} \right)^T \mathbf{\Sigma} \left( \mathbf{s} - \mathbf{s}^{(ref)} \right)$$

where  $\mathbf{s}$  is a column vector containing the values of all the CVs and  $\mathbf{s}^{(ref)}$  is a column vector containing the values of the CVs in the reference configuration.  $\mathbf{\Sigma}$  is then an  $N \times N$  matrix that is specified in the input.

#### Compulsory keywords

<b>TYPE</b>	( default=EUCLIDEAN ) the manner in which the distance should be calculated
<b>REFERENCE</b>	a file in pdb format containing the reference structure. In the PDB file the atomic coordinates and box lengths should be in Angstroms unless you are working with natural units. If you are working with natural units then the coordinates should be in your natural length unit. The charges and masses of the atoms (if required) should be inserted in the beta and occupancy columns respectively. For more details on the PDB file format visit <a href="http://www.wwpdb.org/docs.html">http://www.wwpdb.org/docs.html</a>

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
------------------------------	----------------------------------------------------------------------------

#### Examples

The following input calculates the distance between a reference configuration and the instantaneous position of the system in the trajectory. The position of the reference configuration is specified by providing the values of the distance between atoms 1 and 2 and atoms 3 and 4.

```
d1: DISTANCE ATOMS=1,2
d2: DISTANCE ATOMS=3,4
t1: TARGET REFERENCE=myref.pdb TYPE=EUCLIDEAN
PRINT ARG=t1 FILE=colvar
```

The contents of the file containing the reference structure (myref.pdb) is shown below. As you can see you must provide information on the labels of the CVs that are being used to define the position of the reference configuration in this file together with the values that these quantities take in the reference configuration.

```
DESCRIPTION: a reference point.
REMARK WEIGHT=1.0
REMARK ARG=d1,d2
REMARK d1=1.0 d2=1.0
END
```

## 5.4 Functions

When performing biased dynamics or analysing a trajectory you may wish to analyse/bias the value of some function of a set of collective variables rather than the values of the collective variables directly. You can do this with PLUMED by using any one of the following list of functions.

Notice that in many functions you should explicitly say to PLUMED whether the result is a periodic variable or not using the keyword `PERIODIC`. This is crucial to allow a variable to be properly biased. To know if a function is periodic or not you should answer to the following question:

- Can my function change with a discontinuity when I move my atoms in a continuous manner?

In case the answer is no, then you should use `PERIODIC=NO`. In case the answer is yes, then you should consider the following question:

- Are the values of the function at the discontinuity always the same or do they change?

In case the answer is that they are the same, you should use `PERIODIC=A, B` where A is the smallest value and B is the largest value. In case the answer is that the values at the discontinuity are not always the same, then you cannot construct a variable that can be biased with PLUMED. Consider the following examples:

```
t: TORSION ATOMS=1,2,3,4
# When atoms are moved, t could jump suddenly from -pi to +pi

c: MATHEVAL ARG=t FUNC=x*x*x PERIODIC=-31.0062766802998,31.0062766802998
# When atoms are moved, c could jump suddenly from -pi**3 to +pi**3

# equivalently, we could have used:
# c: COMBINE ARG=t POWERS=3 PERIODIC=-31.0062766802998,31.0062766802998

# compute x/y/z components of the distance between atoms 1 and 10
d: DISTANCE ATOMS=1,10 COMPONENTS

# make a new variable equal to d.z but with the correct periodicity
dz: COMBINE ARG=d.z PERIODIC=-10,10
# here we assumed the system is in a orthorhombic box with z side = 20
```

<b>COMBINE</b>	Calculate a polynomial combination of a set of other variables.
<b>ENSEMBLE</b>	Calculates the replica averaging of a collective variable over multiple replicas.
<b>FUNCPATHMSD</b>	This function calculates path collective variables.
<b>FUNCSUMHILLS</b>	This function is intended to be called by the command line tool <code>sum_hills</code> and it is meant to integrate a HILLS file or an HILLS file interpreted as a histogram in a variety of ways. Therefore it is not expected that you use this during your dynamics (it will crash!)
<b>LOCALENSEMBLE</b>	Calculates the average over multiple arguments.
<b>MATHEVAL</b>	Calculate a combination of variables using a matheval expression.
<b>PIECEWISE</b>	Compute a piecewise straight line through its arguments that passes through a set of ordered control points.
<b>SORT</b>	This function can be used to sort colvars according to their magnitudes.
<b>STATS</b>	Calculates statistical properties of a set of collective variables with respect to a set of reference values. In particular it calculates and store as components the sum of the squared deviations, the correlation, the slope and the intercept of a linear fit.

## 5.4.1 COMBINE

<b>This is part of the function <a href="#">module</a></b>
------------------------------------------------------------

Calculate a polynomial combination of a set of other variables.

The functional form of this function is

$$C = \sum_{i=1}^{N_{arg}} c_i (x_i - a_i)^{p_i}$$

The coefficients c, the parameters a and the powers p are provided as vectors.

Notice that COMBINE is not able to predict which will be periodic domain of the computed value automatically. The user is thus forced to specify it explicitly. Use PERIODIC=NO if the resulting variable is not periodic, and PERIODIC=A,B where A and B are the two boundaries if the resulting variable is periodic.

## Compulsory keywords

<b>PERIODIC</b>	if the output of your function is periodic then you should specify the periodicity of the function. If the output is not periodic you must state this using PERIODIC=NO
<b>COEFFICIENTS</b>	( default=1.0 ) the coefficients of the arguments in your function
<b>PARAMETERS</b>	( default=0.0 ) the parameters of the arguments in your function
<b>POWERS</b>	( default=1.0 ) the powers to which you are raising each of the arguments in your function

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NORMALIZE</b>	( default=off ) normalize all the coefficients so that in total they are equal to one
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...

## Examples

The following input tells plumed to print the distance between atoms 3 and 5 its square (as computed from the x,y,z components) and the distance again as computed from the square root of the square.

```
DISTANCE LABEL=dist      ATOMS=3,5 COMPONENTS
COMBINE  LABEL=distance2 ARG=dist.x,dist.y,dist.z POWERS=2,2,2 PERIODIC=NO
COMBINE  LABEL=distance  ARG=distance2 POWERS=0.5 PERIODIC=NO
PRINT ARG=distance,distance2
```

(See also [PRINT](#) and [DISTANCE](#)).

The following input tells plumed to add a restraint on the cube of a dihedral angle. Notice that since the angle has a periodic domain  $-\pi, \pi$  its cube has a domain  $-\pi^3, \pi^3$ .

```
t: TORSION ATOMS=1,3,5,7
c: COMBINE ARG=t POWERS=3 PERIODIC=-31.0062766802998,31.0062766802998
RESTRAINT ARG=c KAPPA=10 AT=0
```

## 5.4.2 ENSEMBLE

	This is part of the function <a href="#">module</a>
--	-----------------------------------------------------

Calculates the replica averaging of a collective variable over multiple replicas.

Each collective variable is averaged separately and stored in a component labelled *label.cvlabel*.

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>REWEIGHT</b>	( default=off ) simple REWEIGHT using the latest ARG as energy
<b>CENTRAL</b>	( default=off ) calculate a central moment instead of a standard moment
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>TEMP</b>	the system temperature - this is only needed if you are reweighting
<b>MOMENT</b>	the moment you want to calculate in alternative to the mean or the variance
<b>POWER</b>	the power of the mean (and moment)

## Examples

The following input tells plumed to calculate the distance between atoms 3 and 5 and the average it over the available replicas.

```
dist: DISTANCE ATOMS=3,5
ens: ENSEMBLE ARG=dist
PRINT ARG=dist,ens.dist
```

(See also [PRINT](#) and [DISTANCE](#)).

## 5.4.3 FUNCPATHMSD

	This is part of the function <a href="#">module</a>
--	-----------------------------------------------------

This function calculates path collective variables.

This is the Path Collective Variables implementation ( see [\[17\]](#) ). This variable computes the progress along a given set of frames that is provided in input ("s" component) and the distance from them ("z" component). It is a function of MSD that are obtained by the joint use of MSD variable and SQUARED flag (see below).

## Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>s</b>	the position on the path
<b>z</b>	the distance from the path

## Compulsory keywords

<b>LAMBDA</b>	the lambda parameter is needed for smoothing, is in the units of plumed
---------------	-------------------------------------------------------------------------

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
------------------------------	----------------------------------------------------------------------------

<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>NEIGH_SIZE</b>	size of the neighbor list
<b>NEIGH_STRIDE</b>	how often the neighbor list needs to be calculated in time units

### Examples

Here below is a case where you have defined three frames and you want to calculate the progress along the path and the distance from it in p1

```
t1: RMSD REFERENCE=frame_1.dat TYPE=OPTIMAL SQUARED
t2: RMSD REFERENCE=frame_21.dat TYPE=OPTIMAL SQUARED
t3: RMSD REFERENCE=frame_42.dat TYPE=OPTIMAL SQUARED
p1: FUNCPATHMSD ARG=t1,t2,t3 LAMBDA=500.0
PRINT ARG=t1,t2,t3,p1.s,p1.z STRIDE=1 FILE=colvar FMT=%8.4f
```

In this second example is shown how to define a PATH in the [CONTACTMAP](#) space:

```
CONTACTMAP ...
ATOMS1=1,2 REFERENCE1=0.1
ATOMS2=3,4 REFERENCE2=0.5
ATOMS3=4,5 REFERENCE3=0.25
ATOMS4=5,6 REFERENCE4=0.0
SWITCH={RATIONAL R_0=1.5}
LABEL=c1
CMDIST
... CONTACTMAP
```

```
CONTACTMAP ...
ATOMS1=1,2 REFERENCE1=0.3
ATOMS2=3,4 REFERENCE2=0.9
ATOMS3=4,5 REFERENCE3=0.45
ATOMS4=5,6 REFERENCE4=0.1
SWITCH={RATIONAL R_0=1.5}
LABEL=c2
CMDIST
... CONTACTMAP
```

```
CONTACTMAP ...
ATOMS1=1,2 REFERENCE1=1.0
ATOMS2=3,4 REFERENCE2=1.0
ATOMS3=4,5 REFERENCE3=1.0
ATOMS4=5,6 REFERENCE4=1.0
SWITCH={RATIONAL R_0=1.5}
LABEL=c3
CMDIST
... CONTACTMAP
```

```
p1: FUNCPATHMSD ARG=c1,c2,c3 LAMBDA=500.0
PRINT ARG=c1,c2,c3,p1.s,p1.z STRIDE=1 FILE=colvar FMT=%8.4f
```



## 5.4.4 FUNCSUMHILLS

This is part of the function <a href="#">module</a>
-----------------------------------------------------

This function is intended to be called by the command line tool `sum_hills` and it is meant to integrate a HILLS file or an HILLS file interpreted as a histogram in a variety of ways. Therefore it is not expected that you use this during your dynamics (it will crash!)

In the future one could implement periodic integration during the metadynamics or straightforward MD as a tool to check convergence

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ISCLTOOL</b>	( default=on ) use via plumed commandline: calculate at read phase and then go
<b>PARALLELREAD</b>	( default=off ) read parallel HILLS file
<b>NEGBIAS</b>	( default=off ) dump negative bias ( -bias ) instead of the free energy: needed in welltempered with flexible hills
<b>NOHISTORY</b>	( default=off ) to be used with INITSTRIDE: it splits the bias/histogram in pieces without previous history
<b>MINTOZERO</b>	( default=off ) translate the resulting bias/histogram to have the minimum to zero
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the proceeding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>HILLSFILES</b>	source file for hills creation(may be the same as HILLS)
<b>HISTOFILES</b>	source file for histogram creation(may be the same as HILLS)
<b>HISTOSIGMA</b>	sigmas for binning when the histogram correction is needed
<b>PROJ</b>	only with sumhills: the projection on the cvs
<b>KT</b>	only with sumhills: the kt factor when projection on cvs
<b>GRID_MIN</b>	the lower bounds for the grid
<b>GRID_MAX</b>	the upper bounds for the grid
<b>GRID_BIN</b>	the number of bins for the grid
<b>GRID_SPACING</b>	the approximate grid spacing (to be used as an alternative or together with GRID_BIN)
<b>INTERVAL</b>	set monodimensional INTERVAL
<b>OUTHILLS</b>	output file for hills

<b>OUTHISTO</b>	output file for histogram
<b>INITSTRIDE</b>	stride if you want an initial dump
<b>STRIDE</b>	stride when you do it on the fly
<b>FMT</b>	the format that should be used to output real numbers

### Examples

There are currently no examples for this keyword.

### 5.4.5 LOCALENSEMBLE

	This is part of the function <a href="#">module</a>
--	-----------------------------------------------------

Calculates the average over multiple arguments.

If more than one collective variable is given for each argument then they are averaged separately. The average is stored in a component labelled *label.cvlabel*.

### Compulsory keywords

<b>NUM</b>	the number of local replicas
------------	------------------------------

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...



## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three components x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>VAR</b>	the names to give each of the arguments in the function. If you have up to three arguments in your function you can use x, y and z to refer to them. Otherwise you must use this flag to give your variables names.

## Examples

The following input tells plumed to perform a metadynamics using as a CV the difference between two distances.

```
dAB: DISTANCE ATOMS=10,12
dAC: DISTANCE ATOMS=10,15
diff: MATHEVAL ARG=dAB,dAC FUNC=y-x PERIODIC=NO
# notice: the previous line could be replaced with the following
# diff: COMBINE ARG=dAB,dAC COEFFICIENTS=-1,1
METAD ARG=diff WIDTH=0.1 HEIGHT=0.5 BIASFACTOR=10 PACE=100
```

(see also [DISTANCE](#), [COMBINE](#), and [METAD](#)). Notice that forces applied to diff will be correctly propagated to atoms 10, 12, and 15. Also notice that since MATHEVAL is used without the VAR option the two arguments should be referred to as x and y in the expression FUNC. For simple functions such as this one it is possible to use [COMBINE](#), which does not require libmatheval to be installed on your system.

The following input tells plumed to print the angle between vectors identified by atoms 1,2 and atoms 2,3 its square (as computed from the x,y,z components) and the distance again as computed from the square root of the square.

```
DISTANCE LABEL=d1 ATOMS=1,2 COMPONENTS
DISTANCE LABEL=d2 ATOMS=2,3 COMPONENTS
MATHEVAL ...
  LABEL=theta
  ARG=d1.x,d1.y,d1.z,d2.x,d2.y,d2.z
  VAR=ax,ay,az,bx,by,bz
  FUNC=acos((ax*bx+ay*by+az*bz)/sqrt((ax*ax+ay*ay+az*az)*(bx*bx+by*by+bz*bz)))
  PERIODIC=NO
... MATHEVAL
PRINT ARG=theta
```

(See also [PRINT](#) and [DISTANCE](#)).

Notice that the matheval library implements a large number of functions (trigonometric, exp, log, etc). Among the useful functions, have a look at the step function (that is the Heaviside function). `step(x)` is defined as 1 when  $x$  is positive and 0 when  $x$  is negative. This allows for a straightforward implementation of if clauses.

For example, imagine that you want to implement a restraint that only acts when a distance is larger than 0.5. You can do it with

```
d: DISTANCE ATOMS=10,15
m: MATHEVAL ARG=d FUNC=0.5*step(0.5-x)+x*step(x-0.5) PERIODIC=NO
# check the function you are applying:
PRINT ARG=d,n FILE=checkme
RESTRAINT ARG=d AT=0.5 KAPPA=10.0
```

(see also [DISTANCE](#), [PRINT](#), and [RESTRAINT](#))

The meaning of the function  $0.5*\text{step}(0.5-x)+x*\text{step}(x-0.5)$  is:

- If  $x < 0.5$  ( $\text{step}(0.5-x) \neq 0$ ) use 0.5
- If  $x > 0.5$  ( $\text{step}(x-0.5) \neq 0$ ) use  $x$  Notice that the same could have been obtained using an [UPPER\\_WALLS](#) However, with MATHEVAL you can create way more complex definitions.

#### Warning

If you apply forces on the variable (as in the previous example) you should make sure that the variable is continuous! Conversely, if you are just analyzing a trajectory you can safely use discontinuous variables.

A possible continuity check with gnuplot is

```
# this allow to step function to be used in gnuplot:
gnuplot> step(x)=0.5*(erf(x*10000000)+1)
# here you can test your function
gnuplot> p 0.5*step(0.5-x)+x*step(x-0.5)
```

Also notice that you can easily make logical operations on the conditions that you create. The equivalent of the AND operator is the product:  $\text{step}(1.0-x)*\text{step}(x-0.5)$  is only equal to 1 when  $x$  is between 0.5 and 1.0. By combining negation and AND you can obtain an OR. That is,  $1-\text{step}(1.0-x)*\text{step}(x-0.5)$  is only equal to 1 when  $x$  is outside the 0.5-1.0 interval.

MATHEVAL can be used in combination with [DISTANCE](#) to implement variants of the DISTANCE keyword that were present in PLUMED 1.3 and that allowed to compute the distance of a point from a line defined by two other points, or the progression along that line.

```
# take center of atoms 1 to 10 as reference point 1
p1: CENTER ATOMS=1-10
# take center of atoms 11 to 20 as reference point 2
p2: CENTER ATOMS=11-20
# take center of atoms 21 to 30 as reference point 3
p3: CENTER ATOMS=21-30

# compute distances
d12: DISTANCE ATOMS=p1,p2
d13: DISTANCE ATOMS=p1,p3
d23: DISTANCE ATOMS=p2,p3

# compute progress variable of the projection of point p3
# along the vector joining p1 and p2
# notice that progress is measured from the middle point
onaxis: MATHEVAL ARG=d13,d23,d12 FUNC=(0.5*(y^2-x^2)/z) PERIODIC=NO

# compute between point p3 and the vector joining p1 and p2
fromaxis: MATHEVAL ARG=d13,d23,d12,onaxis VAR=x,y,z,o FUNC=(0.5*(y^2+x^2)-o^2-0.25*z^2) PERIODIC=NO

PRINT ARG=onaxis,fromaxis
```

Notice that these equations have been used to combine [RMSD](#) from different snapshots of a protein so as to define progression (S) and distance (Z) variables [24].

## 5.4.6.1 TIME

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

retrieve the time of the simulation to be used elsewhere

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
------------------------------	----------------------------------------------------------------------------

## Examples

```
TIME          LABEL=t1
PRINT ARG=t1
```

(See also [PRINT](#)).

## 5.4.7 PIECEWISE

	This is part of the function <a href="#">module</a>
--	-----------------------------------------------------

Compute a piecewise straight line through its arguments that passes through a set of ordered control points.

For variables less than the first (greater than the last) point, the value of the first (last) point is used.

$$\frac{y_{i+1} - y_i}{x_{i+1} - x_i} (s - x_i) + y_i; \text{ if } x_i < s < x_{i+1}$$

$$y_N; \text{ if } x > x_{N-1}$$

$$y_1; \text{ if } x < x_0$$

Control points are passed using the POINT0=... POINT1=... syntax as in the example below

If one argument is supplied, it results in a scalar quantity. If multiple arguments are supplied, it results in a vector of values. Each value will be named as the name of the original argument with suffix `_pfunc`.

## Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>_pfunc</b>	one or multiple instances of this quantity will be referenceable elsewhere in the input file. These quantities will be named with the arguments of the function followed by the character string <code>_pfunc</code> . These quantities tell the user the values of the piecewise functions of each of the arguments.

## Compulsory keywords

<b>POINT</b>	This keyword is used to specify the various points in the function above. You can use multiple instances of this keyword i.e. POINT1, POINT2, POINT3...
--------------	---------------------------------------------------------------------------------------------------------------------------------------------------------

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...

## Examples

```
dist1: DISTANCE ATOMS=1,10
dist2: DISTANCE ATOMS=2,11

pw: PIECEWISE POINT0=1,10 POINT1=1,PI POINT2=3,10 ARG=dist1
ppww: PIECEWISE POINT0=1,10 POINT1=1,PI POINT2=3,10 ARG=dist1,dist2
PRINT ARG=pw,ppww.dist1_pfunc,ppww.dist2_pfunc
```

(See also [PRINT](#) and [DISTANCE](#)).

## 5.4.8 SORT

This is part of the function <a href="#">module</a>
-----------------------------------------------------

This function can be used to sort colvars according to their magnitudes.

## Description of components

This function sorts its arguments according to their magnitudes. The lowest argument will be labelled *label.1*, the second lowest will be labelled *label.2* and so on.

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...

#### Examples

The following input tells plumed to print the distance of the closest and of the farthest atoms to atom 1, chosen among atoms from 2 to 5

```
d12: DISTANCE ATOMS=1,2
d13: DISTANCE ATOMS=1,3
d14: DISTANCE ATOMS=1,4
d15: DISTANCE ATOMS=1,5
sort: SORT ARG=d12,d13,d14,d15
PRINT ARG=sort.1,sort.4
```

(See also [PRINT](#) and [DISTANCE](#)).

### 5.4.9 STATS

This is part of the function <a href="#">module</a>
-----------------------------------------------------

Calculates statistical properties of a set of collective variables with respect to a set of reference values. In particular it calculates and store as components the sum of the squared deviations, the correlation, the slope and the intercept of a linear fit.

The reference values can be either provided as values using PARAMETERS or using value without derivatives from other actions using PARARG (for example using experimental values from collective variables such as [CS2BACKBONE](#), [RDC](#), [NOE](#), [PRE](#)).



## Description of components

The names of the components in this action can be customized by the user in the actions input file. However, in addition to these customizable components the following quantities will always be output

Quantity	Description
<b>sqdevsum</b>	the sum of the squared deviations between arguments and parameters
<b>corr</b>	the correlation between arguments and parameters
<b>slope</b>	the slope of a linear fit between arguments and parameters
<b>intercept</b>	the intercept of a linear fit between arguments and parameters

In addition the following quantities can be calculated by employing the keywords listed below

Quantity	Keyword	Description
<b>sqd</b>	<b>SQDEV</b>	the squared deviations between arguments and parameters

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>SQDEVSUM</b>	( default=off ) calculates only SQDEVSUM
<b>SQDEV</b>	( default=off ) calculates and store the SQDEV as components
<b>UPPERDISTIS</b>	( default=off ) calculates and store the SQDEV as components
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>PARARG</b>	the input for this action is the scalar output from one or more other actions without derivatives.
<b>PARAMETERS</b>	the parameters of the arguments in your function

## Examples

The following input tells plumed to print the distance between three couple of atoms and compare them with three reference distances.

```
d1: DISTANCE ATOMS=10,50
d2: DISTANCE ATOMS=1,100
d3: DISTANCE ATOMS=45,75
st: STATS ARG=d1,d2,d3 PARAMETERS=1.5,4.0,2.0
PRINT ARG=d1,d2,d3,st.*
```

## 5.5 MultiColvar

Oftentimes, when you do not need one of the collective variables described elsewhere in the manual, what you want instead is a function of a distribution of collective variables of a particular type. In other words, you would like to calculate a function something like this:

$$s = \sum_i g[f(\{X\}_i)]$$

In this expression  $g$  is a function that takes in one argument and  $f$  is a function that takes a set of atomic positions as argument. The symbol  $\{X\}_i$  is used to indicate the fact that the function  $f$  is evaluated for a number of different sets of atoms. If you would just like to output the values of all the various  $f$  functions you should use the command [DUMPMULTICOLVAR](#)

This functionality is useful if you need to calculate a minimum distance or the number of coordination numbers greater than a 3.0.

To avoid duplicating the code to calculate an angle or distance many times and to make it easier to implement very complex collective variables PLUMED provides these sort of collective variables using so-called MultiColvars. MultiColvars are named in this way because a single PLUMED action can be used to calculate a number of different collective variables. For instance the [DISTANCES](#) action can be used to calculate the minimum distance, the number of distances less than a certain value, the number of distances within a certain range... A more detailed introduction to multicolvars is provided in this [10-minute video](#). Descriptions of the various multicolvars that are implemented in PLUMED 2 are given below:

<a href="#">ANGLES</a>	Calculate functions of the distribution of angles .
<a href="#">BRIDGE</a>	Calculate the number of atoms that bridge two parts of a structure
<a href="#">COORDINATIONNUMBER</a>	Calculate the coordination numbers of atoms so that you can then calculate functions of the distribution of coordination numbers such as the minimum, the number less than a certain quantity and so on.
<a href="#">DENSITY</a>	Calculate functions of the density of atoms as a function of the box. This allows one to calculate the number of atoms in half the box.
<a href="#">DISTANCES</a>	Calculate the distances between one or many pairs of atoms. You can then calculate functions of the distribution of distances such as the minimum, the number less than a certain quantity and so on.
<a href="#">FCCUBIC</a>	Measure how similar the environment around atoms is to that found in a FCC structure.
<a href="#">INPLANEDISTANCES</a>	Calculate distances in the plane perpendicular to an axis
<a href="#">MOLECULES</a>	Calculate the vectors connecting a pair of atoms in order to represent the orientation of a molecule.
<a href="#">PLANES</a>	Calculate the plane perpendicular to two vectors in order to represent the orientation of a planar molecule.
<a href="#">Q3</a>	Calculate 3rd order Steinhardt parameters.
<a href="#">Q4</a>	Calculate 4th order Steinhardt parameters.
<a href="#">Q6</a>	Calculate 6th order Steinhardt parameters.
<a href="#">SIMPLECUBIC</a>	Calculate whether or not the coordination spheres of atoms are arranged as they would be in a simple cubic structure.
<a href="#">TETRAHEDRAL</a>	Calculate the degree to which the environment about ions has a tetrahedral order.
<a href="#">TORSIONS</a>	Calculate whether or not a set of torsional angles are within a particular range.
<a href="#">XDISTANCES</a>	Calculate the x components of the vectors connecting one or many pairs of atoms. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.

<a href="#">XYDISTANCES</a>	Calculate distance between a pair of atoms neglecting the z-component. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.
<a href="#">XZDISTANCES</a>	Calculate distance between a pair of atoms neglecting the y-component. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.
<a href="#">YDISTANCES</a>	Calculate the y components of the vectors connecting one or many pairs of atoms. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.
<a href="#">YZDISTANCES</a>	Calculate distance between a pair of atoms neglecting the x-component. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.
<a href="#">ZDISTANCES</a>	Calculate the z components of the vectors connecting one or many pairs of atoms. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.

To instruct PLUMED to calculate a multicolvar you give an instruction that looks something like this:

```
NAME <atoms involved> <parameters> <what am I calculating> TOL=0.001 LABEL=label
```

Oftentimes the simplest way to specify the atoms involved is to use multiple instances of the ATOMS keyword i.e. ATOMS1, ATOMS2, ATOMS3,... Separate instances of the quantity specified by NAME are then calculated for each of the sets of atoms. For example if the command issued contains the following:

```
DISTANCES ATOMS1=1,2 ATOMS2=3,4 ATOMS3=5,6
```

The distances between atoms 1 and 2, atoms 3 and 4, and atoms 5 and 6 are calculated. Obviously, generating this sort of input is rather tedious so short cuts are also available many of the collective variables. These are described on the manual pages for the actions.

After specifying the atoms involved you sometimes need to specify some parameters that required in the calculation. For instance, for [COORDINATIONNUMBER](#) - the number of atoms in the first coordination sphere of each of the atoms in the system - you need to specify the parameters for a [switchingfunction](#) that will tell us whether or not an atom is in the first coordination sphere. Details as to how to do this are provided on the manual pages.

One of the most important keywords for multicolvars is the TOL keyword. This specifies that terms in sums that contribute less than a certain value can be ignored. In addition, it is assumed that the derivative with respect to these terms are essentially zero. By increasing the TOL parameter you can increase the speed of the calculation. Be aware, however, that this increase in speed is only possible because you are lowering the accuracy with which you are computing the quantity of interest.

Once you have specified the base quantities that are to be calculated from the atoms involved and any parameters you need to specify what function of these base quantities is to be calculated. For most multicolvars you can calculate the minimum, the number less than a target value, the number within a certain range, the number more than a target value and the average value directly.

### 5.5.1 MultiColvar functions

It is possible to use multicolvars to calculate complicated collective variables by exploiting the fact that the output from one multicolvar can be used as input to a second multicolvar. One simple way of exploiting this functionality is to filter the atoms based on the value they have for a symmetry function. For example you might want to consider only those atoms that with a [COORDINATIONNUMBER](#) higher than a certain threshold when calculating some particularly expensive symmetry function such as [Q6](#). The following methods can thus all be used to filter the values of multicolvars in this way:

<a href="#">MFILTER_BETWEEN</a>	This action can be used to filter the colvar values calculated by a multicolvar so that one can compute the mean and so on for only those multicolvars within a certain range.
<a href="#">MFILTER_LESS</a>	This action can be used to filter the distribution of colvar values in a multicolvar so that one can compute the mean and so on for only those multicolvars less than a tolerance.
<a href="#">MFILTER_MORE</a>	This action can be used to filter the distribution of colvar values in a multicolvar so that one can compute the mean and so on for only those multicolvars more than a tolerance.

An alternative way of filtering atoms is to consider only those atoms in a particular part of the simulation box. This can be done by exploiting the following methods

<a href="#">AROUND</a>	This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a particular, user-specified part of the cell.
<a href="#">CAVITY</a>	This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a box defined by the positions of four atoms.
<a href="#">INCYLINDER</a>	This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a particular, user-specified part of the cell.
<a href="#">INSPIHERE</a>	This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a particular, user-specified part of the cell.
<a href="#">TETRAHEDRALPORE</a>	This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a box defined by the positions of four atoms at the corners of a tetrahedron.

The idea with these methods is that function of the form:

$$s = \sum_i w(\{X\}_i) g[f(\{X\}_i)]$$

can be evaluated where once again  $g$  is a function with one argument and  $g$  is a function of a set of atomic positions. The difference from the more general function described earlier is that we now have a weight  $w$  which is again a function of the atomic positions. This weight varies between zero and one and it is this weight that is calculated in the list of filtering methods and volume methods described in the lists above.

In addition to these volume and filtering methods it is also possible to calculate the local average of a quantities in the manner described in [25] using the [LOCAL\\_AVERAGE](#) method. Furthermore, in many cases [Q6](#), [MOLECULES](#) and [PLANES](#) the symmetry function being evaluated is a vector. You can thus construct a variety of novel collective variables by taking dot products of vectors on adjacent atoms as described below:

<a href="#">GRADIENT</a>	Calculate the gradient of the average value of a multicolvar value
<a href="#">INTERMOLECULARTORSIONS</a>	Calculate torsions between axis of adjacent molecules
<a href="#">LOCAL_AVERAGE</a>	Calculate averages over spherical regions centered on atoms
<a href="#">LOCAL_Q3</a>	Calculate the local degree of order around an atoms by taking the average dot product between the $q_3$ vector on the central atom and the $q_3$ vector on the atoms in the first coordination sphere.
<a href="#">LOCAL_Q4</a>	Calculate the local degree of order around an atoms by taking the average dot product between the $q_4$ vector on the central atom and the $q_4$ vector on the atoms in the first coordination sphere.
<a href="#">LOCAL_Q6</a>	Calculate the local degree of order around an atoms by taking the average dot product between the $q_6$ vector on the central atom and the $q_6$ vector on the atoms in the first coordination sphere.
<a href="#">NLINKS</a>	Calculate number of pairs of atoms/molecules that are "linked"
<a href="#">SMAC</a>	Calculate a variant on the SMAC collective variable discussed in [26]

The final set of functions that you can apply on multicolvars are functions that transform all the colvars calculated using a multicolvar using a function. This can be useful if you are calculating some complicated derived quantity of

some simpler quantity. It is also useful if you are calculating a Willard Chandler surface or a histogram. The actions that you can use to perform these transforms are:

<a href="#">MTRANSFORM_BETWEEN</a>	This action can be used to transform the colvar values calculated by a multi-colvar using a <a href="#">histogrambead</a>
<a href="#">MTRANSFORM_LESS</a>	This action can be used to transform the colvar values calculated by a multi-colvar using a <a href="#">switchingfunction</a>
<a href="#">MTRANSFORM_MORE</a>	This action can be used to transform the colvar values calculated by a multi-colvar using one minus a <a href="#">switchingfunction</a>

### 5.5.2 MultiColvar bias

There may be occasions when you want add restraints on many collective variables. For instance if you are studying a cluster you might want to add a wall on the distances between each of the atoms and the center of mass of the cluster in order to prevent the cluster subliming. Alternatively, you may wish to insist that a particular set of atoms in your system all have a coordination number greater than 2. You can add these sorts of restraints by employing the following biases, which all act on the set of collective variable values calculated by a multicolvar. So for example the following set of commands:

```
COM ATOMS=1-20 LABEL=c1
DISTANCES GROUPA=c1 GROUPB=1-20 LABEL=d1
UWALLS DATA=d1 AT=2.5 KAPPA=0.2 LABEL=sr
```

creates the aforementioned set of restraints on the distances between the 20 atoms in a cluster and the center of mass of the cluster.

The list of biases of this type are as follows:

<a href="#">UWALLS</a>	Add <a href="#">UPPER_WALLS</a> restraints on all the multicolvar values
------------------------	--------------------------------------------------------------------------

Notice that (in theory) you could also use this functionality to add additional terms to your forcefield or to implement your forcefield.

### 5.5.3 ANGLES

<b>This is part of the multicolvar <a href="#">module</a></b>
---------------------------------------------------------------

Calculate functions of the distribution of angles .

You can use this command to calculate functions such as:

$$f(x) = \sum_{ijk} g(\theta_{ijk})$$

Alternatively you can use this command to calculate functions such as:

$$f(x) = \sum_{ijk} s(r_{ij})s(r_{jk})g(\theta_{ijk})$$

where  $s(r)$  is a [switchingfunction](#). This second form means that you can use this to calculate functions of the angles in the first coordination sphere of an atom / molecule [27].

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>less-than</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>more-than</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM←S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>GROUP</b>	Calculate angles for each distinct set of three atoms in the group
--------------	--------------------------------------------------------------------

Or alternatively by using

<b>GROUPA</b>	A group of central atoms about which angles should be calculated
<b>GROUPB</b>	When used in conjunction with GROUPA this keyword instructs plumed to calculate all distinct angles involving one atom from GROUPA and two atoms from GROUPB. The atom from GROUPA is the central atom.

Or alternatively by using

<b>GROUPC</b>	This must be used in conjunction with GROUPA and GROUPB. All angles involving one atom from GROUPA, one atom from GROUPB and one atom from GROUPC are calculated. The GROUPA atoms are assumed to be the central atoms
---------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>SWITCH</b>	A switching function that ensures that only angles between atoms that are within a certain fixed cutoff are calculated. The following provides information on the <a href="#">switchingfunction</a> that are available.
<b>SWITCHA</b>	A switching function on the distance between the atoms in group A and the atoms in group B
<b>SWITCHB</b>	A switching function on the distance between the atoms in group A and the atoms in group B

## Examples

The following example instructs plumed to find the average of two angles and to print it to a file

```
ANGLES ATOMS1=1,2,3 ATOMS2=4,5,6 MEAN LABEL=a1
PRINT ARG=a1.mean FILE=colvar
```

The following example tells plumed to calculate all angles involving at least one atom from GROUPA and two atoms from GROUPB in which the distances are less than 1.0. The number of angles between  $\frac{\pi}{4}$  and  $\frac{3\pi}{4}$  is then output

```
ANGLES GROUPA=1-10 GROUPB=11-100 BETWEEN={GAUSSIAN LOWER=0.25pi UPPER=0.75pi} SWITCH={GAUSSIAN R_0=1.0} LABEL=a1
PRINT ARG=a1.between FILE=colvar
```

This final example instructs plumed to calculate all the angles in the first coordination spheres of the atoms. A discretized-normalized histogram of the distribution is then output

```
ANGLES GROUP=1-38 HISTOGRAM={GAUSSIAN LOWER=0.0 UPPER=pi NBINS=20} SWITCH={GAUSSIAN R_0=1.0} LABEL=a1
PRINT ARG=a1.* FILE=colvar
```

## 5.5.4 BRIDGE

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate the number of atoms that bridge two parts of a structure

This quantity calculates:

$$f(x) = \sum_{ijk} s_A(r_{ij}) s_B(r_{ik})$$

where the sum over  $i$  is over all the "bridging atoms" and  $s_A$  and  $s_B$  are [switchingfunction](#).

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM↔S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using



<b>BRIDGING_ATOMS</b>	The list of atoms that can form the bridge between the two interesting parts of the structure.
<b>GROUPA</b>	The list of atoms that are in the first interesting part of the structure
<b>GROUPB</b>	The list of atoms that are in the second interesting part of the structure

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	The parameters of the two <a href="#">switchingfunction</a> in the above formula
<b>SWITCHA</b>	The <a href="#">switchingfunction</a> on the distance between bridging atoms and the atoms in group A
<b>SWITCHB</b>	The <a href="#">switchingfunction</a> on the distance between the bridging atoms and the atoms in group B

## Examples

The following example instructs plumed to calculate the number of water molecules that are bridging between atoms 1-10 and atoms 11-20 and to print the value to a file

```
BRIDGE BRIDGING_ATOMS=100-200 GROUPA=1-10 GROUPB=11-20 LABEL=w1
PRINT ARG=a1.mean FILE=colvar
```

## 5.5.5 COORDINATIONNUMBER

<b>This is part of the multicolvar <a href="#">module</a></b>
---------------------------------------------------------------

Calculate the coordination numbers of atoms so that you can then calculate functions of the distribution of coordination numbers such as the minimum, the number less than a certain quantity and so on.

To make the calculation of coordination numbers differentiable the following function is used:

$$s = \frac{1 - \left(\frac{r-d_0}{r_0}\right)^n}{1 - \left(\frac{r-d_0}{r_0}\right)^m}$$

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>less-than</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <i>label.mean</i>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>more-than</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↔ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these coordination numbers specifies how many of the atoms specified using SPECIESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

## Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D↔ _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R↔ _0</b>	The r_0 parameter of the switching function

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...

<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-m</i> .
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

### Examples

The following input tells plumed to calculate the coordination numbers of atoms 1-100 with themselves. The minimum coordination number is then calculated.

```
COORDINATIONNUMBER SPECIES=1-100 R_0=1.0 MIN={BETA=0.1}
```

The following input tells plumed to calculate how many atoms from 1-100 are within 3.0 of each of the atoms from 101-110. In the first 101 is the central atom, in the second 102 is the central atom and so on. The number of coordination numbers more than 6 is then computed.

```
COORDINATIONNUMBER SPECIESA=101-110 SPECIESB=1-100 R_0=3.0 MORE_THAN={RATIONAL R_0=6.0 NN=6 MM=12 D_0=0}
```

### 5.5.6 DENSITY

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate functions of the density of atoms as a function of the box. This allows one to calculate the number of atoms in half the box.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↔ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

Examples

The following example calculates the number of atoms in one half of the simulation box.

```
DENSITY SPECIES=1-100 LABEL=d
AROUND ARG=d XLOWER=0.0 XUPPER=0.5 LABEL=d1
PRINT ARG=d1.* FILE=colvar1 FMT=%8.4f
```

### 5.5.7 DISTANCES

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate the distances between one or many pairs of atoms. You can then calculate functions of the distribution of distances such as the minimum, the number less than a certain quantity and so on.

Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>less-than</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <i>label.mean</i>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>more-than</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM←S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>GROUP</b>	Calculate the distance between each distinct pair of atoms in the group
--------------	-------------------------------------------------------------------------

Or alternatively by using

<b>GROUPA</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPB.
<b>GROUPB</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPA.

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less-than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less-than-1</i> , <i>label.less-than-2</i> , <i>label.less-than-3</i> ...

<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .

### Examples

The following input tells plumed to calculate the distances between atoms 3 and 5 and between atoms 1 and 2 and to print the minimum for these two distances.

```
DISTANCES ATOMS1=3,5 ATOMS2=1,2 MIN={BETA=0.1} LABEL=d1
PRINT ARG=d1.min
```

(See also [PRINT](#)).

The following input tells plumed to calculate the distances between atoms 3 and 5 and between atoms 1 and 2 and then to calculate the number of these distances that are less than 0.1 nm. The number of distances less than 0.1 nm is then printed to a file.

```
DISTANCES ATOMS1=3,5 ATOMS2=1,2 LABEL=d1 LESS_THAN={RATIONAL R_0=0.1}
PRINT ARG=d1.lt0.1
```

(See also [PRINT switchingfunction](#)).

The following input tells plumed to calculate all the distances between atoms 1, 2 and 3 (i.e. the distances between atoms 1 and 2, atoms 1 and 3 and atoms 2 and 3). The average of these distances is then calculated.

```
DISTANCES GROUP=1-3 MEAN LABEL=d1
PRINT ARG=d1.mean
```

(See also [PRINT](#)).

The following input tells plumed to calculate all the distances between the atoms in GROUPA and the atoms in GROUPB. In other words the distances between atoms 1 and 2 and the distance between atoms 1 and 3. The number of distances more than 0.1 is then printed to a file.



```
DISTANCES GROUPA=1 GROUPB=2,3 MORE_THAN={RATIONAL R_0=0.1}
PRINT ARG=d1.gt0.1
```

(See also [PRINT switchingfunction](#))

### Calculating minimum distances

To calculate and print the minimum distance between two groups of atoms you use the following commands

```
d1: DISTANCES GROUPA=1-10 GROUPB=11-20 MIN={BETA=500.}
PRINT ARG=d1.min FILE=colvar STRIDE=10
```

(see [DISTANCES](#) and [PRINT](#))

In order to ensure differentiability the minimum is calculated using the following function:

$$s = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$$

where  $\beta$  is a user specified parameter.

This input is used rather than a separate MINDIST colvar so that the same routine and the same input style can be used to calculate minimum coordination numbers (see [COORDINATIONNUMBER](#)), minimum angles (see [ANGLES](#)) and many other variables.

This new way of calculating mindist is part of plumed 2's multicolvar functionality. These special actions allow you to calculate multiple functions of a distribution of simple collective variables. As an example you can calculate the number of distances less than 1.0, the minimum distance, the number of distances more than 2.0 and the number of distances between 1.0 and 2.0 by using the following command:

```
DISTANCES ...
GROUPA=1-10 GROUPB=11-20
LESS_THAN={RATIONAL R_0=1.0}
MORE_THAN={RATIONAL R_0=2.0}
BETWEEN={GAUSSIAN LOWER=1.0 UPPER=2.0}
MIN={BETA=500.}
... DISTANCES
PRINT ARG=d1.less-than,d1.more-than,d1.between,d1.min FILE=colvar STRIDE=10
```

(see [DISTANCES](#) and [PRINT](#))

A calculation performed this way is fast because the expensive part of the calculation - the calculation of all the distances - is only done once per step. Furthermore, it can be made faster by using the TOL keyword to discard those distance that make only a small contributions to the final values together with the NL\_STRIDE keyword, which ensures that the distances that make only a small contribution to the final values aren't calculated at every step.

### 5.5.8 FCCUBIC

	<p><b>This is part of the crystallization <a href="#">module</a></b></p> <p><b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b></p>
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Measure how similar the environment around atoms is to that found in a FCC structure.

This CV was introduced in this article [28] and again in this article [29] This CV essentially determines whether the environment around any given atom is similar to that found in the FCC structure or not. The function that is used to make this determination is as follows:

$$s_i = \frac{\sum_{i \neq j} \sigma(r_{ij}) \left\{ a \left[ \frac{(x_{ij}y_{ij})^4 + (x_{ij}z_{ij})^4 + (y_{ij}z_{ij})^4}{r_{ij}^8} - \frac{\alpha(x_{ij}y_{ij}z_{ij})^4}{r_{ij}^{12}} \right] + b \right\}}{\sum_{i \neq j} \sigma(r_{ij})}$$

In this expression  $x_{ij}$ ,  $y_{ij}$  and  $z_{ij}$  are the  $x$ ,  $y$  and  $z$  components of the vector connecting atom  $i$  to atom  $j$  and  $r_{ij}$  is the magnitude of this vector.  $\sigma(r_{ij})$  is a [switchingfunction](#) that acts on the distance between atom  $i$  and atom  $j$  and its inclusion in the numerator and the denominator of the above expression as well as the fact that we are summing over all of the other atoms in the system ensures that we are calculating an average of the function of  $x_{ij}$ ,  $y_{ij}$  and  $z_{ij}$  for the atoms in the first coordination sphere around atom  $i$ . Lastly,  $\alpha$  is a parameter that can be set by the user, which by default is equal to three. The values of  $a$  and  $b$  are calculated from  $\alpha$  using:

$$a = \frac{80080}{2717 + 16\alpha} \quad \text{and} \quad b = \frac{16(\alpha - 143)}{2717 + 16\alpha}$$

This quantity is once again a multicolvar so you can compute it for multiple atoms using a single PLUMED action and then compute the average value for the atoms in your system, the number of atoms that have an  $s_i$  value that is more than some target and so on. Notice also that you can rotate the reference frame if you are using a non-standard unit cell.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>less-than</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action

<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <code>label.mean</code>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <code>label.moment-2</code> , the third as <code>label.moment-3</code> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↔ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↔ IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R_0</b>	The r_0 parameter of the switching function
<b>PHI</b>	( default=0.0 ) The Euler rotational angle phi
<b>THETA</b>	( default=0.0 ) The Euler rotational angle theta
<b>PSI</b>	( default=0.0 ) The Euler rotational angle psi
<b>ALPHA</b>	( default=3.0 ) The alpha parameter of the angular function

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less than-1</i> , <i>label.less than-2</i> , <i>label.less than-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <code>moment- <math>m</math></code> .
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <code>label.altmin</code> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <code>label.altmin-1</code> , <code>label.altmin-2</code> , <code>label.altmin-3</code> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <code>label.lowest</code>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <code>label.highest</code>

### Examples

The following input calculates the FCCUBIC parameter for the 64 atoms in the system and then calculates and prints the average value for this quantity.

```
FCCUBIC SPECIES=1-64 SWITCH={RATIONAL D_0=3.0 R_0=1.5} MEAN LABEL=d
PRINT ARG=d.* FILE=colv
```

### 5.5.9 INPLANEDISTANCES

This is part of the multicolvar [module](#)

Calculate distances in the plane perpendicular to an axis

Each quantity calculated in this CV uses the positions of two atoms, this indices of which are specified using the VECTORSTART and VECTOREND keywords, to specify the orientation of a vector, **n**. The perpendicular distance between this vector and the position of some third atom is then computed using:

$$x_j = |\mathbf{r}_j| \sin(\theta_j)$$

where  $\mathbf{r}_j$  is the distance between one of the two atoms that define the vector **n** and a third atom (atom  $j$ ) and where  $\theta_j$  is the angle between the vector **n** and the vector  $\mathbf{r}_j$ . The  $x_j$  values for each of the atoms specified using the GROUP keyword are calculated. Keywords such as MORE\_THAN and LESS\_THAN can then be used to calculate the number of these quantities that are more or less than a given cutoff.

### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
altmin	ALT_MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
between	BETWEEN	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
highest	HIGHEST	the lowest of the quantities calculated by this action
less-than	LESS_THAN	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
lowest	LOWEST	the lowest of the quantities calculated by this action
max	MAX	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
mean	MEAN	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
min	MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
moment	MOMENTS	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
more-than	MORE_THAN	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

VECTORSTART	The first atom position that is used to define the normal to the plane of interest. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
VECTOREND	The second atom position that is used to define the normal to the plane of interest. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>

Or alternatively by using

GROUP	The set of atoms for which you wish to calculate the in plane distance
-------	------------------------------------------------------------------------

Options

NUMERICAL_DERIVATIVES	( default=off ) calculate the derivatives for these quantities numerically
-----------------------	----------------------------------------------------------------------------

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less-than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less-than-1</i> , <i>label.less-than-2</i> , <i>label.less-than-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.more-than</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.more-than-1</i> , <i>label.more-than-2</i> , <i>label.more-than-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...

<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .

### Examples

The following input can be used to calculate the number of atoms that have indices greater than 3 and less than 101 that are within a cylinder with a radius of 0.3 nm that has its long axis aligned with the vector connecting atoms 1 and 2.

```
dl: INPLANEDISTANCES VECTORSTART=1 VECTOREND=2 GROUP=3-100 LESS_THAN={RATIONAL D_0=0.2 R_0=0.1}
PRINT ARG=dl.less_than FILE=colvar
```

### 5.5.10 MOLECULES

<b>This is part of the crystallization <a href="#">module</a></b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate the vectors connecting a pair of atoms in order to represent the orientation of a molecule.

At its simplest this command can be used to calculate the average length of an internal vector in a collection of different molecules. When used in conjunction with MultiColvarFunctions it can be used to do a variety of more complex tasks.

### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less\_than-1*, *label.less\_than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name



Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean

The atoms involved can be specified using

<b>MOL</b>	The numerical indices of the atoms in the molecule. The orientation of the molecule is equal to the vector connecting the first two atoms specified. If a third atom is specified its position is used to specify where the molecule is. If a third atom is not present the molecule is assumed to be at the center of the vector connecting the first two atoms. You can use multiple instances of this keyword i.e. MOL1, MOL2, MOL3...
------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...

#### Examples

The following input tells plumed to calculate the distances between two of the atoms in a molecule. This is done for the same set of atoms four different molecules and the average separation is then calculated.

```
MOLECULES MOL1=1,2 MOL2=3,4 MOL3=5,6 MOL4=7,8 MEAN LABEL=mm
PRINT ARG=mm.mean FILE=colvar
```

#### 5.5.11 PLANES

	<b>This is part of the crystallization <a href="#">module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate the plane perpendicular to two vectors in order to represent the orientation of a planar molecule.

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean

The atoms involved can be specified using

<b>MOL</b>	The numerical indices of the atoms in the molecule. If three atoms are specified the orientation of the molecule is taken as the normal to the plane containing the vector connecting the first and second atoms and the vector connecting the second and third atoms. If four atoms are specified the orientation of the molecule is taken as the normal to the plane containing the vector connecting the first and second atoms and the vector connecting the third and fourth atoms. The molecule is always assumed to lie at the geometric centre for the three/four atoms. You can use multiple instances of this keyword i.e. MOL1, MOL2, MOL3...
------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...

## Examples

## 5.5.12 Q3

	<b>This is part of the crystallization <a href="#">module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate 3rd order Steinhardt parameters.

The 3rd order Steinhardt parameters allow us to measure the degree to which the first coordination shell around an atom is ordered. The Steinhardt parameter for atom,  $i$  is complex vector whose components are calculated using the following formula:

$$q_{3m}(i) = \frac{\sum_j \sigma(r_{ij}) Y_{3m}(\mathbf{r}_{ij})}{\sum_j \sigma(r_{ij})}$$

where  $Y_{3m}$  is one of the 3rd order spherical harmonics so  $m$  is a number that runs from  $-3$  to  $+3$ . The function  $\sigma(r_{ij})$  is a [switchingfunction](#) that acts on the distance between atoms  $i$  and  $j$ . The parameters of this function should be set so that it the function is equal to one when atom  $j$  is in the first coordination sphere of atom  $i$  and is zero otherwise.

The Steinhardt parameters can be used to measure the degree of order in the system in a variety of different ways. The simplest way of measuring whether or not the coordination sphere is ordered is to simply take the norm of the above vector i.e.

$$Q_3(i) = \sqrt{\sum_{m=-3}^3 q_{3m}(i)^* q_{3m}(i)}$$

This norm is small when the coordination shell is disordered and larger when the coordination shell is ordered. Furthermore, when the keywords LESS\_THAN, MIN, MAX, HISTOGRAM, MEAN and so on are used with this colvar it is the distribution of these normed quantities that is investigated.

Other measures of order can be taken by averaging the components of the individual  $q_3$  vectors individually or by taking dot products of the  $q_3$  vectors on adjacent atoms. More information on these variables can be found in the documentation for [LOCAL\\_Q3](#), [LOCAL\\_AVERAGE](#) and [NLINKS](#).

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↵ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↵ IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

Compulsory keywords

<b>NN</b>	( default=12 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D<sub>↔</sub> _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R<sub>↔</sub> _0</b>	The r_0 parameter of the switching function

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEA <sub>↔</sub> N2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less than-1</i> , <i>label.less than-2</i> , <i>label.less than-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.more than</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THA <sub>↔</sub> N3... The corresponding values are then referenced using <i>label.more than-1</i> , <i>label.more than-2</i> , <i>label.more than-3</i> ...
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculates NBIN quantites like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <code>moment-<math>m</math></code> .
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp(\frac{\beta}{s_i})}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <code>label.min</code> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <code>label.min-1</code> , <code>label.min-2</code> , <code>label.min-3</code> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <code>label.altmin</code> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <code>label.altmin-1</code> , <code>label.altmin-2</code> , <code>label.altmin-3</code> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <code>label.lowest</code>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <code>label.highest</code>

### Examples

The following command calculates the average Q3 parameter for the 64 atoms in a box of Lennard Jones and prints this quantity to a file called colvar:

```
Q3 SPECIES=1-64 D_0=1.3 R_0=0.2 MEAN LABEL=q3
PRINT ARG=q3.mean FILE=colvar
```

The following command calculates the histogram of Q3 parameters for the 64 atoms in a box of Lennard Jones and prints these quantities to a file called colvar:

```
Q3 SPECIES=1-64 D_0=1.3 R_0=0.2 HISTOGRAM={GAUSSIAN LOWER=0.0 UPPER=1.0 NBINS=20 SMEAR=0.1} LABEL=q3
PRINT ARG=q3.* FILE=colvar
```

The following command could be used to measure the Q3 paramters that describe the arrangement of chlorine ions around the sodium atoms in NaCl. The imagined system here is composed of 64 NaCl formula units and the atoms are arranged in the input with the 64 Na<sup>+</sup> ions followed by the 64 Cl<sup>-</sup> ions. Once again the average Q3 paramter is calculated and output to a file called colvar

```
Q3 SPECIESA=1-64 SPECIESB=65-128 D_0=1.3 R_0=0.2 MEAN LABEL=q3
PRINT ARG=q3.mean FILE=colvar
```

### 5.5.13 Q4

<b>This is part of the crystallization <a href="#">module</a></b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>
Generated by Doxygen

Calculate 4th order Steinhardt parameters.

The 4th order Steinhardt parameters allow us to measure the degree to which the first coordination shell around an atom is ordered. The Steinhardt parameter for atom,  $i$  is complex vector whose components are calculated using the following formula:

$$q_{4m}(i) = \frac{\sum_j \sigma(r_{ij}) Y_{4m}(\mathbf{r}_{ij})}{\sum_j \sigma(r_{ij})}$$

where  $Y_{4m}$  is one of the 4th order spherical harmonics so  $m$  is a number that runs from  $-4$  to  $+4$ . The function  $\sigma(r_{ij})$  is a [switchingfunction](#) that acts on the distance between atoms  $i$  and  $j$ . The parameters of this function should be set so that it the function is equal to one when atom  $j$  is in the first coordination sphere of atom  $i$  and is zero otherwise.

The Steinhardt parameters can be used to measure the degree of order in the system in a variety of different ways. The simplest way of measuring whether or not the coordination sphere is ordered is to simply take the norm of the above vector i.e.

$$Q_4(i) = \sqrt{\sum_{m=-4}^4 q_{4m}(i)^* q_{4m}(i)}$$

This norm is small when the coordination shell is disordered and larger when the coordination shell is ordered. Furthermore, when the keywords LESS\_THAN, MIN, MAX, HISTOGRAM, MEAN and so on are used with this colvar it is the distribution of these normed quantities that is investigated.

Other measures of order can be taken by averaging the components of the individual  $q_4$  vectors individually or by taking dot products of the  $q_4$  vectors on adjacent atoms. More information on these variables can be found in the documentation for [LOCAL\\_Q4](#), [LOCAL\\_AVERAGE](#) and [NLINKS](#).

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↵ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↵ IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

Compulsory keywords



<b>NN</b>	( default=12 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D<sub>↔</sub> _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R<sub>↔</sub> _0</b>	The r_0 parameter of the switching function

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEA <sub>↔</sub> N2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less than-1</i> , <i>label.less than-2</i> , <i>label.less than-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.more than</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THA <sub>↔</sub> N3... The corresponding values are then referenced using <i>label.more than-1</i> , <i>label.more than-2</i> , <i>label.more than-3</i> ...
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculates NBIN quantites like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp(\frac{\beta}{s_i})}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

### Examples

The following command calculates the average Q4 parameter for the 64 atoms in a box of Lennard Jones and prints this quantity to a file called colvar:

```
Q4 SPECIES=1-64 D_0=1.3 R_0=0.2 MEAN LABEL=q4
PRINT ARG=q4.mean FILE=colvar
```

The following command calculates the histogram of Q4 parameters for the 64 atoms in a box of Lennard Jones and prints these quantities to a file called colvar:

```
Q4 SPECIES=1-64 D_0=1.3 R_0=0.2 HISTOGRAM={GAUSSIAN LOWER=0.0 UPPER=1.0 NBINS=20 SMEAR=0.1} LABEL=q4
PRINT ARG=q4.* FILE=colvar
```

The following command could be used to measure the Q4 paramters that describe the arrangement of chlorine ions around the sodium atoms in NaCl. The imagined system here is composed of 64 NaCl formula units and the atoms are arranged in the input with the 64 Na<sup>+</sup> ions followed by the 64 Cl<sup>-</sup> ions. Once again the average Q4 paramter is calculated and output to a file called colvar

```
Q4 SPECIESA=1-64 SPECIESB=65-128 D_0=1.3 R_0=0.2 MEAN LABEL=q4
PRINT ARG=q4.mean FILE=colvar
```

### 5.5.14 Q6

<b>This is part of the crystallization <a href="#">module</a></b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate 6th order Steinhardt parameters.

The 6th order Steinhardt parameters allow us to measure the degree to which the first coordination shell around an atom is ordered. The Steinhardt parameter for atom,  $i$  is complex vector whose components are calculated using the following formula:

$$q_{6m}(i) = \frac{\sum_j \sigma(r_{ij}) Y_{6m}(\mathbf{r}_{ij})}{\sum_j \sigma(r_{ij})}$$

where  $Y_{6m}$  is one of the 6th order spherical harmonics so  $m$  is a number that runs from  $-6$  to  $+6$ . The function  $\sigma(r_{ij})$  is a [switchingfunction](#) that acts on the distance between atoms  $i$  and  $j$ . The parameters of this function should be set so that it the function is equal to one when atom  $j$  is in the first coordination sphere of atom  $i$  and is zero otherwise.

The Steinhardt parameters can be used to measure the degree of order in the system in a variety of different ways. The simplest way of measuring whether or not the coordination sphere is ordered is to simply take the norm of the above vector i.e.

$$Q_6(i) = \sqrt{\sum_{m=-6}^6 q_{6m}(i)^* q_{6m}(i)}$$

This norm is small when the coordination shell is disordered and larger when the coordination shell is ordered. Furthermore, when the keywords LESS\_THAN, MIN, MAX, HISTOGRAM, MEAN and so on are used with this colvar it is the distribution of these normed quantities that is investigated.

Other measures of order can be taken by averaging the components of the individual  $q_6$  vectors individually or by taking dot products of the  $q_6$  vectors on adjacent atoms. More information on these variables can be found in the documentation for [LOCAL\\_Q6](#), [LOCAL\\_AVERAGE](#) and [NLINKS](#).

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↵Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↵IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

Compulsory keywords

<b>NN</b>	( default=12 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D<sub>↔</sub> _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R<sub>↔</sub> _0</b>	The r_0 parameter of the switching function

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEA <sub>↔</sub> N2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less-than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less-than-1</i> , <i>label.less-than-2</i> , <i>label.less-than-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.more-than</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THA <sub>↔</sub> N3... The corresponding values are then referenced using <i>label.more-than-1</i> , <i>label.more-than-2</i> , <i>label.more-than-3</i> ...
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculates NBIN quantites like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <code>moment-<math>m</math></code> .
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp(\frac{\beta}{s_i})}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <code>label.min</code> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <code>label.min-1</code> , <code>label.min-2</code> , <code>label.min-3</code> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <code>label.altmin</code> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <code>label.altmin-1</code> , <code>label.altmin-2</code> , <code>label.altmin-3</code> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <code>label.lowest</code>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <code>label.highest</code>

### Examples

The following command calculates the average Q6 parameter for the 64 atoms in a box of Lennard Jones and prints this quantity to a file called colvar:

```
Q6 SPECIES=1-64 D_0=1.3 R_0=0.2 MEAN LABEL=q6
PRINT ARG=q6.mean FILE=colvar
```

The following command calculates the histogram of Q6 parameters for the 64 atoms in a box of Lennard Jones and prints these quantities to a file called colvar:

```
Q6 SPECIES=1-64 D_0=1.3 R_0=0.2 HISTOGRAM={GAUSSIAN LOWER=0.0 UPPER=1.0 NBINS=20 SMEAR=0.1} LABEL=q6
PRINT ARG=q6.* FILE=colvar
```

The following command could be used to measure the Q6 paramters that describe the arrangement of chlorine ions around the sodium atoms in NaCl. The imagined system here is composed of 64 NaCl formula units and the atoms are arranged in the input with the 64 Na<sup>+</sup> ions followed by the 64 Cl<sup>-</sup> ions. Once again the average Q6 paramter is calculated and output to a file called colvar

```
Q6 SPECIESA=1-64 SPECIESB=65-128 D_0=1.3 R_0=0.2 MEAN LABEL=q6
PRINT ARG=q6.mean FILE=colvar
```

### 5.5.15 SIMPLECUBIC

<b>This is part of the crystallization <a href="#">module</a></b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>
Generated by Doxygen

Calculate whether or not the coordination spheres of atoms are arranged as they would be in a simple cubic structure.

We can measure how similar the environment around atom  $i$  to a simple cubic structure is by evaluating the following quantity:

$$s_i = \frac{\sum_{j \neq i} \sigma(r_{ij}) \left[ \frac{x_{ij}^4 + y_{ij}^4 + z_{ij}^4}{r_{ij}^4} \right]}{\sum_{j \neq i} \sigma(r_{ij})}$$

In this expression  $x_{ij}$ ,  $y_{ij}$  and  $z_{ij}$  are the  $x$ ,  $y$  and  $z$  components of the vector connecting atom  $i$  to atom  $j$  and  $r_{ij}$  is the magnitude of this vector.  $\sigma(r_{ij})$  is a [switchingfunction](#) that acts on the distance between atom  $i$  and atom  $j$  and its inclusion in the numerator and the denominator of the above expression as well as the fact that we are summing over all of the other atoms in the system ensures that we are calculating an average of the function of  $x_{ij}$ ,  $y_{ij}$  and  $z_{ij}$  for the atoms in the first coordination sphere around atom  $i$ . This quantity is once again a multicolvar so you can compute it for multiple atoms using a single PLUMED action and then compute the average value for the atoms in your system, the number of atoms that have an  $s_i$  value that is more than some target and so on. Notice also that you can rotate the reference frame if you are using a non-standard unit cell.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
altmin	ALT_MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
between	BETWEEN	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
highest	HIGHEST	the lowest of the quantities calculated by this action
lessthan	LESS_THAN	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
lowest	LOWEST	the lowest of the quantities calculated by this action
max	MAX	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
mean	MEAN	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
min	MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
moment	MOMENTS	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
Generated by Doxygen		
morethan	MORE_THAN	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↔ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↔ IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R_0</b>	The r_0 parameter of the switching function
<b>PHI</b>	( default=0.0 ) The Euler rotational angle phi
<b>THETA</b>	( default=0.0 ) The Euler rotational angle theta
<b>PSI</b>	( default=0.0 ) The Euler rotational angle psi

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.



<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantiles like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

## Examples

The following input tells plumed to calculate the simple cubic parameter for the atoms 1-100 with themselves. The mean value is then calculated.

```
SIMPLECUBIC SPECIES=1-100 R_0=1.0 MEAN
```

The following input tells plumed to look at the ways atoms 1-100 are within 3.0 are arranged about atoms from 101-110. The number of simple cubic parameters that are greater than 0.8 is then output

```
SIMPLECUBIC SPECIESA=101-110 SPECIESB=1-100 R_0=3.0 MORE_THAN={RATIONAL R_0=0.8 NN=6 MM=12 D_0=0}
```

## 5.5.16 TETRAHEDRAL

	<b>This is part of the crystallization <a href="#">module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate the degree to which the environment about ions has a tetrahedral order.

We can measure the degree to which the first coordination shell around any atom,  $i$  is tetrahedrally ordered using the following function.

$$s(i) = \frac{1}{\sum_j \sigma(r_{ij})} \sum_j \sigma(r_{ij}) \left[ \frac{(x_{ij} + y_{ij} + z_{ij})^3}{r_{ij}^3} + \frac{(x_{ij} - y_{ij} - z_{ij})^3}{r_{ij}^3} + \frac{(-x_{ij} + y_{ij} - z_{ij})^3}{r_{ij}^3} + \frac{(-x_{ij} - y_{ij} + z_{ij})^3}{r_{ij}^3} \right]$$

Here  $r_{ij}$  is the magnitude for the vector connecting atom  $i$  to atom  $j$  and  $x_{ij}$ ,  $y_{ij}$  and  $z_{ij}$  are its three components. The function  $\sigma(r_{ij})$  is a [switching function](#) that acts on the distance between atoms  $i$  and  $j$ . The parameters of this function should be set so that the function is equal to one when atom  $j$  is in the first coordination sphere of atom  $i$  and is zero otherwise.

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. `label.less than-1`, `label.less than-2` etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
altmin	ALT_MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
between	BETWEEN	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
highest	HIGHEST	the lowest of the quantities calculated by this action
lessthan	LESS_THAN	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
lowest	LOWEST	the lowest of the quantities calculated by this action
max	MAX	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
mean	MEAN	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
min	MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
moment	MOMENTS	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
morethan	MORE_THAN	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↔ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↔ IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R_0</b>	The r_0 parameter of the switching function
<b>PHI</b>	( default=0.0 ) The Euler rotational angle phi
<b>THETA</b>	( default=0.0 ) The Euler rotational angle theta
<b>PSI</b>	( default=0.0 ) The Euler rotational angle psi

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...

<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

### Examples

The following command calculates the average value of the tetrahedrality parameter for a set of 64 atoms all of the same type and outputs this quantity to a file called colvar.

```
tt: TETRAHEDRAL SPECIES=1-64 SWITCH={RATIONAL D_0=1.3 R_0=0.2} MEAN
PRINT ARG=tt.mean FILE=colvar
```

The following command calculates the number of tetrahedrality parameters that are greater than 0.8 in a set of 10 atoms. In this calculation it is assumed that there are two atom types A and B and that the first coordination sphere of the 10 atoms of type A contains atoms of type B. The formula above is thus calculated for ten different A atoms and within it the sum over  $j$  runs over 40 atoms of type B that could be in the first coordination sphere.

```
tt: TETRAHEDRAL SPECIESA=1-10 SPECIESB=11-40 SWITCH={RATIONAL D_0=1.3 R_0=0.2} MORE_THAN={RATIONAL R_0=0.8}
PRINT ARG=tt.* FILE=colvar
```

### 5.5.17 TORSIONS

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate whether or not a set of torsional angles are within a particular range.

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM←S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

## Examples

The following provides an example of the input for the torsions command

```
TORSIONS ...
ATOMS1=168,170,172,188
ATOMS2=170,172,188,190
ATOMS3=188,190,192,230
LABEL=ab
... TORSIONS
PRINT ARG=ab.* FILE=colvar STRIDE=10
```

Writing out the atoms involved in all the torsions in this way can be rather tedious. Thankfully if you are working with protein you can avoid this by using the [MOLINFO](#) command. PLUMED uses the pdb file that you provide to this command to learn about the topology of the protein molecule. This means that you can specify torsion angles using the following syntax:

```
MOLINFO MOLTYPE=protein STRUCTURE=myprotein.pdb
TORSIONS ...
ATOMS1=@phi-3
ATOMS2=@psi-3
ATOMS3=@phi-4
LABEL=ab
... TORSIONS
PRINT ARG=ab FILE=colvar STRIDE=10
```

Here, @phi-3 tells plumed that you would like to calculate the  $\phi$  angle in the third residue of the protein. Similarly @psi-4 tells plumed that you want to calculate the  $\psi$  angle of the 4th residue of the protein.

### 5.5.18 XDISTANCES

This is part of the multicolvar [module](#)

Calculate the x components of the vectors connecting one or many pairs of atoms. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <code>label.mean</code>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <code>label.moment-2</code> , the third as <code>label.moment-3</code> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like <code>ATOM←S1</code> , <code>ATOMS2</code> , <code>ATOMS3</code> ,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. <code>ATOMS1</code> , <code>ATOMS2</code> , <code>ATOMS3</code> ...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>GROUP</b>	Calculate the distance between each distinct pair of atoms in the group
--------------	-------------------------------------------------------------------------

Or alternatively by using

<b>GROUPA</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPB.
<b>GROUPB</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPA.

Options



<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less than-1</i> , <i>label.less than-2</i> , <i>label.less than-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.more than</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.more than-1</i> , <i>label.more than-2</i> , <i>label.more than-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...

<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .

### Examples

The following input tells plumed to calculate the x-component of the vector connecting atom 3 to atom 5 and the x-component of the vector connecting atom 1 to atom 2. The minimum of these two quantities is then printed

```
XDISTANCES ATOMS1=3,5 ATOMS2=1,2 MIN={BETA=0.1} LABEL=d1
PRINT ARG=d1.min
```

(See also [PRINT](#)).

The following input tells plumed to calculate the x-component of the vector connecting atom 3 to atom 5 and the x-component of the vector connecting atom 1 to atom 2. The number of values that are less than 0.1nm is then printed to a file.

```
XDISTANCES ATOMS1=3,5 ATOMS2=1,2 LABEL=d1 LESS_THAN={RATIONAL R_0=0.1}
PRINT ARG=d1.lt0.1
```

(See also [PRINT switchingfunction](#)).

The following input tells plumed to calculate the x-components of all the distinct vectors that can be created between atoms 1, 2 and 3 (i.e. the vectors between atoms 1 and 2, atoms 1 and 3 and atoms 2 and 3). The average of these quantities is then calculated.

```
XDISTANCES GROUP=1-3 AVERAGE LABEL=d1
PRINT ARG=d1.average
```

(See also [PRINT](#)).

The following input tells plumed to calculate all the vectors connecting the atoms in GROUPA to the atoms in GROUPB. In other words the vector between atoms 1 and 2 and the vector between atoms 1 and 3. The number of values more than 0.1 is then printed to a file.

```
XDISTANCES GROUPA=1 GROUPB=2,3 MORE_THAN={RATIONAL R_0=0.1}
PRINT ARG=d1.gt0.1
```

(See also [PRINT switchingfunction](#))

## 5.5.19 XYDISTANCES

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate distance between a pair of atoms neglecting the z-component. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <i>label.mean</i>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM↵S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>GROUP</b>	Calculate the distance between each distinct pair of atoms in the group
--------------	-------------------------------------------------------------------------

Or alternatively by using

<b>GROUPA</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPB.
<b>GROUPB</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPA.

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp \left( \frac{s_i}{\beta} \right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp (-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...

<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .

### Examples

The following input tells plumed to calculate the projection of the length of the vector connecting atom 3 to atom 5 projected in the xy-plane and the projection of the length of the vector connecting atom 1 to atom 2 in the xy-plane. The minimum of these two quantities is then printed

```
XYDISTANCES ATOMS1=3,5 ATOMS2=1,2 MIN={BETA=0.1} LABEL=d1
PRINT ARG=d1.min
```

(See also [PRINT](#)).

### 5.5.20 XZDISTANCES

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate distance between a pair of atoms neglecting the y-component. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <i>label.mean</i>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM↵S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>GROUP</b>	Calculate the distance between each distinct pair of atoms in the group
--------------	-------------------------------------------------------------------------

Or alternatively by using

<b>GROUPA</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPB.
<b>GROUPB</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPA.

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp \left( \frac{s_i}{\beta} \right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp (-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEA↵N2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...

<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .

### Examples

The following input tells plumed to calculate the projection of the length of the vector connecting atom 3 to atom 5 projected in the xz-plane and the projection of the length of the vector connecting atom 1 to atom 2 in the xz-plane. The minimum of these two quantities is then printed

```
XZDISTANCES ATOMS1=3,5 ATOMS2=1,2 MIN={BETA=0.1} LABEL=d1
PRINT ARG=d1.min
```

(See also [PRINT](#)).

### 5.5.21 YDISTANCES



This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate the y components of the vectors connecting one or many pairs of atoms. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <i>label.mean</i>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM↵S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>GROUP</b>	Calculate the distance between each distinct pair of atoms in the group
--------------	-------------------------------------------------------------------------

Or alternatively by using

<b>GROUPA</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPB.
<b>GROUPB</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPA.

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp \left( \frac{s_i}{\beta} \right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp (-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEA↵N2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...

<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment- <math>m</math></i> .

### Examples

The following input tells plumed to calculate the y-component of the vector connecting atom 3 to atom 5 and the y-component of the vector connecting atom 1 to atom 2. The minimum of these two quantities is then printed

```
YDISTANCES ATOMS1=3,5 ATOMS2=1,2 MIN={BETA=0.1} LABEL=d1
PRINT ARG=d1.min
```

(See also [PRINT](#)).

The following input tells plumed to calculate the y-component of the vector connecting atom 3 to atom 5 and the y-component of the vector connecting atom 1 to atom 2. The number of values that are less than 0.1nm is then printed to a file.

```
YDISTANCES ATOMS1=3,5 ATOMS2=1,2 LABEL=d1 LESS_THAN={RATIONAL R_0=0.1}
PRINT ARG=d1.lt0.1
```

(See also [PRINT switchingfunction](#)).

The following input tells plumed to calculate the y-components of all the distinct vectors that can be created between atoms 1, 2 and 3 (i.e. the vectors between atoms 1 and 2, atoms 1 and 3 and atoms 2 and 3). The average of these quantities is then calculated.

```
YDISTANCES GROUP=1-3 AVERAGE LABEL=d1
PRINT ARG=d1.average
```

(See also [PRINT](#))

The following input tells plumed to calculate all the vectors connecting the atoms in GROUPA to the atoms in GROUPB. In other words the vector between atoms 1 and 2 and the vector between atoms 1 and 3. The number of values more than 0.1 is then printed to a file.

```
YDISTANCES GROUPA=1 GROUPB=2,3 MORE_THAN={RATIONAL R_0=0.1}
PRINT ARG=d1.gt0.1
```

(See also [PRINT switchingfunction](#))

## 5.5.22 YZDISTANCES

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate distance between a pair of atoms neglecting the x-component. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.

### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like ATOM↵S1, ATOMS2, ATOMS3,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. ATOMS1, ATOMS2, ATOMS3...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>GROUP</b>	Calculate the distance between each distinct pair of atoms in the group
--------------	-------------------------------------------------------------------------

Or alternatively by using

<b>GROUPA</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPB.
<b>GROUPB</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPA.

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize

<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less than-1</i> , <i>label.less than-2</i> , <i>label.less than-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.more than</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.more than-1</i> , <i>label.more than-2</i> , <i>label.more than-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

**MOMENTS**

calculate the moments of the distribution of collective variables. The  $m$ th moment of a distribution is calculated using  $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where  $\bar{s}$  is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of  $m$ . The final calculated values can be referenced using moment-  $m$ .

**Examples**

The following input tells plumed to calculate the projection of the length of the vector connecting atom 3 to atom 5 in the yz-plane and the projection of the length of the vector the vector connecting atom 1 to atom 2 in the yz-plane. The minimum of these two quantities is then printed

```
YZDISTANCES ATOMS1=3,5 ATOMS2=1,2 MIN={BETA=0.1} LABEL=d1
PRINT ARG=d1.min
```

(See also [PRINT](#)).

**5.5.23 ZDISTANCES**

This is part of the multicolvar [module](#)

Calculate the z components of the vectors connecting one or many pairs of atoms. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.

**Description of components**

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action

<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <code>label.mean</code>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <code>label.moment-2</code> , the third as <code>label.moment-3</code> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms involved in each of the collective variables you wish to calculate. Keywords like <code>ATOM↵S1</code> , <code>ATOMS2</code> , <code>ATOMS3</code> ,... should be listed and one CV will be calculated for each ATOM keyword you specify (all ATOM keywords should define the same number of atoms). The eventual number of quantities calculated by this action will depend on what functions of the distribution you choose to calculate. You can use multiple instances of this keyword i.e. <code>ATOMS1</code> , <code>ATOMS2</code> , <code>ATOMS3</code> ...
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>GROUP</b>	Calculate the distance between each distinct pair of atoms in the group
--------------	-------------------------------------------------------------------------

Or alternatively by using

<b>GROUPA</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPB.
<b>GROUPB</b>	Calculate the distances between all the atoms in GROUPA and all the atoms in GROUPB. This must be used in conjunction with GROUPA.

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements



<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

**MOMENTS**

calculate the moments of the distribution of collective variables. The  $m$ th moment of a distribution is calculated using  $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where  $\bar{s}$  is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of  $m$ . The final calculated values can be referenced using moment-  $m$ .

**Examples**

The following input tells plumed to calculate the z-component of the vector connecting atom 3 to atom 5 and the z-component of the vector connecting atom 1 to atom 2. The minimum of these two quantities is then printed

```
ZDISTANCES ATOMS1=3,5 ATOMS2=1,2 MIN={BETA=0.1} LABEL=d1
PRINT ARG=d1.min
```

(See also [PRINT](#)).

The following input tells plumed to calculate the z-component of the vector connecting atom 3 to atom 5 and the z-component of the vector connecting atom 1 to atom 2. The number of values that are less than 0.1nm is then printed to a file.

```
ZDISTANCES ATOMS1=3,5 ATOMS2=1,2 LABEL=d1 LESS_THAN={RATIONAL R_0=0.1}
PRINT ARG=d1.lt0.1
```

(See also [PRINT switchingfunction](#)).

The following input tells plumed to calculate the z-components of all the distinct vectors that can be created between atoms 1, 2 and 3 (i.e. the vectors between atoms 1 and 2, atoms 1 and 3 and atoms 2 and 3). The average of these quantities is then calculated.

```
ZDISTANCES GROUP=1-3 AVERAGE LABEL=d1
PRINT ARG=d1.average
```

(See also [PRINT](#))

The following input tells plumed to calculate all the vectors connecting the the atoms in GROUPA to the atoms in GROUPB. In other words the vector between atoms 1 and 2 and the vector between atoms 1 and 3. The number of values more than 0.1 is then printed to a file.

```
ZDISTANCES GROUPA=1 GROUPB=2,3 MORE_THAN={RATIONAL R_0=0.1}
PRINT ARG=d1.gt0.1
```

(See also [PRINT switchingfunction](#))

**5.5.24 MFILTER\_BETWEEN**

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

This action can be used to filter the colvar values calculated by a multicolvar so that one can compute the mean and so on for only those multicolvars within a certain range.

This action can be used to create a dynamic group of atom based on the value of a multicolvar. In this action a multicolvar is within the dynamic group if its value lies in a particular range. In practise a weight,  $w_i$  is ascribed to each colvar,  $s_i$  calculated by a multicolvar and this weight measures the degree to which a colvar is a member of the group. This weight is calculated using a [histogrambead](#) so it is given by:

$$w_i = \int_a^b K \left( \frac{s - s_i}{w} \right)$$

where  $a, b$  and  $w$  are parameters. If one calculates a function of the set of multicolvars these weights are included in the calculation. As such if one calculates the MEAN,  $\mu$  of a filtered multicolvar what is computed is the following:

$$\mu = \frac{\sum_i w_i s_i}{\sum_i w_i}$$

One is thus calculating the mean for those colvars that are within the range of interest.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.

#### Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>LOWER</b>	the lower boundary for the range of interest
<b>UPPER</b>	the upper boundary for the range of interest
<b>SMEAR</b>	( default=0.5 ) the ammount by which to smear the value for kernel density estimation

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

**BEAD**

This keyword is used if you want to employ an alternative to the function defined above. The following provides information on the [histogrambead](#) that are available. When this keyword is present you no longer need the LOWER, UPPER and SMEAR keywords.

**Examples**

The example shown below calculates the mean for those distances that are between 0 and 3 nm in length

```
DISTANCES GROUPA=1 GROUPB=2-50 MEAN LABEL=d1
MFILTER_BETWEEN DATA=d1 LOWER=0 UPPER=3.0 SMEAR=0.0001 MEAN LABEL=d4
```

More complicated things can be done by using the label of a filter as input to a new multicolvar as shown in the example below. Here the coordination numbers of all atoms are computed. The atoms with a coordination number between 4 and 6 are then identified using the filter. This reduced list of atoms is then used as input to a second coordination number calculation. This second coordination number thus measures the number of atoms 4-6 coordinated atoms each of the 4-6 coordination atoms is bound to.

```
c1: COORDINATIONNUMBER SPECIES=1-150 SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0}
cf: MFILTER_BETWEEN DATA=c1 LOWER=4 UPPER=6 SMEAR=0.5 LOWMEM
c2: COORDINATIONNUMBER SPECIES=cf SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0} MORE_THAN={RATIONAL D_0=2.0 R_0=0.1}
```

**5.5.25 MFILTER\_LESS**

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

This action can be used to filter the distribution of colvar values in a multicolvar so that one can compute the mean and so on for only those multicolvars less than a tolerance.

This action can be used to create a dynamic group of atom based on the value of a multicolvar. In this action a multicolvar is within the dynamic group if its value is less than a target. In practise a weight,  $w_i$  is ascribed to each colvar,  $s_i$  calculated by a multicolvar and this weight measures the degree to which a colvar is a member of the group. This weight is a number between 0 and 1 that is calculated using a [switchingfunction](#),  $\sigma$ . If one calculates a function of the set of multicolvars these weights are included in the calculation. As such if one calculates the MEAN,  $\mu$  of a filtered multicolvar what is computed is the following:

$$\mu = \frac{\sum_i w_i s_i}{\sum_i w_i}$$

One is thus calculating the mean for those colvars that are less than the target.

**Description of components**

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.

#### Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R_0</b>	The r_0 parameter of the switching function

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.

### Examples

The example shown below calculates the mean for those distances that less than 1.5 nm in length

```
DISTANCES GROUPA=1 GROUPB=2-50 MEAN LABEL=d1
MFILTER_LESS DATA=d1 SWITCH={GAUSSIAN D_0=1.5 R_0=0.00001} MEAN LABEL=d4
```

### 5.5.26 MFILTER\_MORE

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

This action can be used to filter the distribution of colvar values in a multicolvar so that one can compute the mean and so on for only those multicolvars more than a tolerance.

This action can be used to create a dynamic group of atom based on the value of a multicolvar. In this action a multicolvar is within the dynamic group if its value is greater than a target. In practise a weight,  $w_i$  is ascribed to each colvar,  $s_i$  calculated by a multicolvar and this weight measures the degree to which a colvar is a member of the group. This weight is calculated using a [switchingfunction](#),  $\sigma$  so it is given by:

$$w_i = 1 - \sigma(s_i)$$

If one calculates a function of the set of multicolvars these weights are included in the calculation. As such if one calculates the MEAN,  $\mu$  of a filtered multicolvar what is computed is the following:

$$\mu = \frac{\sum_i w_i s_i}{\sum_i w_i}$$

One is thus calculating the mean for those colvars that are greater than the target.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.



## Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R_0</b>	The r_0 parameter of the switching function

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>

<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.

### Examples

The example shown below calculates the mean for those distances that greater than 1.5 nm in length

```
DISTANCES GROUPA=1 GROUPB=2-50 MEAN LABEL=d1
MFILTER_MORE DATA=d1 SWITCH={GAUSSIAN D_0=1.5 R_0=0.00001} MEAN LABEL=d4
```

More complicated things can be done by using the label of a filter as input to a new multicolvar as shown in the example below. Here the coordination numbers of all atoms are computed. The atoms with a coordination number greater than 2 are then identified using the filter. This reduced list of atoms is then used as input to a second coordination number calculation. This second coordination number thus measures the number of two-coordinated atoms that each of the two-coordinated atoms is bound to.

```
1: COORDINATIONNUMBER SPECIES=1-150 SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0}
cf: MFILTER_MORE DATA=c1 SWITCH={RATIONAL D_0=2.0 R_0=0.1} LOWMEM
c2: COORDINATIONNUMBER SPECIES=cf SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0} MORE_THAN={RATIONAL D_0=2.0 R_0=0.1}
```

### 5.5.27 AROUND

<b>This is part of the multicolvar <a href="#">module</a></b>
---------------------------------------------------------------

This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a particular, user-specified part of the cell.

Each of the base quantities calculated by a multicolvar can be assigned to a particular point in three dimensional space. For example, if we have the coordination numbers for all the atoms in the system each coordination number can be assumed to lie on the position of the central atom. Because each base quantity can be assigned to a particular point in space we can calculate functions of the distribution of base quantities in a particular part of the box by using:

$$\bar{s}_\tau = \frac{\sum_i f(s_i) w(x_i, y_i, z_i)}{\sum_i w(x_i, y_i, z_i)}$$

where the sum is over the collective variables,  $s_i$ , each of which can be thought to be at  $(x_i, y_i, z_i)$ . The function  $w(x_i, y_i, z_i)$  measures whether or not the system is in the subregion of interest. It is equal to:

$$w(x_i, y_i, z_i) = \int_{x_l}^{x_u} \int_{y_l}^{y_u} \int_{z_l}^{z_u} dx dy dz K\left(\frac{x-x_i}{\sigma}\right) K\left(\frac{y-y_i}{\sigma}\right) K\left(\frac{z-z_i}{\sigma}\right)$$

where  $K$  is one of the kernel functions described on [histogrambead](#) and  $\sigma$  is a bandwidth parameter. The function  $(s_i)$  can be any of the usual LESS\_THAN, MORE\_THAN, WITHIN etc that are used in all other multicolvars.

When AROUND is used with the [DENSITY](#) action the number of atoms in the specified region is calculated

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>less-than</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>more-than</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOM</b>	the atom whose vicinity we are interested in examining. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
-------------	------------------------------------------------------------------------------------------------------------------------------------------------------------

## Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>SIGMA</b>	the width of the function to be used for kernel density estimation
<b>KERNEL</b>	( default=gaussian ) the type of kernel function to be used
<b>XLOWER</b>	( default=0.0 ) the lower boundary in x relative to the x coordinate of the atom (0 indicates use full extent of box).
<b>XUPPER</b>	( default=0.0 ) the upper boundary in x relative to the x coordinate of the atom (0 indicates use full extent of box).
<b>YLOWER</b>	( default=0.0 ) the lower boundary in y relative to the y coordinate of the atom (0 indicates use full extent of box).
<b>YUPPER</b>	( default=0.0 ) the upper boundary in y relative to the y coordinate of the atom (0 indicates use full extent of box).

<b>ZLOWER</b>	( default=0.0 ) the lower boundary in z relative to the z coordinate of the atom (0 indicates use full extent of box).
<b>ZUPPER</b>	( default=0.0 ) the upper boundary in z relative to the z coordinate of the atom (0 indicates use full extent of box).

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>OUTSIDE</b>	( default=off ) calculate quantities for colvars that are on atoms outside the region of interest
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

## Examples

The following commands tell plumed to calculate the average coordination number for the atoms that have x (in fractional coordinates) within 2.0 nm of the com of mass c1. The final value will be labeled s.mean.

```
COM ATOMS=1-100 LABEL=c1
COORDINATIONNUMBER SPECIES=1-100 R_0=1.0 LABEL=c
AROUND DATA=c ORIGIN=c1 XLOWER=-2.0 XUPPER=2.0 SIGMA=0.1 MEAN LABEL=s
```

### 5.5.28 CAVITY

<b>This is part of the multicolvar <a href="#">module</a></b>
---------------------------------------------------------------

This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a box defined by the positions of four atoms.

Each of the base quantities calculated by a multicolvar can be assigned to a particular point in three dimensional space. For example, if we have the coordination numbers for all the atoms in the system each coordination number can be assumed to lie on the position of the central atom. Because each base quantity can be assigned to a particular point in space we can calculate functions of the distribution of base quantities in a particular part of the box by using:

$$\bar{s}_\tau = \frac{\sum_i f(s_i) w(u_i, v_i, w_i)}{\sum_i w(u_i, v_i, w_i)}$$

where the sum is over the collective variables,  $s_i$ , each of which can be thought to be at  $(u_i, v_i, z_i)$ . The function  $(s_i)$  can be any of the usual LESS\_THAN, MORE\_THAN, WITHIN etc that are used in all other multicolvars. Notice that here (at variance with what is done in [AROUND](#)) we have transformed from the usual  $(x_i, y_i, z_i)$  position to a position in  $(u_i, v_i, z_i)$ . This is done using a rotation matrix as follows:

$$(u_i, v_i, w_i) = \mathbf{R} (x_i - x_o, y_i - y_o, z_i - z_o)$$

where  $\mathbf{R}$  is a rotation matrix that is calculated by constructing a set of three orthonormal vectors from the reference positions specified by the user. The first of these unit vectors points from the first reference atom to the second. The second is then the normal to the plane containing atoms 1, 2 and 3 and the third is the unit vector orthogonal to these first two vectors.  $(x_o, y_o, z_o)$ , meanwhile, specifies the position of the first reference atom.

In the previous function  $w(u_i, v_i, w_i)$  measures whether or not the system is in the subregion of interest. It is equal to:

$$w(u_i, v_i, w_i) = \int_0^{u'} \int_0^{v'} \int_0^{w'} du dv dw K\left(\frac{u - u_i}{\sigma}\right) K\left(\frac{v - v_i}{\sigma}\right) K\left(\frac{w - w_i}{\sigma}\right)$$

where  $K$  is one of the kernel functions described on [histogrambead](#) and  $\sigma$  is a bandwidth parameter. The vector connecting atom 1 to atom 4 is used to define the extent of the box in each of the  $u$ ,  $v$  and  $w$  directions. Essentially the vector connecting atom 1 to atom 4 is projected onto the three unit vectors described above and the resulting projections determine the  $u'$ ,  $v'$  and  $w'$  parameters in the above expression.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the positions of four atoms that define spatial extent of the cavity. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>SIGMA</b>	the width of the function to be used for kernel density estimation
<b>KERNEL</b>	( default=gaussian ) the type of kernel function to be used

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

<b>OUTSIDE</b>	( default=off ) calculate quantities for colvars that are on atoms outside the region of interest
<b>PRINT_BOX</b>	( default=off ) write out the positions of the corners of the box to an xyz file
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>FILE</b>	the file on which to write out the box coordinates
<b>UNITS</b>	( default=nm ) the units in which to write out the corners of the box

### Examples

The following commands tell plumed to calculate the number of atoms in an ion channel in a protein. The extent of the channel is calculated from the positions of atoms 1, 4, 5 and 11. The final value will be labeled cav.

```
d1: DENSITY SPECIES=20-500
CAVITY DATA=d1 ATOMS=1,4,5,11 SIGMA=0.1 LABEL=cav
```

The following command tells plumed to calculate the coordination numbers (with other water molecules) for the water molecules in the protein channel described above. The average coordination number and the number of coordination numbers more than 4 is then calculated. The values of these two quantities are given the labels cav.mean and cav.morethan

```
d1: COORDINATIONNUMBER SPECIES=20-500
CAVITY DATA=d1 ATOMS=1,4,5,11 SIGMA=0.1 MEAN MORE_THAN={RATIONAL R_0=4} LABEL=cav
```

## 5.5.29 INCYLINDER

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a particular, user-specified part of the cell.

Each of the base quantities calculated by a multicolvar can be assigned to a particular point in three dimensional space. For example, if we have the coordination numbers for all the atoms in the system each coordination number can be assumed to lie on the position of the central atom. Because each base quantity can be assigned to a particular point in space we can calculate functions of the distribution of base quantities in a particular part of the box by using:

$$\bar{s}_r = \frac{\sum_i f(s_i)w(x_i, y_i, z_i)}{\sum_i w(x_i, y_i, z_i)}$$

where the sum is over the collective variables,  $s_i$ , each of which can be thought to be at  $(x_i, y_i, z_i)$ . The function  $w(x_i, y_i, z_i)$  measures whether or not the system is in the subregion of interest. It is equal to:

$$w(x_i, y_i, z_i) =$$

where  $\sigma$  is a [switchingfunction](#). The function  $(s_i)$  can be any of the usual LESS\_THAN, MORE\_THAN, WITHIN etc that are used in all other multicolvars.

When INCYLINDER is used with the [DENSITY](#) action the number of atoms in the specified region is calculated

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.



The atoms involved can be specified using

<b>ATOM</b>	the atom whose vicinity we are interested in examining. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
-------------	------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>KERNEL</b>	( default=gaussian ) the type of kernel function to be used
<b>DIRECTION</b>	the direction of the long axis of the cylinder. Must be x, y or z
<b>RADIUS</b>	a switching function that gives the extent of the cylinder in the plane perpendicular to the direction
<b>LOWER</b>	( default=0.0 ) the lower boundary on the direction parallel to the long axis of the cylinder
<b>UPPER</b>	( default=0.0 ) the upper boundary on the direction parallel to the long axis of the cylinder

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>OUTSIDE</b>	( default=off ) calculate quantities for colvars that are on atoms outside the region of interest
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...

<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>SIGMA</b>	the width of the function to be used for kernel density estimation

### Examples

The input below can be used to calculate the average coordination numbers for those atoms that are within a cylindrical tube of radius 1.5 nm that is centered on the position of atom 101 and that has its long axis parallel to the z-axis.

```
c1: COORDINATIONNUMBER SPECIES=1-100 SWITCH={RATIONAL R_0=0.1}
d2: INCYLINDER ATOM=101 DATA=d1 DIRECTION=Z RADIUS={TANH R_0=1.5} SIGMA=0.1 LOWER=-0.1 UPPER=0.1 MEAN
PRINT ARG=d2.* FILE=colvar
```

### 5.5.30 INSPHERE

<b>This is part of the multicolvar <a href="#">module</a></b>
---------------------------------------------------------------

This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a particular, user-specified part of the cell.

Each of the base quantities calculated by a multicolvar can be assigned to a particular point in three dimensional space. For example, if we have the coordination numbers for all the atoms in the system each coordination number can be assumed to lie on the position of the central atom. Because each base quantity can be assigned to a particular point in space we can calculate functions of the distribution of base quantities in a particular part of the box by using:

$$\bar{s}_r = \frac{\sum_i f(s_i) w(x_i, y_i, z_i)}{\sum_i w(x_i, y_i, z_i)}$$

where the sum is over the collective variables,  $s_i$ , each of which can be thought to be at  $(x_i, y_i, z_i)$ . The function  $w(x_i, y_i, z_i)$  measures whether or not the system is in the subregion of interest. It is equal to:

$$w(x_i, y_i, z_i) =$$

where  $\sigma$  is a [switchingfunction](#). The function  $(s_i)$  can be any of the usual LESS\_THAN, MORE\_THAN, WITHIN etc that are used in all other multicolvars.

When INCYLINDER is used with the [DENSITY](#) action the number of atoms in the specified region is calculated

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>less-than</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>more-than</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOM</b>	the atom whose vicinity we are interested in examining. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
-------------	------------------------------------------------------------------------------------------------------------------------------------------------------------

## Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>KERNEL</b>	( default=gaussian ) the type of kernel function to be used
<b>RADIUS</b>	the switching function that tells us the extent of the spherical region of interest

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>OUTSIDE</b>	( default=off ) calculate quantities for colvars that are on atoms outside the region of interest
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less-than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less-than-1</i> , <i>label.less-than-2</i> , <i>label.less-than-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.more-than</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.more-than-1</i> , <i>label.more-than-2</i> , <i>label.more-than-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

### Examples

The input below can be used to calculate the average coordination numbers for those atoms that are within a sphere of radius 1.5 nm that is centered on the position of atom 101.

```
c1: COORDINATIONNUMBER SPECIES=1-100 SWITCH={RATIONAL R_0=0.1}
d2: INSPHERE ATOM=101 DATA=d1 RADIUS={TANH R_0=1.5} SIGMA=0.1 LOWER=-0.1 UPPER=0.1 MEAN
PRINT ARG=d2.* FILE=colvar
```

### 5.5.31 TETRAHEDRALPORE

This is part of the multicolvar [module](#)

This quantity can be used to calculate functions of the distribution of collective variables for the atoms lie that lie in a box defined by the positions of four atoms at the corners of a tetrahedron.

Each of the base quantities calculated by a multicolvar can be assigned to a particular point in three dimensional space. For example, if we have the coordination numbers for all the atoms in the system each coordination number can be assumed to lie on the position of the central atom. Because each base quantity can be assigned to a particular point in space we can calculate functions of the distribution of base quantities in a particular part of the box by using:

$$\bar{s}_\tau = \frac{\sum_i f(s_i) w(u_i, v_i, w_i)}{\sum_i w(u_i, v_i, w_i)}$$

where the sum is over the collective variables,  $s_i$ , each of which can be thought to be at  $(u_i, v_i, z_i)$ . The function  $(s_i)$  can be any of the usual LESS\_THAN, MORE\_THAN, WITHIN etc that are used in all other multicolvars. Notice that here (at variance with what is done in [AROUND](#)) we have transformed from the usual  $(x_i, y_i, z_i)$  position to a position in  $(u_i, v_i, z_i)$ . This is done using a rotation matrix as follows:

$$(u_i, v_i, w_i) = \mathbf{R} (x_i - x_o, y_i - y_o, z_i - z_o)$$

where  $\mathbf{R}$  is a rotation matrix that is calculated by constructing a set of three orthonormal vectors from the reference positions specified by the user. Initially unit vectors are found by calculating the bisector,  $\mathbf{b}$ , and cross product,  $\mathbf{c}$ , of the vectors connecting atoms 1 and 2. A third unit vector,  $\mathbf{p}$  is then found by taking the cross product between the cross product calculated during the first step,  $\mathbf{c}$  and the bisector,  $\mathbf{b}$ . From this second cross product  $\mathbf{p}$  and the bisector  $\mathbf{b}$  two new vectors are calculated using:

$$v_1 = \cos\left(\frac{\pi}{4}\right) \mathbf{b} + \sin\left(\frac{\pi}{4}\right) \mathbf{p} \quad \text{and} \quad v_2 = \cos\left(\frac{\pi}{4}\right) \mathbf{b} - \sin\left(\frac{\pi}{4}\right) \mathbf{p}$$

In the previous function  $w(u_i, v_i, w_i)$  measures whether or not the system is in the subregion of interest. It is equal to:

$$w(u_i, v_i, w_i) = \int_0^{u'} \int_0^{v'} \int_0^{w'} du dv dw K\left(\frac{u - u_i}{\sigma}\right) K\left(\frac{v - v_i}{\sigma}\right) K\left(\frac{w - w_i}{\sigma}\right)$$

where  $K$  is one of the kernel functions described on [histogrambead](#) and  $\sigma$  is a bandwidth parameter. The values of  $u'$  and  $v'$  are found by finding the projections of the vectors connecting atoms 1 and 2 and 1 and 3  $v_1$  and  $v_2$ . This gives four projections: the largest two projections are used in the remainder of the calculations.  $w'$  is calculated by taking the projection of the vector connecting atoms 1 and 4 on the vector  $\mathbf{c}$ . Notice that the manner by which this box is constructed differs from the way this is done in [CAVITY](#). This is in fact the only point of difference between these two actions.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>ATOMS</b>	the positions of four atoms that define spatial extent of the cavity. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>SIGMA</b>	the width of the function to be used for kernel density estimation
<b>KERNEL</b>	( default=gaussian ) the type of kernel function to be used

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>OUTSIDE</b>	( default=off ) calculate quantities for colvars that are on atoms outside the region of interest
<b>PRINT_BOX</b>	( default=off ) write out the positions of the corners of the box to an xyz file
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...

<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>FILE</b>	the file on which to write out the box coordinates
<b>UNITS</b>	( default=nm ) the units in which to write out the corners of the box

### Examples

The following commands tell plumed to calculate the number of atom inside a tetrahedral cavity. The extent of the tetrahedral cavity is calculated from the positions of atoms 1, 4, 5, and 11, The final value will be labeled cav.

```
d1: DENSITY SPECIES=20-500
TETRAHEDRALPORE DATA=d1 ATOMS=1,4,5,11 SIGMA=0.1 LABEL=cav
```

The following command tells plumed to calculate the coordination numbers (with other water molecules) for the water molecules in the tetrahedral cavity described above. The average coordination number and the number of coordination numbers more than 4 is then calculated. The values of these two quantities are given the labels cav.mean and cav.morethan

```
d1: COORDINATIONNUMBER SPECIES=20-500
CAVITY DATA=d1 ATOMS=1,4,5,11 SIGMA=0.1 MEAN MORE_THAN={RATIONAL R_0=4} LABEL=cav
```

### 5.5.32 GRADIENT

	<b>This is part of the crystallization <a href="#">module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate the gradient of the average value of a multicolvar value

This command allows you to calculate the collective variable discussed in [30].

The atoms involved can be specified using

<b>ORIGIN</b>	we will use the position of this atom as the origin in our calculation. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
---------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>DIR</b>	( default=xyz ) the directions in which we are calculating the gradient. Should be x, y, z, xy, xz, yz or xyz
<b>NBINS</b>	number of bins to use in each direction for the calculation of the gradient
<b>SIGMA</b>	( default=1.0 ) the width of the function to be used for kernel density estimation
<b>KERNEL</b>	( default=gaussian ) the type of kernel function to be used

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

#### Examples

The input below calculates the gradient of the density of atoms in the manner described in [30] in order to detect whether or not atoms are distributed uniformly along the x-axis of the simulation cell.

```
d1: DENSITY SPECIES=1-50
s1: GRADIENT ORIGIN=1 DATA=d1 DIR=x NBINS=4 SIGMA=1.0
PRINT ARG=s1 FILE=colvar
```

The input below calculates the coordination numbers of the 50 atoms in the simulation cell. The gradient of this quantity is then evaluated in the manner described using the equation above to detect whether the average values of the coordination number are uniformly distributed along the x-axis of the simulation cell.

```
d2: COORDINATIONNUMBER SPECIES=1-50 SWITCH={RATIONAL R_0=2.0} MORE_THAN={EXP R_0=4.0}
s2: GRADIENT ORIGIN=1 DATA=d2 DIR=x NBINS=4 SIGMA=1.0
PRINT ARG=s2 FILE=colvar
```



### 5.5.33 INTERMOLECULARTORSIONS

	<b>This is part of the crystallization <a href="#">module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate torsions between axis of adjacent molecules

This command can be used to calculate the intermolecular torsional angles between the orientations of nearby molecules. The orientation of a molecule can be calculated by using either the [MOLECULES](#) or the [PLANES](#) commands. These two commands calculate the orientation of a bond in the molecule or the orientation of a plane containing three of the molecule's atoms. Furthermore, when we use these commands we think of molecules as objects that lie at a point in space and that have an orientation. This command calculates the torsional angles between the orientations of these objects. We can then calculate functions of a large number of these torsional angles that measures things such as the number of torsional angles that are within a particular range. Because it is often useful to only consider the torsional angles between objects that are within a certain distance of each other we can, when calculating these sums, perform a weighted sum and use a [switchingfunction](#) to ensure that we focus on molecules that are close together.

The atoms involved can be specified using

<b>MOLS</b>	The molecules you would like to calculate the torsional angles between. This should be the label/s of <a href="#">MOLECULES</a> or <a href="#">PLANES</a> actions. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
-------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>MOLSA</b>	In this version of the input the torsional angles between all pairs of atoms including one atom from M $\leftrightarrow$ OLA one atom from MOLB will be computed. This should be the label/s of <a href="#">MOLECULES</a> or <a href="#">PLANES</a> actions
<b>MOLSB</b>	In this version of the input the torsional angles between all pairs of atoms including one atom from M $\leftrightarrow$ OLA one atom from MOLB will be computed. This should be the label/s of <a href="#">MOLECULES</a> or <a href="#">PLANES</a> actions

Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D<math>\leftrightarrow</math> _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R<math>\leftrightarrow</math> _0</b>	The r_0 parameter of the switching function

Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.

### Examples

The example input below is necessarily but gives you an idea of what can be achieved using this action. The orientations and positions of four molecules are defined using the [MOLECULES](#) action as the position of the centers of mass of the two atoms specified and the direction of the vector connecting the two atoms that were specified. The torsional angles between the molecules are then calculated by the [INTERMOLECULARTORSIONS](#) command labelled tt\_p. We then compute a [HISTOGRAM](#) that shows the distribution that these torsional angles take in the structure. The weight a given torsional angle contributes to this [HISTOGRAM](#) is determined using a [switchingfunction](#) that acts on the distance between the two molecules. As such the torsional angles between molecules that are close together contribute a high weight to the histogram while the torsional angles between molecules that are far apart does not contribute to the histogram. The histogram is averaged over the whole trajectory and output once all the trajectory frames have been read.

```
m1: MOLECULES MOL1=1,2 MOL2=3,4 MOL3=5,6 MOL4=7,8
tt_p: INTERMOLECULARTORSIONS MOLS=m1 SWITCH={RATIONAL R_0=0.25 D_0=2.0 D_MAX=3.0}
htt_p: HISTOGRAM DATA=tt_p GRID_MIN=-pi GRID_MAX=pi BANDWIDTH=0.1 GRID_BIN=200 STRIDE=1
DUMPGRID GRID=htt_p FILE=myhist.out
```

#### 5.5.34 LOCAL\_AVERAGE

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate averages over spherical regions centered on atoms

As is explained in [this video](#) certain multicolvars calculate one scalar quantity or one vector for each of the atoms in the system. For example [COORDINATIONNUMBER](#) measures the coordination number of each of the atoms in the system and [Q4](#) measures the 4th order Steinhardt parameter for each of the atoms in the system. These quantities provide tell us something about the disposition of the atoms in the first coordination sphere of each of the atoms of interest. Lechner and Dellago [25] have suggested that one can probe local order in a system by taking the average value of such symmetry functions over the atoms within a spherical cutoff of each of these atoms in the systems. When this is done with Steinhardt parameters they claim this gives a coordinate that is better able to distinguish solid and liquid configurations of Lennard-Jones atoms.

You can calculate such locally averaged quantities within plumed by using the LOCAL\_AVERAGE command. This command calculates the following atom-centered quantities:

$$s_i = \frac{c_i + \sum_j \sigma(r_{ij})c_j}{1 + \sum_j \sigma(r_{ij})}$$

where the  $c_i$  and  $c_j$  values can be for any one of the symmetry functions that can be calculated using plumed multicolvars. The function  $\sigma(r_{ij})$  is a [switchingfunction](#) that acts on the distance between atoms  $i$  and  $j$ . Lechner

and Dellago suggest that the parameters of this function should be set so that it the function is equal to one when atom  $j$  is in the first coordination sphere of atom  $i$  and is zero otherwise.

The  $s_i$  quantities calculated using the above command can be again thought of as atom-centred symmetry functions. They thus operate much like multicolvars. You can thus calculate properties of the distribution of  $s_i$  values using MEAN, LESS\_THAN, HISTOGRAM and so on. You can also probe the value of these averaged variables in regions of the box by using the command in tandem with the [AROUND](#) command.

### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>vsum</b>	<b>VSUM</b>	the norm of sum of vectors. The output component can be referred to elsewhere in the input file by using the label.vsum
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↵ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these coordination numbers specifies how many of the atoms specified using SPECIESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

#### Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D↔ _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R↔ _0</b>	The r_0 parameter of the switching function

#### Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN↔N1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN↔HAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BET↔WEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...

<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>VSUM</b>	calculate the norm of the sum of vectors. The final value can be referenced using <i>label.vsum</i> . You can use multiple instances of this keyword i.e. VSUM1, VSUM2, VSUM3... The corresponding values are then referenced using <i>label.vsum-1</i> , <i>label.vsum-2</i> , <i>label.vsum-3</i> ...

### Examples

This example input calculates the coordination numbers for all the atoms in the system. These coordination numbers are then averaged over spherical regions. The number of averaged coordination numbers that are greater than 4 is then output to a file.

```
COORDINATIONNUMBER SPECIES=1-64 D_0=1.3 R_0=0.2 LABEL=d1
LOCAL_AVERAGE ARG=d1 SWITCH={RATIONAL D_0=1.3 R_0=0.2} MORE_THAN={RATIONAL R_0=4} LABEL=la
PRINT ARG=la.* FILE=colvar
```

This example input calculates the  $q_4$  (see [Q4](#)) vectors for each of the atoms in the system. These vectors are then averaged component by component over a spherical region. The average value for this quantity is then outputted to a file. This calculates the quantities that were used in the paper by Lechner and Dellago [\[25\]](#)

```
Q4 SPECIES=1-64 SWITCH={RATIONAL D_0=1.3 R_0=0.2} LABEL=q4
LOCAL_AVERAGE ARG=q4 SWITCH={RATIONAL D_0=1.3 R_0=0.2} MEAN LABEL=la
PRINT ARG=la.* FILE=colvar
```

### 5.5.35 LOCAL\_Q3

<b>This is part of the crystallization <a href="#">module</a></b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate the local degree of order around an atom by taking the average dot product between the  $q_3$  vector on the central atom and the  $q_3$  vector on the atoms in the first coordination sphere.

The [Q3](#) command allows one to calculate one complex vector for each of the atoms in your system that describe the degree of order in the coordination sphere around a particular atom. The difficulty with these vectors comes when combining the order parameters from all of the individual atoms/molecules so as to get a measure of the global degree of order for the system. The simplest way of doing this - calculating the average Steinhardt parameter

- can be problematic. If one is examining nucleation say only the order parameters for those atoms in the nucleus will change significantly when the nucleus forms. The order parameters for the atoms in the surrounding liquid will remain pretty much the same. As such if one models a small nucleus embedded in a very large amount of solution/melt any change in the average order parameter will be negligible. Substantial changes in the value of this average can be observed in simulations of nucleation but only because the number of atoms is relatively small.

When the average Q3 parameter is used to bias the dynamics a problems can occur. These averaged coordinates cannot distinguish between the correct, single-nucleus pathway and a concerted pathway in which all the atoms rearrange themselves into their solid-like configuration simultaneously. This second type of pathway would be impossible in reality because there is a large entropic barrier that prevents concerted processes like this from happening. However, in the finite sized systems that are commonly simulated this barrier is reduced substantially. As a result in simulations where average Steinhardt parameters are biased there are often quite dramatic system size effects

If one wants to simulate nucleation using some form on biased dynamics what is really required is an order parameter that measures:

- Whether or not the coordination spheres around atoms are ordered
- Whether or not the atoms that are ordered are clustered together in a crystalline nucleus

LOCAL\_AVERAGE and NLINKS are variables that can be combined with the Steinhardt parameters allow to calculate variables that satisfy these requirements. LOCAL\_Q3 is another variable that can be used in these sorts of calculations. The LOCAL\_Q3 parameter for a particular atom is a number that measures the extent to which the orientation of the atoms in the first coordination sphere of an atom match the orientation of the central atom. It does this by calculating the following quantity for each of the atoms in the system:

$$s_i = \frac{\sum_j \sigma(r_{ij}) \sum_{m=-3}^3 q_{3m}^*(i) q_{3m}(j)}{\sum_j \sigma(r_{ij})}$$

where  $q_{3m}(i)$  and  $q_{3m}(j)$  are the 3rd order Steinhardt vectors calculated for atom  $i$  and atom  $j$  respectively and the asterix denotes complex conjugation. The function  $\sigma(r_{ij})$  is a [switchingfunction](#) that acts on the distance between atoms  $i$  and  $j$ . The parameters of this function should be set so that the function is equal to one when atom  $j$  is in the first coordination sphere of atom  $i$  and is zero otherwise. The sum in the numerator of this expression is the dot product of the Steinhardt parameters for atoms  $i$  and  $j$  and thus measures the degree to which the orientations of these adjacent atoms is correlated.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
between	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
highest	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
lessthan	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
lowest	<b>LOWEST</b>	the lowest of the quantities calculated by this action
mean	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
min	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
moment	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
morethan	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↔ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↔ IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D↔ _0</b>	( default=0.0 ) The d_0 parameter of the switching function



<b>R↔ _0</b>	The r_0 parameter of the switching function
------------------	---------------------------------------------

## Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN↔N1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less than</i> . You can use multiple instances of this keyword i.e. LESS_THAN↔N1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less than-1</i> , <i>label.less than-2</i> , <i>label.less than-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN↔N1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM↔N1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

## Examples

The following command calculates the average value of the LOCAL\_Q3 parameter for the 64 Lennard Jones atoms in the system under study and prints this quantity to a file called colvar.

```
Q3 SPECIES=1-64 D_0=1.3 R_0=0.2 LABEL=q3
LOCAL_Q3 ARG=q3 SWITCH={RATIONAL D_0=1.3 R_0=0.2} MEAN LABEL=lq3
PRINT ARG=lq3.mean FILE=colvar
```

The following input calculates the distribution of LOCAL\_Q3 parameters at any given time and outputs this information to a file.

```
Q3 SPECIES=1-64 D_0=1.3 R_0=0.2 LABEL=q3
LOCAL_Q3 ARG=q3 SWITCH={RATIONAL D_0=1.3 R_0=0.2} HISTOGRAM={GAUSSIAN LOWER=0.0 UPPER=1.0 NBINS=20 SMEAR=0.1}
PRINT ARG=lq3.* FILE=colvar
```

The following calculates the LOCAL\_Q3 parameters for atoms 1-5 only. For each of these atoms comparisons of the geometry of the coordination sphere are done with those of all the other atoms in the system. The final quantity is the average and is outputted to a file

```
Q3 SPECIESA=1-5 SPECIESB=1-64 D_0=1.3 R_0=0.2 LABEL=q3a
Q3 SPECIESA=6-64 SPECIESB=1-64 D_0=1.3 R_0=0.2 LABEL=q3b

LOCAL_Q3 ARG=q3a,q3b SWITCH={RATIONAL D_0=1.3 R_0=0.2} MEAN LOWMEM LABEL=w3
PRINT ARG=w3.* FILE=colvar
```

### 5.5.36 LOCAL\_Q4

	<p><b>This is part of the crystallization <a href="#">module</a></b></p> <p><b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b></p>
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Calculate the local degree of order around an atoms by taking the average dot product between the  $q_4$  vector on the central atom and the  $q_4$  vector on the atoms in the first coordination sphere.

The **Q4** command allows one to calculate one complex vectors for each of the atoms in your system that describe the degree of order in the coordination sphere around a particular atom. The difficulty with these vectors comes when combining the order parameters from all of the individual atoms/molecules so as to get a measure of the global degree of order for the system. The simplest way of doing this - calculating the average Steinhardt parameter - can be problematic. If one is examining nucleation say only the order parameters for those atoms in the nucleus will change significantly when the nucleus forms. The order parameters for the atoms in the surrounding liquid will remain pretty much the same. As such if one models a small nucleus embedded in a very large amount of solution/melt any change in the average order parameter will be negligible. Substantial changes in the value of this average can be observed in simulations of nucleation but only because the number of atoms is relatively small.

When the average **Q4** parameter is used to bias the dynamics a problems can occur. These averaged coordinates cannot distinguish between the correct, single-nucleus pathway and a concerted pathway in which all the atoms rearrange themselves into their solid-like configuration simultaneously. This second type of pathway would be impossible in reality because there is a large entropic barrier that prevents concerted processes like this from happening. However, in the finite sized systems that are commonly simulated this barrier is reduced substantially. As a result in simulations where average Steinhardt parameters are biased there are often quite dramatic system size effects

If one wants to simulate nucleation using some form on biased dynamics what is really required is an order parameter that measures:

- Whether or not the coordination spheres around atoms are ordered
- Whether or not the atoms that are ordered are clustered together in a crystalline nucleus

[LOCAL\\_AVERAGE](#) and [NLINKS](#) are variables that can be combined with the Steinhardt parameters allow to calculate variables that satisfy these requirements. [LOCAL\\_Q4](#) is another variable that can be used in these sorts of calculations. The [LOCAL\\_Q4](#) parameter for a particular atom is a number that measures the extent to which the orientation of the atoms in the first coordination sphere of an atom match the orientation of the central atom. It does this by calculating the following quantity for each of the atoms in the system:

$$s_i = \frac{\sum_j \sigma(r_{ij}) \sum_{m=-4}^4 q_{4m}^*(i) q_{4m}(j)}{\sum_j \sigma(r_{ij})}$$

where  $q_{4m}(i)$  and  $q_{4m}(j)$  are the 4th order Steinhardt vectors calculated for atom  $i$  and atom  $j$  respectively and the asterisk denotes complex conjugation. The function  $\sigma(r_{ij})$  is a [switchingfunction](#) that acts on the distance between atoms  $i$  and  $j$ . The parameters of this function should be set so that it the function is equal to one when atom  $j$  is in the first coordination sphere of atom  $i$  and is zero otherwise. The sum in the numerator of this expression is the dot product of the Steinhardt parameters for atoms  $i$  and  $j$  and thus measures the degree to which the orientations of these adjacent atoms is correlated.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <i>label.mean</i>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
Generated by Doxygen		

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↔ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↔ IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D↔ _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R↔ _0</b>	The r_0 parameter of the switching function

Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THA↔ N1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...

<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

### Examples

The following command calculates the average value of the LOCAL\_Q4 parameter for the 64 Lennard Jones atoms in the system under study and prints this quantity to a file called colvar.

```
Q4 SPECIES=1-64 D_0=1.3 R_0=0.2 LABEL=q4
LOCAL_Q4 ARG=q4 SWITCH={RATIONAL D_0=1.3 R_0=0.2} MEAN LABEL=lq4
PRINT ARG=lq4.mean FILE=colvar
```

The following input calculates the distribution of LOCAL\_Q4 parameters at any given time and outputs this information to a file.

```
Q4 SPECIES=1-64 D_0=1.3 R_0=0.2 LABEL=q4
LOCAL_Q4 ARG=q4 SWITCH={RATIONAL D_0=1.3 R_0=0.2} HISTOGRAM={GAUSSIAN LOWER=0.0 UPPER=1.0 NBINS=20 SMEAR=0.1}
PRINT ARG=lq4.* FILE=colvar
```

The following calculates the LOCAL\_Q4 parameters for atoms 1-5 only. For each of these atoms comparisons of the geometry of the coordination sphere are done with those of all the other atoms in the system. The final quantity is the average and is outputted to a file

```
Q4 SPECIESA=1-5 SPECIESB=1-64 D_0=1.3 R_0=0.2 LABEL=q4a
Q4 SPECIESA=6-64 SPECIESB=1-64 D_0=1.3 R_0=0.2 LABEL=q4b

LOCAL_Q4 ARG=q4a,q4b SWITCH={RATIONAL D_0=1.3 R_0=0.2} MEAN LOWMEM LABEL=w4
PRINT ARG=w4.* FILE=colvar
```

### 5.5.37 LOCAL\_Q6

	<b>This is part of the crystallization module</b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate the local degree of order around an atoms by taking the average dot product between the  $q_6$  vector on the central atom and the  $q_6$  vector on the atoms in the first coordination sphere.

The **Q6** command allows one to calculate one complex vectors for each of the atoms in your system that describe the degree of order in the coordination sphere around a particular atom. The difficulty with these vectors comes when combining the order parameters from all of the individual atoms/molecules so as to get a measure of the global degree of order for the system. The simplest way of doing this - calculating the average Steinhardt parameter - can be problematic. If one is examining nucleation say only the order parameters for those atoms in the nucleus will change significantly when the nucleus forms. The order parameters for the atoms in the surrounding liquid will remain pretty much the same. As such if one models a small nucleus embedded in a very large amount of solution/melt any change in the average order parameter will be negligible. Substantial changes in the value of this average can be observed in simulations of nucleation but only because the number of atoms is relatively small.

When the average **Q6** parameter is used to bias the dynamics a problems can occur. These averaged coordinates cannot distinguish between the correct, single-nucleus pathway and a concerted pathway in which all the atoms rearrange themselves into their solid-like configuration simultaneously. This second type of pathway would be impossible in reality because there is a large entropic barrier that prevents concerted processes like this from happening. However, in the finite sized systems that are commonly simulated this barrier is reduced substantially. As a result in simulations where average Steinhardt parameters are biased there are often quite dramatic system size effects

If one wants to simulate nucleation using some form on biased dynamics what is really required is an order parameter that measures:

- Whether or not the coordination spheres around atoms are ordered
- Whether or not the atoms that are ordered are clustered together in a crystalline nucleus

**LOCAL\_AVERAGE** and **NLINKS** are variables that can be combined with the Steinhardt parameters allow to calculate variables that satisfy these requirements. **LOCAL\_Q6** is another variable that can be used in these sorts of calculations. The **LOCAL\_Q6** parameter for a particular atom is a number that measures the extent to which the orientation of the atoms in the first coordination sphere of an atom match the orientation of the central atom. It does this by calculating the following quantity for each of the atoms in the system:

$$s_i = \frac{\sum_j \sigma(r_{ij}) \sum_{m=-6}^6 q_{6m}^*(i) q_{6m}(j)}{\sum_j \sigma(r_{ij})}$$

where  $q_{6m}(i)$  and  $q_{6m}(j)$  are the 6th order Steinhardt vectors calculated for atom  $i$  and atom  $j$  respectively and the asterisk denotes complex conjugation. The function  $\sigma(r_{ij})$  is a **switchingfunction** that acts on the distance between atoms  $i$  and  $j$ . The parameters of this function should be set so that it the function is equal to one when atom  $j$  is in the first coordination sphere of atom  $i$  and is zero otherwise. The sum in the numerator of this expression is the dot product of the Steinhardt parameters for atoms  $i$  and  $j$  and thus measures the degree to which the orientations of these adjacent atoms is correlated.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <i>label.mean</i>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↵ Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↵ IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
-----------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword
-----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------

## Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D↔ _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R↔ _0</b>	The r_0 parameter of the switching function

## Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN↔N1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN↔HAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BET↔WEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculates NBIN quantites like BETWEEN. The final value can be referenced using <i>la↔bel.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGR↔AM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...

<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using moment- $m$ .
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

### Examples

The following command calculates the average value of the LOCAL\_Q6 parameter for the 64 Lennard Jones atoms in the system under study and prints this quantity to a file called colvar.

```
Q6 SPECIES=1-64 D_0=1.3 R_0=0.2 LABEL=q6
LOCAL_Q6 ARG=q6 SWITCH={RATIONAL D_0=1.3 R_0=0.2} MEAN LABEL=lq6
PRINT ARG=lq6.mean FILE=colvar
```

The following input calculates the distribution of LOCAL\_Q6 parameters at any given time and outputs this information to a file.

```
Q6 SPECIES=1-64 D_0=1.3 R_0=0.2 LABEL=q6
LOCAL_Q6 ARG=q6 SWITCH={RATIONAL D_0=1.3 R_0=0.2} HISTOGRAM={GAUSSIAN LOWER=0.0 UPPER=1.0 NBINS=20 SMEAR=0.1}
PRINT ARG=lq6.* FILE=colvar
```

The following calculates the LOCAL\_Q6 parameters for atoms 1-5 only. For each of these atoms comparisons of the geometry of the coordination sphere are done with those of all the other atoms in the system. The final quantity is the average and is outputted to a file

```
Q6 SPECIESA=1-5 SPECIESB=1-64 D_0=1.3 R_0=0.2 LABEL=q6a
Q6 SPECIESA=6-64 SPECIESB=1-64 D_0=1.3 R_0=0.2 LABEL=q6b

LOCAL_Q6 ARG=q4a,q4b SWITCH={RATIONAL D_0=1.3 R_0=0.2} MEAN LOWMEM LABEL=w4
PRINT ARG=w6.* FILE=colvar
```

### 5.5.38 NLINKS

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Calculate number of pairs of atoms/molecules that are "linked"

In its simplest guise this coordinate calculates a coordination number. Each pair of atoms is assumed "linked" if they are within some cutoff of each other. In more complex applications each entity is a vector and this quantity measures whether pairs of vectors are (a) within a certain cutoff and (b) if the two vectors have similar orientations. The vectors on individual atoms could be Steinhardt parameters (see [Q3](#), [Q4](#) and [Q6](#)) or they could describe some internal vector in a molecule.

The atoms involved can be specified using

<b>GROUP</b>	. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>GROUPA</b>	
<b>GROUPB</b>	

Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D<sub>↔</sub> _0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R<sub>↔</sub> _0</b>	The r_0 parameter of the switching function

Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous swiching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.

Examples

The following calculates how many bonds there are in a system containing 64 atoms and outputs this quantity to a file.

```
DENSITY SPECIES=1-64 LABEL=d1
NLINKS ARG=d1 SWITCH={RATIONAL D_0=1.3 R_0=0.2} LABEL=dd
PRINT ARG=dd FILE=colvar
```

The following calculates how many pairs of neighbouring atoms in a system containg 64 atoms have similar dispositions for the atoms in their coordination sphere. This calculation uses the dot product of the Q6 vectors on adjacent atoms to measure whether or not two atoms have the same ``orientation"

```
Q6 SPECIES=1-64 SWITCH={RATIONAL D_0=1.3 R_0=0.2} LABEL=q6
NLINKS ARG=q6 SWITCH={RATIONAL D_0=1.3 R_0=0.2} LABEL=dd
PRINT ARG=dd FILE=colvar
```

### 5.5.39 SMAC

	<b>This is part of the crystallization module</b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=crystallization</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate a variant on the SMAC collective variable discussed in [26]

The SMAC collective variable can be used to study the formation of molecular solids from either the melt or from solution. The idea behind this variable is that what differentiates a molecular solid from a molecular liquid is an alignment of internal vectors in neighboring molecules. In other words, the relative orientation of neighboring molecules is no longer random as it is in a liquid. In a solid particular torsional angles between molecules are preferred. As such this CV calculates the following average:

$$s_i = \frac{\left\{1 - \psi \left[ \sum_{j \neq i} \sigma(r_{ij}) \right] \right\} \sum_{j \neq i} \sigma(r_{ij}) \sum_n K_n(\theta_{ij})}{\sum_{j \neq i} \sigma(r_{ij})}$$

In this expression  $r_{ij}$  is the distance between molecule  $i$  and molecule  $j$  and  $\sigma(r_{ij})$  is a [switchingfunction](#) that acts on this distance. By including this switching function in the second summation in the numerator and in the denominator we are thus ensuring that we calculate an average over the molecules in the first coordination sphere of molecule  $i$ . All molecules in higher coordination sphere will essentially contribute zero to the sums in the above expression because their  $\sigma(r_{ij})$  will be very small.  $\psi$  is also a switching function. The term including  $\psi$  in the numerator is there to ensure that only those molecules that are attached to a reasonably large number of molecules. It is important to include this "more than" switching function when you are simulating nucleation from solution with this CV. Lastly, the  $K_n$  functions are [kernelfunctions](#) that take the torsion angle,  $\theta_{ij}$ , between the internal orientation vectors for molecules  $i$  and  $j$  as input. These kernel functions should be set so that they are equal to one when the relative orientation of the molecules are as they are in the solid and equal to zero otherwise. The final  $s_i$  quantity thus measures whether (on average) the molecules in the first coordination sphere around molecule  $i$  are oriented as they would be in the solid. Furthermore, this Action is a multicolvar so you can calculate the  $s_i$  values for all the molecules in your system simultaneously and then determine the average, the number less than and so on.

### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
between	BETWEEN	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
highest	HIGHEST	the lowest of the quantities calculated by this action
lessthan	LESS_THAN	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
lowest	LOWEST	the lowest of the quantities calculated by this action
mean	MEAN	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
min	MIN	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
moment	MOMENTS	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
morethan	MORE_THAN	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

The atoms involved can be specified using

<b>SPECIES</b>	this keyword is used for colvars such as coordination number. In that context it specifies that plumed should calculate one coordination number for each of the atoms specified. Each of these coordination numbers specifies how many of the other specified atoms are within a certain cutoff of the central atom. You can specify the atoms here as another multicolvar action or using a MultiColvar↔Filter or ActionVolume action. When you do so the quantity is calculated for those atoms specified in the previous multicolvar. This is useful if you would like to calculate the Steinhardt parameter for those atoms that have a coordination number more than four for example
----------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>SPECIESA</b>	this keyword is used for colvars such as the coordination number. In that context it species that plumed should calculate one coordination number for each of the atoms specified in SPECIESA. Each of these cooordination numbers specifies how many of the atoms specifies using SPEC↔IESB is within the specified cutoff. As with the species keyword the input can also be specified using the label of another multicolvar
<b>SPECIESB</b>	this keyword is used for colvars such as the coordination number. It must appear with SPECIESA. For a full explanation see the documentation for that keyword

Compulsory keywords

<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R_0</b>	The r_0 parameter of the switching function

<b>KERNEL</b>	The kernels used in the function of the angle You can use multiple instances of this keyword i.e. KERNEL1, KERNEL2, KERNEL3...
<b>SWITCH_COORD</b>	This keyword is used to define the coordination switching function.

## Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less than-1</i> , <i>label.less than-2</i> , <i>label.less than-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

## Examples

In the example below the orientation of the molecules in the system is determined by calculating the vector that connects a pair of atoms. SMAC is then used to determine whether the molecules are sitting in a solid or liquid like environment. We can determine whether the environment is solid or liquid like because in the solid the torsional angle between the bond vectors on adjacent molecules is close to 0 or  $\pi$ . The final quantity that is output to the colvar file measures the number of molecules that have a SMAC parameter that is greater than 0.7. N.B. By using the indices of three atoms for each of the MOL keywords below we are telling PLUMED to use the first two numbers to determine the orientation of the molecule that will ultimately be used when calculating the  $\theta_{ij}$  terms in the formula above. The atom with the third index meanwhile is used when we calculate  $r_{ij}$ .

```
MOLECULES ...
MOL1=9, 10, 9
MOL2=89, 90, 89
MOL3=473, 474, 473
MOL4=1161, 1162, 1161
MOL5=1521, 1522, 1521
MOL6=1593, 1594, 1593
MOL7=1601, 1602, 1601
MOL8=2201, 2202, 2201
LABEL=m3
... MOLECULES

SMAC ...
SPECIES=m3 LOWMEM
KERNEL1={GAUSSIAN CENTER=0 SIGMA=0.480} KERNEL2={GAUSSIAN CENTER=pi SIGMA=0.480}
SWITCH={RATIONAL R_0=0.6} MORE_THAN={RATIONAL R_0=0.7} SWITCH_COORD={EXP R_0=4}
LABEL=s2
... SMAC

PRINT ARG=s2.* FILE=colvar
```

This second example works in a way that is very similar to the previous command. Now, however, the orientation of the molecules is determined by finding the plane that contains the positions of three atoms.

```
PLANES ...
MOL1=9, 10, 11
MOL2=89, 90, 91
MOL3=473, 474, 475
MOL4=1161, 1162, 1163
MOL5=1521, 1522, 1523
MOL6=1593, 1594, 1595
MOL7=1601, 1602, 1603
MOL8=2201, 2202, 2203
VMEAN
LABEL=m3
... PLANES

SMAC ...
SPECIES=m3 LOWMEM
KERNEL1={GAUSSIAN CENTER=0 SIGMA=0.480} KERNEL2={GAUSSIAN CENTER=pi SIGMA=0.480}
SWITCH={RATIONAL R_0=0.6} MORE_THAN={RATIONAL R_0=0.7} SWITCH_COORD={EXP R_0=3.0}
LABEL=s2
... SMAC

PRINT ARG=s2.* FILE=colvar
```

## 5.5.40 MTRANSFORM\_BETWEEN

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

This action can be used to transform the colvar values calculated by a multicolvar using a [histogrambead](#)

In this action each colvar,  $s_i$ , calculated by multicolvar is transformed by a [histogrambead](#) function that is equal to one if the colvar is within a certain range and which is equal to zero otherwise. In other words, we compute:

$$f_i = \int_a^b K\left(\frac{s - s_i}{w}\right)$$

where  $a, b$  and  $w$  are parameters.

It is important to understand the distinction between what is done here and what is done by [MFILTER\\_BETWEEN](#). In [MFILTER\\_BETWEEN](#) a weight,  $w_i$  for the colvar is calculated using the [histogrambead](#). If one calculates the MEAN for [MFILTER\\_BETWEEN](#) one is thus calculating:

$$\mu = \frac{\sum_i f_i s_i}{\sum_i f_i}$$

In this action by contrast the colvar is being transformed by the [histogrambead](#). If one thus calculates a MEAN for this action one computes:

$$\mu = \frac{\sum_{i=1}^N f_i}{N}$$

In other words, you are calculating the mean for the transformed colvar.

### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be refer-



## Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>LOWER</b>	the lower boundary for the range of interest
<b>UPPER</b>	the upper boundary for the range of interest
<b>SMEAR</b>	( default=0.5 ) the ammount by which to smear the value for kernel density estimation

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>

<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>BEAD</b>	This keywords is used if you want to employ an alternative to the function defined above. The following provides information on the <a href="#">histogrambead</a> that are available. When this keyword is present you no longer need the LOWER, UPPER and SMEAR keywords.

### Examples

The following input gives an example of how a MTRANSFORM\_BETWEEN action can be used to duplicate functionality that is elsewhere in PLUMED.

```
DISTANCES ...
  GROUPA=1-10 GROUPB=11-20
  LABEL=d1
... DISTANCES
MTRANSFORM_BETWEEN DATA=d1 LOWER=1.0 UPPER=2.0 SMEAR=0.5
```

In this case you can achieve the same result by using:

```
DISTANCES ...
  GROUPA=1-10 GROUPB=11-20
  BETWEEN={GAUSSIAN LOWER=1.0 UPPER=2.0}
... DISTANCES
```

(see [DISTANCES](#))

The advantage of MTRANSFORM\_BETWEEN comes, however, if you want to use transformed colvars as input for [MULTICOLVARENS](#)

### 5.5.41 MTRANSFORM\_LESS

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

This action can be used to transform the colvar values calculated by a multicolvar using a [switchingfunction](#)

In this action each colvar,  $s_i$ , calculated by multicolvar is transformed by a [switchingfunction](#) function that is equal to one if the colvar is less than a certain target value and which is equal to zero otherwise. It is important to understand the distinction between what is done here and what is done by [MFILTER\\_LESS](#). In [MFILTER\\_LESS](#) a weight,  $w_i$  for the colvar is calculated using the [switchingfunction](#). If one calculates the MEAN for [MFILTER\\_LESS](#) one is thus calculating:

$$\mu = \frac{\sum_i \sigma(s_i) s_i}{\sum_i (s_i)}$$

where  $\sigma$  is the [switchingfunction](#). In this action by contrast the colvar is being transformed by the [switchingfunction](#). If one thus calculates a MEAN for this action one computes:

$$\mu = \frac{\sum_{i=1}^N (s_i)}{N}$$

In other words, you are calculating the mean for the transformed colvar.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.

#### Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R_0</b>	The r_0 parameter of the switching function

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-m</i> .
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.

### Examples

The following input gives an example of how a MTRANSFORM\_LESS action can be used to duplicate functionality that is elsewhere in PLUMED.

```
DISTANCES ...
  GROUPA=1-10 GROUPB=11-20
  LABEL=d1
... DISTANCES
MTRANSFORM_LESS DATA=d1 SWITCH={GAUSSIAN D_0=1.5 R_0=0.00001}
```

In this case you can achieve the same result by using:

```
DISTANCES ...
  GROUPA=1-10 GROUPB=11-20
  LESS_THAN={GAUSSIAN D_0=1.5 R_0=0.00001}
... DISTANCES
```

(see [DISTANCES](#))

The advantage of MTRANSFORM\_LESS comes, however, if you want to use transformed colvars as input for [MULTICOLVARDENS](#)

### 5.5.42 MTRANSFORM\_MORE

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

This action can be used to transform the colvar values calculated by a multicolvar using one minus a [switchingfunction](#)

In this action each colvar,  $s_i$ , calculated by multicolvar is transformed by a [switchingfunction](#) function that is equal to one if the colvar is greater than a certain target value and which is equal to zero otherwise. It is important to understand the distinction between what is done here and what is done by [MFILTER\\_MORE](#). In [MFILTER\\_MORE](#) a weight,  $w_i$  for the colvar is calculated using the [histogrambead](#). If one calculates the MEAN for [MFILTER\\_MORE](#) one is thus calculating:

$$\mu = \frac{\sum_i [1 - \sigma(s_i)] s_i}{\sum_i [1 - \sigma(s_i)]}$$

where  $\sigma$  is the [switchingfunction](#). In this action by contrast the colvar is being transformed by the [switchingfunction](#). If one thus calculates a MEAN for this action one computes:

$$\mu = \frac{\sum_{i=1}^N 1 - \sigma(s_i)}{N}$$

In other words, you are calculating the mean for the transformed colvar.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the <code>label.vmean</code>
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <code>label.mean</code>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <code>label.moment-2</code> , the third as <code>label.moment-3</code> , etc.

#### Compulsory keywords

<b>DATA</b>	The multicolvar that calculates the set of base quantities that we are interested in
<b>NN</b>	( default=6 ) The n parameter of the switching function
<b>MM</b>	( default=0 ) The m parameter of the switching function; 0 implies 2*NN
<b>D_0</b>	( default=0.0 ) The d_0 parameter of the switching function
<b>R_0</b>	The r_0 parameter of the switching function

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <code>label.vmean</code> . You can use multiple instances of this keyword i.e. <code>VMEAN1</code> , <code>VMEAN2</code> , <code>VMEAN3</code> ... The corresponding values are then referenced using <code>label.vmean-1</code> , <code>label.vmean-2</code> , <code>label.vmean-3</code> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <code>label.mean</code> . You can use multiple instances of this keyword i.e. <code>MEAN1</code> , <code>MEAN2</code> , <code>MEAN3</code> ... The corresponding values are then referenced using <code>label.mean-1</code> , <code>label.mean-2</code> , <code>label.mean-3</code> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a lists of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <code>moment-<math>m</math></code> .

<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.

### Examples

The following input gives an example of how a MTRANSFORM\_MORE action can be used to duplicate functionality that is elsewhere in PLUMED.

```
DISTANCES ...
  GROUPA=1-10 GROUPB=11-20
  LABEL=d1
... DISTANCES
MTRANSFORM_MORE DATA=d1 SWITCH={GAUSSIAN D_0=1.5 R_0=0.00001}
```

In this case you can achieve the same result by using:

```
DISTANCES ...
  GROUPA=1-10 GROUPB=11-20
  MORE_THAN={GAUSSIAN D_0=1.5 R_0=0.00001}
... DISTANCES
```

(see [DISTANCES](#))

The advantage of MTRANSFORM\_MORE comes, however, if you want to use transformed colvars as input for [MULTICOLVARENS](#)

## 5.5.43 UWALLS



	This is part of the <a href="#">manyrestraints module</a>
	It is only available if you configure PLUMED with <code>./configure --enable-modules=manyrestraints</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.

Add [UPPER\\_WALLS](#) restraints on all the multicolvar values

This action takes the set of values calculated by the colvar specified by label in the DATA keyword and places a restraint on each quantity,  $x$ , with the following functional form:

$$k((x - a + o)/s)^e$$

$k$  (KAPPA) is an energy constant in internal unit of the code,  $s$  (EPS) a rescaling factor and  $e$  (EXP) the exponent determining the power law. By default: EXP = 2, EPS = 1.0, OFF = 0.

#### Description of components

By default the value of the calculated quantity can be referenced elsewhere in the input file by using the label of the action. Alternatively this Action can be used to calculate the following quantities by employing the keywords listed below. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potentials

In addition the following quantities can be calculated by employing the keywords listed below

Quantity	Keyword	Description
<b>gradient</b>	<b>GRADIENT</b>	the gradient
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>vsum</b>	<b>VSUM</b>	the norm of sum of vectors. The output component can be referred to elsewhere in the input file by using the label.vsum
<b>spath</b>	<b>SPATH</b>	the position on the path
<b>zpath</b>	<b>ZPATH</b>	the distance from the path
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean

<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>sum</b>	<b>SUM</b>	the sum of values

### Compulsory keywords

<b>DATA</b>	certain actions in plumed work by calculating a list of variables and summing over them. This particular action can be used to calculate functions of these base variables or prints them to a file. This keyword thus takes the label of one of those such variables as input.
<b>AT</b>	the radius of the sphere
<b>KAPPA</b>	the force constant for the wall. The $k_i$ in the expression for a wall.
<b>OFFSET</b>	( default=0.0 ) the offset for the start of the wall. The $o_i$ in the expression for a wall.
<b>EXP</b>	( default=2.0 ) the powers for the walls. The $e_i$ in the expression for a wall.
<b>EPS</b>	( default=1.0 ) the values for $s_i$ in the expression for a wall

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

### Examples

The following set of commands can be used to stop a cluster composed of 20 atoms subliming. The position of the centre of mass of the cluster is calculated by the [COM](#) command labelled c1. The [DISTANCES](#) command labelled d1 is then used to calculate the distance between each of the 20 atoms in the cluster and the center of mass of the cluster. These distances are then passed to the UWALLS command, which adds a [UPPER\\_WALLS](#) restraint on each of them and thereby prevents each of them from moving very far from the centre of mass of the cluster.

```
COM ATOMS=1-20 LABEL=c1
DISTANCES GROUPA=c1 GROUPB=1-20 LABEL=d1
UWALLS DATA=d1 AT=2.5 KAPPA=0.2 LABEL=sr
```

## 5.6 Exploiting contact matrices

A contact matrix is an  $N \times N$  matrix in which the  $i$ th,  $j$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are adjacent or not. There are various ways of defining

whether a pair of atoms/molecules are adjacent or not. For example we can say two atoms are adjacent if the distance between them is less than some cutoff. Alternatively, if we have a pair of molecules, we might state they are adjacent if their centre's of mass are within a certain cutoff and if the two molecules have the same orientation. Two electronegative atoms might be said to be adjacent if there is a hydrogen bond between them. For these reasons then PLUMED contains all of the following methods for calculating an adjacency matrix

<a href="#">ALIGNED_MATRIX</a>	Adjacency matrix in which two molecule are adjacent if they are within a certain cutoff and if they have the same orientation.
<a href="#">CONTACT_MATRIX</a>	Adjacency matrix in which two atoms are adjacent if they are within a certain cutoff.
<a href="#">HBOND_MATRIX</a>	Adjacency matrix in which two atoms are adjacent if there is a hydrogen bond between them.
<a href="#">SMAC_MATRIX</a>	Adjacency matrix in which two molecules are adjacent if they are within a certain cutoff and if the angle between them is within certain ranges.

Once you have calculated an adjacency matrix you can then perform any one of the following operations on this object in order to reduce it to a scalar number or a set of connected components.

<a href="#">CLUSTER_WITHSURFACE</a>	Take a connected component that was found using a clustering algorithm and create a new cluster that contains those atoms that are in the cluster together with those atoms that are within a certain cutoff of the cluster.
<a href="#">COLUMNSUMS</a>	Sum the columns of a contact matrix
<a href="#">DFSCLUSTERING</a>	Find the connected components of the matrix using the depth first search clustering algorithm.
<a href="#">ROWSUMS</a>	Sum the rows of a adjacency matrix.
<a href="#">SPRINT</a>	Calculate SPRINT topological variables from an adjacency matrix.

If the function you have chosen reduces your contact matrix to a set of connected components you then need a method to convert these connected components into a scalar number or to output this information to a file. The various things that you can do with a set of connected components are listed below:

<a href="#">CLUSTER_DIAMETER</a>	Print out the diameter of one of the connected components
<a href="#">CLUSTER_DISTRIBUTION</a>	Calculate functions of the distribution of properties in your connected components.
<a href="#">CLUSTER_NATOMS</a>	Gives the number of atoms in the connected component
<a href="#">CLUSTER_PROPERTIES</a>	Calculate properties of the distribution of some quantities that are part of a connected component
<a href="#">OUTPUT_CLUSTER</a>	Output the indices of the atoms in one of the clusters identified by a clustering object

### 5.6.1 ALIGNED\_MATRIX

	<b>This is part of the <a href="#">adjmat module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Adjacency matrix in which two molecule are adjacent if they are within a certain cutoff and if they have the same orientation.

As discussed in the section of the manual on [Exploiting contact matrices](#) a useful tool for developing complex collective variables is the notion of the so called adjacency matrix. An adjacency matrix is an  $N \times N$  matrix in which

the  $i$ th,  $j$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are adjacent or not. These matrices can then be further analysed using a number of other algorithms as is detailed in [31].

For this action the elements of the adjacency matrix are calculated using:

$$a_{ij} = \sigma_1(|\mathbf{r}_{ij}|)\sigma_2(\mathbf{v}_i \cdot \mathbf{v}_j)$$

This form of adjacency matrix can only be used if the input species are objects that lie at a point in space and that have an orientation,  $\mathbf{v}$ . These orientations might represent the orientation of a molecule, which could be calculated using [MOLECULES](#) or [PLANES](#), or it might be the complex vectors calculated using the Steinhardt parameters [Q3](#), [Q4](#) or [Q6](#). In the expression above  $\mathbf{r}_{ij}$  is the vector connecting the points in space where objects  $i$  and  $j$  find themselves and  $\sigma_1$  is a [switchingfunction](#) that acts upon the magnitude of this vector.  $\sigma_2$  is a second [switchingfunction](#) that acts on the dot product of the directors of the vectors that define the orientations of objects  $i$  and  $j$ .

The atoms involved can be specified using

<b>ATOMS</b>	The list of molecules for which you would like to calculate the contact matrix. The molecules involved must have an orientation so your list will be a list of the labels of <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a> as PLUMED calculates the orientations of molecules within these operations. Please note also that the majority of <a href="#">MultiColvar</a> and <a href="#">MultiColvar functions</a> do not calculate a molecular orientation.. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Or alternatively by using

<b>ATOMSA</b>	The list of molecules that you would like to use for the rows of the contact matrix. The molecules involved must have an orientation so your list will be a list of the labels of <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a> as PLUMED calculates the orientations of molecules within these operations. Please note also that the majority of <a href="#">MultiColvar</a> and <a href="#">MultiColvar functions</a> do not calculate a molecular orientation.
<b>ATOMSB</b>	The list of molecules that you would like to use for the columns of the contact matrix. The molecules involved must have an orientation so your list will be a list of the labels of <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a> as PLUMED calculates the orientations of molecules within these operations. Please note also that the majority of <a href="#">MultiColvar</a> and <a href="#">MultiColvar functions</a> do not calculate a molecular orientation.

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>HIGHMEM</b>	( default=off ) use a more memory intensive version of this collective variable

<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords. You can use multiple instances of this keyword i.e. SWITCH1, SWITCH2, SWITCH3...
<b>ORIENTATION_SWITCH</b>	A switching function that transforms the dot product of the input vectors. You can use multiple instances of this keyword i.e. ORIENTATION_SWITCH1, ORIENTATION_SWITCH2, ORIENTATION_SWITCH3...

### Examples

The example input below is necessarily but gives you an idea of what can be achieved using this action. The orientations and positions of four molecules are defined using the [MOLECULES](#) action as the position of the centers of mass of the two atoms specified and the direction of the vector connecting the two atoms that were specified. A  $4 \times 4$  matrix is then computed using the formula above. The  $ij$ -element of this matrix tells us whether or not atoms  $i$  and  $j$  are within 0.1 nm of each other and whether or not the dot-product of their orientation vectors is greater than 0.5. The sum of the rows of this matrix are then computed. The sums of the  $i$ th row of this matrix tells us how many of the molecules that are within the first coordination sphere of molecule  $i$  have an orientation that is similar to that of molecule  $i$ . We thus calculate the number of these "coordination numbers" that are greater than 1.0 and output this quantity to a file.

```
m1: MOLECULES MOL1=1,2 MOL2=3,4 MOL3=5,6 MOL4=7,8
mat: ALIGNED_MATRIX ATOMS=m1 SWITCH={RATIONAL R_0=0.1} ORIENTATION_SWITCH={RATIONAL R_0=0.1 D_MAX=0.5}
rr: ROWSUMS MATRIX=mat MORE_THAN={RATIONAL D_0=1.0 R_0=0.1}
PRINT ARG=rr.* FILE=colvar
```

### 5.6.2 CONTACT\_MATRIX

<b>This is part of the adjmat <a href="#">module</a></b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Adjacency matrix in which two atoms are adjacent if they are within a certain cutoff.

As discussed in the section of the manual on [Exploiting contact matrices](#) a useful tool for developing complex collective variables is the notion of the so called adjacency matrix. An adjacency matrix is an  $N \times N$  matrix in which the  $ij$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are adjacent or not. These matrices can then be further analysed using a number of other algorithms as is detailed in [\[31\]](#).

For this action the elements of the contact matrix are calculated using:

$$a_{ij} = \sigma(|\mathbf{r}_{ij}|)$$

where  $|\mathbf{r}_{ij}|$  is the magnitude of the vector connecting atoms  $i$  and  $j$  and where  $\sigma$  is a [switchingfunction](#).

The atoms involved can be specified using

<b>ATOMS</b>	The list of atoms for which you would like to calculate the contact matrix. The atoms involved must be specified as a list of labels of <a href="#">MultiColvar</a> or labels of a <a href="#">MultiColvar functions</a> actions. If you would just like to use the atomic positions you can use a <a href="#">DENSITY</a> command to specify a group of atoms. Specifying your atomic positions using labels of other <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a> commands is useful, however, as you can then exploit a much wider variety of functions of the contact matrix as described in <a href="#">Exploiting contact matrices</a> . For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>HIGHMEM</b>	( default=off ) use a more memory intensive version of this collective variable
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords. You can use multiple instances of this keyword i.e. SWITCH1, SWITCH2, SWITCH3...

## Examples

The input shown below calculates a  $6 \times 6$  matrix whose elements are equal to one if atom  $i$  and atom  $j$  are within 0.3 nm of each other and which is zero otherwise. The columns in this matrix are then summed so as to give the coordination number for each atom. The final quantity output in the colvar file is thus the average coordination number.

```
aa: CONTACT_MATRIX ATOMS=1-6 SWITCH={EXP D_0=0.2 R_0=0.1 D_MAX=0.66}
COLUMNSUMS MATRIX=mat MEAN LABEL=csums
PRINT ARG=csums.* FILE=colvar
```

## 5.6.3 HBOND\_MATRIX

<b>This is part of the adjmat <a href="#">module</a></b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Adjacency matrix in which two atoms are adjacent if there is a hydrogen bond between them.

As discussed in the section of the manual on [Exploiting contact matrices](#) a useful tool for developing complex collective variables is the notion of the so called adjacency matrix. An adjacency matrix is an  $N \times N$  matrix in which the  $i$ th,  $j$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are

adjacent or not. These matrices can then be further analysed using a number of other algorithms as is detailed in [31].

For this action the elements of the adjacency matrix are calculated using:

$$a_{ij} = \sigma_{oo}(|\mathbf{r}_{ij}|) \sum_{k=1}^N \sigma_{oh}(|\mathbf{r}_{ik}|) \sigma_{\theta}(\theta_{kij})$$

This expression was derived by thinking about how to detect if there is a hydrogen bond between atoms  $i$  and  $j$ . The notion is that if the hydrogen bond is present atoms  $i$  and  $j$  should be within a certain cutoff distance. In addition, there should be a hydrogen within a certain cutoff distance of atom  $i$  and this hydrogen should lie on or close to the vector connecting atoms  $i$  and  $j$ . As such  $\sigma_{oo}(|\mathbf{r}_{ij}|)$  is a [switchingfunction](#) that acts on the modulus of the vector connecting atom  $i$  to atom  $j$ . The sum over  $k$  then runs over all the hydrogen atoms that are specified using using HYDROGEN keyword.  $\sigma_{oh}(|\mathbf{r}_{ik}|)$  is a [switchingfunction](#) that acts on the modulus of the vector connecting atom  $i$  to atom  $k$  and  $\sigma_{\theta}(\theta_{kij})$  is a [switchingfunction](#) that acts on the angle between the vector connecting atoms  $i$  and  $j$  and the vector connecting atoms  $i$  and  $k$ .

It is important to note that hydrogen bonds, unlike regular bonds, are asymmetric. In other words, the hydrogen atom does not sit at the mid point between the two other atoms in this three-center bond. As a result of this adjacency matrices calculated using [HBOND\\_MATRIX](#) are not symmetric like those calculated by [CONTACT\\_MATRIX](#). One consequence of this fact is that the quantities found by performing [ROWSUMS](#) and [COLUMNSUMS](#) on a square [HBOND\\_MATRIX](#) are not the same as they would be if you performed [ROWSUMS](#) and [COLUMNSUMS](#) on a square [CONTACT\\_MATRIX](#).

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
sum	SUM	the sum of values

The atoms involved can be specified using

<b>ATOMS</b>	The list of atoms which can be part of a hydrogen bond. When this command is used the set of atoms that can donate a hydrogen bond is assumed to be the same as the set of atoms that can form hydrogen bonds. The atoms involved must be specified as a list of labels of <a href="#">MultiColvar</a> or labels of a <a href="#">MultiColvar functions</a> actions. If you would just like to use the atomic positions you can use a <a href="#">DENSITY</a> command to specify a group of atoms. Specifying your atomic positions using labels of other <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a> commands is useful, however, as you can then exploit a much wider variety of functions of the contact matrix as described in <a href="#">Exploiting contact matrices</a> . For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
<b>HYDROGENS</b>	The list of hydrogen atoms that can form part of a hydrogen bond. The atoms must be specified using a comma separated list, an index range or by using a <a href="#">GROUP</a> . A list of hydrogen atoms is always required even if you specify the other atoms using DONORS and ACCEPTORS as described below.. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>

Or alternatively by using

<b>DONORS</b>	The list of atoms which can donate a hydrogen bond. The atoms involved must be specified as a list of labels of <a href="#">MultiColvar</a> or labels of a <a href="#">MultiColvar functions</a> actions. If you would just like to use the atomic positions you can use a <a href="#">DENSITY</a> command to specify a group of atoms. Specifying your atomic positions using labels of other <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a> commands is useful, however, as you can then exploit a much wider variety of functions of the contact matrix as described in <a href="#">Exploiting contact matrices</a>
<b>ACCEPTORS</b>	The list of atoms which can accept a hydrogen bond. The atoms involved must be specified as a list of labels of <a href="#">MultiColvar</a> or labels of a <a href="#">MultiColvar functions</a> actions. If you would just like to use the atomic positions you can use a <a href="#">DENSITY</a> command to specify a group of atoms. Specifying your atomic positions using labels of other <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a> commands is useful, however, as you can then exploit a much wider variety of functions of the contact matrix as described in <a href="#">Exploiting contact matrices</a>

Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>HIGHMEM</b>	( default=off ) use a more memory intensive version of this collective variable
<b>SWITCH</b>	The <a href="#">switchingfunction</a> that specifies how close a pair of atoms must be together for there to be a hydrogen bond between them You can use multiple instances of this keyword i.e. SWITCH1, SWITCH2, SWITCH3...
<b>HSWITCH</b>	The <a href="#">switchingfunction</a> that specifies how close the hydrogen must be to the donor atom of the hydrogen bond for it to be considered a hydrogen bond You can use multiple instances of this keyword i.e. HSWITCH1, HSWITCH2, HSWITCH3...
<b>ASWITCH</b>	A <a href="#">switchingfunction</a> that is used to specify what the angle between the vector connecting the donor atom to the acceptor atom and the vector connecting the donor atom to the hydrogen must be in order for it considered to be a hydrogen bond You can use multiple instances of this keyword i.e. ASWITCH1, ASWITCH2, ASWITCH3...



<b>SUM</b>	calculate the sum of all the quantities. The final value can be referenced using <i>label.sum</i> . You can use multiple instances of this keyword i.e. SUM1, SUM2, SUM3... The corresponding values are then referenced using <i>label.sum-1</i> , <i>label.sum-2</i> , <i>label.sum-3</i> ...
------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Examples

The following input can be used to analyse the number of hydrogen bonds each of the oxygen atoms in a box of water participates in. Each water molecule can participate in a hydrogen bond in one of two ways. It can either donate its hydrogens to the neighbouring oxygen or it can accept a bond between the hydrogen of a neighboring water molecule and its own oxygen. The input below allows you to output information on the number of hydrogen bonds each of the water molecules donates and accepts. This information is output in two xyz files which each contain five columns of data. The first four of these columns are a label for the atom and the x, y and z position of the oxygen. The last column is then the number of accepted/donated hydrogen bonds.

```
mat: HBOND_MATRIX ATOMS=1-192:3 HYDROGENS=2-192:3,3-192:3 SWITCH={RATIONAL R_0=3.20} HSWITCH={RATIONAL R_0=2.3}
rsums: ROWSUMS MATRIX=mat MEAN
csums: COLUMNSUMS MATRIX=mat MEAN
DUMPMULTICOLVAR DATA=rsums FILE=donors.xyz
DUMPMULTICOLVAR DATA=csums FILE=acceptors.x
```

#### 5.6.4 SMAC\_MATRIX

<b>This is part of the adjmat module</b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Adjacency matrix in which two molecules are adjacent if they are within a certain cutoff and if the angle between them is within certain ranges.

In this case the elements of the adjacency matrix are calculated using:

$$A_{ij} = \sigma(r_{ij}) \sum_n K_n(\theta_{ij})$$

In this expression  $r_{ij}$  is the distance between molecule  $i$  and molecule  $j$  and  $\sigma(r_{ij})$  is a [switching function](#) that acts on this distance. The  $K_n$  functions are [kernel functions](#) that take the torsion angle,  $\theta_{ij}$ , between the internal orientation vectors for molecules  $i$  and  $j$  as input. These kernel functions should be set so that they are equal to one when the relative orientation of the molecules are as they are in the solid and equal to zero otherwise. As the above matrix element is a product of functions it is only equal to one when the centers of mass of molecules  $i$  and  $j$  are with a certain distance of each other and when the molecules are aligned in some desirable way.

The atoms involved can be specified using

<b>ATOMS</b>	The list of molecules for which you would like to calculate the contact matrix. The molecules involved must have an orientation so your list will be a list of the labels of <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a>
<b>Generated by Doxygen</b>	PLUMED calculates the orientations of molecules within these operations. Please note also that the majority of <a href="#">MultiColvar</a> and <a href="#">MultiColvar functions</a> do not calculate a molecular orientation.. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>

Or alternatively by using

<b>ATOMSA</b>	The list of molecules that you would like to use for the rows of the contact matrix. The molecules involved must have an orientation so your list will be a list of the labels of <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a> as PLUMED calculates the orientations of molecules within these operations. Please note also that the majority of <a href="#">MultiColvar</a> and <a href="#">MultiColvar functions</a> do not calculate a molecular orientation.
<b>ATOMSB</b>	The list of molecules that you would like to use for the columns of the contact matrix. The molecules involved must have an orientation so your list will be a list of the labels of <a href="#">MultiColvar</a> or <a href="#">MultiColvar functions</a> as PLUMED calculates the orientations of molecules within these operations. Please note also that the majority of <a href="#">MultiColvar</a> and <a href="#">MultiColvar functions</a> do not calculate a molecular orientation.

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>HIGHMEM</b>	( default=off ) use a more memory intensive version of this collective variable
<b>SWITCH</b>	This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the <a href="#">switchingfunction</a> that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords. You can use multiple instances of this keyword i.e. SWITCH1, SWITCH2, SWITCH3...
<b>KERNEL</b>	The various kernels that are used to determine whether or not the molecules are aligned You can use multiple instances of this keyword i.e. KERNEL1, KERNEL2, KERNEL3...

### Examples

In the following example an adjacency matrix is constructed in which the  $(i, j)$  element is equal to one if molecules  $i$  and  $j$  are within 6 angstroms of each other and if the torsional angle between the orientations of these molecules is close to 0 or  $\pi$ . The various connected components of this matrix are determined using the [DFSCCLUSTERING](#) algorithm and then the size of the largest cluster of connectes molecules is output to a colvar file

```

UNITS LENGTH=A

MOLECULES ...
MOL1=1, 2, 1
MOL2=5, 6, 5
MOL3=9, 10, 9
MOL4=13, 14, 13
MOL5=17, 18, 17
LABEL=m1
... MOLECULES

SMAC_MATRIX ...
```

```

ATOMS=m1 SWITCH={RATIONAL D_0=5.99 R_0=0.1 D_MAX=6.0}
KERNEL1={TRIANGULAR CENTER=0 SIGMA=1.0} KERNEL2={TRIANGULAR CENTER=pi SIGMA=0.6}
LABEL=smacm
... SMAC_MATRIX

dfs1: DFSCUSTERING MATRIX=smacm
cc2: CLUSTER_NATOMS CLUSTERS=dfs1 CLUSTER=1
PRINT ARG=smac.*,cc1.*,cc2 FILE=colvar

```

### 5.6.5 CLUSTER\_WITHSURFACE

	<b>This is part of the <a href="#">adjmat module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Take a connected component that was found using a clustering algorithm and create a new cluster that contains those atoms that are in the cluster together with those atoms that are within a certain cutoff of the cluster.

As discussed in the section of the manual on [Exploiting contact matrices](#) a useful tool for developing complex collective variables is the notion of the so called adjacency matrix. An adjacency matrix is an  $N \times N$  matrix in which the  $i$ th,  $j$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are adjacent or not. When analysing these matrix we can treat them as a graph and find connected components using some clustering algorithm. This action is used in tandem with this form of analysis and takes one of the connected components that was found during this analysis and creates a new cluster that includes all the atoms within the connected component that was found together that were within a certain cutoff distance of the atoms in the connected component. This form of analysis has been used successfully in the forward flux sampling simulations described in this paper [32]

#### Compulsory keywords

<b>CLUSTERS</b>	the label of the action that does the clustering
<b>RCUT_SURF</b>	you also have the option to find the atoms on the surface of the cluster. An atom must be within this distance of one of the atoms of the cluster in order to be considered a surface atom

#### Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

#### Examples

The following input uses PLUMED to calculate a adjacency matrix that connects a pair of atoms if they both have a coordination number that is less than 13.5 and if they are within 0.38 nm of each other. Depth first search clustering

is used to find the connected components in this matrix. The number of atoms with indices that are between 1 and 1996 and that are either in the second largest cluster or that are within 0.3 nm of one of the atoms within the second largest cluster are then counted and this number of atoms is output to a file called size. In addition the indices of the atoms that were counted are output to a file called dfs2.dat.

```
c1: COORDINATIONNUMBER SPECIES=1-1996 SWITCH={CUBIC D_0=0.34 D_MAX=0.38}
cf: MFILTER_LESS DATA=c1 SWITCH={CUBIC D_0=13 D_MAX=13.5}
mat: CONTACT_MATRIX ATOMS=cf SWITCH={CUBIC D_0=0.34 D_MAX=0.38}
dfs: DFSCUSTERING MATRIX=mat
clust2a: CLUSTER_WITHSURFACE CLUSTERS=dfs RCUT_SURF=0.3
size2a: CLUSTER_NATOMS CLUSTERS=clust2a CLUSTER=2
PRINT ARG=size2a FILE=size FMT=%8.4f
OUTPUT_CLUSTER CLUSTERS=clust2a CLUSTER=2 FILE=dfs2.dat
```

### 5.6.6 COLUMNSUMS

	<b>This is part of the <a href="#">adjmat module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Sum the columns of a contact matrix

As discussed in the section of the manual on [Exploiting contact matrices](#) a useful tool for developing complex collective variables is the notion of the so called adjacency matrix. An adjacency matrix is an  $N \times N$  matrix in which the  $i$ th,  $j$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are adjacent or not. This action allows you to calculate the sum of the columns in this adjacency matrix and to then calculate further functions of these quantities.

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action

<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <i>label.mean</i>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

## Compulsory keywords

<b>MAT↔ RIX</b>	the action that calculates the adjacency matrix vessel we would like to analyse
---------------------	---------------------------------------------------------------------------------

## Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BE↔TA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...

<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less-than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less-than-1</i> , <i>label.less-than-2</i> , <i>label.less-than-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.more-than</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.more-than-1</i> , <i>label.more-than-2</i> , <i>label.more-than-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .

## Examples

The first instruction in the following input file tells PLUMED to compute a  $10 \times 10$  matrix in which the  $ij$ -element tells you whether atoms  $i$  and  $j$  are within 1.0 nm of each other. The numbers in each of these rows are then added together and the average value is computed. As such the following input provides an alternative method for calculating the coordination numbers of atoms 1 to 10.

```
mat: CONTACT_MATRIX ATOMS=1-10 SWITCH={RATIONAL R_0=1.0}
rsums: COLUMNSUMS MATRIX=mat MEAN
PRINT ARG=rsums.* FILE=colvar
```

The following input demonstrates another way that an average coordination number can be computed. This input calculates the number of atoms with indices between 1 and 5 that are within the first coordination spheres of each of the atoms within indices between 6 and 15. The average coordination number is then calculated from these fifteen coordination numbers and this quantity is output to a file.

```
mat2: CONTACT_MATRIX ATOMSA=1-5 ATOMSB=6-15 SWITCH={RATIONAL R_0=1.0}
rsums: COLUMNSUMS MATRIX=mat2 MEAN
PRINT ARG=rsums.* FILE=colvar
```

## 5.6.7 DFSCLUSTERING

	<b>This is part of the adjmat module</b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Find the connected components of the matrix using the depth first search clustering algorithm.

As discussed in the section of the manual on [Exploiting contact matrices](#) a useful tool for developing complex collective variables is the notion of the so called adjacency matrix. An adjacency matrix is an  $N \times N$  matrix in which the  $i$ th,  $j$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are adjacent or not. As detailed in [31] these matrices provide a representation of a graph and can thus be analysed using tools from graph theory. This particular action performs a depth first search clustering to find the connected components of this graph. You can read more about depth first search here:

[https://en.wikipedia.org/wiki/Depth-first\\_search](https://en.wikipedia.org/wiki/Depth-first_search)

This action is useful if you are looking at a phenomenon such as nucleation where the aim is to detect the sizes of the crystalline nuclei that have formed in your simulation cell.

#### Compulsory keywords

<b>MATRIX</b>	the action that calculates the adjacency matrix vessel we would like to analyse
<b>MAXCONNECT</b>	( default=0 ) maximum number of connections that can be formed by any given node in the graph. By default this is set equal to zero and the number of connections is set equal to the number of nodes. You only really need to set this if you are working with a very large system and memory is at a premium

#### Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

#### Examples

The input below calculates the coordination numbers of atoms 1-100 and then computes the an adjacency matrix whose elements measures whether atoms  $i$  and  $j$  are within 0.55 nm of each other. The action labelled `dfs` then treats the elements of this matrix as zero or ones and thus thinks of the matrix as defining a graph. This `dfs` action then finds the largest connected component in this graph. The sum of the coordination numbers for the atoms in this largest connected component are then computed and this quantity is output to a colvar file. The way this input can be used is described in detail in [31].

```
lq: COORDINATIONNUMBER SPECIES=1-100 SWITCH={CUBIC D_0=0.45 D_MAX=0.55} LOWMEM
cm: CONTACT_MATRIX ATOMS=lq SWITCH={CUBIC D_0=0.45 D_MAX=0.55}
dfs: DFSCUSTERING MATRIX=cm
clust1: CLUSTER_PROPERTIES CLUSTERS=dfs CLUSTER=1 SUM
PRINT ARG=clust1.* FILE=colvar
```

## 5.6.8 ROWSUMS

	<b>This is part of the adjmat <a href="#">module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Sum the rows of a adjacency matrix.

As discussed in the section of the manual on [Exploiting contact matrices](#) a useful tool for developing complex collective variables is the notion of the so called adjacency matrix. An adjacency matrix is an  $N \times N$  matrix in which the  $i$ th,  $j$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are adjacent or not. This action allows you to calculate the sum of the rows in this adjacency matrix and to then calculate further functions of these quantities.

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the <i>label.mean</i>
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.



## Compulsory keywords

<b>MAT</b> ↔ <b>RIX</b>	the action that calculates the adjacency matrix vessel we would like to analyse
----------------------------	---------------------------------------------------------------------------------

## Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BE↔TA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_T↔HAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THA↔N1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BET↔WEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...

<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .

### Examples

The first instruction in the following input file tells PLUMED to compute a  $10 \times 10$  matrix in which the  $ij$ -element tells you whether atoms  $i$  and  $j$  are within 1.0 nm of each other. The numbers in each of these rows are then added together and the average value is computed. As such the following input provides an alternative method for calculating the coordination numbers of atoms 1 to 10.

```
mat: CONTACT_MATRIX ATOMS=1-10 SWITCH={RATIONAL R_0=1.0}
rsums: ROWSUMS MATRIX=mat MEAN
PRINT ARG=rsums.* FILE=colvar
```

The following input demonstrates another way that an average coordination number can be computed. This input calculates the number of atoms with indices between 6 and 15 that are within the first coordination spheres of each of the atoms within indices between 1 and 5. The average coordination number is then calculated from these five coordination numbers and this quantity is output to a file.

```
mat2: CONTACT_MATRIX ATOMSA=1-5 ATOMSB=6-15 SWITCH={RATIONAL R_0=1.0}
rsums: ROWSUMS MATRIX=mat2 MEAN
PRINT ARG=rsums.* FILE=colvar
```

### 5.6.9 SPRINT

<b>This is part of the adjmat module</b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate SPRINT topological variables from an adjacency matrix.

The SPRINT topological variables are calculated from the largest eigenvalue,  $\lambda$  of an  $n \times n$  adjacency matrix and its corresponding eigenvector,  $\mathbf{V}$ , using:

$$s_i = \sqrt{n} \lambda v_i$$

You can use different quantities to measure whether or not two given atoms/molecules are adjacent or not in the adjacency matrix. The simplest measure of adjacency is whether two atoms/molecules are within some cutoff of each other. Further complexity can be added by insisting that two molecules are adjacent if they are within a certain distance of each other and if they have similar orientations.

## Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>coord</b>	all $n$ sprint coordinates are calculated and then stored in increasing order. the smallest sprint coordinate will be labelled <i>label.coord-1</i> , the second smallest will be labelled <i>label.coord-1</i> and so on

In addition the following quantities can be calculated by employing the keywords listed below

Quantity	Keyword	Description
<b>gradient</b>	<b>GRADIENT</b>	the gradient
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>vsum</b>	<b>VSUM</b>	the norm of sum of vectors. The output component can be referred to elsewhere in the input file by using the label.vsum
<b>spath</b>	<b>SPATH</b>	the position on the path
<b>zpath</b>	<b>ZPATH</b>	the distance from the path
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>sum</b>	<b>SUM</b>	the sum of values

## Compulsory keywords

<b>MAT↔ RIX</b>	the action that calculates the adjacency matrix vessel we would like to analyse
---------------------	---------------------------------------------------------------------------------

## Options

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

## Examples

This example input calculates the 7 SPRINT coordinates for a 7 atom cluster of Lennard-Jones atoms and prints their values to a file. In this input the SPRINT coordinates are calculated in the manner described in ?? so two atoms are adjacent if they are within a cutoff:

```
DENSITY SPECIES=1-7 LABEL=d1
CONTACT_MATRIX ATOMS=d1 SWITCH={RATIONAL R_0=0.1} LABEL=mat
SPRINT MATRIX=mat LABEL=ss
PRINT ARG=ss.* FILE=colvar
```

This example input calculates the 14 SPRINT coordinates for a molecule composed of 7 hydrogen and 7 carbon atoms. Once again two atoms are adjacent if they are within a cutoff:

```
DENSITY SPECIES=1-7 LABEL=c
DENSITY SPECIES=8-14 LABEL=h

CONTACT_MATRIX ...
  ATOMS=c,h
  SWITCH11={RATIONAL R_0=2.6 NN=6 MM=12}
  SWITCH12={RATIONAL R_0=2.2 NN=6 MM=12}
  SWITCH22={RATIONAL R_0=2.2 NN=6 MM=12}
  LABEL=mat
... CONTACT_MATRIX

SPRINT MATRIX=mat LABEL=ss

PRINT ARG=ss.* FILE=colvar
```

## 5.6.10 CLUSTER\_DIAMETER

<b>This is part of the adjmat <a href="#">module</a></b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Print out the diameter of one of the connected components

As discussed in the section of the manual on [Exploiting contact matrices](#) a useful tool for developing complex collective variables is the notion of the so called adjacency matrix. An adjacency matrix is an  $N \times N$  matrix in which the  $i$ th,  $j$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are adjacent or not. When analysing these matrix we can treat them as a graph and find connected components using some clustering algorithm. This action is used in tandem with this form of analysis to output the largest of

the distances between the pairs of atoms that are connected together in a particular connected component. It is important to note that the quantity that is output by this action is not differentiable. As such it cannot be used as a collective variable in a biased simulation.

#### Compulsory keywords

<b>CLUSTERS</b>	the label of the action that does the clustering
<b>CLUSTER</b>	( default=1 ) which cluster would you like to look at 1 is the largest cluster, 2 is the second largest, 3 is the the third largest and so on.

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation

#### Examples

The following input uses PLUMED to calculate a adjacency matrix that connects a pair of atoms if they both have a coordination number that is greater than 2.0 and if they are within 6.0 nm of each other. Depth first search clustering is used to find the connected components in this matrix. The distance between every pair of atoms that are within the largest of the clusters found is then calculated and the largest of these distances is output to a file named colvar.

```
# Calculate coordination numbers
c1: COORDINATIONNUMBER SPECIES=1-512 SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0}
# Select coordination numbers that are more than 2.0
cf: MFILTER_MORE DATA=c1 SWITCH={RATIONAL D_0=2.0 R_0=0.1} LOWMEM
# Build a contact matrix
mat: CONTACT_MATRIX ATOMS=cf SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0}
# Find largest cluster
dfs: DFSCLUSTERING MATRIX=mat
clust1: CLUSTER_PROPERTIES CLUSTERS=dfs CLUSTER=1
dia: CLUSTER_DIAMETER CLUSTERS=dfs CLUSTER=1
PRINT ARG=dia FILE=colvar
```

#### 5.6.11 CLUSTER\_DISTRIBUTION

<b>This is part of the adjmat <a href="#">module</a></b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate functions of the distribution of properties in your connected components.

This collective variable was developed for looking at nucleation phenomena, where you are interested in using studying the behavior of atoms in small aggregates or nuclei. In these sorts of problems you might be interested in the distribution of the sizes of the clusters in your system. A detailed description of this CV can be found in [31].

#### Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.less-than-1*, *label.less-than-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>less-than</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>more-than</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.

#### Compulsory keywords

<b>CLUSTERS</b>	the label of the action that does the clustering
<b>TRANSFORM</b>	( default=none ) the switching function to use to convert the crystallinity parameter to a number between zero and one

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
------------------------------	----------------------------------------------------------------------------

<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>INVERSE_TRANSFORM</b>	( default=off ) when TRANSFORM appears alone the input symmetry functions, $x$ are transformed used $1 - s(x)$ where $s(x)$ is a switching function. When this option is used you instead transform using $s(x)$ only.
<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.less than</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.less than-1</i> , <i>label.less than-2</i> , <i>label.less than-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ The value of $\beta$ in this function is specified using (BETA= $\beta$ ) The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...

## Examples

The input provided below calculates the local q6 Steinhardt parameter on each atom. The coordination number that atoms with a high value for the local q6 Steinhardt parameter have with other atoms that have a high value for the local q6 Steinhardt parameter is then computed. A contact matrix is then computed that measures whether atoms  $i$  and  $j$  have a high value for this coordination number and if they are within 3.6 nm of each other. The connected components of this matrix are then found using a depth first clustering algorithm on the corresponding graph. The number of components in this graph that contain more than 27 atoms is then computed. As discussed in [31] this input was used to analyse the formation of a polycrystal of GeTe from amorphous GeTe.

```
q6: Q6 SPECIES=1-32768 SWITCH={GAUSSIAN D_0=5.29 R_0=0.01 D_MAX=5.3} LOWMEM
lq6: LOCAL_Q6 SPECIES=q6 SWITCH={GAUSSIAN D_0=5.29 R_0=0.01 D_MAX=5.3} LOWMEM
flq6: MFILTER_MORE DATA=lq6 SWITCH={GAUSSIAN D_0=0.19 R_0=0.01 D_MAX=0.2}
cc: COORDINATIONNUMBER SPECIES=flq6 SWITCH={GAUSSIAN D_0=3.59 R_0=0.01 D_MAX=3.6}
fcc: MFILTER_MORE DATA=cc SWITCH={GAUSSIAN D_0=5.99 R_0=0.01 D_MAX=6.0}
mat: CONTACT_MATRIX ATOMS=fcc SWITCH={GAUSSIAN D_0=3.59 R_0=0.01 D_MAX=3.6}
dfs: DFSCUSTERING MATRIX=mat
nclust: CLUSTER_DISTRIBUTION CLUSTERS=dfs TRANSFORM={GAUSSIAN D_0=5.99 R_0=0.01 D_MAX=6.0} MORE_THAN={GAUSSIAN
PRINT ARG=nclust.* FILE=colvar
```

### 5.6.12 CLUSTER\_NATOMS

	<b>This is part of the adjmat module</b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code>. Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Gives the number of atoms in the connected component

As discussed in the section of the manual on [Exploiting contact matrices](#) a useful tool for developing complex collective variables is the notion of the so called adjacency matrix. An adjacency matrix is an  $N \times N$  matrix in which the  $i$ th,  $j$ th element tells you whether or not the  $i$ th and  $j$ th atoms/molecules from a set of  $N$  atoms/molecules are adjacent or not. When analysing these matrix we can treat them as a graph and find connected components using some clustering algorithm. This action is used in tandem with this form of analysis to output the number of atoms that are connected together in a particular connected component. It is important to note that the quantity that is output by this action is not differentiable. As such it cannot be used as a collective variable in a biased simulation.

#### Compulsory keywords

<b>CLUSTERS</b>	the label of the action that does the clustering
<b>CLUSTER</b>	( default=1 ) which cluster would you like to look at 1 is the largest cluster, 2 is the second largest, 3 is the the third largest and so on.

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation



## Examples

The following input uses PLUMED to calculate a adjacency matrix that connects a pair of atoms if they both have a coordination number that is greater than 2.0 and if they are within 6.0 nm of each other. Depth first search clustering is used to find the connected components in this matrix and then the number of atoms in the largest cluster is found. This quantity is then output to a file called colvar

```
# Calculate coordination numbers
cl: COORDINATIONNUMBER SPECIES=1-512 SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0}
# Select coordination numbers that are more than 2.0
cf: MFILTER_MORE DATA=cl SWITCH={RATIONAL D_0=2.0 R_0=0.1} LOWMEM
# Build a contact matrix
mat: CONTACT_MATRIX ATOMS=cf SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0}
# Find largest cluster
dfs: DFSCLUSTERING MATRIX=mat
clust1: CLUSTER_PROPERTIES CLUSTERS=dfs CLUSTER=1
nat: CLUSTER_NATOMS CLUSTERS=dfs CLUSTER=1
PRINT ARG=nat FILE=COLVAR
```

## 5.6.13 CLUSTER\_PROPERTIES

	<b>This is part of the <a href="#">adjmat module</a></b>
	<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Calculate properties of the distribution of some quantities that are part of a connected component

This collective variable was developed for looking at nucleation phenomena, where you are interested in using studying the behavior of atoms in small aggregates or nuclei. In these sorts of problems you might be interested in the degree the atoms in a nucleus have adopted their crystalline structure or (in the case of heterogenous nucleation of a solute from a solvent) you might be interested in how many atoms are present in the largest cluster [31].

## Description of components

When the label of this action is used as the input for a second you are not referring to a scalar quantity as you are in regular collective variables. The label is used to reference the full set of quantities calculated by the action. This is usual when using [MultiColvar functions](#). Generally when doing this the previously calculated multicolvar will be referenced using the DATA keyword rather than ARG.

This Action can be used to calculate the following scalar quantities directly. These quantities are calculated by employing the keywords listed below. These quantities can then be referenced elsewhere in the input file by using this Action's label followed by a dot and the name of the quantity. Some amongst them can be calculated multiple times with different parameters. In this case the quantities calculated can be referenced elsewhere in the input by using the name of the quantity followed by a numerical identifier e.g. *label.lessthan-1*, *label.lessthan-2* etc. When doing this and, for clarity we have made the label of the components customizable. As such by using the LABEL keyword in the description of the keyword input you can customize the component name

Quantity	Keyword	Description
<b>vmean</b>	<b>VMEAN</b>	the norm of the mean vector. The output component can be referred to elsewhere in the input file by using the label.vmean
<b>vsum</b>	<b>VSUM</b>	the norm of sum of vectors. The output component can be referred to elsewhere in the input file by using the label.vsum
<b>altmin</b>	<b>ALT_MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>between</b>	<b>BETWEEN</b>	the number/fraction of values within a certain range. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>highest</b>	<b>HIGHEST</b>	the lowest of the quantities calculated by this action
<b>lessthan</b>	<b>LESS_THAN</b>	the number of values less than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>lowest</b>	<b>LOWEST</b>	the lowest of the quantities calculated by this action
<b>max</b>	<b>MAX</b>	the maximum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>mean</b>	<b>MEAN</b>	the mean value. The output component can be referred to elsewhere in the input file by using the label.mean
<b>min</b>	<b>MIN</b>	the minimum value. This is calculated using the formula described in the description of the keyword so as to make it continuous.
<b>moment</b>	<b>MOMENTS</b>	the central moments of the distribution of values. The second moment would be referenced elsewhere in the input file using <i>label.moment-2</i> , the third as <i>label.moment-3</i> , etc.
<b>morethan</b>	<b>MORE_THAN</b>	the number of values more than a target value. This is calculated using one of the formula described in the description of the keyword so as to make it continuous. You can calculate this quantity multiple times using different parameters.
<b>sum</b>	<b>SUM</b>	the sum of values

## Compulsory keywords

<b>CLUSTERS</b>	the label of the action that does the clustering
<b>CLUSTER</b>	( default=1 ) which cluster would you like to look at 1 is the largest cluster, 2 is the second largest, 3 is the the third largest and so on.

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOPBC</b>	( default=off ) ignore the periodic boundary conditions when calculating distances
<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>MEAN</b>	take the mean of these variables. The final value can be referenced using <i>label.mean</i> . You can use multiple instances of this keyword i.e. MEAN1, MEAN2, MEAN3... The corresponding values are then referenced using <i>label.mean-1</i> , <i>label.mean-2</i> , <i>label.mean-3</i> ...

<b>MORE_THAN</b>	calculate the number of variables more than a certain target value. This quantity is calculated using $\sum_i 1.0 - \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.morethan</i> . You can use multiple instances of this keyword i.e. MORE_THAN1, MORE_THAN2, MORE_THAN3... The corresponding values are then referenced using <i>label.morethan-1</i> , <i>label.morethan-2</i> , <i>label.morethan-3</i> ...
<b>LESS_THAN</b>	calculate the number of variables less than a certain target value. This quantity is calculated using $\sum_i \sigma(s_i)$ , where $\sigma(s)$ is a <a href="#">switchingfunction</a> . The final value can be referenced using <i>label.lessthan</i> . You can use multiple instances of this keyword i.e. LESS_THAN1, LESS_THAN2, LESS_THAN3... The corresponding values are then referenced using <i>label.lessthan-1</i> , <i>label.lessthan-2</i> , <i>label.lessthan-3</i> ...
<b>VMEAN</b>	calculate the norm of the mean vector. The final value can be referenced using <i>label.vmean</i> . You can use multiple instances of this keyword i.e. VMEAN1, VMEAN2, VMEAN3... The corresponding values are then referenced using <i>label.vmean-1</i> , <i>label.vmean-2</i> , <i>label.vmean-3</i> ...
<b>VSUM</b>	calculate the norm of the sum of vectors. The final value can be referenced using <i>label.vsum</i> . You can use multiple instances of this keyword i.e. VSUM1, VSUM2, VSUM3... The corresponding values are then referenced using <i>label.vsum-1</i> , <i>label.vsum-2</i> , <i>label.vsum-3</i> ...
<b>BETWEEN</b>	calculate the number of values that are within a certain range. These quantities are calculated using kernel density estimation as described on <a href="#">histogrambead</a> . The final value can be referenced using <i>label.between</i> . You can use multiple instances of this keyword i.e. BETWEEN1, BETWEEN2, BETWEEN3... The corresponding values are then referenced using <i>label.between-1</i> , <i>label.between-2</i> , <i>label.between-3</i> ...
<b>HISTOGRAM</b>	calculate a discretized histogram of the distribution of values. This shortcut allows you to calculate NBIN quantities like BETWEEN. The final value can be referenced using <i>label.histogram</i> . You can use multiple instances of this keyword i.e. HISTOGRAM1, HISTOGRAM2, HISTOGRAM3... The corresponding values are then referenced using <i>label.histogram-1</i> , <i>label.histogram-2</i> , <i>label.histogram-3</i> ...
<b>MOMENTS</b>	calculate the moments of the distribution of collective variables. The $m$ th moment of a distribution is calculated using $\frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^m$ , where $\bar{s}$ is the average for the distribution. The moments keyword takes a list of integers as input or a range. Each integer is a value of $m$ . The final calculated values can be referenced using <i>moment-<math>m</math></i> .
<b>ALT_MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = -\frac{1}{\beta} \log \sum_i \exp(-\beta s_i)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.altmin</i> . You can use multiple instances of this keyword i.e. ALT_MIN1, ALT_MIN2, ALT_MIN3... The corresponding values are then referenced using <i>label.altmin-1</i> , <i>label.altmin-2</i> , <i>label.altmin-3</i> ...
<b>MIN</b>	calculate the minimum value. To make this quantity continuous the minimum is calculated using $\min = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{s_i}\right)}$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.min</i> . You can use multiple instances of this keyword i.e. MIN1, MIN2, MIN3... The corresponding values are then referenced using <i>label.min-1</i> , <i>label.min-2</i> , <i>label.min-3</i> ...
<b>MAX</b>	calculate the maximum value. To make this quantity continuous the maximum is calculated using $\max = \beta \log \sum_i \exp\left(\frac{s_i}{\beta}\right)$ . The value of $\beta$ in this function is specified using (BETA= $\beta$ ). The final value can be referenced using <i>label.max</i> . You can use multiple instances of this keyword i.e. MAX1, MAX2, MAX3... The corresponding values are then referenced using <i>label.max-1</i> , <i>label.max-2</i> , <i>label.max-3</i> ...

<b>SUM</b>	calculate the sum of all the quantities. The final value can be referenced using <i>label.sum</i> . You can use multiple instances of this keyword i.e. SUM1, SUM2, SUM3... The corresponding values are then referenced using <i>label.sum-1</i> , <i>label.sum-2</i> , <i>label.sum-3</i> ...
<b>LOWEST</b>	this flag allows you to recover the lowest of these variables. The final value can be referenced using <i>label.lowest</i>
<b>HIGHEST</b>	this flag allows you to recover the highest of these variables. The final value can be referenced using <i>label.highest</i>

### Examples

The input below calculates the coordination numbers of atoms 1-100 and then computes the an adjacency matrix whose elements measures whether atoms  $i$  and  $j$  are within 0.55 nm of each other. The action labelled dfs then treats the elements of this matrix as zero or ones and thus thinks of the matrix as defining a graph. This dfs action then finds the largest connected component in this graph. The sum of the coordination numbers for the atoms in this largest connected component are then computed and this quantity is output to a colvar file. The way this input can be used is described in detail in [31].

```
lq: COORDINATIONNUMBER SPECIES=1-100 SWITCH={CUBIC D_0=0.45 D_MAX=0.55} LOWMEM
cm: CONTACT_MATRIX ATOMS=lq SWITCH={CUBIC D_0=0.45 D_MAX=0.55}
dfs: DFSCUSTERING MATRIX=cm
clust1: CLUSTER_PROPERTIES CLUSTERS=dfs CLUSTER=1 SUM
PRINT ARG=clust1.* FILE=colvar
```

### 5.6.14 OUTPUT\_CLUSTER

<b>This is part of the adjmat module</b>
<b>It is only available if you configure PLUMED with <code>./configure --enable-modules=adjmat</code> . Furthermore, this feature is still being developed so take care when using it and report any problems on the mailing list.</b>

Output the indices of the atoms in one of the clusters identified by a clustering object

This action provides one way of getting output from a [DFSCUSTERING](#) calculation. The output in question here is either

- a file that contains a list of the atom indices that form part of one of the clusters that was identified using [DFSCUSTERING](#)
- an xyz file containing the positions of the atoms in one of the the clusters that was identified using [DFSCUSTERING](#)

Notice also that if you choose to output an xyz file you can ask PLUMED to try to reconstruct the cluster taking the periodic boundary conditions into account by using the `MAKE_WHOLE` flag.

### Compulsory keywords

<b>CLUSTERS</b>	the action that performed the clustering
<b>CLUSTER</b>	( default=1 ) which cluster would you like to look at 1 is the largest cluster, 2 is the second largest, 3 is the the third largest and so on
<b>STRIDE</b>	( default=1 ) the frequency with which you would like to output the atoms in the cluster
<b>FILE</b>	the name of the file on which to output the details of the cluster
<b>MAXDEPTH</b>	( default=6 ) maximum depth for searches over paths to reconstruct clusters for PBC
<b>MAXGOES</b>	( default=200 ) number of times to run searches to reconstuct clusters

#### Options

<b>MAKE_WHOLE</b>	( default=off ) reconstruct the clusters and remove all periodic boundary conditions.
-------------------	---------------------------------------------------------------------------------------

#### Examples

The input shown below identifies those atoms with a coordination number less than 13 and then constructs a contact matrix that describes the connectivity between the atoms that satisfy this criteria. The DFS algorithm is then used to find the connected components in this matrix and the indices of the atoms in the largest connected component are then output to a file.

```
c1: COORDINATIONNUMBER SPECIES=1-1996 SWITCH={CUBIC D_0=0.34 D_MAX=0.38}
cf: MFILTER_LESS DATA=c1 SWITCH={CUBIC D_0=13 D_MAX=13.5}
mat: CONTACT_MATRIX ATOMS=cf SWITCH={CUBIC D_0=0.34 D_MAX=0.38}
dfs: DFSCLUSTERING MATRIX=mat
OUTPUT_CLUSTER CLUSTERS=dfs CLUSTER=1 FILE=dfs.dat
```



## Chapter 6

# Analysis

PLUMED can be used to analyse trajectories either on the fly during an MD run or via postprocessing a trajectory using [driver](#). About the simplest form of analysis that PLUMED can perform involves printing information to a file. PLUMED can output various different kinds of information to files as described below:

<a href="#">COMMITTOR</a>	Does a committor analysis.
<a href="#">DUMPATOMS</a>	Dump selected atoms on a file.
<a href="#">DUMPDERIVATIVES</a>	Dump the derivatives with respect to the input parameters for one or more objects (generally CVs, functions or biases).
<a href="#">DUMPFORCES</a>	Dump the force acting on one of a values in a file.
<a href="#">DUMPMASSCHARGE</a>	Dump masses and charges on a selected file.
<a href="#">DUMPMULTICOLVAR</a>	Dump atom positions and multicolvar on a file.
<a href="#">DUMPPROJECTIONS</a>	Dump the derivatives with respect to the input parameters for one or more objects (generally CVs, functions or biases).
<a href="#">PRINT</a>	Print quantities to a file.
<a href="#">UPDATE_IF</a>	Conditional update of other actions.

The [UPDATE\\_IF](#) action allows you to do more complex things using the above print commands. As detailed in the documentation for [UPDATE\\_IF](#) when you put any of the above actions within an UPDATE\_IF block then data will only be output to the file if colvars are within particular ranges. In other words, the above printing commands, in tandem with [UPDATE\\_IF](#), allow you to identify the frames in your trajectory that satisfy some particular criteria and output information on those frames only.

Another useful command is the [COMMITTOR](#) command. This command can only be used when running an molecular dynamics trajectory - it cannot be used when analysing a trajectory using [driver](#). As detailed in the documentation for [COMMITTOR](#) this command tells PLUMED (and the underlying MD code) to stop the calculation one some criteria is satisfied.

A number of more complicated forms of analysis can be performed that take a number of frames from the trajectory as input. In all these commands the STRIDE keyword is used to tell PLUMED how frequently to collect data from the trajectory. In all these methods the output from the analysis is a form of enseble average. If you are running with a bias it is thus likely that you may want to reweight the trajectory frames in order to remove the effect the bias has on the static behavoir of the system. The following methods can thus be used to calculate weights for the various trajectory frames so that the final ensemble average is an average for the cannonical ensemble at the appropriate temperature.

<a href="#">REWEIGHT_BIAS</a>	Calculate weights for ensemble averages that negate the effect the bias has on the region of phase space explored
-------------------------------	-------------------------------------------------------------------------------------------------------------------

<a href="#">REWEIGHT_METAD</a>	Calculate the weights configurations should contribute to the histogram in a simulation in which a metadynamics bias acts upon the system.
<a href="#">REWEIGHT_TEMP</a>	Calculate weights for ensemble averages allow for the computing of ensemble averages at temperatures lower/higher than that used in your original simulation.

You can then calculate ensemble averages using the following actions.

<a href="#">AVERAGE</a>	Calculate the ensemble average of a collective variable
<a href="#">HISTOGRAM</a>	Accumulate the average probability density along a few CVs from a trajectory.
<a href="#">MULTICOLVARDENS</a>	Evaluate the average value of a multicolvar on a grid.

For many of the above commands data is accumulated on the grids. These grids can be further analysed using one of the actions detailed below at some time.

<a href="#">CONVERT_TO_FES</a>	Convert a histogram, $H(x)$ , to a free energy surface using $F(x) = -k_B T \ln H(x)$ .
<a href="#">DUMPCUBE</a>	Output a three dimensional grid using the Gaussian cube file format.
<a href="#">DUMPGRID</a>	Output the function on the grid to a file with the PLUMED grid format.
<a href="#">FIND_CONTOUR_SURFACE</a>	Find an isocontour by searching along either the x, y or z direction.
<a href="#">FIND_CONTOUR</a>	Find an isocontour in a smooth function.
<a href="#">FIND_SPHERICAL_CONTOUR</a>	Find an isocontour in a three dimensional grid by searching over a Fibonacci sphere.
<a href="#">FOURIER_TRANSFORM</a>	Compute the Discrete Fourier Transform (DFT) by means of FFTW of data stored on a 2D grid.
<a href="#">INTERPOLATE_GRID</a>	Interpolate a smooth function stored on a grid onto a grid with a smaller grid spacing.

As an example the following set of commands instructs PLUMED to calculate the distance between atoms 1 and 2 for every 5th frame in the trajectory and to accumulate a histogram from this data which will be output every 100 steps (i.e. when 20 distances have been added to the histogram).

```
x: DISTANCE ATOMS=1,2
h: HISTOGRAM ARG=x GRID_MIN=0.0 GRID_MAX=3.0 GRID_BIN=100 BANDWIDTH=0.1 STRIDE=5
DUMPGRID GRID=h FILE=histo STRIDE=100
```

It is important to note when using commands such as the above the first frame in the trajectory is assumed to be the initial configuration that was input to the MD code. It is thus ignored. Furthermore, if you are running with driver and you would like to analyse the whole trajectory (without specifying its length) and then print the result you simply call [DUMPGRID](#) (or any of the commands above) without a STRIDE keyword as shown in the example below.

```
x: DISTANCE ATOMS=1,2
h: HISTOGRAM ARG=x GRID_MIN=0.0 GRID_MAX=3.0 GRID_BIN=100 BANDWIDTH=0.1 STRIDE=5
DUMPGRID GRID=h FILE=histo
```

Please note that even with this calculation the first frame in the trajectory is ignored when computing the histogram.

Notice that all the commands for calculating smooth functions described above calculate some sort of average. There are two ways that you may wish to average the data in your trajectory:



- You might want to calculate block averages in which the first  $NN$  frames in your trajectory are averaged separately to the second block of  $N$  frames. If this is the case you should use the keyword `CLEAR` in the input to the action that calculates the smooth function. This keyword is used to specify how frequently you are clearing the stored data.
- You might want to calculate an accumulate an average over the whole trajectory and output the average accumulated at step  $N$ , step  $2N$ ... This is what PLUMED does by default so you do not need to use `CLEAR` in this case.

## 6.1 Dimensionality Reduction

The remainder of the analysis tools within PLUMED allow one to do some form of dimensionality reduction as detailed below.

<code>CLASSICAL_MDS</code>	Create a low-dimensional projection of a trajectory using the classical multidimensional-scaling algorithm.
<code>PCA</code>	Perform principal component analysis (PCA) using either the positions of the atoms a large number of collective variables as input.

As with the grids described previously the `STRIDE` keyword tells PLUMED how frequently to collect data from the trajectory. The `RUN` keyword then tells PLUMED how frequently to do the dimensionality reduction. As described above if `RUN` is not present and you are analysing trajectories using `driver` all the data in the trajectory (with the exception of the first frame) will be analysed.

## 6.2 COMMITTOR

	This is part of the analysis <code>module</code>
--	--------------------------------------------------

Does a committor analysis.

Compulsory keywords

<b>BASIN_LL</b>	List of lower limits for basin # You can use multiple instances of this keyword i.e. <code>BASIN_LL1</code> , <code>BASIN_LL2</code> , <code>BASIN_LL3</code> ...
<b>BASIN_UL</b>	List of upper limits for basin # You can use multiple instances of this keyword i.e. <code>BASIN_UL1</code> , <code>BASIN_UL2</code> , <code>BASIN_UL3</code> ...
<b>STRIDE</b>	( default=1 ) the frequency with which the CVs are analysed

Options

<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *. appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three components x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>FILE</b>	the name of the file on which to output the reached basin
<b>FMT</b>	the format that should be used to output real numbers

### Examples

The following input monitors two torsional angles during a simulation, defines two basins (A and B) as a function of the two torsions and stops the simulation when it falls in one of the two. In the log file will be shown the latest values for the CVs and the basin reached.

```
TORSION ATOMS=1,2,3,4 LABEL=r1
TORSION ATOMS=2,3,4,5 LABEL=r2
COMMITTOR ...
  ARG=r1,r2
  STRIDE=10
  BASIN_LL1=0.15,0.20
  BASIN_UL1=0.25,0.40
  BASIN_LL2=-0.25,-0.40
  BASIN_UL2=-0.15,-0.20
... COMMITTOR
```

## 6.3 DUMPATOMS

This is part of the generic <a href="#">module</a>
----------------------------------------------------

Dump selected atoms on a file.

This command can be used to output the positions of a particular set of atoms. The atoms required are output in a xyz or gro formatted file. If PLUMED has been compiled with xdrfile support, then also xtc and trr files can be written. To this aim one should install xdrfile library ([http://www.gromacs.org/Developer\\_Zone/Programming\\_Guide/XTC\\_Library](http://www.gromacs.org/Developer_Zone/Programming_Guide/XTC_Library)). If the xdrfile library is installed properly the PLUMED configure script should be able to detect it and enable it. The type of file is automatically detected from the file extension, but can be also enforced with TYPE. Importantly, if your input file contains actions that edit the atoms position (e.g. [WHOLEMOLECULES](#)) and the DUMPATOMS command appears after this instruction, then the edited atom positions are output. You can control the buffering of output using the [FLUSH](#) keyword on a separate line.

Units of the printed file can be controlled with the UNITS keyword. By default PLUMED units as controlled in the [UNITS](#) command are used, but one can override it e.g. with UNITS=A. Notice that gro/xtc/trr files can only contain coordinates in nm.

The atoms involved can be specified using

<b>ATOMS</b>	the atom indices whose positions you would like to print out. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the atoms should be output
<b>FILE</b>	file on which to output coordinates; extension is automatically detected
<b>UNITS</b>	( default=PLUMED ) the units in which to print out the coordinates. PLUMED means internal PLUMED units

#### Options

<b>PRECISION</b>	The number of digits in trajectory file
<b>TYPE</b>	file type, either xyz, gro, xtc, or trr, can override an automatically detected file extension
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UNTIL</b>	Only update this action until this time

#### Examples

The following input instructs plumed to print out the positions of atoms 1-10 together with the position of the center of mass of atoms 11-20 every 10 steps to a file called file.xyz.

```
COM ATOMS=11-20 LABEL=c1
DUMPATOMS STRIDE=10 FILE=file.xyz ATOMS=1-10,c1
```

(see also [COM](#))

The following input is very similar but dumps a .gro (gromacs) file, which also contains atom and residue names.

```
# this is required to have proper atom names:
MOLINFO STRUCTURE=reference.pdb
# if omitted, atoms will have "X" name...

COM ATOMS=11-20 LABEL=c1
DUMPATOMS STRIDE=10 FILE=file.gro ATOMS=1-10,c1
# notice that last atom is a virtual one and will not have
# a correct name in the resulting gro file
```

(see also [COM](#) and [MOLINFO](#))

## 6.4 DUMPDERIVATIVES

	<b>This is part of the generic <a href="#">module</a></b>
--	-----------------------------------------------------------

Dump the derivatives with respect to the input parameters for one or more objects (generally CVs, functions or biases).

For a CV this line in input instructs plumed to print the derivative of the CV with respect to the atom positions and the cell vectors (virial-like form). In contrast, for a function or bias the derivative with respect to the input "CVs" will be output. This command is most often used to test whether or not analytic derivatives have been implemented correctly. This can be done by outputting the derivatives calculated analytically and numerically. You can control the buffering of output using the [FLUSH](#) keyword.

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the derivatives should be output
<b>FILE</b>	the name of the file on which to output the derivatives
<b>FMT</b>	( default=%15.10f ) the format with which the derivatives should be output

#### Options

<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the <a href="#">PLUMED Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. A↔RG1, ARG2, ARG3...
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UN↔TIL</b>	Only update this action until this time

#### Examples

The following input instructs plumed to write a file called deriv that contains both the analytical and numerical derivatives of the distance between atoms 1 and 2.

```
DISTANCE ATOM=1,2 LABEL=distance
DISTANCE ATOM=1,2 LABEL=distanceN NUMERICAL_DERIVATIVES
DUMPPERIVATIVES ARG=distance,distanceN STRIDE=1 FILE=deriv
```

(See also [DISTANCE](#))

## 6.5 DUMPFORCES

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

Dump the force acting on one of a values in a file.

For a CV this command will dump the force on the CV itself. Be aware that in order to have the forces on the atoms you should multiply the output from this argument by the output from DUMPDERIVATIVES. Furthermore, also note that you can output the forces on multiple quantities simultaneously by specifying more than one argument. You can control the buffering of output using the [FLUSH](#) keyword.

### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the forces should be output
<b>FILE</b>	the name of the file on which to output the forces
<b>FMT</b>	( default=%15.10f ) the format with which the derivatives should be output

### Options

<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the <a href="#">PLUMED Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. A↵RG1, ARG2, ARG3...
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UN↵TIL</b>	Only update this action until this time

### Examples

The following input instructs plumed to write a file called forces that contains the force acting on the distance between atoms 1 and 2.

```
DISTANCE ATOM=1,2 LABEL=distance
DUMPFORCES ARG=distance STRIDE=1 FILE=forces
```

(See also [DISTANCE](#))

## 6.6 DUMPMASSCHARGE

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

Dump masses and charges on a selected file.

This command dumps a file containing charges and masses. It does so only once in the simulation (at first step). File can be recycled in the [driver](#) tool.

Notice that masses and charges are only written once at the beginning of the simulation. In case no atom list is provided, charges and masses for all atoms are written.

The atoms involved can be specified using

<b>ATOMS</b>	the atom indices whose positions you would like to print out. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the atoms should be output
<b>FILE</b>	file on which to output coordinates. .gro extension is automatically detected

Examples

You can add the DUMPMASSCHARGE action at the end of the plumed.dat file that you use during an MD simulations:

```
c1: COM ATOMS=1-10
c2: COM ATOMS=11-20
PRINT ARG=c1,c2 FILE=colvar STRIDE=100

DUMPMASSCHARGE FILE=mcfile
```

(see also [COM](#) and [PRINT](#))

In this way, you will be able to use the same masses while processing a trajectory from the [driver](#) . To do so, you need to add the `--mc` flag on the driver command line, e.g.

```
plumed driver --mc mcfile --plumed plumed.dat --ixyz traj.xyz
```

With the following input you can dump only the charges for a specific group.

```
solute_ions: GROUP ATOMS=1-121,200-2012
DUMPATOMS FILE=traj.gro ATOMS=solute_ions STRIDE=100
DUMPMASSCHARGE FILE=mcfile ATOMS=solute_ions
```

Notice however that if you want to process the charges with the driver (e.g. reading traj.gro) you have to fix atom numbers first, e.g. with the script

```
awk 'BEGIN{c=0}{
    if(match($0,"#")) print ; else {print c,$2,$3; c++}
}' < mc > newmc
}'
```

then

```
plumed driver --mc newmc --plumed plumed.dat --ixyz traj.gro
```

## 6.7 DUMPMULTICOLVAR

This is part of the multicolvar <a href="#">module</a>
--------------------------------------------------------

Dump atom positions and multicolvar on a file.

The atoms involved can be specified using

<b>ORIGIN</b>	You can use this keyword to specify the position of an atom as an origin. The positions output will then be displayed relative to that origin. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
---------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>DATA</b>	certain actions in plumed work by calculating a list of variables and summing over them. This particular action can be used to calculate functions of these base variables or prints them to a file. This keyword thus takes the label of one of those such variables as input.
<b>STRIDE</b>	( default=1 ) the frequency with which the atoms should be output
<b>FILE</b>	file on which to output coordinates
<b>UNITS</b>	( default=PLUMED ) the units in which to print out the coordinates. PLUMED means internal PLUMED units

Options

<b>PRECISION</b>	The number of digits in trajectory file
------------------	-----------------------------------------

Examples

In this examples we calculate the distances between the atoms of the first and the second group and we write them in the file MULTICOLVAR.xyz. For each couple it writes the coordinates of their geometric center and their distance.

```
pos:  GROUP ATOMS=220,221,235,236,247,248,438,439,450,451,534,535
neg:  GROUP ATOMS=65,68,138,182,185,267,270,291,313,316,489,583,621,711
DISTANCES GROUPA=pos GROUPB=neg LABEL=slt

DUMPMULTICOLVAR DATA=slt FILE=MULTICOLVAR.xyz
```

(see also [DISTANCES](#))

## 6.8 DUMPPROJECTIONS

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

Dump the derivatives with respect to the input parameters for one or more objects (generally CVs, functions or biases).

### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the derivatives should be output
<b>FILE</b>	the name of the file on which to output the derivatives
<b>FMT</b>	( default=%15.10f ) the format with which the derivatives should be output

### Options

<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the <a href="#">PLUMED Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. A↵RG1, ARG2, ARG3...
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UN↵TIL</b>	Only update this action until this time

### Examples

Compute the distance between two groups and write on a file the derivatives of this distance with respect to all the atoms of the two groups



```
x1: CENTER ATOMS=1-10
x2: CENTER ATOMS=11-20
d: DISTANCE ATOMS=x1,x2
DUMPPROJECTIONS ARG=d FILE=proj STRIDE=20
```

## 6.9 PRINT

This is part of the generic <a href="#">module</a>
----------------------------------------------------

Print quantities to a file.

This directive can be used multiple times in the input so you can print files with different strides or print different quantities to different files. You can control the buffering of output using the [FLUSH](#) keyword.

### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the quantities of interest should be output
---------------	------------------------------------------------------------------------------------

### Options

<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the <a href="#">PLUMED Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. A↵RG1, ARG2, ARG3...
<b>FILE</b>	the name of the file on which to output these quantities
<b>FMT</b>	the format that should be used to output real numbers
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UN↵TIL</b>	Only update this action until this time

### Examples

The following input instructs plumed to print the distance between atoms 3 and 5 on a file called COLVAR every 10 steps, and the distance and total energy on a file called COLVAR\_ALL every 1000 steps.

```
DISTANCE ATOMS=2,5 LABEL=distance
ENERGY LABEL=energy
PRINT ARG=distance STRIDE=10 FILE=COLVAR
PRINT ARG=distance,energy STRIDE=1000 FILE=COLVAR_ALL
```

(See also [DISTANCE](#) and [ENERGY](#)).

### 6.9.1 FLUSH

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

This command instructs plumed to flush all the open files with a user specified frequency. Notice that all files are flushed anyway every 10000 steps.

This is useful for preventing data loss that would otherwise arise as a consequence of the code storing data for printing in the buffers. Notice that wherever it is written in the plumed input file, it will flush all the open files.

#### Compulsory keywords

<b>STRIDE</b>	the frequency with which all the open files should be flushed
---------------	---------------------------------------------------------------

#### Examples

A command like this in the input will instruct plumed to flush all the output files every 100 steps

```
d1: DISTANCE ATOMS=1,10
PRINT ARG=d1 STRIDE=5 FILE=colvar1

FLUSH STRIDE=100

d2: DISTANCE ATOMS=2,11
# also this print is flushed every 100 steps:
PRINT ARG=d2 STRIDE=10 FILE=colvar2
```

(see also [DISTANCE](#) and [PRINT](#)).

### 6.10 UPDATE\_IF

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

Conditional update of other actions.

This action can be used to enable and disable the update step for the following actions depending on the value of its arguments. This allows for example to extract snapshots with value of some CVs in a given range.

When called with MORE\_THAN and/or LESS\_THAN keywords, this action starts an if block. The block is executed if all the arguments are less than all the respective values in the LESS\_THAN keyword (if present) and all the arguments are more than all the respective values in the MORE\_THAN keyword (if present).

When called with the END flag, this action ends the corresponding IF block. Notice that in this case one should also provide the ARG keyword. It is recommended to use the same ARG keyword that was used to begin the block, so as to make the input more readable.

Of course, blocks can be nested at will.

There are many potential usages for this keyword. One might e.g. decide to analyze some variable only when another variable is within a given range.

#### Warning

Notice that not all the possible usage make particular sense. For example, conditionally updating a [METAD](#) keyword (that is: adding hills only if a variable is within a given range) can lead to unexpected results.

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the quantities of interest should be output
---------------	------------------------------------------------------------------------------------

#### Options

<b>END</b>	( default=off ) end
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>LESS_THAN</b>	upper bound
<b>MORE_THAN</b>	lower bound

#### Examples

The following input instructs plumed dump all the snapshots where an atom is in touch with the solute.

```
solute: GROUP ATOMS=1-124
coord: COORDINATION GROUPA=solute GROUPB=500 R_0=0.5
```

```
# A coordination number higher than 0.5 indicate that there is at least one
# atom of group 'solute' at less than 5 A from atom number 500

UPDATE_IF ARG=coord MORE_THAN=0.5
DUMPATOMS ATOMS=solute,500 FILE=output.xyz
UPDATE_IF ARG=coord END
```

(See also [GROUP](#), [COORDINATION](#), and [DUMPATOMS](#))

## 6.11 REWEIGHT\_BIAS

This is part of the <a href="#">bias module</a>
-------------------------------------------------

Calculate weights for ensemble averages that negate the effect the bias has on the region of phase space explored

If a static or pseudo-static bias  $V(x, t')$  is acting on the system we can remove the bias and get the unbiased probability distribution using:

$$\langle P(s', t) \rangle = \frac{\sum_{t'}^t \delta(s(x) - s') \exp\left(+\frac{V(x, t')}{k_B T}\right)}{\sum_t^t \exp\left(+\frac{V(x, t')}{k_B T}\right)}$$

The weights calculated by this action are equal to  $\exp\left(+\frac{V(x, t')}{k_B T}\right)$  these weights can then be used in any action that computes ensemble averages. For example this action can be used in tandem with [HISTOGRAM](#) or [AVERAGE](#).

### Compulsory keywords

<b>ARG</b>	( default=*bias ) the biases that must be taken into account when reweighting
------------	-------------------------------------------------------------------------------

### Options

<b>TEMP</b>	the system temperature. This is not required if your MD code passes this quantity to PLUMED
-------------	---------------------------------------------------------------------------------------------

### Examples

In the following example there is a fixed restraint on the distance between atoms 1 and 2. Clearly, this restraint will have an effect on the region of phase space that will be sampled when an MD simulation is run using this variable. Consequently, when the histogram as a function of the distance,  $x$ , is accumulated, we use reweighting into order to discount the effect of the bias from our final histogram.

```
x: DISTANCE ATOMS=1,2
RESTRAINT ARG=x SLOPE=1.0 AT=0.0
```

```

bias: REWEIGHT_BIAS TEMP=300

HISTOGRAM ...
  ARG=x
  GRID_MIN=0.0
  GRID_MAX=3.0
  GRID_BIN=100
  BANDWIDTH=0.1
  LOGWEIGHTS=bias
  LABEL=hB
... HISTOGRAM

DUMPGRID GRID=hB FILE=histoB STRIDE=1 FMT=%8.4f

```

## 6.12 REWEIGHT\_METAD

This is part of the <a href="#">bias module</a>
-------------------------------------------------

Calculate the weights configurations should contribute to the histogram in a simulation in which a metadynamics bias acts upon the system.

This command allows you to use the reweighting algorithm discussed in [33] when constructing a histogram of the configurations visited during a metadynamics simulation.

### Compulsory keywords

<b>ARG</b>	( default=*.rbias ) the biases that must be taken into account when reweighting
------------	---------------------------------------------------------------------------------

### Options

<b>TEMP</b>	the system temperature. This is not required if your MD code passes this quantity to PLUMED
-------------	---------------------------------------------------------------------------------------------

### Examples

In the following example there is a metadynamics bias acting on the distance between atoms 1 and 2. Clearly, this bias will have an effect on the region of phase space that will be sampled when an MD simulation is run using this variable. Consequently, when the histogram as a function of the angle,  $\alpha$ , is accumulated, we use reweighting into order to discount the effect of the bias from our final histogram. We do not use [REWEIGHT\\_BIAS](#) here, however, as the bias changes with time. We thus use the reweighting algorithm for metadynamics instead. Notice also that we have to specify how often we would like to calculate the  $c(t)$  reweighting factor and the grid over which we calculate  $c(t)$  in the input to the METAD command.

```

a: ANGLE ATOMS=1,2,3
x: DISTANCE ATOMS=1,2
METAD ARG=x PACE=100 SIGMA=0.1 HEIGHT=1.5 BIASFACTOR=5 GRID_MIN=0 GRID_MAX=10 GRID_BIN=100 REWEIGHTING_NGRID=1

bias: REWEIGHT_METAD TEMP=300

```

```

HISTOGRAM ...
  ARG=a
  GRID_MIN=0.0
  GRID_MAX=pi
  GRID_BIN=100
  BANDWIDTH=0.1
  LOGWEIGHTS=bias
  LABEL=hB
... HISTOGRAM

DUMPGRID GRID=hB FILE=histoB STRIDE=1 FMT=%8.4f

```

## 6.13 REWEIGHT\_TEMP

This is part of the <a href="#">bias module</a>
-------------------------------------------------

Calculate weights for ensemble averages allow for the computing of ensemble averages at temperatures lower/higher than that used in your original simulation.

We can use our knowledge of the Boltzmann distribution in the canonical ensemble to reweight the data contained in trajectories. Using this procedure we can take trajectory at temperature  $T_1$  and use it to extract probabilities at a different temperature,  $T_2$ , using:

$$P(s', t) = \frac{\sum_{t'}^t \delta(s(x) - s') \exp \left( + \left( \frac{1}{T_1} - \frac{1}{T_2} \right) \frac{U(x, t')}{k_B} \right)}{\sum_{t'}^t \exp \left( + \left[ \frac{1}{T_1} - \frac{1}{T_2} \right] \frac{U(x, t')}{k_B} \right)}$$

The weights calculated by this action are equal to  $\exp \left( + \left( \frac{1}{T_1} - \frac{1}{T_2} \right) \frac{U(x, t')}{k_B} \right)$  and take the effect the bias has on the system into account. These weights can be used in any action that computes ensemble averages. For example this action can be used in tandem with [HISTOGRAM](#) or [AVERAGE](#).

### Compulsory keywords

<b>REWEIGHT_TEMP</b>	reweight data from a trajectory at one temperature and output the probability distribution at a second temperature. This is not possible during postprocessing.
----------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------

### Options

<b>TEMP</b>	the system temperature. This is not required if your MD code passes this quantity to PLUMED
-------------	---------------------------------------------------------------------------------------------

### Examples

The following input can be used to postprocess a molecular dynamics trajectory calculated at a temperature of 500 K. The [HISTOGRAM](#) as a function of the distance between atoms 1 and 2 that would have been obtained if the

simulation had been run at the lower temperature of 300 K is estimated using the data from the higher temperature trajectory and output to a file.

```
x: DISTANCE ATOMS=1,2
aa: REWEIGHT_TEMP TEMP=500 REWEIGHT_TEMP=300
hB: HISTOGRAM ARG=x GRID_MIN=0.0 GRID_MAX=3.0 GRID_BIN=100 BANDWIDTH=0.1 LOGWEIGHTS=aa
DUMPGRID GRID=hB FILE=histoB
```

## 6.14 AVERAGE

This is part of the analysis <a href="#">module</a>
-----------------------------------------------------

Calculate the ensemble average of a collective variable

The ensemble average for a non-periodic, collective variable,  $s$  is given by the following expression:

$$\langle s \rangle = \frac{\sum_{t'=0}^t w(t') s(t')}{\sum_{t'=0}^t w(t')}$$

Here the sum runs over a the trajectory and  $s(t')$  is used to denote the value of the collective variable at time  $t'$ . The final quantity evaluated is a weighted average as the weights,  $w(t')$ , allow us to negate the effect any bias might have on the region of phase space sampled by the system. This is discussed in the section of the manual on [Analysis](#).

When the variable is periodic (e.g. [TORSION](#)) and has a value,  $s$ , in  $a \leq s \leq b$  the ensemble average is evaluated using:

$$\langle s \rangle = a + \frac{b-a}{2\pi} \arctan \left[ \frac{\sum_{t'=0}^t w(t') \sin \left( \frac{2\pi[s(t')-a]}{b-a} \right)}{\sum_{t'=0}^t w(t') \cos \left( \frac{2\pi[s(t')-a]}{b-a} \right)} \right]$$

### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>CLEAR</b>	( default=0 ) the frequency with which to clear all the accumulated data. The default value of 0 implies that all the data will be used and that the grid will never be cleared

### Options

<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages

<b>ARG</b>	<p>the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three components x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the <a href="#">PLUMED Getting started</a>. Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a>. To use this feature you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...</p>
------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Examples

The following example calculates the ensemble average for the distance between atoms 1 and 2 and output this to a file called COLVAR. In this example it is assumed that no bias is acting on the system and that the weights,  $w(t')$  in the formulae above can thus all be set equal to one.

```
d1: DISTANCE ATOMS=1,2
dla: AVERAGE ARG=d1
PRINT ARG=dla FILE=colvar STRIDE=100
```

The following example calculates the ensemble average for the torsional angle involving atoms 1, 2, 3 and 4. At variance with the previous example this quantity is periodic so the second formula in the above introduction is used to calculate the average. Furthermore, by using the CLEAR keyword we have specified that block averages are to be calculated. Consequently, after 100 steps all the information acquired thus far in the simulation is forgotten and the process of averaging is begun again. The quantities output in the colvar file are thus the block averages taken over the first 100 frames of the trajectory, the block average over the second 100 frames of trajectory and so on.

```
t1: TORSION ATOMS=1,2,3,4
t1a: AVERAGE ARG=t1 CLEAR=100
PRINT ARG=t1a FILE=colvar STRIDE=100
```

This third example incorporates a bias. Notice that the effect the bias has on the ensemble average is removed by taking advantage of the [REWEIGHT\\_BIAS](#) method. The final ensemble averages output to the file are thus block ensemble averages for the unbiased canonical ensemble at a temperature of 300 K.

```
t1: TORSION ATOMS=1,2,3,4
RESTRAINT ARG=t1 AT=pi KAPPA=100.
ww: REWEIGHT_BIAS TEMP=300
t1a: AVERAGE ARG=t1 LOGWEIGHTS=ww CLEAR=100
PRINT ARG=t1a FILE=colvar STRIDE=100
```

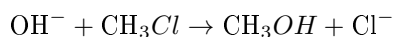
## 6.15 HISTOGRAM

	This is part of the analysis <a href="#">module</a>
--	-----------------------------------------------------



Accumulate the average probability density along a few CVs from a trajectory.

When using this method it is supposed that you have some collective variable  $\zeta$  that gives a reasonable description of some physical or chemical phenomenon. As an example of what we mean by this suppose you wish to examine the following SN2 reaction:



The distance between the chlorine atom and the carbon is an excellent collective variable,  $\zeta$ , in this case because this distance is short for the reactant,  $\text{CH}_3\text{Cl}$ , because the carbon and chlorine are chemically bonded, and because it is long for the product state when these two atoms are not chemically bonded. We thus might want to accumulate the probability density,  $P(\zeta)$ , as a function of this distance as this will provide us with information about the overall likelihood of the reaction. Furthermore, the free energy,  $F(\zeta)$ , is related to this probability density via:

$$F(\zeta) = -k_B T \ln P(\zeta)$$

Accumulating these probability densities is precisely what this Action can be used to do. Furthermore, the conversion of the histogram to the free energy can be achieved by using the method [CONVERT\\_TO\\_FES](#).

We calculate histograms within PLUMED using a method known as kernel density estimation, which you can read more about here:

[https://en.wikipedia.org/wiki/Kernel\\_density\\_estimation](https://en.wikipedia.org/wiki/Kernel_density_estimation)

In PLUMED the value of  $\zeta$  at each discrete instant in time in the trajectory is accumulated. A kernel,  $K(\zeta - \zeta(t'), \sigma)$ , centered at the current value,  $\zeta(t)$ , of this quantity is generated with a bandwidth  $\sigma$ , which is set by the user. These kernels are then used to accumulate the ensemble average for the probability density:

$$\langle P(\zeta) \rangle = \frac{\sum_{t'=0}^t w(t') K(\zeta - \zeta(t'), \sigma)}{\sum_{t'=0}^t w(t')}$$

Here the sums run over a portion of the trajectory specified by the user. The final quantity evaluated is a weighted average as the weights,  $w(t')$ , allow us to negate the effect any bias might have on the region of phase space sampled by the system. This is discussed in the section of the manual on [Analysis](#).

A discrete analogue of kernel density estimation can also be used. In this analogue the kernels in the above formula are replaced by dirac delta functions. When this method is used the final function calculated is no longer a probability density - it is instead a probability mass function as each element of the function tells you the value of an integral between two points on your grid rather than the value of a (continuous) function on a grid.

Additional material and examples can be also found in the tutorial [Belfast tutorial: Analyzing CVs](#).

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>CLEAR</b>	( default=0 ) the frequency with which to clear all the accumulated data. The default value of 0 implies that all the data will be used and that the grid will never be cleared
<b>BANDWIDTH</b>	the bandwidths for kernel density estimation
<b>KERNEL</b>	( default=gaussian ) the kernel function you are using. More details on the kernels available in plumed plumed can be found in <a href="#">kernelfunctions</a> .
<b>GRID_MIN</b>	the lower bounds for the grid
<b>GRID_MAX</b>	the upper bounds for the grid

## Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the <a href="#">PLUMED Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. A↵RG1, ARG2, ARG3...
<b>DATA</b>	input data from action with vessel and compute histogram
<b>GRID_BIN</b>	the number of bins for the grid
<b>GRID_SPACING</b>	the approximate grid spacing (to be used as an alternative or together with GRID_BIN)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UN↵TIL</b>	Only update this action until this time

## Examples

The following input monitors two torsional angles during a simulation and outputs a continuos histogram as a function of them at the end of the simulation.

```
TORSION ATOMS=1,2,3,4 LABEL=r1
TORSION ATOMS=2,3,4,5 LABEL=r2
HISTOGRAM ...
  ARG=r1,r2
  GRID_MIN=-3.14,-3.14
  GRID_MAX=3.14,3.14
  GRID_BIN=200,200
  BANDWIDTH=0.05,0.05
  LABEL=hh
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo
```

The following input monitors two torsional angles during a simulation and outputs a discrete histogram as a function of them at the end of the simulation.

```
TORSION ATOMS=1,2,3,4 LABEL=r1
TORSION ATOMS=2,3,4,5 LABEL=r2
HISTOGRAM ...
  ARG=r1,r2
  KERNEL=DISCRETE
```

```

GRID_MIN=-3.14,-3.14
GRID_MAX=3.14,3.14
GRID_BIN=200,200
LABEL=hh
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo

```

The following input monitors two torsional angles during a simulation and outputs the histogram accumulated thus far every 100000 steps.

```

TORSION ATOMS=1,2,3,4 LABEL=r1
TORSION ATOMS=2,3,4,5 LABEL=r2
HISTOGRAM ...
  ARG=r1,r2
  GRID_MIN=-3.14,-3.14
  GRID_MAX=3.14,3.14
  GRID_BIN=200,200
  BANDWIDTH=0.05,0.05
  LABEL=hh
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo STRIDE=100000

```

The following input monitors two torsional angles during a simulation and outputs a separate histogram for each 100000 steps worth of trajectory. Notice how the CLEAR keyword is used here and how it is not used in the previous example.

```

TORSION ATOMS=1,2,3,4 LABEL=r1
TORSION ATOMS=2,3,4,5 LABEL=r2
HISTOGRAM ...
  ARG=r1,r2 CLEAR=100000
  GRID_MIN=-3.14,-3.14
  GRID_MAX=3.14,3.14
  GRID_BIN=200,200
  BANDWIDTH=0.05,0.05
  GRID_WFILE=histo
  LABEL=hh
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo STRIDE=100000

```

## 6.16 MULTICOLVARDENS

This is part of the <a href="#">multicolvar module</a>
--------------------------------------------------------

Evaluate the average value of a multicolvar on a grid.

This keyword allows one to construct a phase field representation for a symmetry function from an atomistic description. If each atom has an associated order parameter,  $\phi_i$  then a smooth phase field function  $\phi(r)$  can be computed using:

$$\phi(\mathbf{r}) = \frac{\sum_i K(\mathbf{r} - \mathbf{r}_i) \phi_i}{\sum_i K(\mathbf{r} - \mathbf{r}_i)}$$

where  $\mathbf{r}_i$  is the position of atom  $i$ , the sums run over all the atoms input and  $K(\mathbf{r} - \mathbf{r}_i)$  is one of the [kernel functions](#) implemented in plumed. This action calculates the above function on a grid, which can then be used in the input to further actions.

The atoms involved can be specified using

<b>ORIGIN</b>	we will use the position of this atom as the origin. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
---------------	---------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>CLEAR</b>	( default=0 ) the frequency with which to clear all the accumulated data. The default value of 0 implies that all the data will be used and that the grid will never be cleared
<b>BANDWIDTH</b>	the bandwidths for kernel density estimation
<b>KERNEL</b>	( default=gaussian ) the kernel function you are using. More details on the kernels available in plumed can be found in <a href="#">kernelfunctions</a> .
<b>DATA</b>	the multicolvar which you would like to calculate the density profile for
<b>DIR</b>	the direction in which to calculate the density profile

Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>FRACTIONAL</b>	( default=off ) use fractional coordinates for the various axes
<b>XREDUCED</b>	( default=off ) limit the calculation of the density/average to a portion of the x-axis only
<b>YREDUCED</b>	( default=off ) limit the calculation of the density/average to a portion of the y-axis only
<b>ZREDUCED</b>	( default=off ) limit the calculation of the density/average to a portion of the z-axis only
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages
<b>NBINS</b>	the number of bins to use to represent the density profile
<b>SPACING</b>	the approximate grid spacing (to be used as an alternative or together with NBINS)
<b>XLOWER</b>	this is required if you are using XREDUCED. It specifies the lower bound for the region of the x-axis that for which you are calculating the density/average
<b>XUPPER</b>	this is required if you are using XREDUCED. It specifies the upper bound for the region of the x-axis that for which you are calculating the density/average
<b>YLOWER</b>	this is required if you are using YREDUCED. It specifies the lower bound for the region of the y-axis that for which you are calculating the density/average
<b>YUPPER</b>	this is required if you are using YREDUCED. It specifies the upper bound for the region of the y-axis that for which you are calculating the density/average
<b>ZLOWER</b>	this is required if you are using ZREDUCED. It specifies the lower bound for the region of the z-axis that for which you are calculating the density/average
<b>ZUPPER</b>	this is required if you are using ZREDUCED. It specifies the upper bound for the region of the z-axis that for which you are calculating the density/average

## Examples

The following example shows perhaps the simplest way in which this action can be used. The following input computes the density of atoms at each point on the grid and outputs this quantity to a file. In other words this input instructs PLUMED to calculate  $\rho(\mathbf{r}) = \sum_i K(\mathbf{r} - \mathbf{r}_i)$

```
dens: DENSITY SPECIES=1-100
grid: MULTICOLVARDENS DATA=dens ORIGIN=1 DIR=xyz NBINS=100,100,100 BANDWIDTH=0.05,0.05,0.05 STRIDE=1
DUMPGRID GRID=grid STRIDE=500 FILE=density
```

In the above example density is added to the grid on every step. The PRINT\_GRID instruction thus tells PLUMED to output the average density at each point on the grid every 500 steps of simulation. Notice that the grid output on step 1000 is an average over all 1000 frames of the trajectory. If you would like to analyse these two blocks of data separately you must use the CLEAR flag.

This second example computes an order parameter (in this case FCCUBIC) and constructs a phase field model for this order parameter using the equation above.

```
fcc: FCCUBIC SPECIES=1-5184 SWITCH={CUBIC D_0=1.2 D_MAX=1.5} ALPHA=27
dens: MULTICOLVARDENS DATA=fcc ORIGIN=1 DIR=xyz NBINS=14,14,28 BANDWIDTH=1.0,1.0,1.0 STRIDE=1 CLEAR=1
DUMPCUBE GRID=dens STRIDE=1 FILE=dens.cube
```

In this example the phase field model is computed and output to a file on every step of the simulation. Furthermore, because the CLEAR=1 keyword is set on the MULTICOLVARDENS line each Gaussian cube file output is a phase field model for a particular trajectory frame. The average value accumulated thus far is cleared at the start of every single timestep and there is no averaging over trajectory frames in this case.

## 6.17 CONVERT\_TO\_FES

Convert a histogram,  $H(x)$ , to a free energy surface using  $F(x) = -k_B T \ln H(x)$ .

This action allows you to take a free energy surface that was calculated using the HISTOGRAM action and to convert it to a free energy surface. This transformation performed by doing:

$$F(x) = -k_B T \ln H(x)$$

The free energy calculated on a grid is output by this action and can be printed using DUMPGRID

## Compulsory keywords

<b>GRID</b>	the action that creates the input grid you would like to use
-------------	--------------------------------------------------------------

## Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>COMPONENT</b>	if your input is a vector field use this to specify the component of the input vector field for which you wish to use
<b>TEMP</b>	the temperature at which you are operating

### Examples

This is a typical example showing how `CONVERT_TO_FES` might be used when postprocessing a trajectory. The input below calculates the free energy as a function of the distance between atom 1 and atom 2. This is done by accumulating a histogram as a function of this distance using kernel density estimation and the `HISTOGRAM` action. All the data within this trajectory is used in the construction of this `HISTOGRAM`. Finally, once all the data has been read in, the histogram is converted to a free energy using the formula above and the free energy is output to a file called `fes.dat`

```
x: DISTANCE ATOMS=1,2
hA1: HISTOGRAM ARG=x GRID_MIN=0.0 GRID_MAX=3.0 GRID_BIN=100 BANDWIDTH=0.1
ff: CONVERT_TO_FES GRID=hA1 TEMP=300
DUMPGRID GRID=ff FILE=fes.dat
```

## 6.18 DUMPCUBE

Output a three dimensional grid using the Gaussian cube file format.

Suppose you have calculated the value of a function on a three dimensional grid. This function might be a [HISTOGRAM](#) or it might be a free energy energy surface that was calculated from this histogram by using [CONVERT\\_TO\\_FES](#). Alternatively, your function might be a phase-field that was calculated using [MULTICOLVARDENS](#). Whatever the function is, however, you obviously cannot show it using a typical contour plotting program such as `gnuplot` as you have three input variables.

Tools like `VMD` have nice features for plotting these types of three dimensional functions but typically you are required to use a Gaussian cube file format to input the data. This action thus allows you to output a function evaluated on a grid to a Gaussian cube file format.

### Compulsory keywords

<b>GRID</b>	the action that creates the grid you would like to output
<b>STRIDE</b>	( default=0 ) the frequency with which the grid should be output to the file. The default value of 0 ensures that the grid is only output at the end of the trajectory
<b>FILE</b>	( default=density ) the file on which to write the grid.

### Options

<b>FMT</b>	the format that should be used to output real numbers
<b>COMPONENT</b>	if your input is a vector field use this to specify the component of the input vector field for which you wish to output

### Examples

The input below can be used to postprocess a trajectory. A histogram as a function of the distance between atoms 1 and 2, the distance between atom 1 and 3 and the angle between the vector connecting atoms 1 and 2 and 1 and 3 is computed using kernel density estimation. Once all the data contained in the trajectory has been read in and all the kernels have been added the resulting histogram is output to a file called histoA1.cube. This file has the Gaussian cube file format. The histogram can thus be visualized using tools such as VMD.

```
x1: DISTANCE ATOMS=1,2
x2: DISTANCE ATOMS=1,3
x3: ANGLE ATOMS=1,2,3

hA1: HISTOGRAM ARG=x1,x2,x3 GRID_MIN=0.0,0.0,0.0 GRID_MAX=3.0,3.0,3.0 GRID_BIN=10,10,10 BANDWIDTH=1.0,1.0,1.0
DUMPCUBE GRID=hA1 FILE=histoA1.cube
```

## 6.19 DUMPGRID

Output the function on the grid to a file with the PLUMED grid format.

PLUMED provides a number of actions that calculate the values of functions on grids. For instance, whenever you calculate a free energy as a function of a collective variable using [HISTOGRAM](#) and [CONVERT\\_TO\\_FES](#) you will generally want to output the value of the free energy at a number of points on a discrete grid that covers the CV space uniformly. Alternatively you may want to calculate what value some symmetry function takes at different points inside your simulation cell using [MULTICOLVARDENS](#).

This action allows you to output these functions calculated on a grid using a format that can be read in using gnuplot and other such plotting programs. The file output using this action will have a header that contains some essential information about the function plotted and that looks something like this:

```
#! FIELDS x y hA1 dhA1_x dhA1_y
#! SET normalisation 2.0000
#! SET min_x 0.0
#! SET max_x 3.0
#! SET nbins_x 100
#! SET periodic_x false
#! SET min_y 0.0
#! SET max_y 3.0
#! SET nbins_y 100
#! SET periodic_y false
```

The header shown here tells us that we have grid showing the values that a function with two arguments x and y takes at various points in our cell. The lines beneath the first line then tell us a little bit about these two input arguments.

The remaining lines of the file give us information on the positions of our grid points and the value the function and its partial derivatives with respect to x and y. If the header is as above a list of values of the function that have x=0 and 100 values of y between 0.0 and 3.0 will be provided. This block of data will be followed with a blank line. There will then be a second block of values which will all have been evaluated the same value of x and all possible values for y. This block is then followed by a blank line again and this pattern continues until all points of the grid have been covered.

### Compulsory keywords

<b>GRID</b>	the action that creates the grid you would like to output
<b>STRIDE</b>	( default=0 ) the frequency with which the grid should be output to the file. The default value of 0 ensures that the grid is only output at the end of the trajectory
<b>FILE</b>	( default=density ) the file on which to write the grid.

### Options

<b>FMT</b>	the format that should be used to output real numbers
------------	-------------------------------------------------------

### Examples

The following input monitors two torsional angles during a simulation and outputs a continuous histogram as a function of them at the end of the simulation.

```
TORSION ATOMS=1,2,3,4 LABEL=r1
TORSION ATOMS=2,3,4,5 LABEL=r2
HISTOGRAM ...
  ARG=r1,r2
  GRID_MIN=-3.14,-3.14
  GRID_MAX=3.14,3.14
  GRID_BIN=200,200
  BANDWIDTH=0.05,0.05
  LABEL=hh
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo
```

The following input monitors two torsional angles during a simulation and outputs a discrete histogram as a function of them at the end of the simulation.

```
TORSION ATOMS=1,2,3,4 LABEL=r1
TORSION ATOMS=2,3,4,5 LABEL=r2
HISTOGRAM ...
  ARG=r1,r2
  KERNEL=DISCRETE
  GRID_MIN=-3.14,-3.14
  GRID_MAX=3.14,3.14
  GRID_BIN=200,200
  LABEL=hh
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo
```

The following input monitors two torsional angles during a simulation and outputs the histogram accumulated thus far every 100000 steps.

```
TORSION ATOMS=1,2,3,4 LABEL=r1
TORSION ATOMS=2,3,4,5 LABEL=r2
HISTOGRAM ...
  ARG=r1,r2
  GRID_MIN=-3.14,-3.14
  GRID_MAX=3.14,3.14
  GRID_BIN=200,200
  BANDWIDTH=0.05,0.05
  LABEL=hh
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo STRIDE=100000
```



The following input monitors two torsional angles during a simulation and outputs a separate histogram for each 100000 steps worth of trajectory. Notice how the CLEAR keyword is used here and how it is not used in the previous example.

```
TORSION ATOMS=1,2,3,4 LABEL=r1
TORSION ATOMS=2,3,4,5 LABEL=r2
HISTOGRAM ...
  ARG=r1,r2 CLEAR=100000
  GRID_MIN=-3.14,-3.14
  GRID_MAX=3.14,3.14
  GRID_BIN=200,200
  BANDWIDTH=0.05,0.05
  GRID_WFILE=histo
  LABEL=hh
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo STRIDE=100000
```

## 6.20 FIND\_CONTOUR\_SURFACE

Find an isocontour by searching along either the x, y or z direction.

As discussed in the part of the manual on [Analysis](#) PLUMED contains a number of tools that allow you to calculate a function on a grid. The function on this grid might be a [HISTOGRAM](#) as a function of a few collective variables or it might be a phase field that has been calculated using [MULTICOLVARDENS](#). If this function has one or two input arguments it is relatively straightforward to plot the function. If by contrast the data has a three dimensions it can be difficult to visualize.

This action provides one tool for visualizing these functions. It can be used to search for a set of points on a contour where the function takes a particular value. In other words, for the function  $f(x, y, z)$  this action would find a set of points  $\{x_c, y_c, z_c\}$  that have:

$$f(x_c, y_c, z_c) - c = 0$$

where  $c$  is some constant value that is specified by the user. The points on this contour are found by searching along lines that run parallel to the  $x$ ,  $y$  or  $z$  axis of the simulation cell. The result is, therefore, a two dimensional function evaluated on a grid that gives us the height of the interface as a function of two coordinates.

It is important to note that this action can only be used to detect contours in three dimensional functions. In addition, this action will fail to find the full set of contour points if the contour does not have the same topology as an infinite plane. If you are uncertain that the isocontours in your function have the appropriate topology you should use [FIND\\_CONTOUR](#) in place of [FIND\\_CONTOUR\\_SURFACE](#).

### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>CLEAR</b>	( default=0 ) the frequency with which to clear all the accumulated data. The default value of 0 implies that all the data will be used and that the grid will never be cleared
<b>GRID</b>	the action that creates the input grid you would like to use
<b>CONTOUR</b>	the value we would like to draw the contour at in the space
<b>SEARCHDIR</b>	In which directions do you wish to search for the contour.
<b>BUFFER</b> <small>Generated by Doxygen</small>	( default=0 ) number of buffer grid points around location where grid was found on last step. If this is zero the full grid is calculated on each step

## Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages
<b>COMPONENT</b>	if your input is a vector field use this to specify the component of the input vector field for which you wish to use

## Examples

The input shown below was used to analyse the results from a simulation of an interface between solid and molten Lennard Jones. The interface between the solid and the liquid was set up in the plane perpendicular to the  $z$  direction of the simulation cell. The input below calculates something akin to a Willard-Chandler dividing surface [34] between the solid phase and the liquid phase. There are two of these interfaces within the simulation box because of the periodic boundary conditions but we were able to determine that one of these two surfaces lies in a particular part of the simulation box. The input below detects the height profile of one of these two interfaces. It does so by computing a phase field average of the [FCCUBIC](#) symmetry function using the [MULTICOLVARDENS](#) action. Notice that we use the fact that we know roughly where the interface is when specifying how this phase field is to be calculated and specify the region over the  $z$ -axis in which we are going to search for the phase field in the line defining the [MULTICOLVARDENS](#). Once we have calculated the phase field we search for contour points on the lines that run parallel to the  $z$ -direction of the cell box using the `FIND_CONTOUR_SURFACE` command. The final result is a  $14 \times 14$  grid of values for the height of the interface as a function of the  $(x, y)$  position. This grid is then output to a file called `contour2.dat`.

Notice that the commands below calculate the instantaneous position of the surface separating the solid and liquid and that as such the accumulated average is cleared on every step.

```

UNITS NATURAL
FCCUBIC ...
  SPECIES=1-96000 SWITCH={CUBIC D_0=1.2 D_MAX=1.5}
  ALPHA=27 PHI=0.0 THETA=-1.5708 PSI=-2.35619 LABEL=fcc
... FCCUBIC

dens2: MULTICOLVARDENS DATA=fcc ORIGIN=1 DIR=xyz NBINS=14,14,50 ZREDUCED ZLOWER=6.0 ZUPPER=11.0 BANDWIDTH=1.0,
ss2: FIND_CONTOUR_SURFACE GRID=dens2 CONTOUR=0.42 SEARCHDIR=z STRIDE=1 CLEAR=1
DUMPGRID GRID=ss2 FILE=contour2.dat FMT=%8.4f STRIDE=1

```

## 6.21 FIND\_CONTOUR

Find an isocontour in a smooth function.

As discussed in the part of the manual on [Analysis](#) PLUMED contains a number of tools that allow you to calculate a function on a grid. The function on this grid might be a [HISTOGRAM](#) as a function of a few collective variables or it might be a phase field that has been calculated using [MULTICOLVARDENS](#). If this function has one or two input arguments it is relatively straightforward to plot the function. If by contrast the data has a three or more dimensions it can be difficult to visualize.

This action provides one tool for visualizing these functions. It can be used to search for a set of points on a contour where the function takes a particular values. In other words, for the function  $f(x, y)$  this action would find a set of points  $\{x_c, y_c\}$  that have:

$$f(x_c, y_c) - c = 0$$

where  $c$  is some constant value that is specified by the user. The points on this contour are detected using a variant on the marching squares or marching cubes algorithm, which you can find information on here:

[https://en.wikipedia.org/wiki/Marching\\_squares](https://en.wikipedia.org/wiki/Marching_squares) [https://en.wikipedia.org/wiki/Marching\\_cubes](https://en.wikipedia.org/wiki/Marching_cubes)

As such, and unlike [FIND\\_CONTOUR\\_SURFACE](#) or [FIND\\_SPHERICAL\\_CONTOUR](#), the function input to this action can have any dimension. Furthermore, the topology of the contour will be determined by the algorithm and does not need to be specified by the user.

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>CLEAR</b>	( default=0 ) the frequency with which to clear all the accumulated data. The default value of 0 implies that all the data will be used and that the grid will never be cleared
<b>GRID</b>	the action that creates the input grid you would like to use
<b>CONTOUR</b>	the value we would like to draw the contour at in the space
<b>FILE</b>	file on which to output coordinates
<b>UNITS</b>	( default=PLUMED ) the units in which to print out the coordinates. PLUMED means internal PLUMED units
<b>BUFFER</b>	( default=0 ) number of buffer grid points around location where grid was found on last step. If this is zero the full grid is calculated on each step

#### Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages
<b>COMPONENT</b>	if your input is a vector field use this to specify the component of the input vector field for which you wish to use
<b>PRECISION</b>	The number of digits in trajectory file

#### Examples

The input below allows you to calculate something akin to a Willard-Chandler dividing surface [34]. The simulation

cell in this case contains a solid phase and a liquid phase. The Willard-Chandler surface is the surface that separates the parts of the box containing the solid from the parts containing the liquid. To compute the position of this surface the [FCCUBIC](#) symmetry function is calculated for each of the atoms in the system from on the geometry of the atoms in the first coordination sphere of each of the atoms. These quantities are then transformed using a switching function. This procedure generates a single number for each atom in the system and this quantity has a value of one for atoms that are in parts of the box that resemble the solid structure and zero for atoms that are in parts of the box that resemble the liquid. The position of a virtual atom is then computed using [CENTER\\_OF\\_MULTICOLVAR](#) and a phase field model is constructed using [MULTICOLVARDENS](#). These procedure ensures that we have a continuous function that gives a measure of the average degree of solidness at each point in the simulation cell. The Willard-Chandler dividing surface is calculated by finding a a set of points at which the value of this phase field is equal to 0.5. This set of points is output to file called mycontour.dat. A new contour is found on every single step for each frame that is read in.

```

UNITS NATURAL
FCCUBIC ...
  SPECIES=1-96000 SWITCH={CUBIC D_0=1.2 D_MAX=1.5}
  ALPHA=27 PHI=0.0 THETA=-1.5708 PSI=-2.35619 LABEL=fcc
... FCCUBIC

tfcc: MTRANSFORM_MORE DATA=fcc SWITCH={SMAP R_0=0.5 A=8 B=8}
center: CENTER_OF_MULTICOLVAR DATA=tfcc

MULTICOLVARDENS ...
  DATA=tfcc ORIGIN=center DIR=xyz LABEL=dens
  NBINS=80,80,80 BANDWIDTH=1.0,1.0,1.0 STRIDE=25
  LABEL=dens STRIDE=1 CLEAR=1
... MULTICOLVARDENS

FIND_CONTOUR GRID=dens CONTOUR=0.5 FILE=mycontour.dat

```

## 6.22 FIND\_SPHERICAL\_CONTOUR

Find an isocontour in a three dimensional grid by searching over a Fibonacci sphere.

As discussed in the part of the manual on [Analysis](#) PLUMED contains a number of tools that allow you to calculate a function on a grid. The function on this grid might be a [HISTOGRAM](#) as a function of a few collective variables or it might be a phase field that has been calculated using [MULTICOLVARDENS](#). If this function has one or two input arguments it is relatively straightforward to plot the function. If by contrast the data has a three dimensions it can be difficult to visualize.

This action provides one tool for visualizing these functions. It can be used to search for a set of points on a contour wher the function takes a particular value. In other words, for the function  $f(x, y, z)$  this action would find a set of points  $\{x_c, y_c, z_c\}$  that have:

$$f(x_c, y_c, z_c) - c = 0$$

where  $c$  is some constant value that is specified by the user. The points on this contour are find by searching along a set of equally spaced radii of a sphere that centered at on particular, user-specified atom or virtual atom. To ensure that these search radii are equally spaced on the surface of the sphere the search directions are generated by using a fibonacci spiral projected on a sphere. In other words, the search directions are given by:

$$\mathbf{r}_i = \left( \sqrt{1 - y^2} \cos(\phi) \frac{2i}{n} - 1 + \frac{1}{n} \sqrt{1 - y^2} \sin(\phi) \right)$$

where  $y$  is the quantity second component of the vector defined above,  $n$  is the number of directions to look in and  $\phi$  is

$$\phi = (i + R, n)\pi(3 - \sqrt{5})$$

where  $R$  is a random variable between 0 and  $n - 1$  that is generated during the read in of the input file and that is fixed during the whole calculation.

It is important to note that this action can only be used to detect contours in three dimensional functions. In addition, this action will fail to find the full set of contour points if the contour does not have the same topology as a sphere. If you are uncertain that the isocontours in your function have a spherical topology you should use [FIND\\_CONTOUR](#) in place of [FIND\\_SPHERICAL\\_CONTOUR](#).

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>GRID</b>	the action that creates the input grid you would like to use
<b>CONTOUR</b>	the value we would like to draw the contour at in the space
<b>FILE</b>	file on which to output coordinates
<b>UNITS</b>	( default=PLUMED ) the units in which to print out the coordinates. PLUMED means internal PLUMED units
<b>NPOINTS</b>	the number of points for which we are looking for the contour
<b>INNER_RADIUS</b>	the minimum radius on which to look for the contour
<b>OUTER_RADIUS</b>	the outer radius on which to look for the contour
<b>NBINS</b>	( default=1 ) the number of discrete sections in which to divide the distance between the inner and outer radius when searching for a contour

#### Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages
<b>COMPONENT</b>	if your input is a vector field use this to specify the component of the input vector field for which you wish to use
<b>PRECISION</b>	The number of digits in trajectory file

#### Examples

The following input demonstrates how this action can be used. The input here is used to study the shape of a droplet that has been formed during the condensation of Lennard Jones from the vapour. The input below achieves this by calculating the coordination numbers of all the atoms within the gas. Obviously, those atoms within the droplet will have a large value for the coordination number while the isolated atoms in the gas will have a low value.

As such we can detect the sizes of the droplets by constructing a `CONTACT_MATRIX` whose  $ij$  element tells us whether atom  $i$  and atom  $j$  have coordination number that is greater than two. The atoms within the various droplets within the system can then be found by performing a `DFSCCLUSTERING` on this matrix to detect the connected components. We can take the largest of these connected components and find the center of the droplet by exploiting the functionality within `CENTER_OF_MULTICOLVAR`. We can then construct a phase field based on the positions of the atoms in the largest cluster and the values of the coordination numbers of these atoms. The final line in the input then finds the a set of points on the dividing surface that separates the droplet from the surrounding gas. The value of the phase field on this isocontour is equal to 0.75.

```
# Calculate coordination numbers
c1: COORDINATIONNUMBER SPECIES=1-512 SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0}
# Select coordination numbers that are more than 2.0
cf: MFILTER_MORE DATA=c1 SWITCH={RATIONAL D_0=2.0 R_0=0.1} LOWMEM
# Build a contact matrix
mat: CONTACT_MATRIX ATOMS=cf SWITCH={EXP D_0=4.0 R_0=0.5 D_MAX=6.0}
# Find largest cluster
dfs: DFSCCLUSTERING MATRIX=mat
clust1: CLUSTER_PROPERTIES CLUSTERS=dfs CLUSTER=1
# Find center of largest cluster
trans1: MTRANSFORM_MORE DATA=clust1 SWITCH={RATIONAL D_0=2.0 R_0=0.1} LOWMEM
cent: CENTER_OF_MULTICOLVAR DATA=trans1
# Calculate the phase field of the coordination
dens: MULTICOLVARDENS DATA=trans1 ORIGIN=cent DIR=xyz NBINS=30,30,30 BANDWIDTH=2.0,2.0,2.0
# Find the isocontour around the nucleus
FIND_SPHERICAL_CONTOUR GRID=dens CONTOUR=0.85 INNER_RADIUS=10.0 OUTER_RADIUS=40.0 FILE=mysurface.xyz UNITS=A F
```

## 6.23 FOURIER\_TRANSFORM

Compute the Discrete Fourier Transform (DFT) by means of FFTW of data stored on a 2D grid.

This action can operate on any other action that outputs scalar data on a two-dimensional grid.

Up to now, even if the input data are purely real the action uses a complex DFT.

Just as a quick reference, given a 1D array  $\mathbf{X}$  of size  $n$ , this action computes the vector  $\mathbf{Y}$  given by

$$Y_k = \sum_{j=0}^{n-1} X_j e^{2\pi i j k \sqrt{-1}/n}.$$

This can be easily extended to more than one dimension. All the other details can be found at <http://www.↔fftw.org/doc/What-FFTW-Really-Computes.html#What-FFTW-Really-Computes>.

The keyword "FOURIER\_PARAMETERS" deserves just a note on the usage. This keyword specifies how the Fourier transform will be normalized. The keyword takes two numerical parameters ( $a$ ,  $b$ ) that define the normalization according to the following expression

$$\frac{1}{n^{(1-a)/2}} \sum_{j=0}^{n-1} X_j e^{2\pi i j k \sqrt{-1}/n}$$

The default values of these parameters are:  $a = 1$  and  $b = 1$ .

### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>CLEAR</b>	( default=0 ) the frequency with which to clear all the accumulated data. The default value of 0 implies that all the data will be used and that the grid will never be cleared
<b>GRID</b>	the action that creates the input grid you would like to use
<b>FOURIER_PARAMETERS</b>	( default=default ) what kind of normalization is applied to the output and if the Fourier transform in FORWARD or BACKWARD. This keyword takes the form FOURIER_PARAMETERS=A,B, where A and B can be 0, 1 or -1. The default values are A=1 (no normalization at all) and B=1 (forward FFT). Other possible choices for A are: A=-1: normalize by the number of data, A=0: normalize by the square root of the number of data (one forward and followed by backward FFT recover the original data).

## Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages
<b>COMPONENT</b>	if your input is a vector field use this to specify the component of the input vector field for which you wish to use
<b>FT_TYPE</b>	choose what kind of data you want as output on the grid. Possible values are: ABS = compute the complex modulus of Fourier coefficients (DEFAULT); NORM = compute the norm (i.e. $ABS^2$ ) of Fourier coefficients; COMPLEX = store the FFTW complex output on the grid (as a vector).

## Examples

The following example tells Plumed to compute the complex 2D 'backward' Discrete Fourier Transform by taking the data saved on a grid called 'density', and normalizing the output by  $\frac{1}{\sqrt{N_x N_y}}$ , where  $N_x$  and  $N_y$  are the number of data on the grid (it can be the case that  $N_x \neq N_y$ ):

```
FOURIER_TRANSFORM STRIDE=1 GRID=density FT_TYPE=complex FOURIER_PARAMETERS=0,-1 FILE=fourier.dat
```

## 6.24 INTERPOLATE\_GRID

Interpolate a smooth function stored on a grid onto a grid with a smaller grid spacing.

This action takes a function evaluated on a grid as input and can be used to interpolate the values of that function on to a finer grained grid. The interpolation within this algorithm is done using splines.

## Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>CLEAR</b>	( default=0 ) the frequency with which to clear all the accumulated data. The default value of 0 implies that all the data will be used and that the grid will never be cleared
<b>GRID</b>	the action that creates the input grid you would like to use

### Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages
<b>COMPONENT</b>	if your input is a vector field use this to specify the component of the input vector field for which you wish to use
<b>GRID_BIN</b>	the number of bins for the grid
<b>GRID_SPACING</b>	the approximate grid spacing (to be used as an alternative or together with GRID_BIN)

### Examples

The input below can be used to postprocess a trajectory. It calculates a **HISTOGRAM** as a function the distance between atoms 1 and 2 using kernel density estimation. During the calculation the values of the kernels are evaluated at 100 points on a uniform grid between 0.0 and 3.0. Prior to outputting this function at the end of the simulation this function is interpolated onto a finer grid of 200 points between 0.0 and 3.0.

```
x: DISTANCE ATOMS=1,2
hA1: HISTOGRAM ARG=x GRID_MIN=0.0 GRID_MAX=3.0 GRID_BIN=100 BANDWIDTH=0.1
ii: INTERPOLATE_GRID GRID=hA1 GRID_BIN=200
DUMPGRID GRID=ii FILE=histo.dat
```

## 6.25 CLASSICAL\_MDS

This is part of the analysis <b>module</b>
--------------------------------------------

Create a low-dimensional projection of a trajectory using the classical multidimensional scaling algorithm.

Multidimensional scaling (MDS) is similar to what is done when you make a map. You start with distances between London, Belfast, Paris and Dublin and then you try to arrange points on a piece of paper so that the (suitably scaled) distances between the points in your map representing each of those cities are related to the true distances between the cities. Stating this more mathematically MDS endeavors to find an **isometry** between points distributed in a high-dimensional space and a set of points distributed in a low-dimensional plane. In other words, if we have  $M$   $D$ -dimensional points,  $\mathbf{X}$ , and we can calculate dissimilarities between pairs them,  $D_{ij}$ , we can, with an MDS calculation, try to create  $M$  projections,  $\mathbf{x}$ , of the high dimensionality points in a  $d$ -dimensional linear space by trying



to arrange the projections so that the Euclidean distances between pairs of them,  $d_{ij}$ , resemble the dissimilarities between the high dimensional points. In short we minimize:

$$\chi^2 = \sum_{i \neq j} (D_{ij} - d_{ij})^2$$

where  $D_{ij}$  is the distance between point  $X^i$  and point  $X^j$  and  $d_{ij}$  is the distance between the projection of  $X^i$ ,  $x^i$ , and the projection of  $X^j$ ,  $x^j$ . A tutorial on this approach can be used to analyse simulations can be found in the tutorial [Belfast tutorial: Adaptive variables II](#) and in the following [short video](#).

The atoms involved can be specified using

<b>ATOMS</b>	the atoms whose positions we are tracking for the purpose of analysing the data. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>CLEAR</b>	( default=0 ) the frequency with which to clear all the accumulated data. The default value of 0 implies that all the data will be used and that the grid will never be cleared
<b>METRIC</b>	( default=EUCLIDEAN ) how are we measuring the distances between configurations
<b>RUN</b>	( default=0 ) the frequency with which to run the analysis algorithm. The default value of zero assumes you want to analyse the whole trajectory
<b>LANDMARKS</b>	( default=ALL ) only use a subset of the data that was collected. For more information on the landmark selection algorithms that are available in plumed see <a href="#">landmarkselection</a> .
<b>NLOW_DIM</b>	number of low-dimensional coordinates required
<b>OUTPUT_FILE</b>	file on which to output the final embedding coordinates
<b>EMBEDDING_OFILE</b>	( default=dont output ) file on which to output the embedding in plumed input format

Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>WRITE_CHECKPOINT</b>	( default=off ) write out a checkpoint so that the analysis can be restarted in a later run
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages

<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>FMT</b>	the format that should be used in analysis output files
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UNTIL</b>	Only update this action until this time

### Examples

The following command instructs plumed to construct a classical multidimensional scaling projection of a trajectory. The RMSD distance between atoms 1-256 have moved is used to measure the distances in the high-dimensional space.

```
CLASSICAL_MDS ...
  ATOMS=1-256
  METRIC=OPTIMAL-FAST
  NLOW_DIM=2
  OUTPUT_FILE=rmsd-embed
... CLASSICAL_MDS
```

The following section is for people who are interested in how this method works in detail. A solid understanding of this material is not necessary to use MDS.

#### 6.25.1 Method of optimisation

The stress function can be minimized using a standard optimization algorithm such as conjugate gradients or steepest descent. However, it is more common to do this minimization using a technique known as classical scaling. Classical scaling works by recognizing that each of the distances  $D_{ij}$  in the above sum can be written as:

$$D_{ij}^2 = \sum_{\alpha} (X_{\alpha}^i - X_{\alpha}^j)^2 = \sum_{\alpha} (X_{\alpha}^i)^2 + (X_{\alpha}^j)^2 - 2X_{\alpha}^i X_{\alpha}^j$$

We can use this expression and matrix algebra to calculate multiple distances at once. For instance if we have three points,  $\mathbf{X}$ , we can write distances between them as:

$$\begin{aligned}
D^2(\mathbf{X}) &= \begin{bmatrix} 0 & d_{12}^2 & d_{13}^2 \\ d_{12}^2 & 0 & d_{23}^2 \\ d_{13}^2 & d_{23}^2 & 0 \end{bmatrix} \\
&= \sum_{\alpha} \begin{bmatrix} (X_{\alpha}^1)^2 & (X_{\alpha}^1)^2 & (X_{\alpha}^1)^2 \\ (X_{\alpha}^2)^2 & (X_{\alpha}^2)^2 & (X_{\alpha}^2)^2 \\ (X_{\alpha}^3)^2 & (X_{\alpha}^3)^2 & (X_{\alpha}^3)^2 \end{bmatrix} + \sum_{\alpha} \begin{bmatrix} (X_{\alpha}^1)^2 & (X_{\alpha}^2)^2 & (X_{\alpha}^3)^2 \\ (X_{\alpha}^1)^2 & (X_{\alpha}^2)^2 & (X_{\alpha}^3)^2 \\ (X_{\alpha}^1)^2 & (X_{\alpha}^2)^2 & (X_{\alpha}^3)^2 \end{bmatrix} - 2 \sum_{\alpha} \begin{bmatrix} X_{\alpha}^1 X_{\alpha}^1 & X_{\alpha}^1 X_{\alpha}^2 & X_{\alpha}^1 X_{\alpha}^3 \\ X_{\alpha}^2 X_{\alpha}^1 & X_{\alpha}^2 X_{\alpha}^2 & X_{\alpha}^2 X_{\alpha}^3 \\ X_{\alpha}^3 X_{\alpha}^1 & X_{\alpha}^3 X_{\alpha}^2 & X_{\alpha}^3 X_{\alpha}^3 \end{bmatrix} \\
&= \mathbf{c}\mathbf{1}^T + \mathbf{1}\mathbf{c}^T - 2 \sum_{\alpha} \mathbf{x}_{\alpha} \mathbf{x}_{\alpha}^T = \mathbf{c}\mathbf{1}^T + \mathbf{1}\mathbf{c}^T - 2\mathbf{X}\mathbf{X}^T
\end{aligned}$$

This last equation can be extended to situations when we have more than three points. In it  $\mathbf{X}$  is a matrix that has one high-dimensional point on each of its rows and  $\mathbf{X}^T$  is its transpose.  $\mathbf{1}$  is an  $M \times 1$  vector of ones and  $\mathbf{c}$  is a vector with components given by:

$$c_i = \sum_{\alpha} (x_{\alpha}^i)^2$$

These quantities are the diagonal elements of  $\mathbf{X}\mathbf{X}^T$ , which is a dot product or Gram Matrix that contains the dot product of the vector  $X_i$  with the vector  $X_j$  in element  $i, j$ .

In classical scaling we introduce a centering matrix  $\mathbf{J}$  that is given by:

$$\mathbf{J} = \mathbf{I} - \frac{1}{M} \mathbf{1}\mathbf{1}^T$$

where  $\mathbf{I}$  is the identity. Multiplying the equations above from the front and back by this matrix and a factor of a  $-\frac{1}{2}$  gives:

$$\begin{aligned}
-\frac{1}{2} \mathbf{J} D^2(\mathbf{X}) \mathbf{J} &= -\frac{1}{2} \mathbf{J} (\mathbf{c}\mathbf{1}^T + \mathbf{1}\mathbf{c}^T - 2\mathbf{X}\mathbf{X}^T) \mathbf{J} \\
&= -\frac{1}{2} \mathbf{J} \mathbf{c}\mathbf{1}^T \mathbf{J} - \frac{1}{2} \mathbf{J} \mathbf{1}\mathbf{c}^T \mathbf{J} + \frac{1}{2} \mathbf{J} (2\mathbf{X}\mathbf{X}^T) \mathbf{J} \\
&= \mathbf{J}\mathbf{X}\mathbf{X}^T \mathbf{J} = \mathbf{X}\mathbf{X}^T
\end{aligned}$$

The first two terms in this expression disappear because  $\mathbf{1}^T \mathbf{J} = \mathbf{J} \mathbf{1} = \mathbf{0}$ , where  $\mathbf{0}$  is a matrix containing all zeros. In the final step meanwhile we use the fact that the matrix of squared distances will not change when we translate all the points. We can thus assume that the mean value,  $\mu_{\alpha}$ , for each of the components,  $\alpha$ :

$$\mu_{\alpha} = \frac{1}{M} \sum_{i=1}^N \mathbf{X}_{\alpha}^i$$

is equal to 0 so the columns of  $\mathbf{X}$  add up to 0. This in turn means that each of the columns of  $\mathbf{X}\mathbf{X}^T$  adds up to zero, which is what allows us to write  $\mathbf{J}\mathbf{X}\mathbf{X}^T \mathbf{J} = \mathbf{X}\mathbf{X}^T$ .

The matrix of squared distances is symmetric and positive-definite we can thus use the spectral decomposition to decompose it as:

$$\Phi = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$$

Furthermore, because the matrix we are diagonalizing,  $\mathbf{X}\mathbf{X}^T$ , is the product of a matrix and its transpose we can use this decomposition to write:

$$\mathbf{X} = \mathbf{V} \mathbf{\Lambda}^{\frac{1}{2}}$$

Much as in PCA there are generally a small number of large eigenvalues in  $\mathbf{\Lambda}$  and many small eigenvalues. We can safely use only the large eigenvalues and their corresponding eigenvectors to express the relationship between the coordinates  $\mathbf{X}$ . This gives us our set of low-dimensional projections.

This derivation makes a number of assumptions about the how the low dimensional points should best be arranged to minimise the stress. If you use an iterative optimization algorithm such as SMACOF you may thus be able to find a better (lower-stress) projection of the points. For more details on the assumptions made see [this website](#).

## 6.26 PCA

This is part of the analysis <a href="#">module</a>
-----------------------------------------------------

Perform principal component analysis (PCA) using either the positions of the atoms a large number of collective variables as input.

Principal component analysis is a statistical technique that uses an orthogonal transformation to convert a set of observations of poorly correlated variables into a set of linearly uncorrelated variables. You can read more about the specifics of this technique here: [https://en.wikipedia.org/wiki/Principal\\_component\\_analysis](https://en.wikipedia.org/wiki/Principal_component_analysis)

When used with molecular dynamics simulations a set of frames taken from the trajectory,  $\{X_i\}$ , or the values of a number of collective variables which are calculated from the trajectory frames are used as input. In this second instance your input to the PCA analysis algorithm is thus a set of high-dimensional vectors of collective variables. However, if collective variables are calculated from the positions of the atoms or if the positions are used directly the assumption is that this input trajectory is a set of poorly correlated (high-dimensional) vectors. After principal component analysis has been performed the output is a set of orthogonal vectors that describe the directions in which the largest motions have been seen. In other words, principal component analysis provides a method for lowering the dimensionality of the data contained in a trajectory. These output directions are some linear combination of the  $x$ ,  $y$  and  $z$  positions if the positions were used as input or some linear combination of the input collective variables if a high-dimensional vector of collective variables was used as input.

As explained on the Wikipedia page you must calculate the average and covariance for each of the input coordinates. In other words, you must calculate the average structure and the amount the system fluctuates around this average structure. The problem in doing so when the  $x$ ,  $y$  and  $z$  coordinates of a molecule are used as input is that the majority of the changes in the positions of the atoms comes from the translational and rotational degrees of freedom of the molecule. The first six principal components will thus, most likely, be uninteresting. Consequently, to remedy this problem PLUMED provides the functionality to perform an RMSD alignment of the all the structures to be analysed to the first frame in the trajectory. This can be used to effectively remove translational and/or rotational motions from consideration. The resulting principal components thus describe vibrational motions of the molecule.

If you wish to calculate the projection of a trajectory on a set of principal components calculated from this PCA action then the output can be used as input for the [PCAVARS](#) action.

The atoms involved can be specified using

<b>ATOMS</b>	the atoms whose positions we are tracking for the purpose of analysing the data. For more information on how to specify lists of atoms see <a href="#">Groups and Virtual Atoms</a>
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the data should be collected and added to the quantity being averaged
<b>CLEAR</b>	( default=0 ) the frequency with which to clear all the accumulated data. The default value of 0 implies that all the data will be used and that the grid will never be cleared
<b>METRIC</b>	( default=EUCLIDEAN ) how are we measuring the distances between configurations
<b>RUN</b>	( default=0 ) the frequency with which to run the analysis algorithm. The default value of zero assumes you want to analyse the whole trajectory

<b>NLOW_DIM</b>	number of PCA coordinates required
<b>OFFILE</b>	the file on which to output the eigenvectors

## Options

<b>SERIAL</b>	( default=off ) do the calculation in serial. Do not parallelize
<b>LOWMEM</b>	( default=off ) lower the memory requirements
<b>TIMINGS</b>	( default=off ) output information on the timings of the various parts of the calculation
<b>UNORMALIZED</b>	( default=off ) output the unaveraged quantity/quantities.
<b>WRITE_CHECKPOINT</b>	( default=off ) write out a checkpoint so that the analysis can be restarted in a later run
<b>LOGWEIGHTS</b>	list of actions that calculates log weights that should be used to weight configurations when calculating averages
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>FMT</b>	the format that should be used in analysis output files
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UNTIL</b>	Only update this action until this time

## Examples

The following input instructs PLUMED to perform a principal component analysis in which the covariance matrix is calculated from changes in the positions of the first 22 atoms. The TYPE=OPTIMAL instruction ensures that translational and rotational degrees of freedom are removed from consideration. The first two principal components will be output to a file called pca-comp.pdb. Trajectory frames will be collected on every step and the PCA calculation will be performed at the end of the simulation.

```
PCA METRIC=OPTIMAL ATOMS=1-22 STRIDE=1 NLOW_DIM=2 OFFILE=pca-comp.pdb
```

The following input instructs PLUMED to perform a principal component analysis in which the covariance matrix is calculated from chnages in the six distances seen in the previous lines. Notice that here the TYPE=EUCLID←EAN keyword is used to indicate that no alignment has to be done when calculating the various elements of the covariance matrix from the input vectors. In this calculation the first two principal components will be output to a file called pca-comp.pdb. Trajectory frames will be collected every five steps and the PCA calculation is performed every

1000 steps. Consequently, if you run a 2000 step simulation the PCA analysis will be performed twice. The REWEIGHT\_BIAS keyword in this input tells PLUMED that rather than ascribing a weight of one to each of the frames when calculating averages and covariances a reweighting should be performed based on each frame's weight in these calculations should be determined based on the current value of the instantaneous bias (see [REWEIGHT\\_BIAS](#)).

```
d1: DISTANCE ATOMS=1,2
d2: DISTANCE ATOMS=1,3
d3: DISTANCE ATOMS=1,4
d4: DISTANCE ATOMS=2,3
d5: DISTANCE ATOMS=2,4
d6: DISTANCE ATOMS=3,4
```

```
PCA ARG=d1,d2,d3,d4,d5,d6 METRIC=EUCLIDEAN STRIDE=5 RUN=1000 NLOW_DIM=2 REWEIGHT_BIAS OFILE=pca-comp.pdb
```

## Chapter 7

# Bias

PLUMED allows you to run a number of enhanced sampling algorithms. The list of enhanced sampling algorithms contained in PLUMED is as follows:

<a href="#">ABMD</a>	Adds a ratchet-and-pawl like restraint on one or more variables.
<a href="#">BIASVALUE</a>	Takes the value of one variable and use it as a bias
<a href="#">EXTENDED_LAGRANGIAN</a>	Add extended Lagrangian.
<a href="#">EXTERNAL</a>	Calculate a restraint that is defined on a grid that is read during start up
<a href="#">LOWER_WALLS</a>	Defines a wall for the value of one or more collective variables, which limits the region of the phase space accessible during the simulation.
<a href="#">METAD</a>	Used to performed MetaDynamics on one or more collective variables.
<a href="#">METAINTERFERENCE</a>	Calculate the MetaInference Score for a set of back calculated experimental data.
<a href="#">MOVINGRESTRAINT</a>	Add a time-dependent, harmonic restraint on one or more variables.
<a href="#">PBMETAD</a>	Used to performed Parallel Bias MetaDynamics.
<a href="#">RESTRAINT</a>	Adds harmonic and/or linear restraints on one or more variables.
<a href="#">UPPER_WALLS</a>	Defines a wall for the value of one or more collective variables, which limits the region of the phase space accessible during the simulation.

Methods, such as [METAD](#) or [PBMETAD](#), that work by introducing a history dependent bias can be restarted using the [RESTART](#) keyword

### 7.1 ABMD

This is part of the bias [module](#)

Adds a ratchet-and-pawl like restraint on one or more variables.

This action can be used to evolve a system towards a target value in CV space using an harmonic potential moving with the thermal fluctuations of the CV [35] [36] [37]. The biasing potential in this method is as follows:

$$V(\rho(t)) = \begin{cases} \frac{K}{2} (\rho(t) - \rho_m(t))^2, & \rho(t) > \rho_m(t) \\ 0, & \rho(t) \leq \rho_m(t), \end{cases}$$

where

$$\rho(t) = (CV(t) - TO)^2$$

and

$$\rho_m(t) = \min_{0 \leq \tau \leq t} \rho(\tau) + \eta(t).$$

The method is based on the introduction of a biasing potential which is zero when the system is moving towards the desired arrival point and which damps the fluctuations when the system attempts to move in the opposite direction. As in the case of the ratchet and pawl system, propelled by thermal motion of the solvent molecules, the biasing potential does not exert work on the system.  $\eta(t)$  is an additional white noise acting on the minimum position of the bias.

### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential
<b>force2</b>	the instantaneous value of the squared force due to this bias potential
<b>_min</b>	one or multiple instances of this quantity will be refereceable elsewhere in the input file. These quantities will be named with the arguments of the bias followed by the character string _min. These quantities tell the user the minimum value assumed by rho_m(t).

### Compulsory keywords

<b>TO</b>	The array of target values
<b>KAPPA</b>	The array of force constants.

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...



<b>MIN</b>	Array of starting values for the bias (set rho_m(t), otherwise it is set using the current value of ARG)
<b>NOISE</b>	Array of white noise intensities (add a temperature to the ABMD)
<b>SEED</b>	Array of seeds for the white noise (add a temperature to the ABMD)

### Examples

The following input sets up two biases, one on the distance between atoms 3 and 5 and another on the distance between atoms 2 and 4. The two target values are defined using TO and the two strength using KAPPA. The total energy of the bias is printed.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
ABMD ARG=d1,d2 TO=1.0,1.5 KAPPA=5.0,5.0 LABEL=abmd
PRINT ARG=abmd.bias,abmd.d1_min,abmd.d2_min
```

(See also [DISTANCE](#) and [PRINT](#)).

## 7.2 BIASVALUE

This is part of the bias <a href="#">module</a>
-------------------------------------------------

Takes the value of one variable and use it as a bias

This is the simplest possible bias: the bias potential is equal to a collective variable. It is useful to create custom biasing potential, e.g. applying a function (see [Functions](#)) to some collective variable then using the value of this function directly as a bias.

### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential
<b>_bias</b>	one or multiple instances of this quantity will be refereceable elsewhere in the input file. these quantities will named with the arguments of the bias followed by the character string _bias. These quantities tell the user how much the bias is due to each of the colvars.

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
------------------------------	----------------------------------------------------------------------------

**ARG**

the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If `*` or `.*` appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a [DISTANCE](#) action labelled `dist` may have three components `x`, `y` and `z`. To take just the `x` component you should use `dist.x`, if you wish to take all three components then use `dist.*`. More information on the referencing of Actions can be found in the section of the manual on the PLUMED [Getting started](#). Scalar values can also be referenced using POSIX regular expressions as detailed in the section on [Regular Expressions](#). To use this feature you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. `ARG1`, `ARG2`, `ARG3`...

**Examples**

The following input tells plumed to use the value of the distance between atoms 3 and 5 and the value of the distance between atoms 2 and 4 as biases. It then tells plumed to print the energy of the restraint

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=3,6 LABEL=d2
BIASVALUE ARG=d1,d2 LABEL=b
PRINT ARG=d1,d2,b.d1,b.d2
```

(See also [DISTANCE](#) and [PRINT](#)).

Another thing one can do is asking one system to follow a circle in sin/cos according a time dependence

```
t: TIME
# this just print cos and sin of time
cos: MATHEVAL ARG=t VAR=t FUNC=cos(t) PERIODIC=NO
sin: MATHEVAL ARG=t VAR=t FUNC=sin(t) PERIODIC=NO
c1: COM ATOMS=1,2
c2: COM ATOMS=3,4
d: DISTANCE COMPONENTS ATOMS=c1,c2
PRINT ARG=t,cos,sin,d.x,d.y,d.z STRIDE=1 FILE=colvar FMT=%8.4f
# this calculates sine and cosine of a projected component of distance
mycos: MATHEVAL ARG=d.x,d.y VAR=x,y FUNC=x/sqrt(x*x+y*y) PERIODIC=NO
mysin: MATHEVAL ARG=d.x,d.y VAR=x,y FUNC=y/sqrt(x*x+y*y) PERIODIC=NO
# this creates a moving spring so that the system follows a circle-like dynamics
# but it is not a bias, it is a simple value now
vv1: MATHEVAL ARG=mycos,mysin,cos,sin VAR=mc,ms,c,s FUNC=100*((mc-c)^2+(ms-s)^2) PERIODIC=NO
# this takes the value calculated with matheval and uses as a bias
cc: BIASVALUE ARG=vv1
# some printout
PRINT ARG=t,cos,sin,d.x,d.y,d.z,mycos,mysin,cc.bias,vv1 STRIDE=1 FILE=colvar FMT=%8.4f
```

(see also [TIME](#), [MATHEVAL](#), [COM](#), [DISTANCE](#), and [PRINT](#)).

## 7.3 EXTENDED\_LAGRANGIAN

Add extended Lagrangian.

This action can be used to create fictitious collective variables coupled to the real ones. Given  $x_i$  the  $i$ -th argument of this bias potential, potential and kinetic contributions are added to the energy of the system as

$$V = \sum_i \frac{k_i}{2} (x_i - s_i)^2 + \sum_i \frac{\dot{s}_i^2}{2m_i}$$

The resulting potential is thus similar to a [RESTRAINT](#), but the restraint center moved with time following Hamiltonian dynamics with mass  $m_i$ .

This bias potential accepts thus vectorial keywords (one element per argument) to define the coupling constant (KAPPA) and a relaxation time *tau* (TAU). The mass is then computed as  $m = k(\frac{\tau}{2\pi})^2$ .

Notice that this action creates several components. The ones named XX\_fict are the fictitious coordinates. It is possible to add further forces on them by means of other bias potential, e.g. to obtain an indirect [METAD](#) as in [38]. Also notice that the velocities of the fictitious coordinates are reported (XX\_vfict). However, printed velocities are the ones at the previous step.

It is also possible to provide a non-zero friction (one value per component). This is then used to implement a Langevin thermostat, so as to implement TAMD/dAFED method [39] [40]. Notice that here a massive Langevin thermostat is used, whereas usually TAMD employs an overamped Langevin dynamics and dAFED a Gaussian thermostat.

### Warning

The bias potential is reported in the component bias. Notice that this bias potential, although formally compatible with replica exchange framework, probably does not work as expected in that case. Indeed, since fictitious coordinates are not swapped upon exchange, acceptance can be expected to be extremely low unless (by chance) two neighboring replicas have the fictitious variables located properly in space.

[RESTART](#) is not properly supported by this action. Indeed, at every start the position of the fictitious variable is reset to the value of the real variable, and its velocity is set to zero. This is not expected to introduce big errors, but certainly is introducing a small inconsistency between a single long run and many shorter runs.

### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential
<b>_fict</b>	one or multiple instances of this quantity will be refereceable elsewhere in the input file. These quantities will named with the arguments of the bias followed by the character string <code>_tilde</code> . It is possible to add forces on these variable.
<b>_vfict</b>	one or multiple instances of this quantity will be refereceable elsewhere in the input file. These quantities will named with the arguments of the bias followed by the character string <code>_tilde</code> . It is NOT possible to add forces on these variable.

## Compulsory keywords

<b>KAPPA</b>	specifies that the restraint is harmonic and what the values of the force constants on each of the variables are
<b>TAU</b>	specifies that the restraint is harmonic and what the values of the force constants on each of the variables are
<b>FRICTION</b>	( default=0.0 ) add a friction to the variable

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>TEMP</b>	the system temperature - needed when FRICTION is present. If not provided will be taken from MD code (if available)

## Examples

The following input tells plumed to perform a metadynamics with an extended Lagrangian on two torsional angles.

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
ex: EXTENDED_LAGRANGIAN ARG=phi,psi KAPPA=20,20.0 TAU=0.1,0.1
METAD ARG=ex.phi_fict,ex.psi_fict PACE=100 SIGMA=0.35,0.35 HEIGHT=0.1
# monitor the two variables
PRINT STRIDE=10 ARG=phi,psi,ex.phi_fict,ex.psi_fict FILE=COLVAR
```

(See also [TORSION](#), [METAD](#), and [PRINT](#)).

The following input tells plumed to perform a TAMD (or dAFED) calculation on two torsional angles, keeping the two variables at a fictitious temperature of 3000K with a Langevin thermostat with friction 10

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
ex: EXTENDED_LAGRANGIAN ARG=phi,psi KAPPA=20,20.0 TAU=0.1,0.1 FRICTION=10,10 TEMP=3000
# monitor the two variables
PRINT STRIDE=10 ARG=phi,psi,ex.phi_fict,ex.psi_fict FILE=COLVAR
```

(See also [TORSION](#) and [PRINT](#))

## 7.4 EXTERNAL

	This is part of the <a href="#">bias module</a>
--	-------------------------------------------------

Calculate a restraint that is defined on a grid that is read during start up

### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential

### Compulsory keywords

<b>FILE</b>	the name of the file containing the external potential.
-------------	---------------------------------------------------------

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>NOSPLINE</b>	( default=off ) specifies that no spline interpolation is to be used when calculating the energy and forces due to the external potential
<b>SPARSE</b>	( default=off ) specifies that the external potential uses a sparse grid
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...

### Examples

The following is an input for a calculation with an external potential that is defined in the file bias.dat and that acts on the distance between atoms 3 and 5.

```
DISTANCE ATOMS=3,5 LABEL=d1
EXTERNAL ARG=d1 FILE=bias.dat LABEL=external
```

(See also [DISTANCE PRINT](#)).

The header in the file bias.dat should read:

```
#! FIELDS d1 external.bias der_d1
#! SET min_d1 0.0
#! SET max_d1 1.0
#! SET nbins_d1 100
#! SET periodic_d1 false
```

This should then be followed by the value of the potential and its derivative at 100 equally spaced points along the distance between 0 and 1.

You can also include grids that are a function of more than one collective variable. For instance the following would be the input for an external potential acting on two torsional angles:

```
TORSION ATOMS=4,5,6,7 LABEL=t1
TORSION ATOMS=6,7,8,9 LABEL=t2
EXTERNAL ARG=t1,t2 FILE=bias.dat LABEL=ext
```

The header in the file bias.dat for this calculation would read:

```
#! FIELDS t1 t2 ext.bias der_t1 der_t2
#! SET min_t1 -pi
#! SET max_t1 +pi
#! SET nbins_t1 100
#! SET periodic_t1 true
#! SET min_t2 -pi
#! SET max_t2 +pi
#! SET nbins_t2 100
#! SET periodic_t2 true
```

This would be then followed by 100 blocks of data. In the first block of data the value of t1 (the value in the first column) is kept fixed and the value of the function is given at 100 equally spaced values for t2 between  $-pi$  and  $+pi$ . In the second block of data t1 is fixed at  $-pi + \frac{2pi}{100}$  and the value of the function is given at 100 equally spaced values for t2 between  $-pi$  and  $+pi$ . In the third block of data the same is done but t1 is fixed at  $-pi + \frac{4pi}{100}$  and so on until you get to the 100th block of data where t1 is fixed at  $+pi$ .

Please note the order that the order of arguments in the plumed.dat file must be the same as the order of arguments in the header of the grid file.

## 7.5 LOWER\_WALLS

This is part of the <a href="#">bias module</a>
-------------------------------------------------

Defines a wall for the value of one or more collective variables, which limits the region of the phase space accessible during the simulation.

The restraining potential starts acting on the system when the value of the CV is greater (in the case of UPPER\_WALLS) or lower (in the case of LOWER\_WALLS) than a certain limit  $a_i$  (AT) minus an offset  $o_i$  (OFFSET). The expression for the bias due to the wall is given by:

$$\sum_i k_i ((x_i - a_i + o_i)/s_i)^{e_i}$$

$k_i$  (KAPPA) is an energy constant in internal unit of the code,  $s_i$  (EPS) a rescaling factor and  $e_i$  (EXP) the exponent determining the power law. By default: EXP = 2, EPS = 1.0, OFFSET = 0.

#### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential
<b>force2</b>	the instantaneous value of the squared force due to this bias potential

#### Compulsory keywords

<b>AT</b>	the positions of the wall. The $a_i$ in the expression for a wall.
<b>KAPPA</b>	the force constant for the wall. The $k_i$ in the expression for a wall.
<b>OFFSET</b>	( default=0.0 ) the offset for the start of the wall. The $o_i$ in the expression for a wall.
<b>EXP</b>	( default=2.0 ) the powers for the walls. The $e_i$ in the expression for a wall.
<b>EPS</b>	( default=1.0 ) the values for $s_i$ in the expression for a wall

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...

#### Examples

The following input tells plumed to add both a lower and an upper walls on the distance between atoms 3 and 5 and

the distance between atoms 2 and 4. The lower and upper limits are defined at different values. The strength of the walls is the same for the four cases. It also tells plumed to print the energy of the walls.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
UPPER_WALLS ARG=d1,d2 AT=1.0,1.5 KAPPA=150.0,150.0 EXP=2,2 EPS=1,1 OFFSET=0,0 LABEL=uwall
LOWER_WALLS ARG=d1,d2 AT=0.0,1.0 KAPPA=150.0,150.0 EXP=2,2 EPS=1,1 OFFSET=0,0 LABEL=lwall
PRINT ARG=uwall.bias,lwall.bias
```

(See also [DISTANCE](#) and [PRINT](#)).

## 7.6 METAD

This is part of the <a href="#">bias module</a>
-------------------------------------------------

Used to performed MetaDynamics on one or more collective variables.

In a metadynamics simulations a history dependent bias composed of intermittently added Gaussian functions is added to the potential [\[41\]](#).

$$V(\vec{s}, t) = \sum_{k\tau < t} W(k\tau) \exp \left( - \sum_{i=1}^d \frac{(s_i - s_i^{(0)}(k\tau))^2}{2\sigma_i^2} \right).$$

This potential forces the system away from the kinetic traps in the potential energy surface and out into the unexplored parts of the energy landscape. Information on the Gaussian functions from which this potential is composed is output to a file called HILLS, which is used both the restart the calculation and to reconstruct the free energy as a function of the CVs. The free energy can be reconstructed from a metadynamics calculation because the final bias is given by:

$$V(\vec{s}) = -F(\vec{s})$$

During post processing the free energy can be calculated in this way using the [sum\\_hills](#) utility.

In the simplest possible implementation of a metadynamics calculation the expense of a metadynamics calculation increases with the length of the simulation as one has to, at every step, evaluate the values of a larger and larger number of Gaussians. To avoid this issue you can store the bias on a grid. This approach is similar to that proposed in [\[42\]](#) but has the advantage that the grid spacing is independent on the Gaussian width. Notice that you should provide either the number of bins for every collective variable (GRID\_BIN) or the desired grid spacing (GRID\_SPACING). In case you provide both PLUMED will use the most conservative choice (highest number of bins) for each dimension. In case you do not provide any information about bin size (neither GRID\_BIN nor GRID\_SPACING) and if Gaussian width is fixed PLUMED will use 1/5 of the Gaussian width as grid spacing. This default choice should be reasonable for most applications.

Metadynamics can be restarted either from a HILLS file as well as from a GRID, in this second case one can first save a GRID using GRID\_WFILE (and GRID\_WSTRIDE) and at a later stage read it using GRID\_RFILE.

Another option that is available in plumed is well-tempered metadynamics [\[43\]](#). In this variant of metadynamics the heights of the Gaussian hills are rescaled at each step so the bias is now given by:



$$V(s, t) = \sum_{t'=0, \tau_G, 2\tau_G, \dots}^{t' < t} W e^{-V(s(q(t'), t')/\Delta T} \exp \left( - \sum_{i=1}^d \frac{(s_i(q) - s_i(q(t'))^2}{2\sigma_i^2} \right),$$

This method ensures that the bias converges more smoothly. It should be noted that, in the case of well-tempered metadynamics, in the output printed the Gaussian height is re-scaled using the bias factor. Also notice that with well-tempered metadynamics the HILLS file does not contain the bias, but the negative of the free-energy estimate. This choice has the advantage that one can restart a simulation using a different value for the  $\Delta T$ . The applied bias will be scaled accordingly.

Note that you can use here also the flexible gaussian approach [44] in which you can adapt the gaussian to the extent of Cartesian space covered by a variable or to the space in collective variable covered in a given time. In this case the width of the deposited gaussian potential is denoted by one value only that is a Cartesian space (ADAPTIVE=GEOM) or a time (ADAPTIVE=DIFF). Note that a specific integration technique for the deposited gaussians should be used in this case. Check the documentation for utility `sum_hills`.

With the keyword `INTERVAL` one changes the metadynamics algorithm setting the bias force equal to zero outside boundary [45]. If, for example, metadynamics is performed on a CV  $s$  and one is interested only to the free energy for  $s > s_w$ , the history dependent potential is still updated according to the above equations but the metadynamics force is set to zero for  $s < s_w$ . Notice that Gaussians are added also if  $s < s_w$ , as the tails of these Gaussians influence  $V_G$  in the relevant region  $s > s_w$ . In this way, the force on the system in the region  $s > s_w$  comes from both metadynamics and the force field, in the region  $s < s_w$  only from the latter. This approach allows obtaining a history-dependent bias potential  $V_G$  that fluctuates around a stable estimator, equal to the negative of the free energy far enough from the boundaries. Note that:

- It works only for one-dimensional biases;
- It works both with and without `GRID`;
- The interval limit  $s_w$  in a region where the free energy derivative is not large;
- If in the region outside the limit  $s_w$  the system has a free energy minimum, the `INTERVAL` keyword should be used together with a `UPPER_WALLS` or `LOWER_WALLS` at  $s_w$ .

As a final note, since version 2.0.2 when the system is outside of the selected interval the force is set to zero and the bias value to the value at the corresponding boundary. This allows acceptances for replica exchange methods to be computed correctly.

Multiple walkers [46] can also be used. See below the examples.

The  $c(t)$  reweighting factor can also be calculated on the fly using the equations presented in [3]. The expression used to calculate  $c(t)$  follows directly from using Eq. 12 in Eq. 3 in [3] and gives smoother results than equivalent Eqs. 13 and Eqs. 14 in that paper. The  $c(t)$  is given by the `rct` component while the bias normalized by  $c(t)$  is given by the `rbias` component (`rbias=bias-c(t)`) which can be used to obtain a reweighted histogram. The calculation of  $c(t)$  is enabled by using the keyword `REWEIGHTING_NGRID` where the grid used for the calculation is specified. This grid should have a size that is equal or larger than the grid given in `GRID_BIN`. By default  $c(t)$  is updated every 50 Gaussian hills but this can be changed by using the `REWEIGHTING_NHILLS` keyword. This option can only be employed together with Well-Tempered Metadynamics and requires that a grid is used.

Additional material and examples can be also found in the tutorials:

- [Belfast tutorial: Metadynamics](#)
- [Belfast tutorial: Replica exchange I](#)
- [Belfast tutorial: Replica exchange II and Multiple walkers](#)

Notice that at variance with PLUMED 1.3 it is now straightforward to apply concurrent metadynamics as done e.g. in Ref. [47]. This indeed can be obtained by using the METAD action multiple times in the same input file.

#### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential
<b>work</b>	accumulator for work

In addition the following quantities can be calculated by employing the keywords listed below

Quantity	Keyword	Description
<b>rbias</b>	<b>REWEIGHTING_NGRID</b>	the instantaneous value of the bias normalized using the $c(t)$ reweighting factor [ $rbias=bias-c(t)$ ]. This component can be used to obtain a reweighted histogram.
<b>rct</b>	<b>REWEIGHTING_NGRID</b>	the reweighting factor $c(t)$ .
<b>acc</b>	<b>ACCELERATION</b>	the metadynamics acceleration factor

#### Compulsory keywords

<b>SIGMA</b>	the widths of the Gaussian hills
<b>PACE</b>	the frequency for hill addition
<b>FILE</b>	( default=HILLS ) a file in which the list of added hills is stored

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>GRID_SPARSE</b>	( default=off ) use a sparse grid to store hills
<b>GRID_NOSPLINE</b>	( default=off ) don't use spline interpolation with grids
<b>STORE_GRIDS</b>	( default=off ) store all the grid files the calculation generates. They will be deleted if this keyword is not present
<b>WALKERS_MPI</b>	( default=off ) Switch on MPI version of multiple walkers - not compatible with other WALKERS_* options
<b>ACCELERATION</b>	( default=off ) Set to TRUE if you want to compute the metadynamics acceleration factor.

<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>HEIGHT</b>	the heights of the Gaussian hills. Compulsory unless TAU and either BIASF↔ACTOR or DAMPFACTOR are given
<b>FMT</b>	specify format for HILLS files (useful for decrease the number of digits in regtests)
<b>BIASFACTOR</b>	use well tempered metadynamics and use this biasfactor. Please note you must also specify temp
<b>DAMPFACTOR</b>	damp hills with $\exp(-\max(V)/(k_B T * DAMPFACTOR))$
<b>TARGET</b>	target to a predefined distribution
<b>TEMP</b>	the system temperature - this is only needed if you are doing well-tempered metadynamics
<b>TAU</b>	in well tempered metadynamics, sets height to $(k_B * \Delta T * \text{pace} * \text{timestep}) / \tau$
<b>GRID_MIN</b>	the lower bounds for the grid
<b>GRID_MAX</b>	the upper bounds for the grid
<b>GRID_BIN</b>	the number of bins for the grid
<b>GRID_SPACING</b>	the approximate grid spacing (to be used as an alternative or together with GRID_BIN)
<b>REWEIGHTING_NGRID</b>	calculate the $c(t)$ reweighting factor and use that to obtain the normalized bias [ $\text{rbias} = \text{bias} - c(t)$ ]. Here you should specify the number of grid points required in each dimension. The number of grid points should be equal or larger to the number of grid points given in GRID_BIN. This method is not compatible with metadynamics not on a grid.
<b>REWEIGHTING_NHILLS</b>	how many Gaussian hills should be deposited between calculating the $c(t)$ reweighting factor. The default is to do this every 50 hills.
<b>GRID_WSTRIDE</b>	write the grid to a file every N steps
<b>GRID_WFILE</b>	the file on which to write the grid
<b>GRID_RFILE</b>	a grid file from which the bias should be read at the initial step of the simulation
<b>ADAPTIVE</b>	use a geometric (=GEOM) or diffusion (=DIFF) based hills width scheme. Sigma is one number that has distance units or timestep dimensions
<b>WALKERS_ID</b>	walker id
<b>WALKERS_N</b>	number of walkers
<b>WALKERS_DIR</b>	shared directory with the hills files from all the walkers
<b>WALKERS_RSTRIDE</b>	stride for reading hills files
<b>INTERVAL</b>	monodimensional lower and upper limits, outside the limits the system will not feel the biasing force.
<b>SIGMA_MAX</b>	the upper bounds for the sigmas (in CV units) when using adaptive hills. Negative number means no bounds
<b>SIGMA_MIN</b>	the lower bounds for the sigmas (in CV units) when using adaptive hills. Negative number means no bounds
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)

<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UNTIL</b>	Only update this action until this time

### Examples

The following input is for a standard metadynamics calculation using as collective variables the distance between atoms 3 and 5 and the distance between atoms 2 and 4. The value of the CVs and the metadynamics bias potential are written to the COLVAR file every 100 steps.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
METAD ARG=d1,d2 SIGMA=0.2,0.2 HEIGHT=0.3 PACE=500 LABEL=restraint
PRINT ARG=d1,d2,restraint.bias STRIDE=100 FILE=COLVAR
```

(See also [DISTANCE PRINT](#)).

If you use adaptive Gaussians, with diffusion scheme where you use a Gaussian that should cover the space of 20 timesteps in collective variables. Note that in this case the histogram correction is needed when summing up hills.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
METAD ARG=d1,d2 SIGMA=20 HEIGHT=0.3 PACE=500 LABEL=restraint ADAPTIVE=DIFF
PRINT ARG=d1,d2,restraint.bias STRIDE=100 FILE=COLVAR
```

If you use adaptive Gaussians, with geometrical scheme where you use a Gaussian that should cover the space of 0.05 nm in Cartesian space. Note that in this case the histogram correction is needed when summing up hills.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
METAD ARG=d1,d2 SIGMA=0.05 HEIGHT=0.3 PACE=500 LABEL=restraint ADAPTIVE=GEOM
PRINT ARG=d1,d2,restraint.bias STRIDE=100 FILE=COLVAR
```

When using adaptive Gaussians you might want to limit how the hills width can change. You can use `SIGMA_MIN` and `SIGMA_MAX` keywords. The sigmas should be specified in terms of CV so you should use the CV units. Note that if you use a negative number, this means that the limit is not set. Note also that in this case the histogram correction is needed when summing up hills.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
METAD ...
  ARG=d1,d2 SIGMA=0.05 HEIGHT=0.3 PACE=500 LABEL=restraint ADAPTIVE=GEOM
  SIGMA_MIN=0.2,0.1 SIGMA_MAX=0.5,1.0
... METAD
PRINT ARG=d1,d2,restraint.bias STRIDE=100 FILE=COLVAR
```

Multiple walkers can be also used as in [46]. These are enabled by setting the number of walkers used, the id of the current walker which interprets the input file, the directory where the hills containing files reside, and the frequency to read the other walkers. Here is an example

```
DISTANCE ATOMS=3,5 LABEL=d1
METAD ...
  ARG=d1 SIGMA=0.05 HEIGHT=0.3 PACE=500 LABEL=restraint
  WALKERS_N=10
  WALKERS_ID=3
  WALKERS_DIR=../
  WALKERS_RSTRIDE=100
... METAD
```

where WALKERS\_N is the total number of walkers, WALKERS\_ID is the id of the present walker (starting from 0 ) and the WALKERS\_DIR is the directory where all the walkers are located. WALKERS\_RSTRIDE is the number of step between one update and the other. Since version 2.2.5, hills files are automatically flushed every WALKERS\_RSTRIDE steps.

The c(t) reweighting factor can be calculated on the fly using the equations presented in [3] as described above. This is enabled by using the keyword REWEIGHTING\_NGRID where the grid used for the calculation is set. The number of grid points given in REWEIGHTING\_NGRID should be equal or larger than the number of grid points given in GRID\_BIN.

```
METAD ...
  LABEL=metad
  ARG=phi,psi SIGMA=0.20,0.20 HEIGHT=1.20 BIASFACTOR=5 TEMP=300.0 PACE=500
  GRID_MIN=-pi,-pi GRID_MAX=pi,pi GRID_BIN=150,150
  REWEIGHTING_NGRID=150,150
  REWEIGHTING_NHILLS=20
... METAD
```

Here we have asked that the calculation is performed every 20 hills by using REWEIGHTING\_NHILLS keyword. If this keyword is not given the calculation will by default be performed every 50 hills. The c(t) reweighting factor will be given in the rct component while the instantaneous value of the bias potential normalized using the c(t) reweighting factor is given in the rbias component [rbias=bias-c(t)] which can be used to obtain a reweighted histogram or free energy surface using the HISTOGRAM analysis.

The kinetics of the transitions between basins can also be analysed on the fly as in [48]. The flag ACCELERATION turn on accumulation of the acceleration factor that can then be used to determine the rate. This method can be used together with COMMITTOR analysis to stop the simulation when the system get to the target basin. It must be used together with Well-Tempered Metadynamics.

You can also provide a target distribution using the keyword TARGET [49] [50] [51] The TARGET should be a grid containing a free-energy (i.e. the  $-k_B T \log$  of the desired target distribution). Gaussians will then be scaled by a factor

$$e^{\beta(\tilde{F}(s) - \tilde{F}_{max})}$$

Here  $\tilde{F}(s)$  is the free energy defined on the grid and  $\tilde{F}_{max}$  its maximum value. Notice that we here used the maximum value as in ref [51] This choice allows to avoid exceedingly large Gaussians to be added. However, it could make the Gaussian too small. You should always choose carefully the HEIGHT parameter in this case. The grid file should be similar to other PLUMED grid files in that it should contain both the target free-energy and its derivatives.

Notice that if you wish your simulation to converge to the target free energy you should use the DAMPFACOR command to provide a global tempering [52] Alternatively, if you use a BIASFACTOR your simulation will converge to a free energy that is a linear combination of the target free energy and of the intrinsic free energy determined by the original force field.

```
DISTANCE ATOMS=3,5 LABEL=d1
METAD ...
  LABEL=t1
  ARG=d1 SIGMA=0.05 TAU=200 DAMPFACOR=100 PACE=250
  GRID_MIN=0 GRID_MAX=2 GRID_BIN=200
  TARGET=dist.dat
... METAD

PRINT ARG=d1,t1.bias STRIDE=100 FILE=COLVAR
```

The header in the file dist.dat for this calculation would read:

```
#! FIELDS d1 t1.target der_d1
#! SET min_d1 0
#! SET max_d1 2
#! SET nbins_d1 200
#! SET periodic_d1 false
```

## 7.7 METAINFERENCE

	This is part of the bias <a href="#">module</a>
--	-------------------------------------------------

Calculate the Metainference Score for a set of back calculated experimental data.

The back calculated data, that are expected to be averages over replicas ( $NR=1,2,...,N$ ) The functional form of this bias can be chosen between three variants selected with `NOISE=GAUSS,MGAUSS,OUTLIERS` which correspond to modelling the noise for the arguments as a single gaussian common to all the data points, a gaussian per data point or a single long-tailed gaussian common to all the data points.

As from Metainference theory there are two sigma values: `SIGMA_MEAN` represent the error of calculating an average quantity using a finite set of replica and should be set as small as possible following the guidelines for replica-averaged simulations in the framework of the Maximum Entropy Principle. `SIGMA_BIAS` is an uncertainty parameter, sampled by a MC algorithm in the bounded interval defined by `SIGMA_MIN` and `SIGMA_MAX`. The initial value is set at `SIGMA0`. The MC move is a random displacement of maximum value equal to `DSIGMA`.

### Description of components

The names of the components in this action can be customized by the user in the actions input file. However, in addition to these customizable components the following quantities will always be output

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential
<b>sigma</b>	uncertainty parameter
<b>scale</b>	scale parameter
<b>accept</b>	MC acceptance

### Compulsory keywords

<b>NOISETYPE</b>	functional form of the noise (GAUSS,MGAUSS,OUTLIERS)
<b>SCALE0</b>	initial value of the uncertainty parameter
<b>SCALE_MIN</b>	minimum value of the uncertainty parameter
<b>SCALE_MAX</b>	maximum value of the uncertainty parameter
<b>DSCALE</b>	maximum MC move of the uncertainty parameter
<b>SIGMA0</b>	initial value of the uncertainty parameter
<b>SIGMA_MIN</b>	minimum value of the uncertainty parameter
<b>SIGMA_MAX</b>	maximum value of the uncertainty parameter
<b>DSIGMA</b>	maximum MC move of the uncertainty parameter
<b>SIGMA_MEAN</b>	starting value for the uncertainty in the mean estimate

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>SCALEDATA</b>	( default=off ) Set to TRUE if you want to sample a scaling factor common to all values and replicas
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>PARARG</b>	reference values for the experimental data, these can be provided as arguments without derivatives
<b>PARAMETERS</b>	reference values for the experimental data
<b>TEMP</b>	the system temperature - this is only needed if code doesnt' pass the temperature to plumed
<b>MC_STEPS</b>	number of MC steps
<b>MC_STRIDE</b>	MC stride

### Examples

In the following example we calculate a set of [RDC](#), take the replica-average of them and comparing them with a set of experimental values. RDCs are compared with the experimental data but for a multiplication factor SCALE that is also sampled by MC on-the-fly

```

RDC ...
LABEL=rdc
SCALE=0.0001
GYROM=-72.5388
ATOMS1=22,23
ATOMS2=25,27
ATOMS3=29,31
ATOMS4=33,34
... RDC

ardc: ENSEMBLE ARG=rdc.*

METAINFERENCE ...
ARG=ardc.*
NOISETYPE=MGAUSS
PARAMETERS=1.9190,2.9190,3.9190,4.9190
SCALEDATA SCALE0=1 SCALE_MIN=0.00001 SCALE_MAX=3 DSCALE=0.00
SIGMA0=0.01 SIGMA_MIN=0.00001 SIGMA_MAX=3 DSIGMA=0.00
SIGMA_MEAN=0.001
TEMP=300
LABEL=spe
... METAINFERENCE

PRINT ARG=spe.bias FILE=BIAS STRIDE=1

```

in the following example instead of using one uncertainty parameter per data point we use a single uncertainty value in a long-tailed gaussian to take into account for outliers.

```
METAINFERENCE ...
ARG=ardc.*
NOISETYPE=OUTLIERS
PARAMETERS=1.9190,2.9190,3.9190,4.9190
SCALEDATA SCALE0=1 SCALE_MIN=0.00001 SCALE_MAX=3 DSCALE=0.00
SIGMA0=0.01 SIGMA_MIN=0.00001 SIGMA_MAX=3 DSIGMA=0.00
SIGMA_MEAN=0.001
TEMP=300
LABEL=spe
... METAINFERENCE
```

(See also [RDC](#) and [ENSEMBLE](#)).

## 7.8 MOVINGRESTRAINT

This is part of the bias [module](#)

Add a time-dependent, harmonic restraint on one or more variables.

This form of bias can be used to performed steered MD [\[53\]](#) and Jarzynski sampling [\[54\]](#).

The harmonic restraint on your system is given by:

$$V(\vec{s}, t) = \frac{1}{2} \kappa(t) (\vec{s} - \vec{s}_0(t))^2$$

The time dependence of  $\kappa$  and  $\vec{s}_0$  are specified by a list of STEP, KAPPA and AT keywords. These keywords tell plumed what values  $\kappa$  and  $\vec{s}_0$  should have at the time specified by the corresponding STEP keyword. Inbetween these times the values of  $\kappa$  and  $\vec{s}_0$  are linearly interpolated.

Additional material and examples can be also found in the tutorial [Belfast tutorial: Out of equilibrium dynamics](#)

### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential
<b>work</b>	the total work performed changing this restraint
<b>force2</b>	the instantaneous value of the squared force due to this bias potential
<b>_cntr</b>	one or multiple instances of this quantity will be refereceable elsewhere in the input file. these quantities will named with the arguments of the bias followed by the character string _cntr. These quantities give the instantaneous position of the center of the harmonic potential.
<b>_work</b>	one or multiple instances of this quantity will be refereceable elsewhere in the input file. These quantities will named with the arguments of the bias followed by the character string _work. These quantities tell the user how much work has been done by the potential in dragging the system along
	the various colvar axis. <span style="float: right;">Generated by Doxygen</span>
<b>_kappa</b>	one or multiple instances of this quantity will be refereceable elsewhere in the input file. These quantities will named with the arguments of the bias followed by the character string _kappa. These quantities tell the user the time dependent value of kappa.



## Compulsory keywords

<b>VERSE</b>	( default=B ) Tells plumed whether the restraint is only acting for CV larger (U) or smaller (L) than the restraint or whether it is acting on both sides (B)
<b>STEP</b>	This keyword appears multiple times as STEP <sub>x</sub> with x=0,1,2,...,n. Each value given represents the MD step at which the restraint parameters take the values KAPPA <sub>x</sub> and AT <sub>x</sub> . You can use multiple instances of this keyword i.e. STEP1, STEP2, STEP3...
<b>AT</b>	AT <sub>x</sub> is equal to the position of the restraint at time STEP <sub>x</sub> . For intermediate times this parameter is linearly interpolated. If no AT <sub>x</sub> is specified for STEP <sub>x</sub> then the values of AT are kept constant during the interval of time between STEP <sub>x-1</sub> and STEP <sub>x</sub> . You can use multiple instances of this keyword i.e. AT1, AT2, AT3...
<b>KAPPA</b>	KAPPA <sub>x</sub> is equal to the value of the force constants at time STEP <sub>x</sub> . For intermediate times this parameter is linearly interpolated. If no KAPPA <sub>x</sub> is specified for STEP <sub>x</sub> then the values of KAPPA <sub>x</sub> are kept constant during the interval of time between STEP <sub>x-1</sub> and STEP <sub>x</sub> . You can use multiple instances of this keyword i.e. KAPPA1, KAPPA2, KAPPA3...

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...

## Examples

The following input is dragging the distance between atoms 2 and 4 from 1 to 2 in the first 1000 steps, then back in the next 1000 steps. In the following 500 steps the restraint is progressively switched off.

```
DISTANCE ATOMS=2,4 LABEL=d
MOVINGRESTRAINT ...
  ARG=d
  STEP0=0      AT0=1.0 KAPPA0=100.0
  STEP1=1000  AT1=2.0
  STEP2=2000  AT2=1.0
  STEP3=2500      KAPPA3=0.0
... MOVINGRESTRAINT
```

The following input is progressively building restraints distances between atoms 1 and 5 and between atoms 2 and 4 in the first 1000 steps. Afterwards, the restraint is kept static.

```
DISTANCE ATOMS=1,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
MOVINGRESTRAINT ...
  ARG=d1,d2
  STEP0=0      AT0=1.0,1.5 KAPPA0=0.0,0.0
  STEP1=1000 AT1=1.0,1.5 KAPPA1=1.0,1.0
... MOVINGRESTRAINT
```

The following input is progressively bringing atoms 1 and 2 close to each other with an upper wall

```
DISTANCE ATOMS=1,2 LABEL=d1
MOVINGRESTRAINT ...
  ARG=d1
  VERSE=U
  STEP0=0      AT0=1.0 KAPPA0=10.0
  STEP1=1000 AT1=0.0
... MOVINGRESTRAINT
```

By default the Action is issuing some values which are the work on each degree of freedom, the center of the harmonic potential, the total bias deposited

(See also [DISTANCE](#)).

#### Attention

Work is not computed properly when KAPPA is time dependent.

## 7.9 PBMETAD

This is part of the <a href="#">bias module</a>
-------------------------------------------------

Used to performed Parallel Bias MetaDynamics.

This action activate Parallel Bias MetaDynamics (PBMetaD) [55], a version of MetaDynamics [41] in which multiple low-dimensional bias potentials are applied in parallel. In the current implementation, these have the form of mono-dimensional MetaDynamics bias potentials:

$$V(s_1, t), \dots, V(s_N, t)$$

where:

$$V(s_i, t) = \sum_{k\tau < t} W_i(k\tau) \exp \left( -\frac{(s_i - s_i^{(0)}(k\tau))^2}{2\sigma_i^2} \right).$$

To ensure the convergence of each mono-dimensional bias potential to the corresponding free energy, at each deposition step the Gaussian heights are multiplied by the so-called conditional term:

$$W_i(k\tau) = W_0 \frac{\exp \left( -\frac{V(s_i, k\tau)}{k_B T} \right)}{\sum_{i=1}^N \exp \left( -\frac{V(s_i, k\tau)}{k_B T} \right)}$$

where  $W_0$  is the initial Gaussian height.

The PBMetaD bias potential is defined by:

$$V_{PB}(\vec{s}, t) = -k_B T \log \sum_{i=1}^N \exp \left( -\frac{V(s_i, t)}{k_B T} \right).$$

Information on the Gaussian functions that build each bias potential are printed to multiple HILLS files, which are used both to restart the calculation and to reconstruct the mono-dimensional free energies as a function of the corresponding CVs. These can be reconstructed using the [sum\\_hills](#) utility because the final bias is given by:

$$V(s_i) = -F(s_i)$$

Currently, only a subset of the [METAD](#) options are available in PBMetaD.

The bias potentials can be stored on a grid to increase performances of long PBMetaD simulations. You should provide either the number of bins for every collective variable (GRID\_BIN) or the desired grid spacing (GRID\_SPACING). In case you provide both PLUMED will use the most conservative choice (highest number of bins) for each dimension. In case you do not provide any information about bin size (neither GRID\_BIN nor GRID\_SPACING) and if Gaussian width is fixed PLUMED will use 1/5 of the Gaussian width as grid spacing. This default choice should be reasonable for most applications.

Another option that is available is well-tempered metadynamics [43]. In this variant of PBMetaD the heights of the Gaussian hills are rescaled at each step by the additional well-tempered metadynamics term. This ensures that each bias converges more smoothly. It should be noted that, in the case of well-tempered metadynamics, in the output printed the Gaussian height is re-scaled using the bias factor. Also notice that with well-tempered metadynamics the HILLS files do not contain the bias, but the negative of the free-energy estimate. This choice has the advantage that one can restart a simulation using a different value for the  $\Delta T$ . The applied bias will be scaled accordingly.

With the keyword INTERVAL one changes the metadynamics algorithm setting the bias force equal to zero outside boundary [45]. If, for example, metadynamics is performed on a CV  $s$  and one is interested only to the free energy for  $s > sw$ , the history dependent potential is still updated according to the above equations but the metadynamics force is set to zero for  $s < sw$ . Notice that Gaussians are added also if  $s < sw$ , as the tails of these Gaussians influence VG in the relevant region  $s > sw$ . In this way, the force on the system in the region  $s > sw$  comes from both metadynamics and the force field, in the region  $s < sw$  only from the latter. This approach allows obtaining a history-dependent bias potential VG that fluctuates around a stable estimator, equal to the negative of the free energy far enough from the boundaries. Note that:

- It works only for one-dimensional biases;
- It works both with and without GRID;
- The interval limit  $sw$  in a region where the free energy derivative is not large;
- If in the region outside the limit  $sw$  the system has a free energy minimum, the INTERVAL keyword should be used together with a [UPPER\\_WALLS](#) or [LOWER\\_WALLS](#) at  $sw$ .

Multiple walkers [46] can also be used, in the MPI implementation only. See below the examples.

#### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential

#### Compulsory keywords

<b>SIGMA</b>	the widths of the Gaussian hills
<b>PACE</b>	the frequency for hill addition, one for all biases

#### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>GRID_SPARSE</b>	( default=off ) use a sparse grid to store hills
<b>GRID_NOSPLINE</b>	( default=off ) don't use spline interpolation with grids
<b>MULTIPLE_WALKERS</b>	( default=off ) Switch on MPI version of multiple walkers
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...
<b>FILE</b>	files in which the lists of added hills are stored
<b>HEIGHT</b>	the height of the Gaussian hills, one for all biases. Compulsory unless TAU, TEMP and BIASFACTOR are given
<b>FMT</b>	specify format for HILLS files (useful for decrease the number of digits in regtests)
<b>BIASFACTOR</b>	use well tempered metadynamics with this biasfactor, one for all biases. Please note you must also specify temp
<b>TEMP</b>	the system temperature - this is only needed if you are doing well-tempered metadynamics
<b>TAU</b>	in well tempered metadynamics, sets height to $(kb \cdot \Delta T \cdot \text{pace} \cdot \text{timestep}) / \tau$
<b>GRID_RFILES</b>	read grid for the bias
<b>GRID_WFILES</b>	dump grid for the bias
<b>GRID_WSTRIDE</b>	frequency for dumping the grid
<b>GRID_MIN</b>	the lower bounds for the grid
<b>GRID_MAX</b>	the upper bounds for the grid
<b>GRID_BIN</b>	the number of bins for the grid
<b>GRID_SPACING</b>	the approximate grid spacing (to be used as an alternative or together with GRID_BIN)

<b>INTERVAL_MIN</b>	monodimensional lower limits, outside the limits the system will not feel the biasing force.
<b>INTERVAL_MAX</b>	monodimensional upper limits, outside the limits the system will not feel the biasing force.
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UNTIL</b>	Only update this action until this time

### Examples

The following input is for PBMetaD calculation using as collective variables the distance between atoms 3 and 5 and the distance between atoms 2 and 4. The value of the CVs and the PBMetaD bias potential are written to the COLVAR file every 100 steps.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
PBMETAD ARG=d1,d2 SIGMA=0.2,0.2 HEIGHT=0.3 PACE=500 LABEL=pb FILE=HILLS_d1,HILLS_d2
PRINT ARG=d1,d2,pb.bias STRIDE=100 FILE=COLVAR
```

(See also [DISTANCE](#) and [PRINT](#)).

If you use well-tempered metadynamics, you should specify a single biasfactor and initial Gaussian height.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
PBMETAD ...
ARG=d1,d2 SIGMA=0.2,0.2 HEIGHT=0.3
PACE=500 BIASFACTOR=8 LABEL=pb
FILE=HILLS_d1,HILLS_d2
... PBMETAD
PRINT ARG=d1,d2,pb.bias STRIDE=100 FILE=COLVAR
```

Only the MPI version of multiple-walkers is currently implemented.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
PBMETAD ...
ARG=d1,d2 SIGMA=0.2,0.2 HEIGHT=0.3
PACE=500 BIASFACTOR=8 LABEL=pb
FILE=HILLS_d1,HILLS_d2
MULTIPLE_WALKERS
... PBMETAD
PRINT ARG=d1,d2,pb.bias STRIDE=100 FILE=COLVAR
```

## 7.10 RESTRAINT

	This is part of the bias <a href="#">module</a>
--	-------------------------------------------------

Adds harmonic and/or linear restraints on one or more variables.

Either or both of SLOPE and KAPPA must be present to specify the linear and harmonic force constants respectively.

The resulting potential is given by:

$$\sum_i \frac{k_i}{2} (x_i - a_i)^2 + m_i * (x_i - a_i)$$

The number of components for any vector of force constants must be equal to the number of arguments to the action.

Additional material and examples can be also found in the tutorial [Belfast tutorial: Umbrella sampling](#)

### Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential
<b>force2</b>	the instantaneous value of the squared force due to this bias potential

### Compulsory keywords

<b>SLOPE</b>	( default=0.0 ) specifies that the restraint is linear and what the values of the force constants on each of the variables are
<b>KAPPA</b>	( default=0.0 ) specifies that the restraint is harmonic and what the values of the force constants on each of the variables are
<b>AT</b>	the position of the restraint

### Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*.More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...

## Examples

The following input tells plumed to restrain the distance between atoms 3 and 5 and the distance between atoms 2 and 4, at different equilibrium values, and to print the energy of the restraint

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
RESTRAINT ARG=d1,d2 AT=1.0,1.5 KAPPA=150.0,150.0 LABEL=restraint
PRINT ARG=restraint.bias
```

(See also [DISTANCE](#) and [PRINT](#)).

## 7.11 UPPER\_WALLS

This is part of the <a href="#">bias module</a>
-------------------------------------------------

Defines a wall for the value of one or more collective variables, which limits the region of the phase space accessible during the simulation.

The restraining potential starts acting on the system when the value of the CV is greater (in the case of UPPER\_WALLS) or lower (in the case of LOWER\_WALLS) than a certain limit  $a_i$  (AT) minus an offset  $o_i$  (OFFSET). The expression for the bias due to the wall is given by:

$$\sum_i k_i ((x_i - a_i + o_i) / s_i)^{e_i}$$

$k_i$  (KAPPA) is an energy constant in internal unit of the code,  $s_i$  (EPS) a rescaling factor and  $e_i$  (EXP) the exponent determining the power law. By default: EXP = 2, EPS = 1.0, OFFSET = 0.

## Description of components

By default this Action calculates the following quantities. These quantities can be referenced elsewhere in the input by using this Action's label followed by a dot and the name of the quantity required from the list below.

Quantity	Description
<b>bias</b>	the instantaneous value of the bias potential
<b>force2</b>	the instantaneous value of the squared force due to this bias potential

## Compulsory keywords

<b>AT</b>	the positions of the wall. The $a_i$ in the expression for a wall.
<b>KAPPA</b>	the force constant for the wall. The $k_i$ in the expression for a wall.
<b>OFFSET</b>	( default=0.0 ) the offset for the start of the wall. The $o_i$ in the expression for a wall.
<b>EXP</b>	( default=2.0 ) the powers for the walls. The $e_i$ in the expression for a wall.
<b>EPS</b>	( default=1.0 ) the values for $s_i$ in the expression for a wall

## Options

<b>NUMERICAL_DERIVATIVES</b>	( default=off ) calculate the derivatives for these quantities numerically
<b>ARG</b>	the input for this action is the scalar output from one or more other actions. The particular scalars that you will use are referenced using the label of the action. If the label appears on its own then it is assumed that the Action calculates a single scalar value. The value of this scalar is thus used as the input to this new action. If * or *.* appears the scalars calculated by all the preceding actions in the input file are taken. Some actions have multi-component outputs and each component of the output has a specific label. For example a <a href="#">DISTANCE</a> action labelled dist may have three componets x, y and z. To take just the x component you should use dist.x, if you wish to take all three components then use dist.*. More information on the referencing of Actions can be found in the section of the manual on the PLUMED <a href="#">Getting started</a> . Scalar values can also be referenced using POSIX regular expressions as detailed in the section on <a href="#">Regular Expressions</a> . To use this feature you you must compile PLUMED with the appropriate flag. You can use multiple instances of this keyword i.e. ARG1, ARG2, ARG3...

## Examples

The following input tells plumed to add both a lower and an upper walls on the distance between atoms 3 and 5 and the distance between atoms 2 and 4. The lower and upper limits are defined at different values. The strength of the walls is the same for the four cases. It also tells plumed to print the energy of the walls.

```
DISTANCE ATOMS=3,5 LABEL=d1
DISTANCE ATOMS=2,4 LABEL=d2
UPPER_WALLS ARG=d1,d2 AT=1.0,1.5 KAPPA=150.0,150.0 EXP=2,2 EPS=1,1 OFFSET=0,0 LABEL=uwall
LOWER_WALLS ARG=d1,d2 AT=0.0,1.0 KAPPA=150.0,150.0 EXP=2,2 EPS=1,1 OFFSET=0,0 LABEL=lwall
PRINT ARG=uwall.bias,lwall.bias
```

(See also [DISTANCE](#) and [PRINT](#)).

## 7.12 RESTART

This is part of the setup <a href="#">module</a>
--------------------------------------------------

Activate restart.

This is a Setup directive and, as such, should appear at the beginning of the input file. It influences the way PLUMED treat files open for writing (see also [Files](#)).

Notice that it is also possible to enable or disable restart on a per-action basis using the RESTART keyword on a single action. In this case, the keyword should be assigned a value. RESTART=AUTO means that global settings are used, RESTART=YES or RESTART=NO respectively enable and disable restart for that single action.



**Attention**

This directive can have also other side effects, e.g. on [METAD](#) and [PBMETAD](#) and on some analysis action.

**Options**

<b>NO</b>	( default=off ) switch off restart - can be used to override the behavior of the MD engine
-----------	--------------------------------------------------------------------------------------------

**Examples**

Using the following input:

```
d: DISTANCE ATOMS=1,2
PRINT ARG=d FILE=out
```

(See also [DISTANCE](#) and [PRINT](#)). a new 'out' file will be created. If an old one is on the way, it will be automatically backed up.

On the other hand, using the following input:

```
RESTART
d: DISTANCE ATOMS=1,2
PRINT ARG=d FILE=out
```

the file 'out' will be appended. (See also [DISTANCE](#) and [PRINT](#)).

In the following case, file out1 will be backed up and file out2 will be concatenated

```
RESTART
d1: DISTANCE ATOMS=1,2
d2: DISTANCE ATOMS=1,2
PRINT ARG=d1 FILE=out1 RESTART=NO
PRINT ARG=d2 FILE=out2
```

(See also [DISTANCE](#) and [PRINT](#)).



## Chapter 8

# Command Line Tools

PLUMED contains a number of simple command line tools. To use one of these tools you issue a command something like:

```
plumed <toolname> <list of input flags for that tool>
```

The following is a list of the various standalone tools that PLUMED contains.

<a href="#">driver-float</a>	Equivalent to <a href="#">driver</a> , but using single precision reals.
<a href="#">driver</a>	driver is a tool that allows one to to use plumed to post-process an existing trajectory.
<a href="#">gentemplate</a>	gentemplate is a tool that you can use to construct template inputs for the variousactions
<a href="#">info</a>	This tool allows you to obtain information about your plumed version
<a href="#">kt</a>	Print out the value of $k_B T$ at a particular temperature
<a href="#">manual</a>	manual is a tool that you can use to construct the manual page fora particular action
<a href="#">simplemd</a>	simplemd allows one to do molecular dynamics on systems of Lennard-Jones atoms.
<a href="#">sum_hills</a>	sum_hills is a tool that allows one to to use plumed to post-process an existing hills/colvar file

For all these tools and to use PLUMED as a plugin in an MD calculation you will need an input file.

### 8.1 driver-float

This is part of the <a href="#">cltools module</a>
----------------------------------------------------

Equivalent to [driver](#), but using single precision reals.

The purpose of this tool is just to test what PLUMED does when linked from a single precision code.

The input trajectory is specified using one of the following

---

<b>--xyz</b>	the trajectory in xyz format
<b>--gro</b>	the trajectory in gro format
<b>--xtc</b>	the trajectory in xtc format (xdrfile implementation)
<b>--trr</b>	the trajectory in trr format (xdrfile implementation)
<b>--mf_dcd</b>	molfile: the trajectory in dcd format
<b>--mf_crd</b>	molfile: the trajectory in crd format
<b>--mf_crdbox</b>	molfile: the trajectory in crdbox format
<b>--mf_gro</b>	molfile: the trajectory in gro format
<b>--mf_g96</b>	molfile: the trajectory in g96 format
<b>--mf_trr</b>	molfile: the trajectory in trr format
<b>--mf_trj</b>	molfile: the trajectory in trj format
<b>--mf_xtc</b>	molfile: the trajectory in xtc format
<b>--mf_pdb</b>	molfile: the trajectory in pdb format

The following must be present

<b>--plumed</b>	( default=plumed.dat ) specify the name of the plumed input file
<b>--timestep</b>	( default=1.0 ) the timestep that was used in the calculation that produced this trajectory in picoseconds
<b>--trajectory-stride</b>	( default=1 ) the frequency with which frames were output to this trajectory during the simulation (0 means that the number of the step is read from the trajectory file, currently working only for xtc/trr files read with <code>-ixtc/-trr</code> )
<b>--multi</b>	( default=0 ) set number of replicas for multi environment (needs mpi)

The following options are available

<b>--help/-h</b>	( default=off ) print this help
<b>--help-debug</b>	( default=off ) print special options that can be used to create regtests
<b>--noatoms</b>	( default=off ) don't read in a trajectory. Just use colvar files as specified in plumed.dat
<b>--dump-full-virial</b>	( default=off ) with <code>--dump-forces</code> , it dumps the 9 components of the virial
<b>--length-units</b>	units for length, either as a string or a number
<b>--mass-units</b>	units for mass in pdb and mc file, either as a string or a number
<b>--charge-units</b>	units for charge in pdb and mc file, either as a string or a number
<b>--dump-forces</b>	dump the forces on a file
<b>--dump-forces-fmt</b>	( default=%f ) the format to use to dump the forces
<b>--pdb</b>	provides a pdb with masses and charges
<b>--mc</b>	provides a file with masses and charges as produced with DUMPMASCHARGE
<b>--box</b>	comma-separated box dimensions (3 for orthorombic, 9 for generic)
<b>--natoms</b>	provides number of atoms - only used if file format does not contain number of atoms
<b>--debug-forces</b>	output a file containing the forces due to the bias evaluated using numerical derivatives and using the analytical derivatives implemented in plumed

## Examples

```
plumed driver-float --plumed plumed.dat --ixyz trajectory.xyz
```

See also examples in [driver](#)

## 8.2 driver

This is part of the <a href="#">cltools module</a>
----------------------------------------------------

driver is a tool that allows one to use plumed to post-process an existing trajectory.

The input to driver is specified using the command line arguments described below.

In addition, you can use the special [READ](#) command inside your plumed input to read in colvar files that were generated during your MD simulation. The values read in can then be treated like calculated colvars.

### Warning

Notice that by default the driver has no knowledge about the masses and charges of your atoms! Thus, if you want to compute quantities depending charges (e.g. [DHENERGY](#)) or masses (e.g. [COM](#)) you should pass the proper information to the driver. You can do it either with the `-pdb` option or with the `-mc` option. The latter will read a file produced by [DUMPMASSCHARGE](#).

The input trajectory is specified using one of the following

<code>--ixyz</code>	the trajectory in xyz format
<code>--igro</code>	the trajectory in gro format
<code>--ixtc</code>	the trajectory in xtc format (xdrfile implementation)
<code>--itr</code>	the trajectory in trr format (xdrfile implementation)
<code>--mf_dcd</code>	molfile: the trajectory in dcd format
<code>--mf_crd</code>	molfile: the trajectory in crd format
<code>--mf_crdbox</code>	molfile: the trajectory in crdbox format
<code>--mf_gro</code>	molfile: the trajectory in gro format
<code>--mf_g96</code>	molfile: the trajectory in g96 format
<code>--mf_trr</code>	molfile: the trajectory in trr format
<code>--mf_trj</code>	molfile: the trajectory in trj format
<code>--mf_xtc</code>	molfile: the trajectory in xtc format
<code>--mf_pdb</code>	molfile: the trajectory in pdb format

The following must be present

<b>--plumed</b>	( default=plumed.dat ) specify the name of the plumed input file
<b>--timestep</b>	( default=1.0 ) the timestep that was used in the calculation that produced this trajectory in picoseconds
<b>--trajectory-stride</b>	( default=1 ) the frequency with which frames were output to this trajectory during the simulation (0 means that the number of the step is read from the trajectory file, currently working only for xtc/trr files read with -ixtc/-trr)
<b>--multi</b>	( default=0 ) set number of replicas for multi environment (needs mpi)

The following options are available

<b>--help/-h</b>	( default=off ) print this help
<b>--help-debug</b>	( default=off ) print special options that can be used to create regtests
<b>--noatoms</b>	( default=off ) don't read in a trajectory. Just use colvar files as specified in plumed.dat
<b>--dump-full-virial</b>	( default=off ) with --dump-forces, it dumps the 9 components of the virial
<b>--length-units</b>	units for length, either as a string or a number
<b>--mass-units</b>	units for mass in pdb and mc file, either as a string or a number
<b>--charge-units</b>	units for charge in pdb and mc file, either as a string or a number
<b>--dump-forces</b>	dump the forces on a file
<b>--dump-forces-fmt</b>	( default=%f ) the format to use to dump the forces
<b>--pdb</b>	provides a pdb with masses and charges
<b>--mc</b>	provides a file with masses and charges as produced with DUMPMASCHARGE
<b>--box</b>	comma-separated box dimensions (3 for orthorombic, 9 for generic)
<b>--natoms</b>	provides number of atoms - only used if file format does not contain number of atoms
<b>--debug-forces</b>	output a file containing the forces due to the bias evaluated using numerical derivatives and using the analytical derivatives implemented in plumed

## Examples

The following command tells plumed to postprocess the trajectory contained in `trajectory.xyz` by performing the actions described in the input file `plumed.dat`. If an action that takes the stride keyword is given a stride equal to  $n$  then it will be performed only on every  $n$ th frames in the trajectory file.

```
plumed driver --plumed plumed.dat --ixyz trajectory.xyz
```

Notice that `xyz` files are expected to be in internal PLUMED units, that is by default nm. You can change this behavior by using the `--length-units` option:

```
plumed driver --plumed plumed.dat --ixyz trajectory.xyz --length-units A
```

The strings accepted by the `--length-units` options are the same ones accepted by the [UNITS](#) action. Other file formats typically have their default coordinates (e.g., `gro` files are always in nm) and it thus should not be necessary to use the `--length-units` option. Additionally, consider that the units used by the `driver` might be different by the units used in the PLUMED input file `plumed.dat`. For instance consider the command:

```
plumed driver --plumed plumed.dat --ixyz trajectory.xyz --length-units A
```

where `plumed.dat` is

```
# no explicit UNITS action here
d: DISTANCE ATOMS=1,2
PRINT ARG=d FILE=colvar
```

In this case, the driver reads the `xyz` file assuming it to contain coordinates in Angstrom units. However, the resulting `colvar` file contains a distance expressed in nm.

The following command tells `plumed` to postprocess the trajectory contained in `trajectory.xyz`. by performing the actions described in the input file `plumed.dat`.

```
plumed driver --plumed plumed.dat --ixyz trajectory.xyz --trajectory-stride 100 --timestep 0.001
```

Here though `--trajectory-stride` is set equal to the frequency with which frames were output during the trajectory and the `--timestep` is equal to the simulation timestep. As such the `STRIDE` parameters in the `plumed.dat` files are referred to the original timestep and any files output resemble those that would have been generated had we run the calculation we are running with driver when the MD simulation was running.

PLUMED can read natively `xyz` files (in PLUMED units) and `gro` files (in nm). In addition, PLUMED includes by default support for a subset of the trajectory file formats supported by VMD, e.g. `xtc` and `dcd`:

```
plumed driver --plumed plumed.dat --pdb diala.pdb --mf_xtc traj.xtc --trajectory-stride 100 --timestep 0.001
```

where `--mf_` prefixes the extension of one of the accepted molfile plugin format. If PLUMED has been [installed](#) with full molfile support, other formats will be available. Just type `plumed driver --help` to see which plugins are available.

Molfile plugin require periodic cell to be triangular (i.e. first vector oriented along x and second vector in xy plane). This is true for many MD codes. However, it could be false if you rotate the coordinates in your trajectory before reading them in the driver. Also notice that some formats (e.g. `amber crd`) do not specify atom number. In this case you can use the `--natoms` option:

```
plumed driver --plumed plumed.dat --imf_crd trajectory.crd --natoms 128
```

Check the available molfile plugins and limitations at [this link](#).

Additionally, you can use the `xdrfile` implementation of `xtc` and `trr`. To this aim, just download and install properly the `xdrfile` library (see [this link](#)). If the `xdrfile` library is installed properly the PLUMED configure script should be able to detect it and enable it. Notice that the `xdrfile` implementation of `xtc` and `trr` is more robust than the molfile one, since it provides support for generic cell shapes. In addition, it allows [DUMPATOMS](#) to write compressed `xtc` files.

### 8.2.1 READ

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

Read quantities from a `colvar` file.

This Action can be used with driver to read in a `colvar` file that was generated during an MD simulation

### Description of components

The READ command will read those fields that are labelled with the text string given to the VALUE keyword. It will also read in any fields that are labelled with the text string given to the VALUE keyword followed by a dot and a further string. If a single Value is read in this value can be referenced using the label of the Action. Alternatively, if multiple quantities are read in, they can be referenced elsewhere in the input by using the label for the Action followed by a dot and the character string that appeared after the dot in the title of the field.

### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which the file should be read.
<b>EVERY</b>	( default=1 ) only read every ith line of the colvar file. This should be used if the colvar was written more frequently than the trajectory.
<b>VALUES</b>	the values to read from the file
<b>FILE</b>	the name of the file from which to read these quantities

### Options

<b>IGNORE_TIME</b>	( default=off ) ignore the time in the colvar file. When this flag is not present read will be quite strict about the start time of the simulation and the stride between frames
<b>IGNORE_FORCES</b>	( default=off ) use this flag if the forces added by any bias can be safely ignored. As an example forces can be safely ignored if you are doing postprocessing that does not involve outputting forces
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UNTIL</b>	Only update this action until this time

### Examples

This input reads in data from a file called input\_colvar.data that was generated in a calculation that involved PLU↔MED. The first command reads in the data from the column headed phi1 while the second reads in the data from the column headed phi2.

```
rphi1:      READ FILE=input_colvar.data  VALUES=phi1
rphi2:      READ FILE=input_colvar.data  VALUES=phi2
PRINT ARG=rphi1,rphi2 STRIDE=500  FILE=output_colvar.data
```

## 8.3 gentemplate

This is part of the <a href="#">cltools module</a>
----------------------------------------------------

gentemplate is a tool that you can use to construct template inputs for the various actions



The templates generated by this tool are primarily for use with Toni Giorgino's vmd gui. It may be useful however to use this tool as a quick aid memoir.

#### Options

<b>--help/-h</b>	( default=off ) print this help
<b>--list</b>	( default=off ) print a list of the available actions
<b>--include-optional</b>	( default=off ) also print optional modifiers
<b>--action</b>	print the template for this particular action

#### Examples

The following generates template input for the action DISTANCE.

```
plumed gentemplate --action DISTANCE
```

## 8.4 info

This is part of the <a href="#">cltools module</a>
----------------------------------------------------

This tool allows you to obtain information about your plumed version

You can specify the information you require using the following command line arguments

#### Options

<b>--help/-h</b>	( default=off ) print this help
<b>--configuration</b>	( default=off ) prints the configuration file
<b>--root</b>	( default=off ) print the location of the root directory for the plumed source
<b>--user-doc</b>	( default=off ) print the location of user manual (html)
<b>--developer-doc</b>	( default=off ) print the location of user manual (html)
<b>--version</b>	( default=off ) print the version number
<b>--long-version</b>	( default=off ) print the version number (long version)
<b>--git-version</b>	( default=off ) print the version number (git version, if available)

#### Examples

The following command returns the root directory for your plumed distribution.

```
plumed info --root
```

## 8.5 kt

	This is part of the <a href="#">cltools module</a>
--	----------------------------------------------------

Print out the value of  $k_B T$  at a particular temperature

### Compulsory keywords

<b>--temp</b>	print the manual for this particular action
<b>--units</b>	( default=kj/mol ) the units of energy can be kj/mol, kcal/mol, j/mol, eV or the conversion factor from kj/mol

### Options

<b>--help/-h</b>	( default=off ) print this help
------------------	---------------------------------

### Examples

The following command will tell you the value of  $k_B T$  when T is equal to 300 K in eV

```
plumed kt --temp 300 --units eV
```

## 8.6 manual

	This is part of the <a href="#">cltools module</a>
--	----------------------------------------------------

manual is a tool that you can use to construct the manual page for a particular action

The manual constructed by this action is in html. In all probability you will never need to use this tool. However, it is used within the scripts that generate plumed's html manual. If you need to use this tool outside those scripts the input is specified using the following command line arguments.

### Compulsory keywords

<b>--action</b>	print the manual for this particular action
-----------------	---------------------------------------------

## Options

<b>--help/-h</b>	( default=off ) print this help
<b>--vim</b>	( default=off ) print the keywords in vim syntax

## Examples

The following generates the html manual for the action DISTANCE.

```
plumed manual --action DISTANCE
```

## 8.7 simplemd

This is part of the <a href="#">cltools module</a>
----------------------------------------------------

simplemd allows one to do molecular dynamics on systems of Lennard-Jones atoms.

The input to simplemd is spcified in an input file. Configurations are input and output in xyz format. The input file should contain one directive per line. The directives available are as follows:

## Compulsory keywords

<b>nstep</b>	The number of steps of dynamics you want to run
<b>temperature</b>	( default=NVE ) the temperature at which you wish to run the simulation in LJ units
<b>friction</b>	( default=off ) The friction (in LJ units) for the langevin thermostat that is used to keep the temperature constant
<b>tstep</b>	( default=0.005 ) the integration timestep in LJ units
<b>inputfile</b>	An xyz file containing the initial configuration of the system
<b>forcecutoff</b>	( default=2.5 )
<b>listcutoff</b>	( default=3.0 )
<b>outputfile</b>	An output xyz file containing the final configuration of the system
<b>nconfig</b>	( default=10 ) The frequency with which to write configurations to the trajectory file followed by the name of the trajectory file
<b>nstat</b>	( default=1 ) The frequency with which to write the statistics to the statistics file followed by the name of the statistics file
<b>maxneighbours</b>	( default=10000 ) The maximum number of neighbours an atom can have
<b>idum</b>	( default=0 ) The random number seed
<b>ndim</b>	( default=3 ) The dimensionality of the system (some interesting LJ clusters are two dimensional)
<b>wrapatoms</b>	( default=false ) If true, atomic coordinates are written wrapped in minimal cell

## Examples

You run an MD simulation using simplemd with the following command:

```
plumed simplemd < in
```

The following is an example of an input file for a simplemd calculation. This file instructs simplemd to do 50 steps of MD at a temperature of 0.722

```
nputfile input.xyz
outputfile output.xyz
temperature 0.722
tstep 0.005
friction 1
forcecutoff 2.5
listcutoff 3.0
nstep 50
nconfig 10 trajectory.xyz
nstat 10 energies.dat
```

If you run the following a description of all the directives that can be used in the input file will be output.

```
plumed simplemd --help
```

## 8.8 sum\_hills

This is part of the <a href="#">cltools module</a>
----------------------------------------------------

sum\_hills is a tool that allows one to use plumed to post-process an existing hills/colvar file

### Options

<b>--help/-h</b>	( default=off ) print this help
<b>--help-debug</b>	( default=off ) print special options that can be used to create regtests
<b>--negbias</b>	( default=off ) print the negative bias instead of the free energy (only needed with welltempered runs and flexible hills)
<b>--nohistory</b>	( default=off ) to be used with --stride: it splits the bias/histogram in pieces without previous history
<b>--mintozero</b>	( default=off ) it translate all the minimum value in bias/histogram to zero (usefull to compare results)
<b>--hills</b>	specify the name of the hills file
<b>--histo</b>	specify the name of the file for histogram a colvar/hills file is good
<b>--stride</b>	specify the stride for integrating hills file (default 0=never)
<b>--min</b>	the lower bounds for the grid
<b>--max</b>	the upper bounds for the grid
<b>--bin</b>	the number of bins for the grid

<b>--spacing</b>	grid spacing, alternative to the number of bins
<b>--idw</b>	specify the variables to be used for the free-energy/histogram (default is all). With <code>--hills</code> the other variables will be integrated out, with <code>--histo</code> the other variables won't be considered
<b>--outfile</b>	specify the outputfile for sumhills
<b>--outhisto</b>	specify the outputfile for the histogram
<b>--kt</b>	specify temperature in energy units for integrating out variables
<b>--sigma</b>	a vector that specify the sigma for binning (only needed when doing histogram)
<b>--fmt</b>	specify the output format

### Examples

a typical case is about the integration of a hills file:

```
plumed sum_hills --hills PATHTOMYHILLSFILE
```

The default name for the output file will be `fes.dat`. Note that starting from this version `plumed` will automatically detect the number of the variables you have and their periodicity. Additionally, if you use flexible hills (multivariate gaussians), `plumed` will understand it from the HILLS file.

now `sum_hills` tool accepts als multiple files that will be integrated one after the other

```
plumed sum_hills --hills PATHTOMYHILLSFILE1,PATHTOMYHILLSFILE2,PATHTOMYHILLSFILE3
```

if you want to integrate out some variable you do

```
plumed sum_hills --hills PATHTOMYHILLSFILE --idw t1 --kt 0.6
```

where with `--idw` you define the variables that you want all the others will be integrated out. `--kt` defines the temperature of the system in energy units. (be consistent with the units you have in your hills: `plumed` will not check this for you) If you need more variables then you may use a comma separated syntax

```
plumed sum_hills --hills PATHTOMYHILLSFILE --idw t1,t2 --kt 0.6
```

You can define the output grid only with the number of bins you want while min/max will be detected for you

```
plumed sum_hills --bin 99,99 --hills PATHTOMYHILLSFILE
```

or full grid specification

```
plumed sum_hills --bin 99,99 --min -pi,-pi --max pi,pi --hills PATHTOMYHILLSFILE
```

You can of course use numbers instead of `-pi/pi`.

You can use a `--stride` keyword to have a dump each bunch of hills you read

```
plumed sum_hills --stride 300 --hills PATHTOMYHILLSFILE
```

You can also have, in case of welltempered metadynamics, only the negative bias instead of the free energy through the keyword `--negbias`

```
plumed sum_hills --negbias --hills PATHTOMYHILLSFILE
```

Here the default name will be `negativebias.dat`

From time to time you might need to use HILLS or a COLVAR file as it was just a simple set of points from which you want to build a free energy by using  $-(1/\beta)\log(P)$  then you use `--histo`

```
plumed sum_hills --histo PATHTOMYCOLVARORHILLSFILE --sigma 0.2,0.2 --kt 0.6
```

in this case you need a `--kt` to do the reweighting and then you need also some width (with the `--sigma` keyword) for the histogram calculation (actually will be done with gaussians, so it will be a continuous histogram) Here the default output will be `histo.dat`. Note that also here you can have multiple input files separated by a comma.

Additionally, if you want to do histogram and hills from the same file you can do as this

```
plumed sum_hills --hills --histo PATHTOMYCOLVARORHILLSFILE --sigma 0.2,0.2 --kt 0.6
```

The two files can be eventually the same

Another interesting thing one can do is monitor the difference in blocks as a metadynamics goes on. When the bias deposited is constant over the whole domain one can consider to be at convergence. This can be done with the `--nohistory` keyword

```
plumed sum_hills --stride 300 --hills PATHTOMYHILLSFILE --nohistory
```

and similarly one can do the same for an histogram file

```
plumed sum_hills --histo PATHTOMYCOLVARORHILLSFILE --sigma 0.2,0.2 --kt 0.6 --nohistory
```

just to check the hypothetical free energy calculated in single blocks of time during a simulation and not in a cumulative way

Output format can be controlled via the `--fmt` field

```
plumed sum_hills --hills PATHTOMYHILLSFILE --fmt %8.3f
```

where here we chose a float with length of 8 and 3 digits

The output can be named in a arbitrary way :

```
plumed sum_hills --hills PATHTOMYHILLSFILE --outfile myfes.dat
```

will produce a file `myfes.dat` which contains the free energy.

If you use stride, this keyword is the suffix

```
plumed sum_hills --hills PATHTOMYHILLSFILE --outfile myfes_ --stride 100
```

will produce `myfes_0.dat`, `myfes_1.dat`, `myfes_2.dat` etc.

The same is true for the output coming from histogram

```
plumed sum_hills --histo HILLS --kt 2.5 --sigma 0.01 --outhisto myhisto.dat
```

is producing a file `myhisto.dat` while, when using stride, this is the suffix

```
plumed sum_hills --histo HILLS --kt 2.5 --sigma 0.01 --outhisto myhisto_ --stride 100
```

that gives `myhisto_0.dat`, `myhisto_1.dat`, `myhisto_3.dat` etc..

## Chapter 9

# Miscellaneous

- [Comments](#)
- [Continuation lines](#)
- [Using VIM syntax file](#)
- [Including other files](#)
- [Loading shared libraries](#)
- [Debugging the code](#)
- [Changing exchange patterns in replica exchange](#)
- [List of modules](#)
- [Frequently used tools](#)

### 9.1 Comments

If you are an organised sort of person who likes to remember what the hell you were trying to do when you ran a particular simulation you might find it useful to put comments in your input file. In PLUMED you can do this as comments can be added using a # sign. On any given line everything after the # sign is ignored so erm... yes add lines of comments or trailing comments to your hearts content as shown below (using Shakespeare is optional):

```
# This is the distance between two atoms:
DISTANCE ATOM=1,2 LABEL=d1
UPPER_WALLS ARG=d1 AT=3.0 KAPPA=3.0 LABEL=Snout # In this same interlude it doth befall.
# That I, one Snout by name, present a wall.
```

(see [DISTANCE](#) and [UPPER\\_WALLS](#))

An alternative to including comments in this way is to use line starting ENDPLUMED. Everything in the PLUMED input after this keyword will be ignored.

## 9.2 Continuation lines

If your input lines get very long then editing them using vi and other such text editors becomes a massive pain in the arse.

We at PLUMED are aware of this fact and thus have provided a way of doing line continuations so as to make your life that much easier - aren't we kind? Well no not really, we have to use this code too. Anyway, you can do continuations by using the "..." syntax as this makes this:

```
DISTANCES ATOMS1=1,300 ATOMS2=1,400 ATOMS3=1,500 LABEL=dist
```

(see [DISTANCES](#))

equivalent to this:

```
DISTANCES ...
  LABEL=dist
# we can also insert comments here
  ATOMS1=1,300
# multiple keywords per line are allowed
  ATOMS2=1,400 ATOMS3=1,500
#empty lines are also allowed

... DISTANCES
```

Notice that the closing ... is followed by the word DISTANCES. This is optional, but might be useful to find more easily which is the matching start of the statement. The following is equally correct

```
DISTANCES ...
  LABEL=dist
# we can also insert comments here
  ATOMS1=1,300
# multiple keywords per line are allowed
  ATOMS2=1,400 ATOMS3=1,500
#empty lines are also allowed

...
```

Notice that PLUMED makes a check that the word following the closing ... is actually identical to the first word in the line with the first ... If not, it will throw an error. Also notice that you might put more than one word in the first line. E.g.

```
DISTANCES LABEL=dist ...
# we can also insert comments here
  ATOMS1=1,300
# multiple keywords per line are allowed
  ATOMS2=1,400 ATOMS3=1,500
#empty lines are also allowed

...
```

or, equivalently,

```
dist: DISTANCES ...
# we can also insert comments here
  ATOMS1=1,300
# multiple keywords per line are allowed
  ATOMS2=1,400 ATOMS3=1,500
#empty lines are also allowed

...
```



## 9.3 Using VIM syntax file

For the impatient:

- Add the following to your .vimrc file:

```
" Enable syntax
:syntax on
" This allows including the proper PLUMED syntax file:
:let &runtimepath.=',$$PLUMED_VIMPATH
" The former command requires PLUMED_VIMPATH to be set. Alternatively, use this:
" let &runtimepath.=',$$usr/local/lib/plumed/vim'
" properly adjusted to the path where PLUMED is installed.
" This makes autocompletion work in the expected way:
:set completeopt=longest,menuone
" This enables bindings of F2/F3/F4 to plumed specific commands:
:let plumed_shortcuts=1
```

- When you open a PLUMED input file, you can enable syntax highlighting with:

```
:set ft=plumed
```

This will also enable autocompletion. Use <CTRL-X><CTRL-O> to autocomplete a word.

- If you want to fold multiline statements, type

```
:setlocal foldmethod=syntax
```

- While editing a plumed input file, you can use command :PHelp (or shortcut <F2>) to show in a split window a short help about the action defined in the line where the cursor is. Typing :PHelp again (or pushing <F2>) you will close that window. With <CTRL-W><CTRL-W> you go back and forth between the two windows.
- When you open a file starting with #! FIELDS, VIM will automatically understand it is a PLUMED output file (VIM filetype = plumedf) and will color fields and data columns with alternating colors. Typing :PPlus and :PMinus (or pushing <F3> and <F4>) you can move a highlighted column.

See below for more detailed instructions.

### Configuration

When PLUMED is compiled, directories `help` and `syntax` will appear in `builddir/vim`. They contain a VIM plugin that can be used to highlight proper PLUMED instructions in a PLUMED input file and to quickly retrieve help. There is also a file `builddir/vim/scripts.vim` that helps VIM in recognizing PLUMED output files.

### Warning

Notice that these files do not appear if you are cross compiling. In this case, you must copy the plugin files from another machine.

To make VIM aware of these files, you should copy them to your `$HOME/.vim` directory. Later you can enable plumed syntax with the command

```
:set ft=plumed
```

If you work in an environment where several PLUMED versions are installed (e.g. using env modules), we recommend the following procedure:

- Install PLUMED
- Add to your `.vimrc` file the following line:

```
:let &runtimepath.=',$$PLUMED_VIMPATH
```

The modulefile provided with PLUMED should set the `PLUMED_VIMPATH` environment variable to the proper path. Thus, when working with a given PLUMED module loaded, you should be able to enable proper syntax by just typing

```
:set ft=plumed
```

in VIM. Notice that the variable `PLUMED_VIMPATH` is also set in the `sourceme.sh` script in the build directory. Thus, if you modify your `.vimrc` file as suggested, you will be able to use the correct syntax both when using an installed PLUMED and when running from a just compiled copy. Finally, in case you have both a preinstalled PLUMED **and** you have your development version the following command would give you the optimal flexibility:

```
:let &runtimepath.=',$$PLUMED_VIMPATH.',/opt/local/lib/plumed/vim/'
```

The environment variable `PLUMED_VIMPATH`, if set, will take the precedence. Otherwise, vim will resort to the hardcoded path. In this case we assumed that there is a PLUMED installed in `/opt/local/` (e.g. using MacPorts), but you can override it sourcing a `sourceme.sh` file in the compilation directory or loading a PLUMED module with `module load plumed`.

If you are tired of typing `:set ft=plumed`, you can use a modeline. Add to your `.vimrc` file the following commands

```
:set modeline
:set modelines=5
```

Then, at the beginning of your PLUMED input file, put the following comment:

```
# vim:ft=plumed
d: DISTANCE ATOMS=1,2
RESTRAINT ARG=d AT=0.0 KAPPA=1.0
```

Now, every time you open this file, you will see it highlighted.

### Syntax highlighting

The syntax file contains a definition of all possible PLUMED actions and keywords. It is designed to allow for a quick validation of the PLUMED input file before running it. As such, all the meaningful words in the input should be highlighted:

- Valid action names (such as `METAD`) and labels (such as `metad:` or `LABEL=metad`) will be highlighted in the brightest way (Type in VIM). Those are the most important words.
- Keyword and flag names (such as `ATOMS=` or `COMPONENTS` when part of the action `DISTANCE`) will be highlighted with a different color (Statement in VIM).
- Values provided by users (such as the number of the atoms following `ATOMS=`) will be highlighted with a different color (String in VIM).
- Comments (see [Comments](#)) will be highlighted as comments (Comment in VIM).

If you see something that is not highlighted and appears in black, this is likely going to result in an error at runtime. Think of this as a sort of preliminary spell-check. For this checks to be effective, we recommend to use a syntax file generated with exactly the same version of PLUMED that you are using. In case you find that parts of an input file that is valid are not highlighted, then please report it as a bug. On the contrary, you cannot expect the VIM syntax file to recognize all possible errors in a PLUMED input. Thus, a file for which the highlighting looks correct might still contain errors.

### Multi-line folding

Notice that syntax highlighting also allow VIM to properly fold multi-line actions. Try to do the following:

- Open a PLUMED input file
- Enable PLUMED syntax

```
:set ft=plumed
```

- Enable syntax-based folding

```
:setlocal foldmethod=syntax
```

Now look at what happened to all the multi-line statements in PLUMED (i.e. those using [Continuation lines](#)). As you can see, they will be folded into single lines. Folded lines can be expanded with `zo` and folded with `zc`. Look at VIM documentation to learn more. In case you want to use this feature, we suggest you to put both label and action type on the first line of multi-line statements. E.g.

```
m: METAD ...
  ARG=d
  HEIGHT=1.0
  SIGMA=0.5
  PACE=100
...
```

will be folded to

```
+-- 6 lines: m: METAD ...-----
```

and

```
METAD LABEL=m ...
  ARG=d
  HEIGHT=1.0
  SIGMA=0.5
  PACE=100
...
```

will be folded to

```
+-- 6 lines: METAD LABEL=m ...-----
```

This will allow you to easily identify the folded lines by seeing the most important information, that is the action type (METAD) and its label (m). This feature is convenient if you want to browse files that contain a lot of actions defined on multiple lines.

## Autocompletion

Another VIM feature that comes when you load PLUMED syntax is autocompletion of PLUMED actions and keywords. Open your favorite PLUMED input file and set it to PLUMED syntax highlighting with

```
:set ft=plumed
```

Now go into insert mode pressing `i` and type `DU` followed by `<CTRL+X><CTRL+O>`. Here `<CTRL+X>` stands for autocompletion and `<CTRL+O>` for omnifunc autocompletion. You will see a short menu listing the following actions

```
DUMPATOMS
DUMPPERIVATIVES
DUMPFORCES
DUMPMASSCHARGE
DUMPMULTICOLVAR
DUMPPROJECTIONS
```

That is, all the actions starting with `DU`. You can navigate it with up and down arrows so as to choose the best match.

Notice that the default behavior of VIM is to use the first match by default. In the first example (`DU<CTRL+X><CTRL+O>`), it would be `DUMPATOMS`. The following settings make it work as most of the people expect:

```
:set completeopt=longest,menuone
```

With these settings, in the first example (`DU<CTRL+X><CTRL+O>`) VIM will only complete up to the longest common part (`DUMP`).

As you can imagine, if you use autocompletion after you have typed the word `DISTANCE` followed by a space you will see a menu listing `LABEL=`, `COMPONENTS`, etc. Basically, all the keywords that are possibly used within a `DISTANCE` line will be shown. This is very useful if you do not remember the exact name of the keywords associated with a given action.

## Quick help

You can also retrieve quick explanation of the input options for a specific action. Try to do the following. Enable plumed syntax:

```
:set ft=plumed
```

Then add the following line

```
DISTANCE
```

Now, in normal mode, go with the cursor on the `DISTANCE` line and type

```
:PHeLp
```

A new split window should appear containing some documentation about the `DISTANCE` collective variable. You can go back and forth between the two windows with `<CTRL+W><CTRL+W>`, as usually in vim. Notice that if you are in the help window and type `:PHeLp` this window will be closed.

To make the navigation easier, you can add a shortcut in your `.vimrc` file. For example, adding:

```
: nmap <F2> : PHelp<CR>
```

you should be able to open and close the manual hitting the F2 key. This is done automatically in the PLUMED syntax file if you add `let plumed_shortcuts=1` to your vimrc file.

### Displaying output files

Most of the PLUMED output files look like this

```
#! FIELDS A B C
1 2 3
```

This is useful since in the header you can see the name of the quantities that are printed in the data lines. However, when you have an output file with many columns it might be a bit error prone to count them. To simplify this, when PLUMED syntax for VIM is configured properly VIM should be able to:

- Detect that this file is a PLUMED output file with fields, automatically setting its type to `plumedf`. If not, just type `:set ft=plumedf`.
- Show this file with syntax highlighting to increase its readability.

Notice that the syntax file for the output files (`plumedf.vim`) is not the same one that is used for the PLUMED input file (`plumed.vim`).

To make output files more readable, vim will show `FIELDS` and `SET` words in a different color, and data columns with alternating colors (e.g. dark/light/dark/light). The colors in the columns are consistent with those shown in the `FIELD` line. In the example above, 1, 2, and 3 will be of the same color as A, B, and C respectively. This should make it much easier to find which columns correspond to a given quantity.

It is also possible to highlight a specific field of the file. Typing

```
:5PCol
```

you will highlight the fifth field. Notice that in the `FIELDS` line (the first line of the file) the 7th word of the line will be highlighted, which is the one containing the name of the field. This allows for easy matching of values shown in the file and tags provided in the `FIELDS` line. The highlighted column can be moved back and forth using `:PPlus` and `:PMinus`. Adding a count to the command will move the highlighted column more. E.g. `:2PPlus` will move the column to the right twice.

If you have a long output file, it might be convenient to split it with `:split` so that one of the two windows will only show the header. The other window can be used to navigate the file.

To make the navigation easier, you can add a shortcut in your `.vimrc` file. For example, adding:

```
: map <F3> :PMinus<CR>
: map <F4> :PPlus<CR>
```

you should be able to move the highlight column using F3 and F4 buttons. This is done automatically in the PLUMED syntax file if you add `let plumed_shortcuts=1` to your vimrc file.

## 9.4 Including other files

If, for some reason, you want to spread your PLUMED input over a number of files you can use [INCLUDE](#) as shown below:

```
INCLUDE FILE=filename
```

So, for example, a single "plumed.dat" file:

```
DISTANCE ATOMS=0,1 LABEL=dist
RESTRAINT ARG=dist
```

(see [DISTANCE](#) and [RESTRAINT](#))

could be split up into two files as shown below:

```
DISTANCE ATOMS=0,1 LABEL=dist
INCLUDE FILE=toBeIncluded.dat
```

plus a "toBeIncluded.dat" file

```
RESTRAINT ARG=dist
```

However, when you do this it is important to recognise that [INCLUDE](#) is a real directive that is only resolved after all the [Comments](#) have been stripped and the [Continuation lines](#) have been unrolled. This means it is not possible to do things like:

```
# this is wrong:
DISTANCE INCLUDE FILE=options.dat
RESTRAINT ARG=dist
```

### 9.4.1 INCLUDE

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

Includes an external input file, similar to "#include" in C preprocessor.

Useful to split very large plumed.dat files.

#### Compulsory keywords

<b>FILE</b>	file to be included
-------------	---------------------

#### Examples

This input

```
c1: COM ATOMS=1-100
c2: COM ATOMS=101-202
d: DISTANCE ATOMS=c1,c2
PRINT ARG=d
```

can be replaced with

```
INCLUDE FILE=pippo.dat
d: DISTANCE ATOMS=c1,c2
PRINT ARG=d
```

where the content of file pippo.dat is

```
c1: COM ATOMS=1-100
c2: COM ATOMS=101-202
```

(see also [COM](#), [DISTANCE](#), and [PRINT](#)).

## 9.5 Loading shared libraries

You can introduce new functionality into PLUMED by placing it directly into the src directory and recompiling the PLUMED libraries. Alternatively, if you want to keep your code independent from the rest of PLUMED (perhaps so you can release it independently - we won't be offended), then you can create your own dynamic library. To use this in conjunction with PLUMED you can then load it at runtime by using the [LOAD](#) keyword as shown below:

```
LOAD FILE=library.so
```

N.B. If your system uses a different suffix for dynamic libraries (e.g. macs use .dylib) then PLUMED will try to automatically adjust the suffix accordingly.

### 9.5.1 LOAD

This is part of the setup <a href="#">module</a>
--------------------------------------------------

Loads a library, possibly defining new actions.

It is available only on systems allowing for dynamic loading. It can also be fed with a cpp file, in which case the file is compiled first.

**Compulsory keywords**

<b>FILE</b>	file to be loaded
-------------	-------------------

## Examples

If you have a shared object named `extensions.so` and want to use the functionalities implemented in it within PLUMED you can load it with the following syntax

```
LOAD FILE=extensions.so
```

As a more practical example, imagine that you want to make a small change to one collective variable that is already implemented in PLUMED, say `DISTANCE`. Copy the file `src/colvar/Distance.cpp` into your work directory, rename it as `Distance2.cpp` and edit it as you wish. It might be better to also replace any occurrence of the string `DISTANCE` within the file with `DISTANCE2`, so that both old and new implementation will be available with different names. Then you can compile it into a shared object using

```
> plumed mklib Distance2.cpp
```

This will generate a file `Distance2.so` (or `Distance2.dylib` on a mac) that can be loaded. Now you can use your new implementation with the following input

```
# load the new library
LOAD FILE=Distance2.so
# compute standard distance
d: DISTANCE ATOMS=1,10
# compute modified distance
d2: DISTANCE2 ATOMS=1,10
# print them on a file
PRINT ARG=d,d2 FILE=compare-them
```

You can even skip the initial step and directly feed PLUMED with the `Distance2.cpp` file: it will be compiled on the fly.

```
# load the new definition
# this is a cpp file so it will be compiled
LOAD FILE=Distance2.cpp
# compute standard distance
d: DISTANCE ATOMS=1,10
# compute modified distance
d2: DISTANCE2 ATOMS=1,10
# print them on a file
PRINT ARG=d,d2 FILE=compare-them
```

This will allow to make quick tests while developing your own variables. Of course, after your implementation is ready you might want to add it to the PLUMED source tree and recompile the whole PLUMED.

## 9.6 Debugging the code

The `DEBUG` action provides some functionality for debugging the code that may be useful if you are doing very intensive development of the code or if you are running on a computer with a strange architecture.

### 9.6.1 DEBUG

	This is part of the generic <code>module</code>
--	-------------------------------------------------



Set some debug options.

Can be used while debugging or optimizing plumed.

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) the frequency with which this action is to be performed
---------------	-----------------------------------------------------------------------

#### Options

<b>logActivity</b>	( default=off ) write in the log which actions are inactive and which are inactive
<b>logRequestedAtoms</b>	( default=off ) write in the log which atoms have been requested at a given time
<b>NOVIRIAL</b>	( default=off ) switch off the virial contribution for the entirety of the simulation
<b>DETAILED_TIMERS</b>	( default=off ) switch on detailed timers
<b>FILE</b>	the name of the file on which to output these quantities

#### Examples

```
# print detailed (action-by-action) timers at the end of simulation
DEBUG DETAILED_TIMERS
# dump every two steps which are the atoms required from the MD code
DEBUG logRequestedAtoms STRIDE=2
```

## 9.7 Changing exchange patterns in replica exchange

Using the [RANDOM\\_EXCHANGES](#) keyword it is possible to make exchanges between randomly chosen replicas. This is useful e.g. for bias exchange metadynamics [56].

### 9.7.1 RANDOM\_EXCHANGES

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

Set random pattern for exchanges.

In this way, exchanges will not be done between replicas with consecutive index, but will be done using a random pattern. Typically used in bias exchange [56].

#### Options

<b>SEED</b>	seed for random exchanges
-------------	---------------------------

### Examples

Using the following three input files one can run a bias exchange metadynamics simulation using a different angle in each replica. Exchanges will be randomly tried between replicas 0-1, 0-2 and 1-2

Here is plumed.0.dat

```
RANDOM_EXCHANGES
t: TORSION ATOMS=1,2,3,4
METAD ARG=t HEIGHT=0.1 PACE=100 SIGMA=0.3
```

Here is plumed.1.dat

```
RANDOM_EXCHANGES
t: TORSION ATOMS=2,3,4,5
METAD ARG=t HEIGHT=0.1 PACE=100 SIGMA=0.3
```

Here is plumed.2.dat

```
RANDOM_EXCHANGES
t: TORSION ATOMS=3,4,5,6
METAD ARG=t HEIGHT=0.1 PACE=100 SIGMA=0.3
```

### Warning

Multi replica simulations are presently only working with gromacs.  
The directive should appear in input files for every replicas. In case SEED is specified, it should be the same in all input files.

## 9.8 List of modules

The functionality in PLUMED 2 is divided into a small number of modules. Some users may only wish to use a subset of the functionality available within the code while others may wish to use some of PLUMED's more complicated features. For this reason the plumed source code is divided into modules, which users can activate or deactivate to their hearts content.

You can activate a module at configure time using the keyword `--enable-modules`. For example:

```
./configure --enable-modules=modulename
```

will enable module called modulename. A module that is on by default can be disabled using the following syntax

```
./configure --enable-modules=-modulename
```

To enable or disable multiple modules one should provide them as a : separated list. Notice that `+modulename` and `modulename` both activate the module, whereas `-modulename` deactivates it. E.g.

```
./configure --enable-modules=+crystallization:-colvar
```

will disable the colvar module and enable the crystallization module. Also notice that `:` can be omitted when using `+` or `-`. Thus, the same can be obtained with

```
./configure --enable-modules=+crystallization-colvar
```

If you repeat the `--enable-modules` keyword only the last instance will be used. Thus `./configure --enable-modules=crystallization --enable-modules=-colvar` will *not* do what you expect!

There are also some shortcuts available:

- `./configure --enable-modules=all` to enable all optional modules. This includes the maximum number of features in PLUMED, including modules that might not be properly functional.
- `./configure --enable-modules=none` or `./configure --disable-modules` to disable all optional modules. This produces a minimalistic PLUMED which can be used as a library but has no command line tools and no collective variables or biasing methods.
- `./configure --enable-modules=reset` or `./configure --enable-modules` to enable the default modules.

The two kinds of syntax can be combined and, for example, `./configure --enable-modules=none --enable-modules=:colvar` will result in a PLUMED with all the modules disabled with the exception of the colvar module.

Some modules are active by default in the version of PLUMED 2 that you download from the website while others are inactive. The following lists all of the modules that are available in plumed and tells you whether or not they are active by default.

Module name	Default behavior
adjmat	off
analysis	on
bias	on
cltools	on
colvar	on
crystallization	off
function	on
generic	on
manyrestraints	off
mapping	on
molfile	on
multicolvar	on
secondarystructure	on
setup	on
vatom	on

Until PLUMED 2.2, it was also possible to switch on or off modules by adding files in the `plumed2/src` directory. Since PLUMED 2.3 this is discouraged, since any choice made in this manner will be overwritten next time `./configure` is used.

## 9.9 Frequently used tools

<a href="#">histogrambead</a>	INTERNAL	A function that can be used to calculate whether quantities are between fixed upper and lower bounds.
<a href="#">kernelfunctions</a>	INTERNAL	Functions that are used to construct histograms
<a href="#">landmarkselection</a>	INTERNAL	This is currently a filler page.
<a href="#">pdbreader</a>	INTERNAL	PLUMED can use the PDB format in several places
<a href="#">switchingfunction</a>	INTERNAL	Functions that measure whether values are less than a certain quantity.

<a href="#">Regular Expressions</a>		POSIX regular expressions can be used to select multiple actions when using ARG (i.e. <a href="#">PRINT</a> ).
<a href="#">Files</a>		Dealing with Input/Output

### 9.9.1 histogrambead

A function that can be used to calculate whether quantities are between fixed upper and lower bounds. A function that can be used to calculate whether quantities are between fixed upper and lower bounds.

If we have multiple instances of a variable we can estimate the probability distribution (pdf) for that variable using a process called kernel density estimation:

$$P(s) = \sum_i K\left(\frac{s - s_i}{w}\right)$$

In this equation  $K$  is a symmetric function that must integrate to one that is often called a kernel function and  $w$  is a smearing parameter. From a pdf calculated using kernel density estimation we can calculate the number/fraction of values between an upper and lower bound using:

$$w(s) = \int_a^b \sum_i K\left(\frac{s - s_i}{w}\right)$$

All the input to calculate a quantity like  $w(s)$  is generally provided through a single keyword that will have the following form:

KEYWORD={TYPE UPPER=  $a$  LOWER=  $b$  SMEAR=  $\frac{w}{b-a}$ }

This will calculate the number of values between  $a$  and  $b$ . To calculate the fraction of values you add the word NORM to the input specification. If the function keyword SMEAR is not present  $w$  is set equal to  $0.5(b - a)$ . Finally, type should specify one of the kernel types that is present in plumed. These are listed in the table below:

TYPE	FUNCTION
GAUSSIAN	$\frac{1}{\sqrt{2\pi}w} \exp\left(-\frac{(s-s_i)^2}{2w^2}\right)$
TRIANGULAR	$\frac{1}{2w} \left(1 - \left \frac{s-s_i}{w}\right \right) \quad \frac{s-s_i}{w} < 1$

Some keywords can also be used to calculate a discretized version of the histogram. That is to say the number of values between  $a$  and  $b$ , the number of values between  $b$  and  $c$  and so on. A keyword that specifies this sort of calculation would look something like

KEYWORD={TYPE UPPER=  $a$  LOWER=  $b$  NBINS=  $n$  SMEAR=  $\frac{w}{n(b-a)}$ }

This specification would calculate the following vector of quantities:

$$w_j(s) = \int_{a+\frac{j-1}{n}(b-a)}^{a+\frac{j}{n}(b-a)} \sum_i K\left(\frac{s-s_i}{w}\right)$$

### 9.9.2 kernelfunctions

Functions that are used to construct histograms Functions that are used to construct histograms

Constructing histograms is something you learnt to do relatively early in life. You perform an experiment a number of times, count the number of times each result comes up and then draw a bar graph that describes how often each of the results came up. This only works when there are a finite number of possible results. If the result a number between 0 and 1 the bar chart is less easy to draw as there are as many possible results as there are numbers between zero and one - an infinite number. To resolve this problem we replace probability,  $P$  with probability density,  $\pi$ , and write the probability of getting a number between  $a$  and  $b$  as:

$$P = \int_a^b dx \pi(x)$$

To calculate probability densities from a set of results we use a process called kernel density estimation. Histograms are accumulated by adding up kernel functions,  $K$ , with finite spatial extent, that integrate to one. These functions are centered on each of the  $n$ -dimensional data points,  $\mathbf{x}_i$ . The overall effect of this is that each result we obtain in our experiments contributes to the probability density in a finite sized region of the space.

Expressing all this mathematically in kernel density estimation we write the probability density as:

$$\pi(\mathbf{x}) = \sum_i K[(\mathbf{x} - \mathbf{x}_i)^T \Sigma (\mathbf{x} - \mathbf{x}_i)]$$

where  $\Sigma$  is an  $n \times n$  matrix called the bandwidth that controls the spatial extent of the kernel. Whenever we accumulate a histogram (e.g. in [HISTOGRAM](#) or in [METAD](#)) we use this technique.

There is thus some flexibility in the particular function we use for  $K[\mathbf{r}]$  in the above. The following variants are available.

TYPE	FUNCTION
gaussian	$f(r) = \frac{1}{(2\pi)^n \sqrt{ \Sigma^{-1} }} \exp(-0.5r^2)$
truncated-gaussian	$f(r) = \frac{1}{(2\pi)^n \sqrt{ \Sigma^{-1} } \left( \frac{(-6.25/sqrt{2}) - (-6.25/sqrt{2})}{2} \right)^n} \exp(-0.5r^2)$
triangular	$f(r) = \frac{3}{V} (1 -  r ) H(1 -  r )$
uniform	$f(r) = \frac{1}{V} H(1 -  r )$

In the above  $H(y)$  is a function that is equal to one when  $y > 0$  and zero when  $y \leq 0$ .  $n$  is the dimensionality of the vector  $\mathbf{x}$  and  $V$  is the volume of an ellipse in an  $n$  dimensional space which is given by:

$$V = |\Sigma^{-1}|^{\frac{n}{2}} \left( \frac{n}{2} \right)! \quad \text{for even } n$$

$$V = |\Sigma^{-1}| \frac{2^{\frac{n+1}{2}} \pi^{\frac{n-1}{2}}}{n!!}$$

In [METAD](#) the normalization constants are ignored so that the value of the function at  $r = 0$  is equal to one. In addition in [METAD](#) we must be able to differentiate the bias in order to get forces. This limits the kernels we can use in this method. Notice also that Gaussian kernels should have infinite support. When used with grids, however, they are assumed to only be non-zero over a finite range. The difference between the truncated-gaussian and regular gaussian is that the truncated gaussian is scaled so that its integral over the grid is equal to one when it is normalised. The integral of a regular gaussian when it is evaluated on a grid will be slightly less than one because of the truncation of a function that should have infinite support.

### 9.9.3 landmarkselection

	This is part of the analysis <a href="#">module</a>
--	-----------------------------------------------------

This is currently a filler page. This is currently a filler page.

Just use LANDMARKS=ALL. More complex versions will appear in later versions.

### 9.9.4 pdbreader

PLUMED can use the PDB format in several places PLUMED can use the PDB format in several places

- To read molecular structure ([MOLINFO](#)).
- To read reference conformations ([RMSD](#), but also many other methods in [Distances from reference configurations](#), [FIT\\_TO\\_TEMPLATE](#), etc).

The implemented PDB reader expects a file formatted correctly according to the [PDB standard](#). In particular, the following columns are read from ATOM records

columns	content
1-6	record name (ATOM or HETATM)
7-11	serial number of the atom (starting from 1)
13-16	atom name
18-20	residue name
22	chain id
23-26	residue number
31-38	x coordinate
39-46	y coordinate
47-54	z coordinate
55-60	occupancy
61-66	beta factor

PLUMED parser is slightly more permissive than the official PDB format in the fact that the format of real numbers is not fixed. In other words, any parsable real number is ok and the dot can be placed anywhere. However, **columns are interpret strictly**. A sample PDB should look like the following

ATOM	2	CH3	ACE	1	12.932	-14.718	-6.016	1.00	1.00
ATOM	5	C	ACE	1	21.312	-9.928	-5.946	1.00	1.00
ATOM	9	CA	ALA	2	19.462	-11.088	-8.986	1.00	1.00

Notice that serial numbers need not to be consecutive. In the three-line example above, only the coordinates of three atoms are provided. This is perfectly legal and indicates PLUMED that information about these atoms only is available. This could be both for structural information in [MOLINFO](#), where the other atoms would have no name assigned, and for reference structures used in [RMSD](#), where only the provided atoms would be used to compute RMSD.

### Occupancy and beta factors

PLUMED reads also occupancy and beta factors that however are given a very special meaning. In cases where the PDB structure is used as a reference for an alignment (that's the case for instance in [RMSD](#) and in [FIT\\_TO\\_TEMPLATE](#)), the occupancy column is used to provide the weight of each atom in the alignment. In cases where, perhaps after alignment, the displacement between running coordinates and the provided PDB is computed, the beta factors are used as weight for the displacement. Since setting the weights to zero is the same as **not** including an atom in the alignment or displacement calculation, the two following reference files would be equivalent when used in an [RMSD](#) calculation. First file:

```
ATOM      2  CH3  ACE      1      12.932 -14.718  -6.016  1.00  1.00
ATOM      5   C   ACE      1      21.312  -9.928  -5.946  1.00  1.00
ATOM      9  CA   ALA      2      19.462 -11.088  -8.986  0.00  0.00
```

Second file:

```
ATOM      2  CH3  ACE      1      12.932 -14.718  -6.016  1.00  1.00
ATOM      5   C   ACE      1      21.312  -9.928  -5.946  1.00  1.00
```

However notice that many extra atoms with zero weight might slow down the calculation, so removing lines is better than setting their weights to zero. In addition, weights for alignment need not to be equivalent to weights for displacement.

### 9.9.5 switchingfunction

Functions that measure whether values are less than a certain quantity. Functions that measure whether values are less than a certain quantity.

Switching functions  $s(r)$  take a minimum of one input parameter  $d_0$ . For  $r \leq d_0$   $s(r) = 1.0$  while for  $r > d_0$  the function decays smoothly to 0. The various switching functions available in plumed differ in terms of how this decay is performed.

Where there is an accepted convention in the literature (e.g. [COORDINATION](#)) on the form of the switching function we use the convention as the default. However, the flexibility to use different switching functions is always present generally through a single keyword. This keyword generally takes an input with the following form:

```
KEYWORD={TYPE <list of parameters>}
```

The following table contains a list of the various switching functions that are available in plumed 2 together with an example input.

TYPE	FUNCTION	EXAMPLE INPUT	DEFAULT PARAMETERS
RATIONAL	$s(r) = \frac{1 - \left(\frac{r-d_0}{r_0}\right)^n}{1 - \left(\frac{r-d_0}{r_0}\right)^m}$	{RATIONAL R_0= $r_0$ D_0= $d_0$ NN= $n$ MM= $m$ }	$d_0 = 0.0, n = 6, m = 2n$
EXP	$s(r) = \exp\left(-\frac{r-d_0}{r_0}\right)$	{EXP R_0= $r_0$ D_0= $d_0$ }	$d_0 = 0.0$
GAUSSIAN	$s(r) = \exp\left(-\frac{(r-d_0)^2}{2r_0^2}\right)$	{GAUSSIAN R_0= $r_0$ D_0= $d_0$ }	$d_0 = 0.0$
SMAP	$s(r) = \frac{1}{\left[1 + (2^{a/b} - 1) \left(\frac{r-d_0}{r_0}\right)^a\right]^{-b/a}}$	{SMAP R_0= $r_0$ D_0= $d_0$ A= $a$ B= $b$ }	$d_0 = 0.0$
Q	$s(r) = \frac{1}{1 + \exp(\beta(r_{ij} - \lambda r_{ij}^0))}$	{Q REF= $r_{ij}^0$ BETA= $\beta$ LAM= $\lambda$ BDA= $\lambda$ }	$\lambda = 1.8, \beta = 50nm^{-1}$ (all-atom) $\lambda = 1.5, \beta = 50nm^{-1}$ (coarse-grained)
CUBIC	$s(r) = (y - 1)^2(1 + 2y)$ where $y = \frac{r-r_1}{r_0-r_1}$	{CUBIC D_0= $r_1$ D_MAX= $r_0$ }	
TANH	$s(r) = 1 - \tanh\left(\frac{r-d_0}{r_0}\right)$	{TANH R_0= $r_0$ D_0= $d_0$ }	
MATHEVAL	$s(r) = FUNC$	{MATHEVAL FU= $FUNC$ NC=1/(1+x^6) R_0= $r_0$ D_0= $d_0$ }	

#### Attention

Similarly to the [MATHEVAL](#) function, the MATHEVAL switching function only works if libmatheval is installed on the system and PLUMED has been linked to it. Also notice that using MATHEVAL is much slower than using e.g. RATIONAL. Thus, the MATHEVAL switching function is useful to perform quick tests on switching functions with arbitrary form before proceeding to their implementation in C++.

For all the switching functions in the above table one can also specify a further (optional) parameter using the parameter keyword D\_MAX to assert that for  $r > d_{\max}$  the switching function can be assumed equal to zero. In this case the function is brought smoothly to zero by stretching and shifting it.

```
KEYWORD={RATIONAL R_0=1 D_MAX=3}
```

the resulting switching function will be  $s(r) = \frac{s'(r) - s'(d_{\max})}{s'(0) - s'(d_{\max})}$  where  $s'(r) = \frac{1-r^6}{1-r^{12}}$ . Since PLUMED 2.2 this is the default. The old behavior (no stretching) can be obtained with the NOSTRETCH flag. The NOSTRETCH keyword is only provided for backward compatibility and might be removed in the future. Similarly, the STRETCH keyword is still allowed but has no effect.

Notice that switching functions defined with the simplified syntax are never stretched for backward compatibility. This might change in the future.

### 9.9.6 Regular Expressions

When you use a collective variable that has many calculated components and you want to refer to them as arguments you can use regular expressions.

Since version 2.1, plumed takes advantage of a configuration scripts that detects libraries installed on your system. If regex library is found, then you will be able to use regular expressions to refer to collective variables or function names.

Regular expressions are enclosed in round braces and must not contain spaces (the components names have no spaces indeed, so why use them?).

As an example then command



```
d1: DISTANCE ATOMS=1,2 COMPONENTS
PRINT ARG=(d1\[.xy]) STRIDE=100 FILE=colvar FMT=%8.4f
```

will cause both the d1.x and d1.y components of the DISTANCE action to be printed out in the order that they are created by plumed. The "." character must be escaped in order to interpret it as a literal ".". An unescaped dot is a wildcard which is matched by any character, So as an example

```
d1: DISTANCE ATOMS=1,2 COMPONENTS
dxy: DISTANCE ATOMS=1,3

# this will match d1.x,d1.y,dxy
PRINT ARG=(d1\[xy]) STRIDE=100 FILE=colvar FMT=%8.4f

# while this will match d1.x,d1.y only
PRINT ARG=(d1\[.xy]) STRIDE=100 FILE=colvar FMT=%8.4f
```

You can include more than one regular expression by using comma separated regular expressions

```
t1: TORSION ATOMS=5,7,9,15
t2: TORSION ATOMS=7,9,15,17
d1: DISTANCE ATOMS=7,17 COMPONENTS
PRINT ARG=(d1\[xy]),(t[0-9]) STRIDE=100 FILE=colvar FMT=%8.4f
```

(this selects t1,t2,d1.x and d2.x) Be aware that if you have overlapping selection they will be duplicated so it a better alternative is to use the "or" operator "|".

```
t1: TORSION ATOMS=5,7,9,15
t2: TORSION ATOMS=7,9,15,17
d1: DISTANCE ATOMS=7,17 COMPONENTS
PRINT ARG=(d1\[xy]|t[0-9]) STRIDE=100 FILE=colvar FMT=%8.4f
```

this selects the same set of arguments as the previous example.

You can check the log to see whether or not your regular expression is picking the set of components you desire.

For more information on regular expressions visit <http://www.regular-expressions.info/reference.html>.

### 9.9.7 Files

We tried to design PLUMED in such a manner that input/output is done consistently irrespectively of the file type. Most of the files written or read by PLUMED thus follow the very same conventions discussed below.

#### 9.9.7.1 Restart

Whenever the **RESTART** option is used, all the files written by PLUMED are appended. This makes it easy to analyze results of simulations performed as a chain of several sub-runs. Notice that most of the PLUMED textual files have a header. The header is repeated at every restart. Additionally, several files have time in the first column. PLUMED just takes the value of the physical time from the MD engine. As such, you could have that time starts again from zero upon restart or not.

An exception from this behavior is given by files which are not growing as the simulation proceeds. For example, grids written with **METAD** with **GRID\_WFILE** are overwritten by default during the simulation. As such, when restarting, there is no point in appending the file. Internally, PLUMED opens the file in append mode but then rewinds it every time a new grid is dumped.

### 9.9.7.2 Backup

Whenever the [RESTART](#) option is not used, PLUMED tries to write new files. If an old file is found in the way, PLUMED takes a backup named "bck.X.filename" where X is a progressive number. Notice that by default PLUMED only allows a maximum of 100 backup copies for a file. This behavior can be changed by setting the environment variable `PLUMED_MAXBACKUP` to the desired number of copies. E.g. `export PLUMED_MAXBACKUP=10` will fail after 10 copies. `PLUMED_MAXBACKUP=-1` will never fail - be careful since your disk might fill up quickly with this setting.

### 9.9.7.3 Replica suffix

When running with multiple replicas (e.g., with GROMACS, `-multi` option) PLUMED adds the replica index as a suffix to all the files. The following command will thus print files named `COLVAR.0`, `COLVAR.1`, etc for the different replicas.

```
d: DISTANCE ATOMS=1,2
PRINT ARG=d FILE=COLVAR
```

(see also [DISTANCE](#) and [PRINT](#)).

When reading a file, PLUMED will try to add the suffix. If the file is not found, it will fall back to the name without suffix. The most important case is the reading of the plumed input file. If you provide a file for each replica (e.g. `plumed.0.dat`, `plumed.1.dat`, etc) you will be able to setup plumed differently on each replica. On the other hand, using a single `plumed.dat` will make all the replicas read the same file.

#### Warning

This rule is true for almost all the files read by PLUMED. Current exceptions are PDB files, where the replica suffix is not added, and trajectories read by [driver](#), for which the replica suffix is always added.

Notice that when PLUMED adds the replica suffix, it recognizes the file extension and add the suffix *before* the extension. Before PLUMED 2.2, the only recognized suffix was ".gz". Since 2.2, any suffix with length less or equal to five letters is recognized.

This means that using in a multireplica context an input such as

```
d: DISTANCE ATOMS=1,2
PRINT ARG=d FILE=COLVAR.gz
METAD ARG=d FILE=test.HILLS SIGMA=0.1 HEIGHT=0.1
```

(see [DISTANCE](#), [PRINT](#), and [METAD](#)) PLUMED will write files named `COLVAR.0.gz`, `COLVAR.1.gz`, `test.0.HILLS`, `test.1.HILLS`, etc etc. This is useful since the preserved extension makes it easy to process the files later.

## Chapter 10

# Performances

In this page we collect hints on how to use the features available in PLUMED to speed up your calculations. Please note that PLUMED performs many different tasks, it can calculate a number of different collective variables, functions of collective variables, bias, on-the-fly analysis, etc in a way that is compatible with a number of different molecular dynamics codes. This means that there cannot be a single strategy to speed up all the possible calculations.

PLUMED makes use of MPI and OpenMP to parallelise some of its functions, try to always compile it with these features enabled. Furthermore, newer compilers with proper optimisation flags can provide a dramatic boost to performances.

PLUMED collects atoms from an external code and sends back forces, so it is key to minimise the effect of  $P \leftrightarrow$  PLUMED on highly parallel calculations to keep to the minimum the number of atoms used by PLUMED at every calculation step. The less is the number of atoms you need to send to PLUMED the less will be the overhead in the communication between PLUMED and the code.

In the following you can find specific strategies for specific calculations, these could help in taking the most by using PLUMED for your simulations.

- [GROMACS and PLUMED with GPU](#)
- [Metadynamics](#)
- [Multiple time stepping](#)
- [Multicolvar](#)
- [Neighbour Lists](#)
- [OpenMP](#)
- [Secondary Structure](#)
- [Time your Input](#)

## 10.1 GROMACS and PLUMED with GPU

Since version 4.6.x GROMACS can run in an hybrid mode making use of both your CPU and your GPU (either using CUDA or OpenCL for newer versions of GROMACS). The calculation of the short-range non-bonded interactions is performed on the GPU while long-range and bonded interactions are at the same time calculated on the CPU. By varying the cut-off for short-range interactions GROMACS can optimise the balance between GPU/CPU loading and obtain amazing performances.

GROMACS patched with PLUMED takes into account PLUMED in its load-balancing, adding the PLUMED timings to the one resulting from bonded interactions and long- range interactions. This means that the CPU/GPU balance will be optimised automatically to take into account PLUMED!

It is important to notice that the optimal setup to use GROMACS alone on the GPU or GROMACS + PLUMED can be different, try to change the number of MPI/OpenMP processes ([OpenMP](#)) used by GROMACS and PLUMED to find optimal performances. Remember that in GROMACS multiple MPI threads can use the same GPU:

i.e. if you have 4 cores and 2 GPU you can:

- use 2 MPI/2GPU/2OPENMP:

```
export PLUMED_NUM_THREADS=2
```

```
mpiexec -np 2 gmx_mpi mdrun -nb gpu -ntomp 2 -pin on -gpu_id 01
```

- use 4 MPI/2GPU:

```
export PLUMED_NUM_THREADS=1
```

```
mpiexec -np 4 gmx_mpi mdrun -nb gpu -ntomp 1 -pin on -gpu_id 0011
```

## 10.2 Metadynamics

Metadynamics can be sped up significantly using grids, which are activated setting the GRID\_MIN and GRID\_MAX keywords of [METAD](#). This makes addition of a hill to the list a bit slower (since the Gaussian has to be evaluated for many grid points) but the evaluation of the potential very fast. Since the former is usually done every few hundred steps, whereas the latter typically at every step, using grids will make the simulation faster in particular for long runs.

Notice that when restarting a simulation the history is read by default from a file and hills are added again to the grid. This allows one to change the grid boundaries upon restart. However, the first step after restart is usually very slow. Since PLUMED 2.3 you can also store the grid on a file and read it upon restart. This can be particularly useful if you perform many restarts and if your hills are large.

For the precise syntax, see [METAD](#)

## 10.3 Multiple time stepping

By setting a STRIDE different from 1, you change how frequently an action is calculated. In the case of actions such as [PRINT](#), this just means how frequently you dump some quantity on the disk. Notice that variables are only computed when necessary. Thus, if a variable is only appearing as the argument of a [PRINT](#) statement with STRIDE=10, it will be computed every 10 steps.

In a similar fashion, the STRIDE keyword can be used in a bias potential so as to apply the bias potential every few steps. In this case, forces from this bias potential are scaled up by a factor equal to STRIDE.

This technique can allow your simulation to run faster if you need to apply a bias potential on some very expensive collective variable. Consider the following input:

```
c1: COM ATOMS=1-1000
c2: COM ATOMS=1001-2000
d:  DISTANCE ATOMS=c1,c2
METAD ARG=d HEIGHT=1 SIGMA=0.1 BIASFACTOR=5 PACE=500
```

This performs a [METAD](#) simulation biasing the distance between two centers of mass. Since computing these centers requires a lot of atoms to be imported from the MD engine, it could slow down significantly the simulation. Notice that whereas the bias is changed every PACE=500 steps, it is applied every STRIDE step, where STRIDE=1 by default. The following input could lead to a significantly faster simulation at the price of a negligible systematic error

```
c1: COM ATOMS=1-1000
c2: COM ATOMS=1001-2000
d:  DISTANCE ATOMS=c1,c2
METAD ARG=d HEIGHT=1 SIGMA=0.1 BIASFACTOR=5 PACE=500 STRIDE=2
```

Similarly, the STRIDE keyword can be used with other biases (e.g. [RESTRAINT](#)).

The technique is discussed in details here [\[57\]](#). See also [EFFECTIVE\\_ENERGY\\_DRIFT](#).

### 10.3.1 EFFECTIVE\_ENERGY\_DRIFT

	This is part of the generic <a href="#">module</a>
--	----------------------------------------------------

Print the effective energy drift described in Ref [\[57\]](#)

#### Compulsory keywords

<b>STRIDE</b>	( default=1 ) should be set to 1. Effective energy drift computation has to be active at each step.
<b>FILE</b>	file on which to output the effective energy drift.
<b>PRINT_STRIDE</b>	frequency to which output the effective energy drift on FILE

#### Options

<b>ENSEMBLE</b>	( default=off ) Set to TRUE if you want to average over multiple replicas.
<b>RESTART</b>	allows per-action setting of restart (YES/NO/AUTO)
<b>UPDATE_FROM</b>	Only update this action from this time
<b>UPDATE_UN↔TIL</b>	Only update this action until this time

### Examples

This is to monitor the effective energy drift for a metadynamics simulation on the Debye-Huckel energy. Since this variable is very expensive, it could be conveniently computed every second step.

```
dh: DHENERGY GROUPA=1-10 GROUPB=11-20 EPSILON=80.0 I=0.1 TEMP=300.0
METAD ARG=dh HEIGHT=0.5 SIGMA=0.1 PACE=500 STRIDE=2
EFFECTIVE_ENERGY_DRIFT PRINT_STRIDE=100 FILE=eff
```

This is to monitor if a restraint is too stiff

```
d: DISTANCE ATOMS=10,20
RESTRAINT ARG=d KAPPA=100000 AT=0.6
EFFECTIVE_ENERGY_DRIFT PRINT_STRIDE=100 FILE=eff
```

## 10.4 Multicolvar

Whenever you have a multicolvar action such as:

```
COORDINATIONNUMBER SPECIES=1-100 SWITCH={RATIONAL R_0=1. D_MAX=3.0} MORE_THAN={RATIONAL R_0=6.0 NN=6 MM=12 D_0=1.0}
```

You will get a colossal speedup by specifying the D\_MAX keyword in all switching functions that act on distances. D\_MAX tells PLUMED that the switching function is strictly zero if the distance is greater than this value. As a result PLUMED knows that it does not need to calculate these zero terms in what are essentially sums with a very large number of terms. In fact when D\_MAX is set PLUMED uses linked lists when calculating these coordination numbers, which is what gives you such a dramatic increase in performance.

## 10.5 Neighbour Lists

Collective variables that can be speed up making us of neighbour lists:

- [COORDINATION](#)
- [DHENERGY](#)
- [PATHMSD](#)

By tuning the cut-off for the neighbour list and the frequency for the recalculation of the list it is possible to balance between accuracy and performances.

Notice that for [COORDINATION](#) and [DHENERGY](#) using a neighbor list could imply that a smaller number of atoms are requested to the host MD engine. This is typically true when considering [COORDINATION](#) of a small number of atoms (e.g. a ligand) against many atoms (e.g. water). When the neighbor list is used, only the water atoms close to the ligand will be requested at each step.

### Warning

Notice that the calculation of the neighbour list is not parallelized for [COORDINATION](#) and [DHENERGY](#). As a consequence, if you run with many processors and/or OpenMP threads, the neighbor list might even make the calculation slower.

## 10.6 OpenMP

PLUMED is partly parallelized using OpenMP. This should be enabled by default if your compiler supports it, and can be disabled with `--disable-openmp`. At runtime, you should set the environment variable `PLUMED_NUM_THREADS` to the number of threads you wish to use with PLUMED. By default (if `PLUMED_NUM_THREADS` is unset) openmp will be disabled at runtime. E.g., to run with gromacs you should do:

```
export PLUMED_NUM_THREADS=8
mdrun -plumed
```

Notice that:

- This option is likely to improve the performance, but could also slow down the code in some case.
- Results could be slightly different because of numerical roundoff and different order in summations. This should be harmless.
- The optimum number of threads is not necessary "all of them", nor should be equal to the number of threads used to parallelize MD.
- Only a few CVs are parallelized with openMP (currently, [COORDINATION](#) and [DHENERGY](#)).
- You might want to tune also the environmental variable `PLUMED_CACHELINE_SIZE`, by default 512, to set the size of cachelines on your machine. This is used by PLUMED to decrease the number of threads to be used in each loop so as to avoid clashes in memory access. This variable is expected to affect performance only, not results.

## 10.7 Secondary Structure

Secondary Structure collective variables ([ALPHARMSD](#), [PARABETARMSD](#) and [ANTIBETARMSD](#)) can be particularly demanding if you want to calculate them for all the residues of a protein. This is particularly true for the calculation of beta structures.

The FIRST thing to speed up [PARABETARMSD](#) and [ANTIBETARMSD](#) is to use the keyword `STRANDS_CUTOFF` (i.e. `STRANDS_CUTOFF=1`), in this way only a subset of possible fragments, the one less than 1. nm apart, are used in the calculation.

The metric used to calculate the distance from ideal secondary structure elements can also influence the performances, try to use `TYPE=OPTIMAL` or `TYPE=OPTIMAL-FAST` instead of `TYPE=DRMSD`.

At last, try to reduce the number of residues in the calculation.

## 10.8 Time your Input

Once you have prepared your plumed input file you can run a test simulation, or use driver, to see which collective variable, function, bias or analysis is consuming more time and can thus be the target for a different definition (use less atoms, change relevant parameters, or just use something else)

To have an accurate timing of your input you can use the [DEBUG](#) `DETAILED_TIMERS`.





# Chapter 11

## Tutorials

The following pages describe how to perform a variety of tasks using PLUMED

<a href="#">Belfast tutorial: Analyzing CVs</a>	This tutorial explains how to use plumed to analyze CVs
<a href="#">Belfast tutorial: Adaptive variables I</a>	How to use path CVs
<a href="#">Belfast tutorial: Adaptive variables II</a>	Dimensionality reduction and sketch maps
<a href="#">Belfast tutorial: Umbrella sampling</a>	Umbrella sampling, reweighting, and weighted histogram
<a href="#">Belfast tutorial: Out of equilibrium dynamics</a>	How to run a steered MD simulations and how to estimate the free energy
<a href="#">Belfast tutorial: Metadynamics</a>	How to run a metadynamics simulation
<a href="#">Belfast tutorial: Replica exchange I</a>	Parallel tempering and Metadynamics, Well-Tempered Ensemble
<a href="#">Belfast tutorial: Replica exchange II and Multiple walkers</a>	Bias exchange and multiple walkers
<a href="#">Belfast tutorial: NMR restraints</a>	NMR restraints
<a href="#">Belfast tutorial: Steinhardt Parameters</a>	Steinhardt Parameters
<a href="#">Cambridge tutorial</a>	A short 2 hours tutorial that introduces Well-Tempered Metadynamics, Bias-Exchange Metadynamics and Replica-Average Metadynamics
<a href="#">Cineca tutorial</a>	A short 2 hours tutorial that introduces analysis, well-tempered metadynamics, and multiple-restraints umbrella sampling.
<a href="#">Using Hamiltonian replica exchange with GROMACS</a>	This tutorial explains how to use Hamiltonian replica exchange in GROMACS
<a href="#">Julich tutorial: Developing CVs in plumed</a>	Implementing new collective variables in plumed
<a href="#">Moving from PLUMED 1 to PLUMED 2</a>	This tutorial explains how plumed 1 input files can be translated into the new plumed 2 syntax.
<a href="#">Munster tutorial</a>	A short 3 hours tutorial that introduces analysis, well-tempered metadynamics, and multiple-restraints umbrella sampling.

In addition, the following websites contain resources that might be helpful

<a href="http://www.youtube.com/watch?v=iDv↔ZmbWE5ps">http://www.youtube.com/watch?v=iDv↔ZmbWE5ps</a>	A short video introduction to the use of multicolvars in PLUMED 2
<a href="http://www.youtube.com/watch?v=PxJ↔P16qNCYs">http://www.youtube.com/watch?v=PxJ↔P16qNCYs</a>	A short video introduction to the syntax of the PLUM↔ED 2 input file

<a href="http://en.wikipedia.org/wiki/Metadynamics">http://en.wikipedia.org/wiki/Metadynamics</a>	A wikipedia article on metadynamics
---------------------------------------------------------------------------------------------------	-------------------------------------

## 11.1 Belfast tutorial: Analyzing CVs

### 11.1.1 Aims

The aim of this tutorial is to introduce the users to the plumed syntax. We will go through the writing of simple collective variable and we will use them to analyse a trajectory in terms of probability distributions and free energy.

### 11.1.2 Learning Outcomes

Once this tutorial is completed students will:

- Know how to write a simple plumed input file
- Know how to analyse a trajectory using plumed

### 11.1.3 Resources

The `tarball` for this project contains the following files:

- `trajectory-short.xyz` : a (short) trajectory for a 16 residue protein in xyz format. All calculations with plumed driver use this trajectory.
- `template.pdb` : a single frame from the trajectory that can be used in conjunction with the `MOLINFO` command

### 11.1.4 Instructions

PLUMED2 is a library that can be accessed by multiple codes adding a relatively simple and well documented interface. Once PLUMED is installed you can run a plumed executable that can be used for multiple purposes:

```
plumed --help
```

some of the listed options report about the plumed available functionalities, other can be used to tell plumed to do something: from analysing a trajectory to patch the source code of a MD code and so on. All the commands have further options that can be seen using plumed command `--help`, i.e.:

```
plumed driver --help
```

In the following we are going to see how to write an input file for plumed2 that can be used to analyse a trajectory.

## 11.1.4.1 A note on units

By default the PLUMED inputs and outputs quantities in the following units:

- Energy - kJ/mol
- Length - nanometers
- Time - picoseconds

If you want to change these units you can do this using the [UNITS](#) keyword.

## 11.1.4.2 Introduction to the PLUMED input file

A typical input file for PLUMED input is composed by specification of one or more CVs, the printout frequency and a termination line. Comments are denoted with a # and the termination of the input for PLUMED is marked with the keyword ENDPLUMED. Whatever it follows is ignored by PLUMED. You can introduce blank lines. They are not interpreted by PLUMED.

In the following input we will analyse the [DISTANCE](#) between the two terminal carbons of a 16 residues peptide, and we will [PRINT](#) the results in file named COLVAR.

```
#my first plumed input:
DISTANCE ATOMS=2,253 LABEL=e2edist

#printout frequency
PRINT ARG=e2edist STRIDE=1 FILE=COLVAR

#endofinput
ENDPLUMED
here I can write what I want it won't be read.
```

Now we can use this simple input file to analyse the trajectory included in the [RESOURCES](#):

```
plumed driver --plumed plumed.dat --ixyz trajectory-short.xyz --length-units 0.1
```

NOTE: --length-units 0.1, xyz files, as well as pdb files, are in Angstrom.

You should have a file COLVAR, if you look at it (i.e. more COLVAR) the first two lines should be:

```
#! FIELDS time e2edist
0.000000 2.5613161
```

NOTE: the first line of the file COLVAR tells you what is the content of each column.

In PLUMED2 the commands defined in the input files are executed in the same order in which they are written, this means that the following input file is wrong:

```
#printout frequency
PRINT ARG=e2edist STRIDE=1 FILE=COLVAR
#my first plumed input:
DISTANCE ATOMS=2,253 LABEL=e2edist
#endofinput
ENDPLUMED
here I can write what I want it won't be read.
```

Try to run it.

Sometimes, when calculating a collective variable, you may not want to use the positions of a number of atoms directly. Instead you may wish to use the position of a virtual atom whose position is generated based on the positions of a collection of other atoms. For example you might want to use the center of a group of atoms ([CENTER](#)):

Since PLUMED executes the input in order you need to define the new Virtual Atom before using it:

```
first: CENTER ATOMS=1,2,3,4,5,6
last: CENTER ATOMS=251-256

e2edist: DISTANCE ATOMS=2,253
comdist: DISTANCE ATOMS=first,last

PRINT ARG=e2edist,comdist STRIDE=1 FILE=COLVAR

ENDPLUMED
```

NOTE: an action (i.e. [CENTER](#) or [DISTANCE](#) here) can be either labeled using [LABEL](#) as we did before or as label: [ACTION](#) as we have just done here.

With the above input this is what happen inside PLUMED with a [STRIDE](#)=1:

1. calculates the position of the Virtual Atom 'first' as the [CENTER](#) of atoms from 1 to 6;
2. calculates the position of the Virtual Atom 'last' as the [CENTER](#) of atoms from 251 to 256;
3. calculates the distance between atoms 2 and 253 and saves it in 'e2edist';
4. calculates the distance between the two atoms 'first' and 'last' and saves it in 'comdist';
5. print the content of 'e2edist' and 'comdist' in the file COLVAR

In the above input we have used two different ways of writing the atoms used in [CENTER](#) calculation:

1. [ATOMS](#)=1,2,3,4,5,6 is the explicit list of the atoms we need
2. [ATOMS](#)=251-256 is the range of atoms needed

ranges of atoms can be defined with a stride which can also be negative:

1. [ATOMS](#)=from,to:by (i.e.: 251-256:2)
2. [ATOMS](#)=to,from:-by (i.e.: 256-251:-2)

Now by plotting the content of the COLVAR file we can compare the behaviour in this trajectory of both the terminal carbons as well as of the centre of masses of the terminal residues.

gnuplot

What do you expect to see now by looking at the trajectory? Let's have a look at it

```
vmd template.pdb trajectory-short.xyz
```

Virtual atoms can be used in place of standard atoms everywhere an atom can be given as input, they can also be used together with standard atoms. So for example we can analyse the [TORSION](#) angle for a set of Virtual and Standard atoms:

```
first: CENTER ATOMS=1-6
last: CENTER ATOMS=251-256
cvtor: TORSION ATOMS=first,102,138,last

PRINT ARG=cvtor STRIDE=1 FILE=COLVAR

ENDPLUMED
```

The above CV don't look smart to learn something about the system we are looking at. In principle CV are used to reduce the complexity of a system by looking at a small number of properties that could be enough to rationalise its behaviour.

Now try to write a collective variable that measures the Radius of Gyration of the system: [GYRATION](#).

NOTE: if what you need for one or more variables is a long list of atoms and not a virtual atom one can use the keyword [GROUP](#). A GROUP can be defined using ATOMS in the same way we saw before, in addition it is also possible to define a GROUP by reading a GROMACS index file.

```
ca: GROUP ATOMS=9,16,31,55,69,90,102,114,124,138,160,174,194,208,224,238
```

Now 'ca' is not a virtual atom but a simple list of atoms.

### 11.1.4.3 MULTICOLVAR

Sometimes it can be useful to calculate properties of many similar collective variables at the same time, for example one can be interested in calculating the properties of the distances between a group of atoms, or properties linked to the distribution of the dihedral angles of a chain and so on. In PLUMED2 this kind of collective variables fall under the name of MULTICOLVAR (cf. [MultiColvar](#).) Here we are going to analyse the distances between CA carbons along the chain:

```
ca: GROUP ATOMS=9,16,31,55,69,90,102,114,124,138,160,174,194,208,224,238
dd: DISTANCES GROUP=ca MEAN MIN={BETA=50} MAX={BETA=0.02} MOMENTS=2

PRINT ARG=dd.mean,dd.min,dd.max,dd.moment-2 STRIDE=1 FILE=COLVAR

ENDPLUMED
```

The above input tells PLUMED to calculate all the distances between CA carbons and then look for the mean distance, the minimum distance, the maximum distance and the variance. In this way we have defined four collective variables that are calculated using the distances. These four collective variables are stored as components of the defined action 'dd': dd.mean, dd.min, dd.max, dd.moment-2.

The infrastructure of multicolvar has been used to develop many PLUMED2 collective variables as for example the set of Secondary Structure CVs ([ANTIBETARMSD](#), [PARABETARMSD](#) and [ALPHARMSD](#)).

```
MOLINFO STRUCTURE=template.pdb
abeta: ANTIBETARMSD RESIDUES=all TYPE=DRMSD LESS_THAN={RATIONAL R_0=0.08 NN=8 MM=12} STRANDS_CUTOFF=1

PRINT ARG=abeta.less-than STRIDE=1 FILE=COLVAR

ENDPLUMED
```

We have now seen how to write the input some of the many CVs available in PLUMED. More complex CVs will be discussed in the next workshop, [Belfast tutorial: Adaptive variables I](#).

#### 11.1.4.4 Analysis of Collective Variables

Collective variables are usually used to visualize the Free Energy of a system. Given a system evolving at fixed temperature, fixed number of particles and fixed volume, it will explore different conformations with a probability

$$P(q) \propto e^{-\frac{U(q)}{k_B T}}$$

where  $q$  are the microscopic coordinates and  $k_B$  is the Boltzmann constant.

It is possible to analyse the above probability as a function of one or more collective variable  $s(q)$ :

$$P(s) \propto \int dq e^{-\frac{U(q)}{k_B T}} \delta(s - s(q))$$

where the  $\delta$  function means that to for a given value  $s$  of the collective variable are counted only those conformations for which the CV is  $s$ . The probability can be recast to a free energy by taking its logarithm:

$$F(s) = -k_B T \log P(s)$$

This means that by estimating the probability distribution of a CV it is possible to know the free energy of a system along that CV. Estimating the probability distribution of the conformations of a system is what is called 'sampling'.

In order to estimate a probability distribution one needs to make [HISTOGRAM](#) from the calculated CVs. PLUMED2 includes the possibility of histogramming data both on the fly as well as a posteriori as we are going to do now.

```
MOLINFO STRUCTURE=template.pdb
abeta: ANTIBETARMSD RESIDUES=all TYPE=DRMSD LESS_THAN={RATIONAL R_0=0.08 NN=8 MM=12} STRANDS_CUTOFF=1
ca: GROUP ATOMS=9,16,31,55,69,90,102,114,124,138,160,174,194,208,224,238
DISTANCES ...
GROUP=ca MEAN MIN={BETA=50} MAX={BETA=0.02} MOMENTS=2 LABEL=dd
... DISTANCES

PRINT ARG=abeta.lessthan,dd.mean,dd.min,dd.max,dd.moment-2 STRIDE=1 FILE=COLVAR

HISTOGRAM ...
ARG=abeta.lessthan,dd.mean
LABEL=hh
KERNEL=DISCRETE
GRID_MIN=0,0.8
GRID_MAX=4,1.2
GRID_BIN=40,40
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo

ENDPLUMED
```

NOTE: HISTOGRAM ... means that what follow is part of the [HISTOGRAM](#) function, the same can be done for any action in PLUMED.

The above input tells PLUMED to accumulate the two collective variables on a GRID. In addition the probability can be converted to a free-energy using the [CONVERT\\_TO\\_FES](#) method and then use [DUMPGRID](#) to write it. Histograms can be accumulated in a smoother way by using a KERNEL function, a kernel is a normalised function, for example a normalised gaussian is the default kernel in PLUMED, that is added to the histogram centered in the position of the data. Estimating a probability density using kernels can in principle give more accurate results, on the other hand in addition to the choice of the binning one has to choose a parameter that is the WIDTH of the kernel function. As a rule of thumb: the grid spacing should be smaller (i.e. one half or less) than the BANDWIDTH and the BANDWIDTH should be smaller (i.e. one order of magnitude) than the variance observed/expected for the variable.

```

HISTOGRAM ...
LABEL=hh
ARG=abeta.lessthan,dd.mean
GRID_MIN=0,0.8
GRID_MAX=4,1.2
GRID_SPACING=0.04,0.004
BANDWIDTH=0.08,0.008
... HISTOGRAM

DUMPGRID GRID=hh FILE=histo

ENDPLUMED

```

If you have time less at the end of the session read the manual and look for alternative collective variables to analyse the trajectory. Furthermore try to play with the [HISTOGRAM](#) parameters to see the effect of using KERNEL in analysing data.

## 11.2 Belfast tutorial: Adaptive variables I

### 11.2.1 Aim

In this section we want to introduce the concept of adaptive collective variables. These are special variables that are knowledge-based in that are built from a pre-existing notion of the mechanism of the transition under study.

### 11.2.2 Resources

Here is the [tarball](#) with the files referenced in the following.

### 11.2.3 What happens when in a complex reaction?

When you deal with a complex conformational transition that you want to analyze (or bias), very often you cannot just describe it with a single order parameter.

As an example in Figure [belfast-2-cdk-fig](#) I consider a large conformational transition like those involved in activating the kinase via open-close transition of the activation loop. In sticks you see the part involved in the large conformational change, the rest is either keeping the structure and just moving a bit or is a badly resolved region in the X-ray structure. This is a complex transition and it is hard to tell which is the order parameter that best describes the transition.

One could identify a distance that can distinguish open from close but

- the plasticity of the loop is such that the same distance can correspond to an almost closed loop and almost open loop. One would like to completely divide these two situations with something which is discriminating what intuitively one would think as open and closed
- the transition state is an important point where one would like to see a certain crucial interaction forming/breaking so to better explain what is really happening. If you capture then hypothetically you would be able to say what is dictating the rate of this conformational transition. A generic distance is a very hybrid measure that is unlikely to capture a salt-bridge formation and a concerted change of many dihedral change or desolvation contribution which are happening while the transition is happening. All these things are potentially important in such transition but none of them is explaining the whole thing.  
So basically in these cases you have to deal with an intrinsic multidimensional collective variable where you would need many dimensions. How would you visualize a 10 dimensional CV where you use many distances, coordinations and dihedrals (ouch, they're periodic too!) ?

Another typical case is the docking of a small molecule in a protein cleft or gorge, which is the mechanism of drug action. This involves specific conformational transition from both the small molecule and the protein as the small molecule approaches the protein cavity. This also might imply a specific desolvation pattern.

Other typical examples are chemical reactions. Nucleophilic attacks typically happen with a role from the solvent (see some nice paper from Nobel-prize winner Arieh Warshel) and sizeable geometric distortions of the neighboring groups.

### 11.2.4 Path collective variables

One possibility to describe many different things that happen in a single reaction is to use a dimensional reduction technique and in plumed the simplest example that may show its usefulness can be considered that of the path collective variables.

In a nutshell, your reaction might be very complex and happening in many degree of freedom but intuitively is a sort of track along which the reaction proceeds. So what we need is a coordinate that, given a conformation, just tells which point along the "reactive track" is closest.

For example, in Fig. [belfast-2-ab-fig](#), you see a typical chemical reaction (hydrolysis of methylphosphate) with the two end-points denoted by A and B. If you are given a third point, just by looking at it, you might find that this is more resemblant to the reactant than the product, so, hypothetically, if you would intuitively give a parameter that would be 1 for a configuration in the A state and 2 for a configuration in the B state, you probably would give it something like 1.3, right?

Path collective variables are the extension to this concept in the case you have many conformation that describe your path, and therefore, instead of an index that goes from 1 to 2 you have an index that goes from 1 to  $N$ , where  $N$  is the number of conformation that you use in input to describe your path.

From a mathematical point of view, that's rather simple. The progress along the path is calculated with the following equation:

$$S(X) = \frac{\sum_{i=1}^N i \exp^{-\lambda|X-X_i|}}{\sum_{i=1}^N \exp^{-\lambda|X-X_i|}}$$

where in [belfast-2-s-eq](#) the  $|X - X_i|$  represents a distance between one configuration  $X$  which is analyzed and another from the set that compose the path  $X_i$ . The parameter  $\lambda$  is a positive value that is tuned in a way explained later. here are a number of things to note to make you think that this is exactly what you want.

- The negative exponential function is something that is 1 whenever the value at the exponent is zero, and is progressively smaller when the value is larger than zero (trivially, the case with the value at the exponent larger than zero never occurs since lambda is a positive quantity and the distance is by definition positive).
- Whenever you sit exactly on a specific images  $X_j$  then all the other terms in the sum disappear (if  $\lambda$  is large enough) and only the value  $j$  survives returning exactly  $S(X) = j$ .

In order to provide a value which is continuous, the parameter  $\lambda$  should be correctly tuned. As a rule of thumb I use the following formula

$$\lambda = \frac{2.3(N-1)}{\sum_{i=1}^{N-1} |X_i - X_{i+1}|}$$

which imply that one should calculate the average distance between consecutive frames composing the path. Note also that this distance should be more or less similar between the frames. Generally I tolerate fluctuation of the order of 10/15 percent tops. If you have larger, then it is better to have a smaller value of  $\lambda$ .



It is important to note that in principle one could even have a specific  $\lambda$  value associated to each frame of the path but this would provide some distortion in the diffusion coefficient which could potentially harm a straightforward interpretation of the free energy landscape.

So, at this point it is better to understand what is meant with "distance" since a distance between two conformations can be calculated in very many ways. The way we refer here is by using mean square deviation after optimal alignment. This means that at each step in which the analysis is performed, a number  $N$  of optimal alignments is performed. Namely what is calculated is  $|X - X_i| = d(X, X_i)^2$  where  $d(X, X_i)$  is the RMSD as defined in what you see here [RMSD](#).

Using the MSD instead of RMSD is sometimes more convenient and more stable (you do not have a denominator that goes to zero in the derivatives when biasing).

Anyway this is a matter of choice. Potentially one could equally employ other metrics like a set of coordinations (this was done in the past), and then you would avoid the problem of rototranslations (well, which is not a problem since you have it already in plumed) but for some applications that might become appealing. So in path collective variables (and in all the dimensional reduction based collective variables) the problem is converted from the side of choosing the collective variable in choosing the right way to calculate distances, also called "metrics".

The discussion of this issue is well beyond the topic of this tutorial, so we can move forward in how to tell plumed to calculate the progress along the path whenever the MSD after optimal alignment is used as distance.

```
p1: PATHMSD REFERENCE=all.pdb LAMBDA=50.0
PRINT ARG=p1.sss,p1.zzz STRIDE=100 FILE=colvar FMT=%8.4f
```

Note that reference contains a set of PDB, appended one after the other, with a END field. Note that there is no need to place all the atoms of the system in the PDB reference file you provide. Just put the atoms that you think might be needed. You can leave out big domains, solvent and ions if you think that is not important for your use.

Additionally, note that the measure units of LAMBDA are in the units of the code. In gromacs they are in  $\text{nm}^2$  while NAMD is  $\text{\AA}^2$ . [PATHMSD](#) produces two arguments that can be printed or used in other ActionWithArguments. One is the progress along the path of [belfast-2-s-eq](#), the other is the distance from the closest point along the path, which is denoted with the zzz component. This is defined as

$$Z(X) = -\frac{1}{\lambda} \log\left(\sum_{i=1}^N \exp^{-\lambda|X-X_i|}\right)$$

It is easy to understand that in case of perfect match of  $X = X_i$  this equation gives back the value of  $|X - X_i|$  provided that the lambda is adjusted correctly.

So, the two variables, put together can be visualized as This variable is important because whenever your simulation is running close to the path (low  $Z$  values), then you know that you are reproducing reliably the path you provided in input but if by chance you find some other path that goes, say, from  $S = 1, Z = 0$  to  $S = N, Z = 0$  via large  $Z$  values, then it might well be that you have just discovered a good alternative pathway. If your path indeed is going from  $S = 1, Z = \text{large}$  to  $S = N, Z = \text{large}$  then it might well be that you do not have your reaction accomplished, since your reaction, by definition should go from the reactant which is located at  $S = 1, Z = 0$  to the product, which is located at  $S = 1, Z = N$  so you should pay attention. This case is exemplified in Fig. [belfast-2-ab-sz-nowhere-fig](#)

### 11.2.5 A note on the path topology

A truly important point is that if you get a trajectory from some form of accelerated dynamics (e.g. simply by heating) this cannot simply be converted into a path. Since it is likely that your trajectory is going stochastically back and forth (not in the case of SMD or US, discussed later), your trajectory might be not topologically suitable. To understand that, suppose you simply collect a reactive trajectory of 100 ps into the reference path you give to the [PATHMSD](#) and simply you assign the frame of 1 ps to index 1 (first frame occurring in the reference file provided to [PATHMSD](#)), the frame of 2 ps to index 2 and so on : it might be that you have two points which are really similar but one is associated to step, say 5 and the other is associated with frame 12. When you analyse the same trajectory, when you are sitting on any of those points then the calculation of  $S$  will be an average like  $S(X) = (5 + 12)/2 = 8.5$  which is an average of the two indexes and is completely misleading since it let you think that you are moving between point 8 and 9, which is not the case. So this evidences that your reference frames should be "topologically consecutive". This means that frame 1 should be the closest to frame 2 and all the other frames should be farther apart. Similarly frame 2 should be equally close (in an [RMSD](#) sense) to 1 and 3 while all the others should be farther apart. Same for frame 3: this should be closest to frame 2 and 4 and farther apart from all the others and so on. This is equivalent to calculate an "RMSD matrix" which can be conveniently done in vmd (this is a good exercise for a number of reasons) with RMSD Trajectory tools, by choosing different reference system along the set of reference frames.

This is shown in Fig. [belfast-2-good-matrix-fig](#) where the matrix has a typical gullwing shape.

On the contrary, whenever you extract the frames from a pdb that you produced via free MD or some biased methods (SMD or Targeted MD for example) then your frame-to-frame distance is rather inhomogeneous and looks something like

Aside from the general shape, which is important to keep the conformation-to-index relation (this, as we will see in the next part is crucial in the multidimensional scaling) the crucial thing is the distance between neighbors.

As a matter of fact, this is not much important in the analysis but is truly crucial in the bias. When biasing a simulation, you generally want to introduce a force that push your system somewhere. In particular, when you add a bias which is based on a path collective variable, most likely you want that your system goes back and forth along your path. The bias is generally applied by an additional term in the hamiltonian, this can be a spring term for Umbrella Sampling, a Gaussian function for Metadynamics or whatever term which is a function of the collective variable  $s$ . Therefore the Hamiltonian  $H(X)$  where  $X$  is the point of in the configurational phase space where your system is takes the following form

$$H'(X) = H(X) + U(S(X))$$

where  $U(S(X))$  is the force term which depends on the collective variable that ultimately is a function of the  $X$ . Now, when you use biased dynamics you need to evolve according the forces that this term produces (this only holds for MD, while not in MC) and therefore you need

$$F_i = -\frac{dH'(X)}{dx_i} = -\frac{dH(X)}{dx_i} - \frac{\partial U(S(X))}{\partial S} \frac{\partial S(X)}{\partial x_i}$$

This underlines the fact that, whenever  $\frac{\partial S(X)}{\partial x_i}$  is zero, then you have no force on the system. Now the derivative of the progress along the path is

$$\frac{\partial S(X)}{\partial x_i} = \frac{\sum_{i=1}^N -\lambda i \frac{\partial |X-X_i|}{\partial x_i} \exp^{-\lambda |X-X_i|}}{\sum_{i=1}^N \exp^{-\lambda |X-X_i|}} - \frac{(\sum_{i=1}^N i \exp^{-\lambda |X-X_i|})(\sum_{j=1}^N -\lambda \frac{\partial |X-X_j|}{\partial x_i} \exp^{-\lambda |X-X_j|})}{(\sum_{i=1}^N \exp^{-\lambda |X-X_i|})^2} = \frac{\sum_{i=1}^N -\lambda i \frac{\partial |X-X_i|}{\partial x_i} \exp^{-\lambda |X-X_i|}}{\sum_{i=1}^N \exp^{-\lambda |X-X_i|}}$$

which can be rewritten as

$$\frac{\partial S(X)}{\partial x_i} = \lambda \frac{\sum_{i=1}^N \frac{\partial |X-X_i|}{\partial x_i} \exp^{-\lambda |X-X_i|} [s(X) - i]}{\sum_{i=1}^N \exp^{-\lambda |X-X_i|}}$$

It is interesting to note that whenever the  $\lambda$  is too small the force will vanish. Additionally, when  $\lambda$  is too large, then it one single exponential term will dominate over the other on a large part of phase space while the other will vanish. This means that the  $S(X)$  will assume almost discrete values that produce zero force. Funny, isn't it?

### 11.2.6 How many frames do I need?

A very common question that comes is the following: "I have my reaction or a model of it. how many frames do I need to properly define a path collective variable?" This is a very important point that requires a bit of thinking. It all depends on the limiting scale in your reaction. For example, if in your process you have a torsion, as the smallest event that you want to capture with path collective variable, then it is important that you mimic that torsion in the path and that this does not contain simply the initial and final point but also some intermediate. Similarly, if you have a concerted bond breaking, it might be that all takes place in the range of an Angstrom or so. In this case you should have intermediate frames that cover the sub-Angstrom scale. If you have both in the same path, then the smallest dominates and you have to mimic also the torsion with sub-Angstrom accuracy.

### 11.2.7 Some tricks of the trade: the neighbors list.

If it happens that you have a very complex and detailed path to use, say that it contains 100 frames with 200 atoms each, then the calculation of a 100 alignment is required every time you need the CV. This can be quite expensive but you can use a trick. If your trajectory is continuous and you are sure that your trajectory does not show jumps where your system suddenly move from the reactant to the product, then you can use a so-called neighbor list. The plumed input shown before then becomes

```
p1: PATHMSD REFERENCE=all.pdb LAMBDA=50.0 NEIGH_STRIDE=100 NEIGH_SIZE=10
PRINT ARG=p1.sss,p1.zzz STRIDE=100 FILE=colvar FMT=%8.4f
```

and in this case only the closest 10 frames from the path will be used for the CV. Then the list of the frames to use is updated every 100 steps. If you are using a biased dynamics this may introduce sudden change in the derivatives, therefore it is better to check the stability of the setup before running production-quality calculations.

### 11.2.8 The molecule of the day: alanine dipeptide

Here and probably in other parts of the tutorial a simple molecule is used as a test case. This is alanine dipeptide in vacuum. This rather simple molecule is useful to make many benchmark that are around for data analysis and free energy methods. It is a rather nice example since it presents two states separated by a high (free) energy barrier.

In Fig. [belfast-2-ala-fig](#) its structure is shown.

The two main metastable states are called  $C_{7eq}$  and  $C_{7ax}$ .

Here metastable states are intended as states which have a relatively low free energy compared to adjacent conformation. At this stage it is not really useful to know what is the free energy, just think in term of internal energy. This is almost the same for such a small system with so few degrees of freedom.

It is conventional use to show the two states in terms of Ramachandran dihedral angles, which are denoted with  $\Phi$  and  $\Psi$  in Fig. [belfast-2-transition-fig](#).

### 11.2.9 Examples

Now as a simple example, I want to show you that plotting some free dynamics trajectories shoot from the saddle point, you get a different plot in the path collective variables if you use the right path or if you use the wrong path.

In Fig. [belfast-2-good-bad-path-fig](#) I show you two example of possible path that join the  $C_{7eq}$  and  $C_{7ax}$  metastable states in alanine dipeptide. You might clearly expect that real (rare) trajectories that move from one basin to the other would rather move along the black line than on the red line.

So, in this example we do a sort of "committor analysis" where we start shooting a number of free molecular dynamics from the saddle point located at  $\Phi = 0$  and  $\Psi = -1$  and we want to see which way do they go. Intuitively, by assigning random velocities every time we should find part of the trajectories that move toward  $C_{7eq}$  and part that move towards  $C_{7ax}$ .

I provided you with two directories, each containing a bash script script.sh whose core (it is a bit more complicated in practice...) consists in:

```
#
# set how many runs you want to do
#
ntests=50
for i in `seq 1 $ntests`
do
    #
    # assign a random velocity at each timestep by initializing the
    #
    sed s/SEED/$RANDOM/ md.mdp >newmd.mdp
    #
    # do the topology: this should write a topol.tpr
    #
    $GROMPP -c start.gro -p topol.top -f newmd.mdp
    $GROMACS_BIN/$MDRUN -plumed plumed.dat
    mv colvar colvar_$i
done
```

This runs 50 short MD runs (few hundreds steps) and just saves the colvar file into a labeled colvar file. In each mdrun plumed is used to plot the collective variables and it is something that reads like the following:

```
# Phi
t1: TORSION ATOMS=5,7,9,15
# Psi
t2: TORSION ATOMS=7,9,15,17
# The right path
p1: PATHMSD REFERENCE=right_path.dat LAMBDA=15100.
# The wrong path
p2: PATHMSD REFERENCE=wrong_path.dat LAMBDA=8244.4
# Just a printout of all the stuff calculated so far
PRINT ARG=* STRIDE=2 FILE=colvar FMT=%12.8f
```

where I just want to plot  $\Phi$ ,  $\Psi$  and the two path collective variables. Note that each path has a different LAMBDA parameters. Here the Ramachandran angles are plotted so you can realize which path is the system walking in a more comfortable projection. This is of course fine in such a small system but whenever you have to deal with larger systems and control hundreds of CVs at the same time, I think that path collective variables produce a more intuitive description for what you want to do.

If you run the script simply with

```
pd@plumed:~> ./script.sh
```

then after a minute or so, you should have a directory which is full of colvar files. Let's revise together how the colvar file is formatted:

```

#! FIELDS time t1 t2 p1.sss p1.zzz p2.sss p2.zzz
#! SET min_t1 -pi
#! SET max_t1 pi
#! SET min_t2 -pi
#! SET max_t2 pi
 0.000000 -0.17752998 -1.01329788 13.87216908 0.00005492 12.00532256 0.00233905
 0.004000 -0.13370142 -1.10611136 13.87613508 0.00004823 12.03390658 0.00255806
 0.008000 -0.15633049 -1.14298481 13.88290617 0.00004511 12.07203319 0.00273764
 0.012000 -0.23856451 -1.12343958 13.89969608 0.00004267 12.12872544 0.00284883
...

```

In first column you have the time, then  $t1$  ( $\Phi$ ), then  $t2$  ( $\Psi$ ) and the path collective variables  $p1$  and  $p2$ . Note that the action PATHMSD calculates both the progress along the path ( $p1.sss$ ) and the distance from it ( $p1.zzz$ ). In PLUMED jargon, these are called "components". So a single Action (a line in plumed input) can calculate many components at the same time. This is not always the case: sometimes by default you have one component but specific flags may enable more components to be calculated (see [DISTANCE](#) for example). Note that the header (all the part of a colvar file that contains # as first character) is telling already what it inside the file and eventually also tells you if a variable is contained in boundaries (for example torsions, are periodic and their codomain is defined in  $-\Pi$  and  $\Pi$ ).

At the end of the script, you also have two additional scripts. One is named `script_rama.gplt` and the other is named `script_path.gplt`. They contain some gnuplot commands that are very handy to visualize all the colvar files without making you load one by one, that would be a pain.

Now, let's visualize the result from the wrong path directory. In order to do so, after having run the calculation, then do

```

pd@plumed:~>gnuplot
gnuplot> load "script_rama.gplt"

```

what you see is that all the trajectories join the reactant and the product state along the low free energy path depicted before.

Now if you try to load the same bunch of trajectories as a function of the progress along the path and the distance from the path in the case of the wrong path then simply do

```

gnuplot> load "script_path_wrong.gplt"

```

What do you see? A number of trajectories move from the middle towards the right bottom side at low distance from the path. In the middle of the progress along the path, you have higher distance. This is expected since the distance zero corresponds to a unlikely, highly-energetic path which is unlikely to occur. Differently, if you now do

```

gnuplot> load "script_path_right.gplt"

```

You see that the path, compared to what happened before, run much closer to small distance from the path. This means that the provided path is highl ressemblant and representative of what hapens in the reactive path.

Note that going at high distances can be also beneficial. It might help you to explore alternative paths that you have not explored before. But beware, the more you get far from the path, the more states are accessible, in a similar way as the fact that the surface of a sphere increase with its radius. The surface ramps up much faster than the radius therefore you have a lots of states there. This means also high entropy, so many systems actually tend to drift to high distances while, on the contrary, zero distance is never reached in practice (zero entropy system is impossible to reach at finite temperature). So you can see by yourself that this can be a big can of worms. In particular, my experience with path collective variables and biological systems tells me that most of time is hopeless to go to high distances to find new path in many cases (for example, in folding). While this is worth whenever you think that the paths are not too many (alternative routes in chemical reaction or enzymatic catalysis).

### 11.2.10 How to format my input?

Very often it is asked how to format a set of pdb to be suitably used with path collective variables. Here are some tricks.

- When you dump the files with vmd or (for gromacs users, using trjcat), the pdb you obtain is reindexed from 1. This is also the case when you select a subensemble of atoms of the path (e.g. the heavy atoms only or the backbone atoms). This is rather unfortunate and you have to fix it somehow. My preference is to dump the whole pdb but water (when I do not need it) and use some awk script to select the atoms I am interested in.
- Pay attention to the last two columns. These are occupancy and beta. With the first (occupancy) you set the atoms which are used to perform the alignment. The atoms which have zero occupancy will not be used in the alignment. The second column is beta and controls which atoms are used for the calculation of the distance after having performed the alignment on the set of atoms which have nonzero occupancy column. In this way you can align all your system by using a part of the system and calculate the distance respect to another set. This is handy in case of protein-ligand. You set the alignment of the protein and you calculate the distance based on the ligand and the part of the protein which is in contact with the protein. This is done for example in [this article](#).
- **Plumed-GUI** (version > 2.0) provides the *Structure->Build reference structure...* function to generate inputs that conform to the above rules from within VMD.
- Note that all the atoms contained in the REFERENCE must be the same. You cannot have a variable number of atoms in each pdb contained in the reference.
- The reference is composed as a set of concatenated PDBs that are interrupted by a TER/END/ENDMDL card. Both HETATM and ATOM cards denote the atoms of the set.
- Include in the reference frames only the needed atoms. For example, if you have a methyl group involved in a conformational transition, it might be that you do not want to include the hydrogen atoms of the methyl since these rotate fast and probably they do not play an relevant role.

### 11.2.11 Fast forward: metadynamics on the path

This section is actually set a bit forward but I included here for completeness now. It is recommended to be read after you have an introduction on Metadynamics and to well-tempered Metadynamics in particular. Here I want to show a couple of concept together.

- Path collective variables can be used for exploring alternative routes. It is effective in highly structure molecules, while it is tricky on complex molecules whenever you have many competing routes
- Path collective variables suffer from problems at the endpoints (as the highly popular coordinates **COORDINATION** for example) that can be cured with flexible hills and an appropriate reweighting procedure within the well-tempered Metadynamics scheme.

Let's go to the last problem. All comes from the derivative [belfast-2-sder-eq](#). Whenever you have a point of phase space which is similar to one of the endpoint than one of the points in the center then you get a  $s(X)$  which is 1 or N (where N is the number of frames composing the path collective variable). In this case that exponential will dominate the others and you are left with a constant (since the derivative of RMSD is a constant since it is linear in space). This means that, no matter what happens here, you have small force. Additionally you have small motion in the CV space. You can move a lot in configuration space but if the closest point is one of the endpoint, your CV value will always be one of the endpoint itself. So, if you use a fixed width of your CV which you retrieve from a middle point in your path, this is not suitable at all at the endpoints where your CV fluctuates much less. On the contrary if you pick the hills width by making a free dynamics on the end states you might pick some sigmas that are smaller than what you might use in the middle of the path. This might give a rough free energy profile and definitely

more time to converge. A possible solution is to use the adaptive gaussian width scheme. In this scheme you adapt the hills to their fluctuation in time. You find more details in [44] Additionally you also adopt a non spherical shape taking into account variable correlation. So in this scheme you do not have to fix one sigma per variable sigma, but just the time in which you calculate this correlation (another possibility is to calculate it from the compression of phase space but will not be covered here). The input of metadynamics might become something like this

```
t1: TORSION ATOMS=5,7,9,15
t2: TORSION ATOMS=7,9,15,17
p1: PATHMSD REFERENCE=right_path.dat LAMBDA=15100.
p2: PATHMSD REFERENCE=wrong_path.dat LAMBDA=8244.4
#
# do a metadynamics on the right path, use adaptive hills whose decay time is 125 steps (250 fs)
# and rather standard WT parameters
#
meta: METAD ARG=p1.sss,p1.zzz ADAPTIVE=DIFF SIGMA=125 HEIGHT=2.4 TEMP=300 BIASFACTOR=12 PACE=125
PRINT ARG=* STRIDE=10 FILE=colvar FMT=%12.8f
```

You can find this example in the directory `BIASED_DYNAMICS`. After you run for a while it is interesting to have a feeling for the change in shape of the hills. That you can do with

```
pd@plumed:~> gnuplot
gnuplot> p "<head -400 HILLS" u 2:3:4:5 w xyer
```

that plots the hills in the progress along the path and the distance from the path and add an error bar which reflects the diagonal width of the flexible hills for the first 400 hills (Hey note the funny trick in gnuplot in which you can manipulate the data like in a bash script directly in gnuplot. That's very handy!).

There are a number of things to observe: first that the path explores high distance since the metadynamics is working also in the distance from the path thus accessing the paths that were not explored before, namely the one that goes from the upper left corner of the ramachandran plot and the one that passes through the lower left corner. So in this way one can also explore other paths. Additionally you can see that the hills are changing size rather considerably. This helps the system to travel faster since at each time you use something that has a nonzero gradient and your forces act on your system in an effective way. Another point is that you can see broad hills covering places which you have not visited in practice. For example you see that hills extend so wide to cover point that have negative progress along the path, which is impossible by using the present definition of the progress along the path. This introduced a problem in calculating the free energy. You actually have to correct for the point that you visited in reality.

You can actually use `sum_hills` to this purpose in a two-step procedure. First you calculate the negative bias on a given range:

```
pd@plumed:~> plumed sum_hills --hills HILLS --negbias --outfile negative_bias.dat --bin 100,100 --min -5,-0.005
```

and then calculate the correction. You can use the same hills file for that purpose. The initial transient time should not matter if your simulation is long enough to see many recrossing and, secondly, you should check that the hills height in the welltempered are small compared to the beginning.

```
pd@plumed:~> plumed sum_hills --histo HILLS --bin 100,100 --min -5,-0.005 --max 25,0.05 --kt 2.5 --sigma 0.5,0.005
```

Note that in the correction you should assign a sigma, that is a "trust radius" in which you think that whenever you have a point somewhere, there in the neighborhood you assign a bin (it is done with Gaussian in reality, but this does not matter much). This is the important point that compensates for the issues you might encounter putting excessive large hills in places that you have not visited. It is nice to have a look to the correction and compare with the hills in the same range.

```

gnuplot> set pm3d
gnuplot> spl "correction.dat" u 1:2:3 w l
gnuplot> set contour
gnuplot> set cntrp lev incremental -20,4.,1000.
gnuplot> set view map
gnuplot> unset clabel
gnuplot> replot

```

You might notice that there are no contour in the unrealistic range, this means that the free energy correction is impossible to calculate since it is too high (see Fig. [belfast-2-metadpath-correction-fig](#) ).

Now the last thing that one has to do to have the most plausible free energy landscape is to sum both the correction and the negative bias produced. This can be easily done in gnuplot as follows:

```

gnuplot> set pm3d
gnuplot> spl "<paste negative_bias.dat correction.dat " u 1:2:($3+$8) w pm3d
gnuplot> set view map
gnuplot> unset key
gnuplot> set xr [-2:23]
gnuplot> set contour
gnuplot> unset clabel
gnuplot> set cbrange [-140:-30]
gnuplot> set cntrp lev incr -140,6,-30

```

So now we can comment a bit on the free energy surface obtained and note that there is a free energy path that connects the two endpoints and runs very close to zero distance from the path. This means that our input path is actually resemblant of what is really happening in the system. Additionally you can see that there are many comparable routes different from the straight path. Can you make a sense of it just by looking at the free energy on the Ramachandran plot?

## 11.3 Belfast tutorial: Adaptive variables II

### 11.3.1 Aims

The aim of this tutorial is to consolidate the material that was covered during [Belfast tutorial: Analyzing CVs](#) and [Belfast tutorial: Adaptive variables I](#) on analysing trajectories using collective variables and path collective variables. We will then build on this material by showing how you can use the multidimensional scaling algorithm to automate the process of finding collective variables.

### 11.3.2 Learning Outcomes

Once this tutorial is completed students will:

- Know how to load colvar data into the GISMO plugin
- Know how to run the multidimensional scaling algorithms on a trajectory
- Be able to explain how we can automate the process of finding collective variables by seeking out an isometry between a high-dimensional and low-dimensional space



### 11.3.3 Resources

The `tarball` for this project contains the following files:

- `trajectory-short.xyz` : a (short) trajectory for a 16 residue protein in xyz format. All calculations with plumed driver use this trajectory.
- `trajectory-short.pdb` : the same trajectory in pdb format, this can be loaded with VMD
- `template.pdb` : a single frame from the trajectory that can be used in conjunction with the `MOLINFO` command

### 11.3.4 Instructions

#### 11.3.4.1 Visualising the trajectory

The aim of this tutorial is to understand the data contained in the trajectory called `trajectory-short.pdb`. This file contains some frames from a simulation from a 16 residue protein. As a start point then let load this trajectory with vmd and have a look at it. Type the following command into the command line:

```
vmd trajectory-short.pdb
```

Look at it with the various representations that vmd offers. If you are at the plumed tutorial try typing the letter m on the keyboard - we have made the new cartoon representation will update automatically for each frame of the trajectory - cool huh!

What are your impressions about this trajectory based on looking at it with VMD? How many basins in the free energy landscape is this trajectory sampling from? What can we tell from looking at this trajectory that we could perhaps put in a paper?

If your answers to the questions at the end of the above paragraph are I don't know that is good. We can tell very little by just looking at a trajectory. In fact the whole point of today has been to find ways of analyzing trajectories precisely so that we are not put in this position of staring at trajectories mystified!

#### 11.3.4.2 Finding collective variables

Right so lets come up with some CVs to analyse this trajectory with. As some of you may know we can understand the conformation of proteins by looking at the Ramachandran angles. For those of you who don't know here is a Wikipedia article:

[http://en.wikipedia.org/wiki/Ramachandran\\_plot](http://en.wikipedia.org/wiki/Ramachandran_plot)

Our protein has 32 ramachandran angles. We'll come back to that. For the time being pick out a couple that you think might be useful and construct a plumed input to calculate them and print them to a file. You will need to use the `TORSION` and `PRINT` commands in order to do this. Once you have created your plumed input use driver to calculate the torsional angles using the following command:

```
plumed driver --plumed plumed.dat --ixyz trajectory.xyz
```

If you have done this correctly you should have an output file containing your torsional angles. We can use vmd+GISMO to visualise the relationship between the ramachandran angles and the atomic configurations. To do this first load the trajectory in VMD:

```
vmd trajectory-short.pdb
```

Then click on Extensions>Analysis>GISMO. A new window should open in this window click on File>Load colvars. You will be asked to select a colvar file. Select the file that was output by the plumed calculation above. Once the file is loaded you should be able to select the labels that you gave to the Ramachandran angles you calculated with plumed. If you do so you will see that this data is plotted in the GISMO window so that you can interact with it and the trajectory.

What can you conclude from this exercise. Do the CV values of the various frames appear in clusters in the plane? Do points in different clusters correspond to structures that look the same or different? Are there similar looking structures clustered together or are they always far apart? What can we conclude about the various basins in the free energy landscape that have been explored in this trajectory? How many are there? Would your estimate be the same if you tried the above estimate with a different pair of ramachandran angles?

### 11.3.4.3 Dimensionality reduction

What we have done for most of today is seek out a function that takes as input the position of all the atoms in the system - a  $3N$  dimensional vector, where  $N$  is the number of atoms. This function then outputs a single number - the value of the collective variable - that tells us where in a low dimensional space we should project that configuration. Problems can arise because this collective-variable function is many-to-one. As you have hopefully seen in the previous exercise markedly different configurations of the protein can actually have quite similar values of a particular ramachandran angle.

We are going to spend the rest of this session introducing an alternative approach to this business of finding collective variables. In this alternative approach we are going to stop trying to seek out a function that can take any configuration of the atoms (any  $3N$ -dimensional vector) and find it's low dimensional projection on the collective variable axis. Instead we are going to take a set of configurations of the atoms (a set of  $3N$ -dimensional vectors of atom positions) and try to find a sensible set of projections for these configurations. We already touched on this idea earlier when we looked at paths. Our assumption, when we introduced this idea, was that we could find an ordered set of high-dimensional configurations that represented the transition pathway the system takes as it crossed a barrier and changed between two particularly interesting configurations. Lets say we have a path defined by four reference configurations - this implies that to travel between the configurations at the start and the end of this path you have to pass through configuration 1, then configuration 2, then configuration 3 and then configuration 4. This ordering means that the numbers 1 through 4 constitute sensible projections of these high-dimensional configurations. The numbers 1 through 4 all lie on a single cartesian axis - a low-dimensional space.

The problem when it comes to applying this idea to the data that we have in the trajectory-short trajectory is that we have no information on the "order" of these points. We have not been told that this trajectory represents the transition between two interesting points in phase space and thus we cannot apply the logic of paths. Hence, to seek out a low dimensional representation we are going to try and find a representation of this data we are going to seek out **an isometry** between the space containing the  $3N$ -dimensional vectors of atom positions and some lower-dimensional space. This idea is explained in more detail in the following **video** .

Let's now generate our isometric embedding. You will need to create a plumed input file that contains the following instructions:

```
CLASSICAL_MDS ...
  ATOMS=1-256
  METRIC=OPTIMAL-FAST
  USE_ALL_DATA
  NLOW_DIM=2
  OUTPUT_FILE=rmsd-embed
... CLASSICAL_MDS
```

You should then run this calculation using the following command:

```
plumed driver --ixyz trajectory-short.xyz --plumed plumed.dat
```

This should generate an output file called `rmsd-embed`. You should now be able to use VMD+GISMO to visualise this output.

Do the CV values of the various frames appear in clusters in the plane? Do points in different clusters correspond to structures that look the same or different? Are there similar looking structures clustered together or are they always far apart? What can we conclude about the various basins in the free energy landscape that have been explored in this trajectory? How many are there? Do you think this gives you a fuller picture of the trajectory than the ones you obtained by considering ramachandran angles?

### 11.3.5 Extensions

As discussed in the previous section this approach to trajectory analysis works by calculating distances between pairs of atomic configurations. Projections corresponding to these configurations are then generated in the low dimensional space in a way that tries to preserve these pairwise distances. There are, however, an infinite number of ways of calculating the distance between two high-dimensional configurations. In the previous section we used the RMSD distance but you could equally use the DRMSD distance. You could even try calculating a large number of collective variables for each of the high-dimensional points and seeing how much these all changed. You can use these different types of distances with the `CLASSICAL_MDS` action that was introduced in the previous section. If you have time less at the end of the session read the manual for the `CLASSICAL_MDS` action and see if you can calculate an MDS projection using an alternative definition of the distances between configurations. Some suggestions to try in order of increasing difficulty: DRMSD, how much the full set of 32 ramachandran angles change and the change in the contact map

### 11.3.6 Further Reading

There is a growing community of people using these ideas to analyse trajectory data. Some start points for investigating their work in more detail are:

- <http://sketchmap.berlios.de>
- <http://www.annualreviews.org/doi/abs/10.1146/annurev-physchem-040412-110006>

## 11.4 Belfast tutorial: Umbrella sampling

### 11.4.1 Aims

In the previous lectures we learned how to compute collective variables (CVs) from atomic positions. We will now learn how one can add a bias potential to enforce the exploration of a particular region of the space. We will also see how it is possible to bias CVs so as to enhance the sampling of events hindered by large free-energy barriers and how to analyze this kind of simulation. This technique is known as "umbrella sampling" and can be used in combination with the weighted-histogram analysis method to compute free-energy landscapes. In this tutorial we will use simple collective variables, but the very same approach can be used with any kind of collective variable.

## 11.4.2 Summary of theory

### 11.4.2.1 Biased sampling

A system at temperature  $T$  samples conformations from the canonical ensemble:

$$P(q) \propto e^{-\frac{U(q)}{k_B T}}$$

. Here  $q$  are the microscopic coordinates and  $k_B$  is the Boltzmann constant. Since  $q$  is a highly dimensional vector, it is often convenient to analyze it in terms of a few collective variables (see [Belfast tutorial: Analyzing CVs](#), [Belfast tutorial: Adaptive variables I](#), and [Belfast tutorial: Adaptive variables II](#)). The probability distribution for a CV  $s$  is

$$P(s) \propto \int dq e^{-\frac{U(q)}{k_B T}} \delta(s - s(q))$$

This probability can be expressed in energy units as a free energy landscape  $F(s)$ :

$$F(s) = -k_B T \log P(s)$$

Now we would like to modify the potential by adding a term that depends on the CV only. That is, instead of using  $U(q)$ , we use  $U(q) + V(s(q))$ . There are several reasons why one would like to introduce this potential. One is to avoid that the system samples some un-desired portion of the conformational space. As an example, imagine that you want to study dissociation of a complex of two molecules. If you perform a very long simulation you will be able to see association and dissociation. However, the typical time required for association will depend on the size of the simulation box. It could be thus convenient to limit the exploration to conformations where the distance between the two molecules is lower than a given threshold. This could be done by adding a bias potential on the distance between the two molecules. Another example is the simulation of a portion of a large molecule taken out from its initial context. The fragment alone could be unstable, and one might want to add additional potentials to keep the fragment in place. This could be done by adding a bias potential on some measure of the distance from the experimental structure (e.g. on root-mean-square deviation).

Whatever CV we decide to bias, it is very important to recognize which is the effect of this bias and, if necessary, remove it a posteriori. The biased distribution of the CV will be

$$P'(s) \propto \int dq e^{-\frac{U(q) + V(s(q))}{k_B T}} \delta(s - s(q)) \propto e^{-\frac{V(s(q))}{k_B T}} P(s)$$

and the biased free energy landscape

$$F'(s) = -k_B T \log P'(s) = F(s) + V(s) + C$$

Thus, the effect of a bias potential on the free energy is additive. Also notice the presence of an undetermined constant  $C$ . This constant is irrelevant for what concerns free-energy differences and barriers, but will be important later when we will learn the weighted-histogram method. Obviously the last equation can be inverted so as to obtain the original, unbiased free-energy landscape from the biased one just subtracting the bias potential

$$F(s) = F'(s) - V(s) + C$$

Additionally, one might be interested in recovering the distribution of an arbitrary observable. E.g., one could add a bias on the distance between two molecules and be willing to compute the unbiased distribution of some torsional angle. In this case there is no straightforward relationship that can be used, and one has to go back to the relationship between the microscopic probabilities:

$$P(q) \propto P'(q) e^{\frac{V(s(q))}{k_B T}}$$

The consequence of this expression is that one can obtain any kind of unbiased information from a biased simulation just by weighting every sampled conformation with a weight

$$w \propto e^{\frac{V(s(q))}{k_B T}}$$

That is, frames that have been explored in spite of a high (disfavoring) bias potential  $V$  will be counted more than frames that have been explored with a less disfavoring bias potential.

## 11.4.2.2 Umbrella sampling

Often in interesting cases the free-energy landscape has several local minima. If these minima have free-energy differences that are on the order of a few times  $k_B T$  they might all be relevant. However, if they are separated by a high saddle point in the free-energy landscape (i.e. a low probability region) then the transition between one and the other will take a lot of time and these minima will correspond to metastable states. The transition between one minimum and the other could require a time scale which is out of reach for molecular dynamics. In these situations, one could take inspiration from catalysis and try to favor in a controlled manner the conformations corresponding to the transition state.

Imagine that you know since the beginning the shape of the free-energy landscape  $F(s)$  as a function of one CV  $s$ . If you perform a molecular dynamics simulation using a bias potential which is exactly equal to  $-F(s)$ , the biased free-energy landscape will be flat and barrierless. This potential acts as an "umbrella" that helps you to safely cross the transition state in spite of its high free energy.

It is however difficult to have an a priori guess of the free-energy landscape. We will see later how adaptive techniques such as metadynamics ([Belfast tutorial: Metadynamics](#)) can be used to this aim. Because of this reason, umbrella sampling is often used in a slightly different manner.

Imagine that you do not know the exact height of the free-energy barrier but you have an idea of where the barrier is located. You could try to just favor the sampling of the transition state by adding a harmonic restraint on the CV, e.g. in the form

$$V(s) = \frac{k}{2}(s - s_0)^2$$

. The sampled distribution will be

$$P'(q) \propto P(q) e^{-\frac{k(s(q)-s_0)^2}{2k_B T}}$$

For large values of  $k$ , only points close to  $s_0$  will be explored. It is thus clear how one can force the system to explore only a predefined region of the space adding such a restraint. By combining simulations performed with different values of  $s_0$ , one could obtain a continuous set of simulations going from one minimum to the other crossing the transition state. In the next section we will see how to combine the information from these simulations.

## 11.4.2.3 Weighted histogram analysis method

Let's now consider multiple simulations performed with restraints located in different positions. In particular, we will consider the  $i$ -th bias potential as  $V_i$ . The probability to observe a given value of the collective variable  $s$  is:

$$P_i(s) = \frac{P(s) e^{-\frac{V_i(s)}{k_B T}}}{\int ds' P(s') e^{-\frac{V_i(s')}{k_B T}}} = \frac{P(s) e^{-\frac{V_i(s)}{k_B T}}}{Z_i}$$

where

$$Z_i = \sum_q e^{-(U(q) + V_i(q))}$$

The likelihood for the observation of a sequence of snapshots  $q_i(t)$  (where  $i$  is the index of the trajectory and  $t$  is time) is just the product of the probability of each of the snapshots. We use here the minus-logarithm of the likelihood (so that the product is converted to a sum) that can be written as

$$\mathcal{L} = - \sum_i \int dt \log P_i(s_i(t)) = \sum_i \int dt \left( -\log P(s_i(t)) + \frac{V_i(s_i(t))}{k_B T} + \log Z_i \right)$$

One can then maximize the likelihood by setting  $\frac{\delta \mathcal{L}}{\delta P(s)} = 0$ . After some boring algebra the following expression can be obtained

$$0 = \sum_i \int dt \left( -\frac{\delta_{s_i(t), s}}{P(s)} + \frac{e^{-\frac{V_i(s)}{k_B T}}}{Z_i} \right)$$

In this equation we aim at finding  $P(s)$ . However, also the list of normalization factors  $Z_i$  is unknown, and they should be found selfconsistently. Thus one can find the solution as

$$P(\mathbf{s}) \propto \frac{N(\mathbf{s})}{\sum_i \int dt \frac{e^{-\frac{V_i(\mathbf{s})}{k_B T}}}{Z_i}}$$

where  $Z$  is selfconsistently determined as

$$Z_i \propto \int d\mathbf{s}' P(\mathbf{s}') e^{-\frac{V_i(\mathbf{s}')}{k_B T}}$$

These are the WHAM equations that are traditionally solved to derive the unbiased probability  $P(s)$  by the combination of multiple restrained simulations. To make a slightly more general implementation, one can compute the weights that should be assigned to each snapshot, that turn out to be:

$$w_i(t) \propto \frac{1}{\sum_j \int dt \frac{e^{-\beta V_j(\mathbf{s}_i(t))}}{Z_j}}$$

The normalization factors can in turn be found from the weights as

$$Z_i \propto \frac{\sum_j \int dt e^{-\beta V_i(\mathbf{s}_j(t))} w_j(t)}{\sum_j \int dt w_j(t)}$$

This allows to straightforwardly compute averages related to other, non-biased degrees of freedom, and it is thus a bit more flexible. It is sufficient to precompute this factors  $w$  and use them to weight every single frame in the trajectory.

### 11.4.3 Learning Outcomes

Once this tutorial is completed students will know how to:

- Setup simulations with restraints.
- Use multiple-restraint umbrella sampling simulations to enhance the transition across a free-energy barrier.
- Analyze the results and compute weighted averages and free-energy profiles.

### 11.4.4 Resources

The `tarball` for this project contains the following files:

- A gromacs topology (topol.top), configuration (conf.gro), and control file (grompp.mdp). They should not be needed.
- A gromacs binary file (topol.tpr). This is enough for running this system.
- A small C++ program that computes WHAM (wham.cpp) and a script that can be used to feed it (wham.sh)

By working in the directory where the topol.tpr file is stored, one can launch gromacs with the command

```
gmx_mpi mdrun -plumed plumed.dat -nsteps 100000
```

(notice that the -nsteps flag allows the number of steps to be changed).

### 11.4.5 Instructions

#### 11.4.5.1 The model system

We here use a a model system alanine dipeptide with CHARM27 all atom force field already seen in the previous section.

#### 11.4.5.2 Restrained simulations

The simplest way in which one might influence a CV is by forcing the system to stay close to a chosen value during the simulation. This is achieved with a restraining potential that PLUMED provides via the directive [RESTRAINT](#). In the umbrella sampling method a bias potential is added so as to favor the exploration of some regions of the conformational space and to disfavor the exploration of other regions [58]. A properly chosen bias potential could allow for example to favor the transition state sampling thus enhancing the transition state for a conformational transition. However, choosing such a potential is not trivial. In a later section we will see how metadynamics can be used to this aim. The simplest way to use umbrella sampling is that to apply harmonic constraints to one or more CVs.

We will now see how to enforce the exploration of a the neighborhood of a selected point the CV space using a [RESTRAINT](#) potential.

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Impose an umbrella potential on CV 1 and CV 2
# with a spring constant of 500 kJoule/mol
# at fixed points on the Ramachandran plot
#
restraint-phi: RESTRAINT ARG=phi KAPPA=500 AT=-0.3
restraint-psi: RESTRAINT ARG=psi KAPPA=500 AT=+0.3

# monitor the two variables and the bias potential from the two restraints
PRINT STRIDE=10 ARG=phi,psi,restraint-phi.bias,restraint-psi.bias FILE=COLVAR
```

(see [TORSION](#), [RESTRAINT](#), and [PRINT](#)).

The syntax for the command [RESTRAINT](#) is rather trivial. The directive is followed by a keyword ARG followed by the label of the CV on which the umbrella potential has to act. The keyword KAPPA determines the hardness of the spring constant and its units are [Energy units]/[Units of the CV]. The additional potential introduced by the UMBRELLA takes the form of a simple Hooke's law:

$$U(s) = \frac{k}{2}(x - x_0)^2$$

.

where  $x_0$  is the value specified following the AT keyword. The choice of AT ( $x_0$ ) is obviously depending on the specific case. KAPPA ( $k$ ) is typically chosen not to affect too much the intrinsic fluctuations of the system. A typical recipe is  $k \approx \frac{k_B T}{\sigma^2}$ , where  $\sigma^2$  is the variance of the CV in a free simulation). In real applications, one must be careful with values of  $k$  larger than  $\frac{k_B T}{\sigma^2}$  because they could break down the molecular dynamics integrator.

The CVs as well as the two bias potentials are shown in the COLVAR file. For this specific input the COLVAR file has in first column the time, in the second the value of  $\phi$ , in the third the value of  $\psi$ , in the fourth the the additional potential introduced by the restraint on  $\phi$  and in the fifth the additional potential introduced by the restraint on  $\psi$ .

It may happen that one wants that a given CV just stays within a given range of values. This is achieved in plumed through the directives [UPPER\\_WALLS](#) and [LOWER\\_WALLS](#) that act on specific collective variables and limit the exploration within given ranges.

### 11.4.5.3 Reweighting the results

Now consider a simulation performed restraining the variable  $\phi$ :

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
restraint-phi: RESTRAINT ARG=phi KAPPA=10.0 AT=-2
PRINT STRIDE=10 ARG=phi,psi,restraint-phi.bias FILE=COLVAR10
```

and compare the result with the one from a single simulation with no restraint

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
# we use a "dummy" restraint with strength zero here
restraint-phi: RESTRAINT ARG=phi KAPPA=0.0 AT=-2
PRINT STRIDE=10 ARG=phi,psi,restraint-phi.bias FILE=COLVAR0
```

Plot the time dependence of  $\phi$  in the two cases and try to understand the difference.

Now let's try to compute the probability that  $\psi$  falls within a given range, say between 1 and 2. This can be done e.g. with this shell script

```
> grep -v \# COLVAR0 | tail -n 80000 |
  awk '{if($3>1 && $3<2)a++; else b++;}END{print a/(a+b)}'
```

Notice that we here considered only the last 80000 frames in the average. Look at the time series for  $\psi$  and guess why. Also notice that the script is removing the initial comments. After this trivial preprocessing, the script is just counting how many times the third column ( $\psi$ ) lies between 1 and 2 and how many times it doesn't. At the end it prints the number of times the variable is between 1 and 2 divided by the total count. The result should be something around 0.40. Now try to do it on trajectories generated with different values of AT. Does the result depend on AT?

We can now try to reweight the result so as to get rid of the bias introduced by the restraint. Since the reweighting factor is just  $\exp(-\frac{V}{k_B T})$  the script should be modified as

```
> grep -v \# COLVAR10 | tail -n 80000 |
  awk '{w=exp($4/2.5); if($3>1 && $3<2)a+=w; else b+=w;}END{print a/(a+b)}'
```

Notice that 2.5 is just  $k_B T$  in kJ/mol units.

Repeat this calculation for different values of AT. Does the result depend on AT?

### 11.4.5.4 A free-energy landscape

One can also count the probability of an angle to be in a precise bin. The logarithm of this quantity, in  $k_B T$  units, is the free-energy associated to that bin. There are several ways to compute histograms, either with PLUMED or with external programs. Here I decided to use awk.



```
grep -v \# COLVAR10 | tail -n 80000 |
awk 'BEGIN{
  min1=-3.14159265358979
  max1=+3.14159265358979
  min2=-3.14159265358979
  max2=+3.14159265358979
  nb1=100;
  nb2=100;
  for(i1=0;i1<nb1;i1++) for(i2=0;i2<nb2;i2++) f[i1,i2]=0.0;
}{
  i1=int(($2-min1)*nb1/(max1-min1));
  i2=int(($3-min2)*nb2/(max2-min2));
# we assume the potential is in the last column, and kbT=2.5 kj/mol
  w=exp($NF/2.5);
  f[i1,i2]+=w;
}
END{
  for(i1=0;i1<nb1;i1++){
    for(i2=0;i2<nb2;i2++) print min1+i1/100.0*(max1-min1), min2+i2/100.0*(max2-min2), -2.5*log(f[i1,i2]);
    print "";
  }
}' > plotme
```

You can then plot the "plotme" file with

```
gnuplot> set pm3d map
gnuplot> splot "plotme"
```

#### 11.4.5.5 Combining multiple restraints

In the last paragraph you have seen how to reweight simulations done with restraints in different positions to obtain virtually the same result. Let's now see how to combine data from multiple restraint simulations. A possible choice is to download and use the WHAM software [here](#), which is well documented. This is probably the best idea for analyzing a real simulation.

For the sake of learning a bit, we will use a different approach here, namely we will use a short C++ program that implements the weight calculation. Notice that whereas people typically use harmonic restraints in this framework, PLUMED offers a very large variety of bias potentials. For this reason we will keep things as general as possible and use an approach that can be in principle used also to combine simulation with restraint on different variables or with complicated bias potential.

The first step is to generate several simulations with different positions of the restraint, gradually going from say -2 to +2. You can obtain them using e.g. the following script:

```
for AT in -2.0 -1.5 -1.0 -0.5 +0.0 +0.5 +1.0 +1.5 +2.0
do

cat >plumed.dat << EOF
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Impose an umbrella potential on CV 1 and CV 2
# with a spring constant of 500 kjoule/mol
# at fixed points on the Ramachandran plot
#
restraint-phi: RESTRAINT ARG=phi KAPPA=40.0 AT=$AT
# monitor the two variables and the bias potential from the two restraints
PRINT STRIDE=10 ARG=phi,psi,restraint-phi.bias FILE=COLVAR$AT
EOF

gmx_mpi mdrun -plumed plumed.dat -nsteps 100000 -x traj$AT.xtc

done
```

Notice that we are here saving separate trajectories for the separate simulation, as well as separate colvar files. In each simulation the restraint is located in a different position. Have a look at the plot of (phi,psi) for the different simulations to understand what is happening.

An often misunderstood fact about WHAM is that data of the different trajectories can be mixed and it is not necessary to keep track of which restraint was used to produce every single frame. Let's get the concatenated trajectory

```
gmx_mpi trjcat -cat -f traj*.xtc -o alltraj.xtc
```

Now we should compute the value of each of the bias potentials on the entire (concatenated) trajectory

```
for AT in -2.0 -1.5 -1.0 -0.5 +0.0 +0.5 +1.0 +1.5 +2.0
do

cat >plumed.dat << EOF
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
restraint-phi: RESTRAINT ARG=phi KAPPA=40.0 AT=$AT

# monitor the two variables and the bias potential from the two restraints
PRINT STRIDE=10 ARG=phi,psi,restraint-phi.bias FILE=ALLCOLVAR$AT
EOF

plumed driver --mf_xtc alltraj.xtc --trajectory-stride=10 --plumed plumed.dat

done
```

It is very important that this script is consistent with the one used to generate the multiple simulations above. Now, single files named ALLCOLVARXX will contain on the fourth column the value of the bias centered in XX computed on the entire concatenated trajectory.

Next step is to compile the C++ program that computes weights self-consistently solving the WHAM equations. This is named wham.cpp and can be compiled with

```
g++ -O3 wham.cpp -o wham.x
```

and can be then used through a wrapper script wham.sh as

```
./wham.sh ALLCOLVAR* > colvar
```

The resulting colvar file will contain 3 columns: time, phi, and psi, plus the weights obtained from WHAM written in logarithmic scale. That is, the file will contain  $k_B T \log w$ .

Try now to use this file to compute the unbiased free-energy landscape as a function of phi and psi. You can use the script that you used earlier to compute histogram.

## 11.4.6 Comments

### 11.4.6.1 How does PLUMED work

The fact that when you add a force on the collective variable PLUMED can force the atoms to do something depends on the fact that the collective variables implemented in PLUMED has analytical derivatives. By biasing the value of a single CV one turns to affect the time evolution of the system itself. Notice that some of the collective variables could be implemented without derivatives (either because the developers were lazy or because the CVs cannot be derived). In this case you might want to have a look at the NUMERICAL\_DERIVATIVES option.

### 11.4.7 Further Reading

Umbrella sampling method is a widely used technique. You can find several resources on the web, e.g.:

- [http://en.wikipedia.org/wiki/Umbrella\\_sampling](http://en.wikipedia.org/wiki/Umbrella_sampling)

## 11.5 Belfast tutorial: Out of equilibrium dynamics

In plumed you can bring a system in a specific state in a collective variable by means of the [MOVINGRESTRAINT](#) directive. This directive is very flexible and allows for a programmed series of draggings and can be used also to sample multiple events within a single simulation. Here I will explain the concepts of it and show some examples

### 11.5.1 Resources

Here is the [tarball with the files referenced in the following](#).

### 11.5.2 Steered MD

Steered MD (SMD) is often used to drag the system from an initial configuration to a final one by pulling one or more CVs. Most of time the aim of such simulations is to prepare the system in a particular state or produce nice snapshots for a cool movie. All the CVs present in PLUMED can be used in SMD.

In SMD the Hamiltonian of the system  $H$  is modified into  $H_\lambda$ . This new Hamiltonian contains now another new term which now depends on time only via a Harmonic potential centered on a point which moves linear with time

$$\begin{aligned} H_\lambda(X, t) &= H(X) + U_\lambda(X, t) \\ &= H(X) + \frac{k(t)}{2} (s(X) - \lambda(t))^2 \\ &= H(X) + \frac{k(t)}{2} (s(X) - s_0 - vt)^2. \end{aligned}$$

This means that if the  $k$  is tight enough the system will follow closely the center of the moving harmonic spring. But be careful, if the spring constant is too hard your equations of motion will now keep up since they are tuned to the fastest motion in your system so if you artificially introduce a higher frequency in your system you'll screw up the dynamics. The same is true for the pulling speed  $v$ . As a matter of fact I never encountered the case where I had to lower the time step and I could all the time be happy just by making a softer spring constant or a slower steering speed. Generally, integrators of equations of motion like velocity-Verlet are very tolerant. Note that one can also make the spring constant depend on time and this, as we will see later in the examples is particularly useful to prepare your state.

In simulations, it is more convenient to adopt a situation where you specify only the starting point, the final point of cvs and the time in which you want to cover the transition. That's why the plumed input is done in such a way.

For example, let's go back to the alanine dipeptide example encountered in [The molecule of the day: alanine dipeptide](#). Let's say that now we want to steer from  $C_{7eq}$  to  $C_{7ax}$ . If you think, just by dragging along the  $\Phi$  dihedral angle from a value of -1 to a value 1 should be enough to the state of interest. Additionally, it might be important to you not to stress the system too much, so you might want first to increase the  $k$  first so to lock the system in  $\Phi = -1$ , then move it gently to  $\Phi = 1$  and then release again your spring constant so to end up to an equilibrated and unconstrained state. This you can program in PLUMED like this

```
# set up two variables for Phi and Psi dihedral angles
# drag this
phi: TORSION ATOMS=5,7,9,15
# this is just to monitor that you end up in the interesting state
psi: TORSION ATOMS=7,9,15,17
# the movingrestraint
restraint: ...
    MOVINGRESTRAINT
    ARG=phi
    AT0=-1.0 STEP0=0          KAPPA0=0
    AT1=-1.0 STEP1=2000      KAPPA1=1000
    AT2=1.0 STEP2=4000      KAPPA2=1000
    AT3=1.0 STEP3=6000      KAPPA3=0
...
# monitor the two variables and various restraint outputs
PRINT STRIDE=10 ARG=* FILE=COLVAR
```

Please note the syntax of **MOVINGRESTRAINT** : You need one (or more) argument(s) and a set of steps denote by ATX, STEPX, KAPPAX where X is a incremental starting from 0 that assign the center and the harness of the spring at step STEPX. What happens in between is a linear interpolation of the AT and KAPPA parameters. If those are identical in two consecutive steps then nothing is happening to that parameter. So if you put the same KAPPA and AT in two different STEPs then this will give you an umbrella potential placed in the same point between the two time intervals defined by STEP. Note that you need to run a bit more than 6000 steps because after this your system has no more restraints so the actual thermalization period starts here.

The COLVAR file produced has the following shape

```
#! FIELDS time phi psi restraint.bias restraint.force2 restraint.phi_cntr restraint.phi_work
#! SET min_phi -pi
#! SET max_phi pi
#! SET min_psi -pi
#! SET max_psi pi
0.000000 -1.409958 1.246193 0.000000 0.000000 -1.000000 0.000000
0.020000 -1.432352 1.256545 0.467321 4.673211 -1.000000 0.441499
0.040000 -1.438652 1.278405 0.962080 19.241592 -1.000000 0.918101
0.060000 -1.388132 1.283709 1.129846 33.895372 -1.000000 1.344595
0.080000 -1.360254 1.275045 1.297832 51.913277 -1.000000 1.691475
...
```

So we have time, phi, psi and the bias from the moving restraint. Note that at step 0 is zero since we imposed this to start from zero and ramp up in the first 2000 steps up to a value of 2000 kJ/mol/rad<sup>2</sup>. It increases immediately since already at step 1 the harmonic potential is going to be increased in bits towards the value of 1000 which is set by KAPPA. The value of restraint.force2 is the squared force (which is a proxy of the force magnitude, despite the direction) on the CV.

$$-\frac{\partial H_{\lambda}(X,t)}{\partial s} = -(s(X) - s_0 - vt)$$

Note that the actual force on an atom of the system involved in a CV is instead

$$\begin{aligned} -\frac{\partial H_{\lambda}(X,t)}{\partial x_i} &= -\frac{\partial H_{\lambda}(X,t)}{\partial s} \frac{\partial s}{\partial x_i} \\ &= -(s(X) - s_0 - vt) \frac{\partial s}{\partial x_i} \end{aligned}$$

This is important because in CVs that have a derivative that change significantly with space then you might have regions in which no force is exerted while in others you might have an enormous force on it. Typically this is the case of sigmoids that are used in coordination numbers in which, in the tails, they are basically flat as a function of

particle positions. Additionally note that this happens on any force-based enhanced-sampling algorithm so keep it in mind. Very often people miss this aspect and complain either that a CV or a enhanced-sampling method does not work. This is another good reason to use tight spring force so to compensate in the force the lack of derivative from the CV.

The other argument in colvar is `restraint.phi_cntr` which is the center of the harmonic potential. This is a useful quantity so you may know how close the system is following the center of harmonic potential (more on this below). The last parameter is `restraint.phi_work`. The work is defined as

$$W = \int_0^{t_s} dt \frac{\partial H_\lambda(t)}{\partial t}$$

so this is changing only when the Hamiltonian is changing with time. There are two time dependent contributions in this integral: one can come from the fact that  $k(t)$  changes with time and another from the fact that the center of the spring potential is changing with time.

$$\begin{aligned} W &= \int_0^{t_s} dt \frac{\partial H_\lambda(t)}{\partial t} \\ &= \int_0^{t_s} dt \frac{\partial H_\lambda(t)}{\partial \lambda} \frac{\partial \lambda(t)}{\partial t} + \int_0^{t_s} dt \frac{\partial H_\lambda(t)}{\partial k} \frac{\partial k}{\partial t} \\ &= \int_0^{t_s} -k(t)(s(X) - \lambda(t)) \frac{\partial \lambda(t)}{\partial t} dt + \int_0^{t_s} \frac{(s(X) - \lambda(t))^2}{2} \frac{\partial k}{\partial t} dt \\ &= \int_0^{t_s} -k(t)(s(X) - \lambda(t)) v dt + \int_0^{t_s} \frac{(s(X) - \lambda(t))^2}{2} \frac{\partial k}{\partial t} dt \\ &\simeq \sum_i -k(t_i)(s(X(t_i)) - \lambda(t_i)) \Delta \lambda(t_i) + \sum_i \frac{(s(X(t_i)) - \lambda(t_i))^2}{2} \Delta k(t_i) \end{aligned}$$

where we denoted  $\Delta \lambda(t_i)$  the difference of the center of the harmonic potential respect to the step before and  $\Delta k(t_i)$  is the difference in spring constant respect to the step before. So in the exercised proposed in the first phase you see only the second part of the work since this is the part connected with the spring constant increase. After this phase you see the increase due to the motion of the center and then you later the release of the spring constant.

### The work profile as function of time when steering ala dipeptide along the $\Phi$ variable.

This you get with gnuplot:

```
pd@plumed:~>gnuplot
gnuplot> p "COLVAR" u 1:7 w lp
```

Another couple of interesting thing that you can check is

- Is my system finally in the *C7ax* ? Plot the two dihedral to have a sense if we are in the right state. You know the target position what should look like, right?
- Is my system moving close to the center of the harmonic potential? This is important and we will see why in a while.

### 11.5.3 Moving on a more complex path

Very often it is useful to use this `movingrestraint` to make a fancier schedule by using nice properties of `MOVINGRESTRAINT`. For example you can plan a schedule to drive multiple CVs at the same time in specific point of the phase space and also to stop for a while in specific using a fixed harmonic potential. This can be handy in case of an umbrella sampling run where you might want to explore a 1-dimensional landscape by acquiring some statistics in one point and then moving to the next to acquire more statistics. With `MOVINGRESTRAINT` you can do it in only one file. To give an example of such capabilities, let's say that we want to move from  $C7eq$  vertically toward  $\Phi = -1.5; \Psi = -1.3$ , stop by for a while (e.g. to acquire a statistics that you might need for an umbrella sampling), then moving toward  $\Phi = 1.3; \Psi = -1.3$  which roughly corresponds to  $C7ax$ .

This can be programmed conveniently with `MOVINGRESTRAINT` by adopting the following schedule

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
# the movingrestraint
restraint: ...
    MOVINGRESTRAINT
    ARG=phi,psi
    AT0=-1.5,1.3 STEP0=0 KAPPA0=0,0
    AT1=-1.5,1.3 STEP1=2000 KAPPA1=1000,1000
    AT2=-1.5,-1.3 STEP2=4000 KAPPA2=1000,1000
    AT3=-1.5,-1.3 STEP3=4000 KAPPA3=1000,1000
    AT4=1.3,-1.3 STEP4=6000 KAPPA4=1000,1000
    AT5=1.3,-1.3 STEP5=8000 KAPPA5=0,0
...
# monitor the two variables and various restraint outputs
PRINT STRIDE=10 ARG=* FILE=COLVAR
```

Note that by adding two arguments for `movingrestraint`, now I am allowed to put two values (separated by comma, as usual for multiple values in `PLUMED`) and correspondingly two `KAPPA` values. One for each variable. Please note that no space must be used between the arguments! This is a very common fault in writing the inputs.

By plotting the instantaneous value of the variables and the value of the center of the harmonic potentials we can inspect the pathways that we make the system walk on the Ramachandran plot. (How to do this? Have a look to the header of `COLVAR` file to plot the right fields)

.png

#### Plot of the double steering schedule using `MOVINGRESTRAINT`

### 11.5.4 Why work is important?

The work as we have seen is the cumulative change of the hamiltonian in time. So it is connected with the change in energy of your system while you move it around. It has also a more important connection with the free energy via the Jarzynski equation which reads

$$\Delta F = -\beta^{-1} \ln \langle \exp^{-\beta W} \rangle$$

This is important and says that potentially you can calculate the free energy even by driving your system very fast and out of equilibrium between two states of your interest. Unfortunately this in practice not possible since the accurate calculation of the quantity  $\langle \exp^{-\beta W} \rangle$  has a huge associated error bar since it involves the average of a noisy quantity (the work) being exponentiated. So, before going wild with SMD, I want to make a small exercise on how tricky that is even for the smallest system.

Now we run, say 30 SMD run and we calculate the free energy difference by using Jarzynski equality and see how this differs from the average. First note that the average  $\langle \exp^{-\beta W} \rangle$  is an average over a number of steered MD runs which start from the same value of CV and reach the final value of CV. So it is important to create initially an

ensemble of states which are compatible with a given value of CVs. Let's assume that we can do this by using a restrained MD in a point (say at  $\Phi = -1.5$ ). In practice the umbrella biases a bit your distribution and the best situation would be to do this with a flat bottom potential and then choosing the snapshot that correspond to the wanted starting value and start from them.

In the directory JARZ/MAKE\_ENSEMBLE you find the script to run. After you generate the constrained ensemble this needs to be translated from xtc format to something that GROMACS is able to read in input, typically a more convenient gro file. To do so just to

```
pd@plumed:~> echo 0 | trjconv_mpi-dp-pl -f traj.xtc -s topol.tpr -o all.gro
pd@plumed:~> awk 'BEGIN{i=1}{if($1=="Generated"){outfile=sprintf("start_%d.gro",i);i++;}print >>outfile; if(NF
```

This will generate a set of numbered gro files. Now copy them in the parallel directory MAKE\_STEER. There you will find a script (script.sh) where you can set the number of runs that you want to go for. Just try 20 and let it run. Will take short time. The script will also produce a script\_rama.gplt that you can use to visualize all the work performed in a single gnuplot session. Just do:

```
pd@plumed:~> gnuplot
gnuplot> load "script_work.gplt"
```

What you see is something like in Fig.

There are a number of interesting fact here. First you see that different starting points may end with very different work values. Not bad, one would say, since Jarzynski equality is saying that we can make an average and everything will be ok. So now calculate the average work by using the following bash one-liner:

```
pd@plumed:~> ntest=20; for i in `seq 1 $ntest`; do tail -1 colvar_$i | awk '{print $7 }' ; done | awk '{g+=ex
```

For my test, what I get is a value of

```
FREE ENERGY ESTIMATE 17.482 STDEV 7.40342
```

and what this is saying is that the only thing that matters is the lowest work that I sampled. This has such an enormous weight over all the other trajectories that will do so that it will be the only other to count, and all the other do not matter much. So it is a kind of a waste of time. Also the standard deviation is rather high and probably it might well be that you obtain a much better result by using a standard umbrella sampling where you can use profitably most of the statistics. Here you waste most of the statistics indeed, since only the lowest work sampled will matter.

Some important point for doing some further exercises:

- How does the work distribution change if you increase the simulation time? Note that you have to increase both the time in the md.mdp file and in the plumed.dat file.
- How the work change if you now use a softer spring constant? And a harder one?
- In particular, what happens when you have softer spring constant, say 10? This does not look like working? Can you guess what is going on there from an analysis of COLVAR files only?
- Have a look of the trajectories in the Ramachandran plot in case of fast simulations and slow simulation. What can you observe? Is there a correlation between steering speed and how often you can go on the low energy path?

### 11.5.5 Targeted MD

Targeted MD can be seen as a special case of steered MD where the RMSD from a reference structure is used as a collective variable. It can be used for example if one wants to prepare the system so that the coordinates of selected atoms are as close as possible to a target pdb structure.

As an example we can take alanine dipeptide again

```
# set up two variables for Phi and Psi dihedral angles
# these variables will be just monitored to see what happens
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
# creates a CV that measures the RMSD from a reference pdb structure
# the RMSD is measured after OPTIMAL alignment with the target structure
rmsd: RMSD REFERENCE=c7ax.pdb TYPE=OPTIMAL
# the movingrestraint
restraint: ...
    MOVINGRESTRAINT
    ARG=rmsd
    AT0=0.0 STEP0=0      KAPPA0=0
    AT1=0.0 STEP1=5000   KAPPA1=10000
...
# monitor the two variables and various restraint outputs
PRINT STRIDE=10 ARG=* FILE=COLVAR
```

(see [TORSION](#), [RMSD](#), [MOVINGRESTRAINT](#), and [PRINT](#)).

Note that [RMSD](#) should be provided a reference structure in pdb format and can contain part of the system but the second column (the index) must reflect that of the full pdb so that PLUMED knows specifically which atom to drag where. The [MOVINGRESTRAINT](#) bias potential here acts on the rmsd, and the other two variables (phi and psi) are untouched. Notice that whereas the force from the restraint should be applied at every step (thus rmsd is computed at every step) the two torsions are computed only every 10 steps. PLUMED automatically detect which variables are used at every step, leading to better performance when complicated and computationally expensive variables are monitored - this is not the case here, since the two torsions are very fast to compute. Note that here the work always increase with time and never gets lower which is somewhat surprising if you think that we are moving in another metastable state. One would expect this to bend and give a signal of approaching a minimum like before. Nevertheless consider what you are doing: we are constraining the system in one specific conformation and this is completely unnatural for a system at 300 kelvin so, even for this small system adopting a specific conformation in which all the heavy atoms are in a precise position is rather unrealistic. This means that this state is an high free energy state.

## 11.6 Belfast tutorial: Metadynamics

### 11.6.1 Aims

The aim of this tutorial is to introduce the users to running a metadynamics simulation with PLUMED. We will set up a simple simulation of alanine dipeptide in vacuum, analyze the output, and estimate free energies from the simulation. We will also learn how to run a well-tempered metadynamics simulation and detect issues related to a bad choice of collective variables.



### 11.6.2 Summary of theory

In metadynamics, an external history-dependent bias potential is constructed in the space of a few selected degrees of freedom  $\vec{s}(q)$ , generally called collective variables (CVs) [41]. This potential is built as a sum of Gaussians deposited along the trajectory in the CVs space:

$$V(\vec{s}, t) = \sum_{k\tau < t} W(k\tau) \exp \left( - \sum_{i=1}^d \frac{(s_i - s_i(q(k\tau)))^2}{2\sigma_i^2} \right).$$

where  $\tau$  is the Gaussian deposition stride,  $\sigma_i$  the width of the Gaussian for the  $i$ th CV, and  $W(k\tau)$  the height of the Gaussian. The effect of the metadynamics bias potential is to push the system away from local minima into visiting new regions of the phase space. Furthermore, in the long time limit, the bias potential converges to minus the free energy as a function of the CVs:

$$V(\vec{s}, t \rightarrow \infty) = -F(\vec{s}) + C.$$

In standard metadynamics, Gaussians of constant height are added for the entire course of a simulation. As a result, the system is eventually pushed to explore high free-energy regions and the estimate of the free energy calculated from the bias potential oscillates around the real value. In well-tempered metadynamics [43], the height of the Gaussian is decreased with simulation time according to:

$$W(k\tau) = W_0 \exp \left( - \frac{V(\vec{s}(q(k\tau)), k\tau)}{k_B \Delta T} \right),$$

where  $W_0$  is an initial Gaussian height,  $\Delta T$  an input parameter with the dimension of a temperature, and  $k_B$  the Boltzmann constant. With this rescaling of the Gaussian height, the bias potential smoothly converges in the long time limit, but it does not fully compensate the underlying free energy:

$$V(\vec{s}, t \rightarrow \infty) = - \frac{\Delta T}{T + \Delta T} F(\vec{s}) + C.$$

where  $T$  is the temperature of the system. In the long time limit, the CVs thus sample an ensemble at a temperature  $T + \Delta T$  which is higher than the system temperature  $T$ . The parameter  $\Delta T$  can be chosen to regulate the extent of free-energy exploration:  $\Delta T = 0$  corresponds to standard molecular dynamics,  $\Delta T \rightarrow \infty$  to standard metadynamics. In well-tempered metadynamics literature and in PLUMED, you will often encounter the term "biasfactor" which is the ratio between the temperature of the CVs ( $T + \Delta T$ ) and the system temperature ( $T$ ):

$$\gamma = \frac{T + \Delta T}{T}.$$

The biasfactor should thus be carefully chosen in order for the relevant free-energy barriers to be crossed efficiently in the time scale of the simulation.

Additional information can be found in the several review papers on metadynamics [59] [60] [61].

### 11.6.3 Learning Outcomes

Once this tutorial is completed students will know how to:

- run a metadynamics simulation using PLUMED
- analyze the output of the simulation
- restart a metadynamics simulation
- calculate free energies from a metadynamics simulation
- run a well-tempered metadynamics simulation using PLUMED
- detect issues with the choice of the collective variables

### 11.6.4 Resources

The `tarball` for this project contains the following directories:

- TOPO: it contains the gromacs topology and configuration files to simulate alanine dipeptide in vacuum
  - Exercise\_1: run a metadynamics simulation with 2 CVs, dihedrals phi and psi, and analyze the output
- Exercise\_2: restart a metadynamics simulation
- Exercise\_3: calculate free energies from a metadynamics simulation and monitor convergence
- Exercise\_4: run a well-tempered metadynamics simulation with 2 CVs, dihedrals phi and psi
- Exercise\_5: run a well-tempered metadynamics simulation with 1 CV, dihedral psi

### 11.6.5 Instructions

#### 11.6.5.1 The model system

Here we use as model system alanine dipeptide in vacuum with AMBER99SB-ILDN all-atom force field.

#### 11.6.5.2 Exercise 1. Setup and run a metadynamics simulation

In this exercise, we will run a metadynamics simulation on alanine dipeptide in vacuum, using as CVs the two backbone dihedral angles phi and psi. In order to run this simulation we need to prepare the PLUMED input file (`plumed.dat`) as follows.

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate metadynamics in phi and psi
# depositing a Gaussian every 500 time steps,
# with height equal to 1.2 kJoule/mol,
# and width 0.35 rad for both CVs.
#
metad: METAD ARG=phi,psi PACE=500 HEIGHT=1.2 SIGMA=0.35,0.35 FILE=HILLS

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR
```

(see [TORSION](#), [METAD](#), and [PRINT](#)).

The syntax for the command [METAD](#) is simple. The directive is followed by a keyword ARG followed by the labels of the CVs on which the metadynamics potential will act. The keyword PACE determines the stride of Gaussian deposition in number of time steps, while the keyword HEIGHT specifies the height of the Gaussian in kJoule/mol. For each CVs, one has to specify the width of the Gaussian by using the keyword SIGMA. Gaussian will be written to the file indicated by the keyword FILE.

Once the PLUMED input file is prepared, one has to run Gromacs with the option to activate PLUMED and read the input file:

```
gmx_mpi mdrun -plumed
```

During the metadynamics simulation, PLUMED will create two files, named COLVAR and HILLS. The COLVAR file contains all the information specified by the PRINT command, in this case the value of the CVs every 10 steps of simulation, along with the current value of the metadynamics bias potential. The HILLS file contains a list of the Gaussians deposited along the simulation. If we give a look at the header of this file, we can find relevant information about its content:

```
#! FIELDS time phi psi sigma_phi sigma_psi height biasf
#! SET multivariate false
#! SET min_phi -pi
#! SET max_phi pi
#! SET min_psi -pi
#! SET max_psi pi
```

The line starting with FIELDS tells us what is displayed in the various columns of the HILLS file: the time of the simulation, the value of phi and psi, the width of the Gaussian in phi and psi, the height of the Gaussian, and the biasfactor. This quantity is relevant only for well-tempered metadynamics simulation (see [Exercise 4. Setup and run a well-tempered metadynamics simulation, part I](#)) and it is equal to 1 in standard metadynamics simulations. We will use the HILLS file later to calculate free-energies from the metadynamics simulation and assess its convergence. For the moment, we can plot the behavior of the CVs during the simulation.

By inspecting Figure [belfast-6-metad-fig](#), we can see that the system is initialized in one of the two metastable states of alanine dipeptide. After a while ( $t=0.3$  ns), the system is pushed by the metadynamics bias potential to visit the other local minimum. As the simulation continues, the bias potential fills the underlying free-energy landscape, and the system is able to diffuse in the entire phase space.

If we use the PLUMED input file described above, the expense of a metadynamics simulation increases with the length of the simulation as one has to evaluate the values of a larger and larger number of Gaussians at every step. To avoid this issue you can store the bias on a grid. In order to use grids, we have to add some additional information to the line of the [METAD](#) directive, as follows.

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate metadynamics in phi and psi
# depositing a Gaussian every 500 time steps,
# with height equal to 1.2 kJoule/mol,
# and width 0.35 rad for both CVs.
# The bias potential will be stored on a grid
# with bin size equal to 0.1 rad for both CVs.
# The boundaries of the grid are -pi and pi, for both CVs.
#
METAD ...
LABEL=metad
ARG=phi,psi
PACE=500
HEIGHT=1.2
```

```

SIGMA=0.35,0.35
FILE=HILLS
GRID_MIN=-pi,-pi
GRID_MAX=pi,pi
GRID_SPACING=0.1,0.1
... METAD

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR

```

The bias potential will be stored on a grid, whose boundaries are specified by the keywords `GRID_MIN` and `GRID_MAX`. Notice that you should provide either the number of bins for every collective variable (`GRID_BIN`) or the desired grid spacing (`GRID_SPACING`). In case you provide both PLUMED will use the most conservative choice (highest number of bins) for each dimension. In case you do not provide any information about bin size (neither `GRID_BIN` nor `GRID_SPACING`) and if Gaussian width is fixed PLUMED will use 1/5 of the Gaussian width as grid spacing. This default choice should be reasonable for most applications.

### 11.6.5.3 Exercise 2. Restart a metadynamics simulation

If we try to run again a metadynamics simulation using the script above in a directory where a `COLVAR` and `HILLS` files are already present, PLUMED will create a backup copy of the old files, and run a new simulation. Instead, if we want to restart a previous simulation, we have to add the keyword `RESTART` to the PLUMED input file (`plumed.dat`), as follows.

```

# restart previous simulation
RESTART

# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate metadynamics in phi and psi
# depositing a Gaussian every 500 time steps,
# with height equal to 1.2 kJoule/mol,
# and width 0.35 rad for both CVs.
#
metad: METAD ARG=phi,psi PACE=500 HEIGHT=1.2 SIGMA=0.35,0.35 FILE=HILLS

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR

```

(see [RESTART](#), [TORSION](#), [METAD](#), and [PRINT](#)).

In this way, PLUMED will read the old Gaussians from the `HILLS` file and append the new information to both `COLVAR` and `HILLS` files.

### 11.6.5.4 Exercise 3. Calculate free-energies and monitor convergence

One can estimate the free energy as a function of the metadynamics CVs directly from the metadynamics bias potential. In order to do so, the utility [sum\\_hills](#) should be used to sum the Gaussians deposited during the simulation and stored in the `HILLS` file.

To calculate the two-dimensional free energy as a function of `phi` and `psi`, it is sufficient to use the following command line:

```
plumed sum_hills --hills HILLS
```

The command above generates a file called `fes.dat` in which the free-energy surface as function of  $\phi$  and  $\psi$  is calculated on a regular grid. One can modify the default name for the free energy file, as well as the boundaries and bin size of the grid, by using the following options of `sum_hills` :

```
--outfile - specify the outputfile for sumhills
--min - the lower bounds for the grid
--max - the upper bounds for the grid
--bin - the number of bins for the grid
--spacing - grid spacing, alternative to the number of bins
```

It is also possible to calculate one-dimensional free energies from the two-dimensional metadynamics simulation. For example, if one is interested in the free energy as a function of the  $\phi$  dihedral alone, the following command line should be used:

```
plumed sum_hills --hills HILLS --idw phi --kt 2.5
```

The result should look like this:

To assess the convergence of a metadynamics simulation, one can calculate the estimate of the free energy as a function of simulation time. At convergence, the reconstructed profiles should be similar, apart from a constant offset. The option `--stride` should be used to give an estimate of the free energy every  $N$  Gaussians deposited, and the option `--mintozero` can be used to align the profiles by setting the global minimum to zero. If we use the following command line:

```
plumed sum_hills --hills HILLS --idw phi --kt 2.5 --stride 500 --mintozero
```

one free energy is calculated every 500 Gaussians deposited, and the global minimum is set to zero in all profiles. The resulting plot should look like the following:

To assess the convergence of the simulation more quantitatively, we can calculate the free-energy difference between the two local minima in the one-dimensional free energy along  $\phi$  as a function of simulation time. We can use the bash script `analyze_FES.sh` to integrate the multiple free-energy profiles in the two basins defined by the following intervals in  $\phi$  space: basin A,  $-3 < \phi < -1$ , basin B,  $0.5 < \phi < 1.5$ .

```
./analyze_FES.sh NFES -3.0 -1.0 0.5 1.5 KBT
```

where `NFES` is the number of profiles (free-energy estimates at different times of the simulation) generated by the option `--stride` of `sum_hills`, and `KBT` is the temperature in energy units (in this case `KBT=2.5`).

This analysis, along with the observation of the diffusive behavior in the CVs space, suggest that the simulation is converged.

#### 11.6.5.5 Exercise 4. Setup and run a well-tempered metadynamics simulation, part I

In this exercise, we will run a well-tempered metadynamics simulation on alanine dipeptide in vacuum, using as CVs the two backbone dihedral angles phi and psi. To activate well-tempered metadynamics, we need to add two keywords to the line of [METAD](#), which specifies the biasfactor and temperature of the simulation. For the first example, we will try a biasfactor equal to 6. Here how the PLUMED input file (plumed.dat) should look like:

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate metadynamics in phi and psi
# depositing a Gaussian every 500 time steps,
# with height equal to 1.2 kJoule/mol,
# and width 0.35 rad for both CVs.
# Well-tempered metadynamics is activated,
# and the biasfactor is set to 6.0
#
metad: METAD ARG=phi,psi PACE=500 HEIGHT=1.2 SIGMA=0.35,0.35 FILE=HILLS BIASFACTOR=6.0 TEMP=300.0

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR
```

(see [TORSION](#), [METAD](#), and [PRINT](#)).

After running the simulation using the instruction described above, we can have a look at the HILLS file. At variance with standard metadynamics, the last two columns of the HILLS file report more useful information. The last column contains the value of the biasfactor used, while the last but one the height of the Gaussian, which is rescaled during the simulation following the well-tempered recipe.

```
#! FIELDS time phi psi sigma_phi sigma_psi height biasf
#! SET multivariate false
#! SET min_phi -pi
#! SET max_phi pi
#! SET min_psi -pi
#! SET max_psi pi
1.0000 -1.3100 0.0525 0.35 0.35 1.4400 6
2.0000 -1.4054 1.9742 0.35 0.35 1.4400 6
3.0000 -1.9997 2.5177 0.35 0.35 1.4302 6
4.0000 -2.2256 2.1929 0.35 0.35 1.3622 6
```

If we carefully look at the height column, we will notice that in the beginning the value reported is higher than the initial height specified in the input file, which should be 1.2 kJoule/mol. In fact, this column reports the height of the Gaussian rescaled by the pre-factor that in well-tempered metadynamics relates the bias potential to the free energy. In this way, when we will use [sum\\_hills](#), the sum of the Gaussians deposited will directly provide the free-energy, without further rescaling needed.

We can plot the time evolution of the CVs along with the height of the Gaussian.

The system is initialized in one of the local minimum where it starts accumulating bias. As the simulation progresses and the bias added grows, the Gaussian height is progressively reduced. After a while ( $t=0.8$  ns), the system is able to escape the local minimum and explore a new region of the phase space. As soon as this happens, the Gaussian height is restored to the initial value and starts to decrease again. In the long time, the Gaussian height becomes smaller and smaller while the system diffuses in the entire CVs space.

We can now try a different biasfactor and see the effect on the simulation. If we choose a biasfactor equal to 1.5, we can notice a faster decrease of the Gaussian height with simulation time, as expected by the well-tempered recipe. We will also conclude from the plot below that this biasfactor is not large enough to allow for the system to escape from the initial local minimum in the time scale of this simulation.

Following the procedure described for standard metadynamics in the previous example, we can estimate the free energy as a function of time and monitor the convergence of the simulations using the `analyze_FES.sh` script. We will do this for the simulation in which the biasfactor was set to 6.0. In this case we will notice that the oscillations observed in standard metadynamics are here damped, and the bias potential converges more smoothly to the underlying free-energy landscape, provided that the biasfactor is sufficiently high for the free-energy barriers of the system under study to be crossed.

## 11.6.5.6 Exercise 5. Setup and run a well-tempered metadynamics simulation, part II

In this exercise, we will study the effect of neglecting a relevant degree of freedom in the choice of metadynamics CVs. We are going to run a well-tempered metadynamics simulation with the psi dihedral alone as CV, using the following PLUMED input file (plumed.dat):

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate metadynamics in psi
# depositing a Gaussian every 500 time steps,
# with height equal to 1.2 kJoule/mol,
# and width 0.35 rad.
# Well-tempered metadynamics is activated,
# and the biasfactor is set to 10.0
#
metad: METAD ARG=psi PACE=500 HEIGHT=1.2 SIGMA=0.35 FILE=HILLS BIASFACTOR=10.0 TEMP=300.0

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR
```

(see [TORSION](#), [METAD](#), and [PRINT](#)).

Let's look at the HILLS file, in particular at the time series of the CV psi and of the Gaussian height.

From this plot, we observe a nice diffusive behavior of the CV psi when the Gaussian height is already quite small. This happens until  $t=3$  ns, when the CV seems to be stuck for a while in a small region of the CV space. This behavior is typical of a situation in which a slow variable is not included in the set of CVs. When something happens in this hidden degree of freedom, the biased CVs typically cannot access anymore regions of the phase space previously visited. To understand this behavior, we need to visualize the time evolution of both phi and psi stored in the COLVAR file.

It is clear from the plot above that what happened around  $t=3$  ns is a jump of the neglected, slow degree of freedom phi from one free-energy basin to another. The dynamics of phi is not biased by any potential, so we need to equilibrate this degree of freedom, i.e. to observe multiple transitions from the two basins, before declaring convergence of our simulation. Or alternatively we can add phi to the set of CVs. This example demonstrates how to declare convergence of a well-tempered metadynamics simulation it is necessary but not sufficient to observe: 1) Gaussians with very small height, 2) a diffusive behavior in the CV space (as in the first 3 ns of this example).

What we should do is repeating the simulation multiple times starting from different initial conformations. If in all simulations, we observe a diffusive behavior in the biased CV when the Gaussian height is very small, and we obtain very similar free-energy surfaces, then we can be quite confident that our simulations are converged to the right value. If this is not the case, a manual inspection of the runs can help us identifying the missing slow degrees of freedom to add to the set of biased CVs.

## 11.7 Belfast tutorial: Replica exchange I

### 11.7.1 Aims

The aim of this tutorial is to introduce the users to running a parallel tempering (PT) simulation using PLUMED. We will set up a simple simulation of alanine dipeptide in vacuum, analyze the output, calculate free energies from the simulation, and detect problems. We will also learn how to run a combined PT-metadynamics simulation (PTMetaD) and introduce the users to the Well-Tempered Ensemble (WTE).

### 11.7.2 Summary of theory

In Replica Exchange Methods [62] (REM), sampling is accelerated by modifying the original Hamiltonian of the system. This goal is achieved by simulating  $N$  non-interacting replicas of the system, each evolving in parallel according to a different Hamiltonian. At fixed intervals, an exchange of configurations between two replicas is attempted. One popular case of REM is PT, in which replicas are simulated using the same potential energy function, but different temperatures. By accessing high temperatures, replicas are prevented from being trapped in local minima. In PT, exchanges are usually attempted between adjacent temperatures with the following acceptance probability:

$$p(i \rightarrow j) = \min\{1, e^{\Delta_{i,j}^{PT}}\},$$

with

$$\Delta_{i,j}^{PT} = \left( \frac{1}{k_B T_i} - \frac{1}{k_B T_j} \right) (U(R_i) - U(R_j)),$$

where  $R_i$  and  $R_j$  are the configurations at temperature  $T_i$  and  $T_j$ , respectively. The equation above suggests that the acceptance probability is ultimately determined by the overlap between the energy distributions of two replicas. The efficiency of the algorithm depends on the benefits provided by sampling at high-temperature. Therefore, an efficient diffusion in temperature space is required and configurational sampling is still limited by entropic barriers. Finally, PT scales poorly with system size. In fact, a sufficient overlap between the potential energy distributions of neighboring temperatures is required in order to obtain a significant diffusion in temperature. Therefore, the number of temperatures needed to cover a given temperature range scales as the square root of the number of degrees of freedom, making this approach prohibitively expensive for large systems.

PT can be easily combined with metadynamics [63]. In the resulting PTMetaD algorithm (16),  $N$  replicas performed in parallel a metadynamics simulation at different temperatures, using the same set of CVs. The PT acceptance probability must be modified in order to account for the presence of a bias potential:

$$\Delta_{i,j}^{PTMetaD} = \Delta_{i,j}^{PT} + \frac{1}{k_B T_i} [V_G^i(s(R_i), t) - V_G^i(s(R_j), t)] + \frac{1}{k_B T_j} [V_G^j(s(R_j), t) - V_G^j(s(R_i), t)],$$

where  $V_G^i$  and  $V_G^j$  are the bias potentials acting on the  $i$ -th and  $j$ -th replicas, respectively.

PTMetaD is particularly effective because it compensates for some of the weaknesses of each method alone. The effect of neglecting a slow degree of freedom in the choice of the metadynamics CVs is alleviated by PT, which allows the system to cross moderately high free-energy barriers on all degrees of freedom. On the other hand, the metadynamics bias potential allows crossing higher barriers on a few selected CVs, in such a way that the sampling efficiency of PTMetaD is greater than that of PT alone.

PTMetaD still suffers from the poor scaling of computational resources with system size. This issue may be circumvented by including the potential energy of the system among the set of well-tempered metadynamics CVs. The well-tempered metadynamics bias leads to the sampling of a well-defined distribution called Well-Tempered Ensemble (WTE) [13]. In this ensemble, the average energy remains close to the canonical value but its fluctuations are enhanced in a tunable way, thus improving sampling.

In the so-called PTMetaD-WTE scheme [64], each replica diffusion in temperature space is enhanced by the increased energy fluctuations at all temperatures.



### 11.7.3 Learning Outcomes

Once this tutorial is completed students will know how to:

- run a PT simulation
- analyze the output of the PT simulation and detect problems
- run a PTMetaD simulation
- run a PT and PTMetaD in the WTE

### 11.7.4 Resources

The `tarball` for this project contains the following directories:

- Exercise\_1: run a PT simulation using 2 replicas and analyze the output
  - Exercise\_2: run a PT simulation using 4 replicas and analyze the output
- Exercise\_3: run a PTMetaD simulation
- Exercise\_4: run a PT, PT-WTE and PTMetaD-WTE simulations

Each directory contains a TOPO subdirectory where topology and configuration files for Gromacs are stored.

### 11.7.5 Instructions

#### 11.7.5.1 The model system

Here we use as model systems alanine dipeptide in vacuum and water with AMBER99SB-ILDN all-atom force field.

#### 11.7.5.2 Exercise 1. Setup and run a PT simulation, part I

In this exercise, we will run a PT simulation of alanine dipeptide in vacuum, using only two replicas, one at 300K, the other at 305K. During this simulation, we will monitor the time evolution of the two dihedral angles phi and psi. In order to do that, we need the following PLUMED input file (plumed.dat).

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

# monitor the two variables
PRINT STRIDE=10 ARG=phi,psi FILE=COLVAR
```

(see `TORSION`, and `PRINT`).

To submit this simulation with Gromacs, we need the following command line.

```
mpirun -np 2 gmx_mpi mdrun -s TOPO/topol -plumed -multi 2 -replex 100
```

This command will execute two MPI processes in parallel, using the topology files `topol0.tpr` and `topol1.tpr` stored in the `TOPO` subdirectory. These two binary files have been created using the usual Gromacs procedure (see Gromacs manual for further details) and setting the temperature of the two simulations at 300K and 305K in the configuration files. An exchange between the configurations of the two simulations will be attempted every 100 steps.

When Gromacs is executed using the `-multi` option and PLUMED is activated, the output files produced by PLUMED will be renamed and a suffix indicating the replica id will be appended. We will start inspecting the output file `COLVAR.0`, which reports the time evolution of the CVs at 300K.

The plot above suggests that during the PT simulation the system is capable to access both the relevant basins in the free-energy surface. This seems to suggest that our simulation is converged. We can use the `COLVAR.0` and `COLVAR.1` along with the tool [sum\\_hills](#) to estimate the free energy as a function of the CV  $\phi$ . We will do this as a function of simulation time to assess convergence more quantitatively, using the following command line:

```
plumed sum_hills --histo COLVAR.0 --idw phi --sigma 0.2 --kt 2.5 --outhisto fes_ --stride 1000
```

As we did in our previous tutorial, we can now use the script `analyze_FES.sh` to calculate the free-energy difference between basin A ( $-3.0 < \phi < -1.0$ ) and basin B ( $0.5 < \phi < 1.5$ ), as a function of simulation time.

The estimate of the free-energy difference between these two basins seems to be converged. This consideration, along with the observation that the system is exploring all the relevant free-energy basins, might lead us to declare convergence and to state that the difference in free energy between basin A and B is roughly 0 kJoule/mol. Unfortunately, in doing so we would make a big mistake.

In PT simulations we have to be a little bit more careful, and examine the time evolution of each replica diffusing in temperature space, before concluding that our simulation is converged. In order to do that, we need to reconstruct the continuous trajectories of the replicas in temperature from the two files (typically `traj0.trr` and `traj1.trr`) which contain the discontinuous trajectories of the system at 300K and 305K. To demux the trajectories, we need to use the following command line:

```
demux.pl md0.log
```

which will create two files, called `replica_temp.xvg` and `replica_index.xvg`. We can plot the first file, which reports the temperature index of each configuration at a given time of the simulation. This file is extremely useful, because it allows us monitoring the replicas diffusion in temperature. As we discussed in [Summary of theory](#), in order for the PT algorithm to be effective, we need an efficient diffusion of the replicas in temperature space. In this case, both replicas are rapidly accessing the highest temperature, so there seems not to be any problem on this side.

We can use the second file to reconstruct the continuous trajectories of each replica in temperature:

```
gmx_mpi trjcat -f traj0.trr traj1.trr -demux replica_index.xvg
```

and the following PLUMED input file (`plumed_demux.dat`) to recalculate the value of the CVs on the demuxed trajectories, typically called `0_trajout.xtc` and `1_trajout.xtc`.

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

# monitor the two variables
PRINT STRIDE=1 ARG=phi,psi FILE=COLVAR_DEMUX
```

(see [TORSION](#), and [PRINT](#)).

For the analysis of the demuxed trajectories, we can use the `-rerun` option of Gromacs, as follows:

```
# rerun Gromacs on replica 0 trajectory
gmx_mpi mdrun -s TOPO/topol0.tpr -plumed plumed_demux.dat -rerun 0_trajout.xtc
# rename the output
mv COLVAR_DEMUX COLVAR_DEMUX.0
# rerun Gromacs on replica 1 trajectory
gmx_mpi mdrun -s TOPO/topol1.tpr -plumed plumed_demux.dat -rerun 1_trajout.xtc
# rename the output
mv COLVAR_DEMUX COLVAR_DEMUX.1
```

We can now plot the time evolution of the two CVs in the two demuxed trajectories.

This plot shows clearly that each replica is sampling only one of the two basins of the free-energy landscape, and it is never able to cross the high barrier that separates them. This means that what we considered an exhaustive exploration of the free energy landscape at 300K (Figure [belfast-7-pt-fig](#)), was instead caused by an efficient exchange of configurations between replicas that were trapped in different free-energy basins. The results of the present simulation were then influenced by the initial conformations of the two replicas. If we had initialized both of them in the same basin, we would have never observed "transitions" to the other basin at 300K.

To declare convergence of a PT simulation, it is thus mandatory to examine the behavior of the replicas diffusing in temperature and check that these are exploring all the relevant basins. Another good practice is repeating the PT simulation starting from different initial conformations, and check that the results are consistent.

### 11.7.5.3 Exercise 2. Setup and run a PT simulation, part II

We will now repeat the previous exercise, but with a different setup. The problem with the previous exercise was that replicas were never able to interconvert between the two metastable states of alanine dipeptide in vacuum. The reason was that the highest temperature used (305K) was too low to accelerate barrier crossing in the time scale of the simulation. We will now use 4 replicas at the following temperatures: 300K, 400K, 600K, and 1000K.

We can use the same PLUMED input file described above (plumed.dat), and execute Gromacs using the following command line:

```
mpirun -np 4 gmx_mpi mdrun -s TOPO/topol -plumed -multi 4 -replex 100
```

At the end of the simulation, we first monitor the diffusion in temperature space of each replica. We need to create the `replica_temp.xvg` and `replica_index.xvg`:

```
demux.pl md0.log
```

and plot the content of `replica_temp.xvg`. Here how it should look for Replica 0:

From this analysis, we can conclude that replicas are diffusing effectively in temperature. Now, we need to monitor the space sampled by each replica while diffusing in temperature space and verify that they are interconverting between the different basins of the free-energy landscape. The demux is carried out as in the previous example:

```
trjcat_mpi -f traj0.trr traj1.trr traj2.trr traj3.trr -demux replica_index.xvg
```

and so is the analysis of the demuxed trajectories using the `-rerun` option of Gromacs and the `plumed_demux.dat` input file. Here is the space sampled by two of the four replicas:

It is clear that in this case replicas are able to interconvert between the two metastable states, while efficiently diffusing in temperature. We can then calculate the free-energy difference between basin A and B as a function of simulation time at 300K:

and conclude that in this case the PT simulation is converged.

#### 11.7.5.4 Exercise 3. Setup and run a PTMetaD simulation

In this exercise we will learn how to combine PT with metadynamics. We will use the setup of the previous exercise, and run a PT simulations with 4 replicas at the following temperatures: 300K, 400K, 600K, and 1000K. Each simulation will perform a well-tempered metadynamics calculation, using the dihedral psi alone as CV and a biasfactor equal to 10 (see [Exercise 5. Setup and run a well-tempered metadynamics simulation, part II](#)).

Previously, we prepared a single PLUMED input file to run a PT simulation. This was enough, since in that case the same task was performed at all temperatures. Here instead we need to have a slightly different PLUMED input file for each simulation, since we need to use the keyword TEMP to specify the temperature on the line of the **METAD** directory. We will thus prepare 4 input files, called plumed.0.dat, plumed.1.dat, plumed.2.dat, and plumed.3.dat, with a different value for the keyword TEMP.

#### Warning

Notice that the rules for replica suffix are changed with version 2.2. With PLUMED versions 2.0 and 2.1 these files should have been named plumed.dat.0, plumed.dat.1, etc.

Here how plumed.3.dat should look like:

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate metadynamics in psi
# depositing a Gaussian every 500 time steps,
# with height equal to 1.2 kJoule/mol,
# and width 0.35 rad.
# Well-tempered metadynamics is activated,
# and the biasfactor is set to 10.0
#
metad: METAD ARG=psi PACE=500 HEIGHT=1.2 SIGMA=0.35 FILE=HILLS BIASFACTOR=10.0 TEMP=1000.0

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR
```

(see [TORSION](#), [METAD](#), and [PRINT](#)).

The PTMetaD simulation is executed in the same way as the PT:

```
mpirun -np 4 gmx_mpi mdrun -s TOPO/topol -plumed -multi 4 -replex 100
```

and it will produce one COLVAR and HILLS file per temperature (COLVAR.0, HILLS.0, ...). The analysis of the results requires what we have learned in the previous exercise for the PT case (analysis of the replica diffusion in temperature and demuxing of each replica trajectory), and the post-processing of a well-tempered metadynamics simulation (FES calculation using [sum\\_hills](#) and convergence analysis).

Since in the previous tutorial we performed the same well-tempered metadynamics simulation without the use of PT (see [Exercise 5. Setup and run a well-tempered metadynamics simulation, part II](#)), here we can focus on the differences with the PTMetaD simulation. Let's compare the behavior of the biased variable psi in the two simulations:

In well-tempered metadynamics (left panel), the biased variable psi looked stuck early in the simulation (t=3 ns). The reason was the transition of the other hidden degree of freedom phi from one free-energy basin to the other. In the PTMetaD case (right panel), this seems not to happen. To better appreciate the difference, we can plot the time evolution of the hidden degree of freedom phi in the two cases.

Thanks to the excursions at high temperature, in the PTMetaD simulation the transition of the CV phi between the two basins is accelerate. As a result, the convergence of the reconstructed free energy in psi will be accelerated. This simple exercise demonstrates how PTMetaD can be used to cure a bad choice of metadynamics CVs and the neglecting of slow degrees of freedom.

## 11.7.5.5 Exercise 4. The Well-Tempered Ensemble

In this exercise we will learn how to run a PT-WTE and PTMetaD-WTE simulations of alanine dipeptide in water. We will start by running a short PT simulation using 4 replicas in the temperature range between 300K and 400K. We will use a geometric distribution of temperatures, which is valid under the assumption that the specific heat of the system is constant across temperatures. Replicas will thus be simulated at  $T=300, 330.2, 363.4$ , and 400K. In this simulation, we will just monitor the two dihedral angles and the total energy of the system, by preparing the following PLUMED input file (plumed\_PT.dat):

```
# set up three variables for Phi and Psi dihedral angles
# and total energy
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
ene: ENERGY

# monitor the three variables
PRINT STRIDE=10 ARG=phi,psi,ene FILE=COLVAR_PT
```

(see [TORSION](#), [ENERGY](#), and [PRINT](#)).

As usual, the simulation is run for 400ps using the following command:

```
mpirun -np 4 gmx_mpi mdrun -s TOPO/topol -plumed plumed_PT.dat -multi 4 -replex 100
```

At the end of the run, we want to analyze the acceptance rate between exchanges. This quantity is reported at the end of the Gromacs output file, typically called md.log, and it can be extracted using the following bash command line:

```
grep -A2 "Repl average probabilities" md0.log
```

From the line above, we will find out that none of the attempted exchanges has been accepted. The reason is that the current setup (4 replicas to cover the temperature range 300-400K) resulted in a poor overlap of the energy distributions at different temperatures. We can easily realize this by plotting the time series of the total energy in the different replicas:

To improve the overlap of the potential energy distributions at different temperatures, we enlarge the fluctuations of the energy by sampling the Well-Tempered Ensemble (WTE). In order to do that, we need to setup a well-tempered metadynamics simulation using energy as CV. In WTE, fluctuations - the standard deviation of the energy time series measured above - will be enhanced by a factor equal to the square root of the biasfactor. In this exercise, we will enhance fluctuations of a factor of 4, thus we will set the biasfactor equal to 16. We need to prepare 4 PLUMED input files (plumed\_PTWTE.dat.0, plumed\_PTWTE.dat.1,...), which will be identical to the following but for the temperature specified in the line of the [METAD](#) directive:

```
# set up three variables for Phi and Psi dihedral angles
# and total energy
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
ene: ENERGY

# Activate metadynamics in ene
# depositing a Gaussian every 250 time steps,
# with height equal to 1.2 kJoule/mol,
# and width 140 kJoule/mol.
# Well-tempered metadynamics is activated,
# and the biasfactor is set to 16.0
#
wte: METAD ARG=ene PACE=250 HEIGHT=1.2 SIGMA=140.0 FILE=HILLS_PTWTE BIASFACTOR=16.0 TEMP=300.0

# monitor the three variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,ene,wte.bias FILE=COLVAR_PTWTE
```

(see [TORSION](#), [ENERGY](#), [METAD](#), and [PRINT](#)).

Here, we use a Gaussian width larger than usual, and of the order of the fluctuations of the potential energy at 300K, as calculated from the preliminary PT run.

We run the simulation following the usual procedure:

```
mpirun -np 4 gmx_mpi mdrun -s TOPO/topol -plumed plumed_PTWTE.dat -multi 4 -replex 100
```

If we analyze the average acceptance probability in this run:

```
grep -A2 "Repl average probabilities" md0.log
```

we will notice that now on average 18% of the exchanges are accepted. To monitor the diffusion of each replica in temperature, we can examine the file `replica_temp.xvg` created by the following command line:

```
demux.pl md0.log
```

This analysis assures us that the system is efficiently diffusing in the entire temperature range and no bottlenecks are present.

Finally, as done in the previous run, we can visualize the time serie of the energy CV at all temperatures:

If we compare this plot with the one obtained in the PT run, we can notice that now the enlarged fluctuations caused by the use of WTE lead to a good overlap between energy distributions at different temperatures, thus increasing the exchange acceptance probability. At this point, we can extend our PT-WTE simulation and for example converge the free energy as a function of the dihedral angles phi and psi. Alternatively, we can accelerate sampling of the phi and psi dihedrals by combining PT-WTE with a metadynamics simulation using phi and psi as CVs (PTMetaD-WTE [64]). This can be achieved by preparing 4 PLUMED input files (`plumed_PTMetaDWTE.dat.0`, `plumed_PTMetaDWTE.dat.1`,...), which will be identical to the following but for the temperature specified in the two lines containing the [METAD](#) directives:

```
# reload WTE bias
RESTART

# set up three variables for Phi and Psi dihedral angles
# and total energy
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
ene: ENERGY

# Activate metadynamics in ene
# Old Gaussians will be reloaded to perform
# the second metadynamics run in WTE.
#
wte: METAD ARG=ene PACE=99999999 HEIGHT=1.2 SIGMA=140.0 FILE=HILLS_PTWTE BIASFACTOR=16.0 TEMP=300.0

# Activate metadynamics in phi and psi
# depositing a Gaussian every 500 time steps,
# with height equal to 1.2 kJoule/mol,
# and width 0.35 rad for both CVs.
# Well-tempered metadynamics is activated,
# and the biasfactor is set to 6.0
#
metad: METAD ARG=phi,psi PACE=500 HEIGHT=1.2 SIGMA=0.35,0.35 FILE=HILLS_PTMetaDWTE BIASFACTOR=6.0 TEMP=300.0

# monitor the three variables, the wte and metadynamics bias potentials
PRINT STRIDE=10 ARG=phi,psi,ene,wte.bias,metad.bias FILE=COLVAR_PTMetaDWTE
```

(see [TORSION](#), [ENERGY](#), [METAD](#), and [PRINT](#)).

These scripts activate two metadynamics simulations. One will use the energy as CV and will reload the Gaussians deposited during the preliminary PT-WTE run. No additional Gaussians on this variable will be deposited during the PTMetaD-WTE simulation, due to the large deposition stride. A second metadynamics simulation will be activated on the dihedral angles. Please note the different parameters and biasfactors in the two metadynamics runs.

The simulation is carried out using the usual procedure:

```
mpirun -np 4 gmx_mpi mdrun -s TOPO/topol -plumed plumed_PTMetaDWTE.dat -multi 4 -replex 100
```

## 11.8 Belfast tutorial: Replica exchange II and Multiple walkers

### 11.8.1 Aims

The aim of this tutorial is to introduce the users to the use of Bias-Exchange Metadynamics. We will go through the writing of the input files for BEMETA for a simple case of three peptide and we will use METAGUI to analyse them. We will compare the results of WT-BEMETA and STANDARD-BEMETA with four independent runs on the four Collective Variables. Finally we will use a simplified version of BEMETA that is Multiple Walkers Metadynamics.

#### 11.8.1.1 Learning Outcomes

Once this tutorial is completed students will:

- Know how to run a Bias-Exchange simulation using PLUMED and GROMACS
- Know how to analyse the results of BEMETA with the help of METAGUI
- Know how to run a Multiple Walker simulation

### 11.8.2 Resources

The `tarball` for this project contains the following files:

- system folder: a starting structure for Val-Ile-Leu system
- WTBX: a run of Well-Tempered Bias-Exchange Metadynamics ready for the analysis

### 11.8.3 Instructions

#### 11.8.3.1 Bias-Exchange Metadynamics

In all variants of metadynamics the free-energy landscape of the system is reconstructed by gradually filling the local minima with gaussian hills. The dimensionality of the landscape is equal to the number of CVs which are biased, and typically a number of CVs smaller than three is employed. The reason for this is that qualitatively, if the CVs are not correlated among them, the simulation time required to fill the free-energy landscape grows exponentially with the number of CVs. This limitation can be severe when studying complex transformations or reactions in which more than say three relevant CVs can be identified.

A possible technique to overcome this limitation is parallel-tempering metadynamics, [Belfast tutorial: Replica exchange I](#). A different solution is performing a bias-exchange simulation: in this approach a relatively large number  $N$  of CVs is chosen to describe the possible transformations of the system (e.g., to study the conformations of a peptide one may consider all the dihedral angles between amino acids). Then,  $N$  metadynamics simulations (replicas) are run on the same system at the same temperature, biasing a different CV in each replica.

Normally, in these conditions, each bias profile would converge very slowly to the equilibrium free-energy, due to hysteresis. Instead, in the bias-exchange approach every fixed number of steps (say 10,000) an exchange is attempted between a randomly selected pair of replicas  $a$  and  $b$ . The probability to accept the exchange is given by a Metropolis rule:

$$\min (1, \exp [\beta (V_G^a(x^a, t) + V_G^b(x^b, t) - V_G^a(x^b, t) - V_G^b(x^a, t))])$$

where  $x^a$  and  $x^b$  are the coordinates of replicas  $a$  and  $b$  and  $V_G^{a(b)}(x, t)$  is the metadynamics potential acting on the replica  $a$  ( $b$ ). Each trajectory evolves through the high dimensional free energy landscape in the space of the CVs sequentially biased by different metadynamics potentials acting on one CV at each time. The results of the simulation are  $N$  one-dimensional projections of the free energy.

In the following example, a bias-exchange simulation is performed on a VIL peptide (zwitterionic form, in vacuum with  $\epsilon = 80$ , force field amber03), using the four backbone dihedral angles as CVs.

Four replicas of the system are employed, each one biased on a different CV, thus four similar Plumed input files are prepared as follows:

- a common input file in which all the collective variables are defined:

```
MOLINFO STRUCTURE=VIL.pdb
RANDOM_EXCHANGES

cv1: TORSION ATOMS=@psi-1
cv2: TORSION ATOMS=@phi-2
cv3: TORSION ATOMS=@psi-2
cv4: TORSION ATOMS=@phi-3
```

NOTE:

1. By using [MOLINFO](#) we can use shortcut to select atoms for dihedral angles (currently @phi, @psi, @omega and @chi1 are available).
  2. We use cv# as labels in order to make the output compatible with METAGUI.
  3. [RANDOM\\_EXCHANGES](#) generates random exchanges list that are sent back to GROMACS.
- four additional input files that [INCLUDE](#) the common input and define the four [METAD](#) along the four CVs, respectively.

```
INCLUDE FILE=plumed-common.dat
be: METAD ARG=cv1 HEIGHT=0.2 SIGMA=0.2 PACE=100 GRID_MIN=-pi GRID_MAX=pi
PRINT ARG=cv1,cv2,cv3,cv4 STRIDE=1000 FILE=COLVAR
```

NOTE:

1. in COLVAR we [PRINT](#) only the four collective variables, always in the same order in such a way that COLVARs are compatible with METAGUI
2. if you want to print additional information, like the [METAD](#) bias it is possible to use additional [PRINT](#) keyword

```
PRINT ARG=cv1,be.bias STRIDE=xxx FILE=BIAS
```

The four replicas start from the same GROMACS topology file replicated four times: topol0.tpr, topol1.tpr, topol2.tpr, topol3.tpr. Finally, GROMACS is launched as a parallel run on 4 cores, with one replica per core, with the command

```
mpirun -np 4 gmx_mpi mdrun -s topol -plumed plumed -multi 4 -replex 2000 >& log &
```

where -replex 2000 indicates that every 2000 molecular-dynamics steps all replicas are randomly paired (e.g. 0-2 and 1-3) and exchanges are attempted between each pair (as printed in the GROMACS \*.log files).

The same simulation can be run using WELLTEMPERED metadynamics.



### 11.8.3.2 Convergence of the Simulations

In the resources for this tutorial you can find the results for a 40ns long Well-Tempered Bias Exchange simulation. First of all we can try to assess the convergence of the simulations by looking at the profiles. In the "convergence" folder there is a script that calculates the free energy from the HILLS.0 file at increasing simulation lengths (i.e. every more 0.8 ns of simulation). The scripts also generate two measures of the evolution of the profiles in time:

1. time-diff.HILLS.0: where it is stored the average deviation between two successive profiles
2. KL.HILLS.0: where it is stored the average deviation between profiles correctly weighted for the free energy of the profiles themselves (Symmetrized Kullback-Lieber divergence)

From both plots one can deduce that after 8 ns the profiles don't change significantly thus suggesting that averaging over the range 8-40ns should result in an accurate profile (we will test this using metagui). Another test is that of looking at the fluctuations of the profiles in a time window instead of looking at successive profiles:

### 11.8.3.3 Bias-Exchange Analysis with METAGUI

In principle Bias-Exchange Metadynamics can give as a results only N 1D free energy profiles. But the information contained in all the replicas can be used to recover multidimensional free energy surfaces in  $\geq N$  dimensions. A simple way to perform this analysis is to use METAGUI. METAGUI performs the following operations:

1. Clusterizes the trajectories on a multidimensional GRID defined by at least the biased coordinates.
2. By using the 1D free energy profiles and the clustering assigns a free energy to the cluster using a WHAM procedure.
3. Lets the user visualize the clusters.
4. Approximates the kinetics among clusters.

METAGUI (Biarnes et. al) is a plugin for VMD that implements the approach developed by Marinelli et. al 2009. It can be downloaded from the PLUMED website.

In order for the colvar and hills file to be compatible with METAGUI their header must be formatted as following:

COLVAR. #:

```
#! FIELDS time cv1 cv2 cv3 cv4
#! ACTIVE 1 1 A
#! ..
...
```

NOTE:

1. the COLVAR. # files should contain ALL the collective variables employed (all those biased in at least one replica plus those additionally analysed). They MUST be named cv1 ... cvN.
2. the COLVAR. # files must be synchronised with the trajectories, this means that for each frame in the trajectory at time t there must be a line in each colvar at time t and viceversa. The best option is usually to analyse the trajectories a posteriori using plumed driver.

3. a keyword `#! ACTIVE NBIASEDCV BIASEDCV LABEL` is needed, where `NBIASEDCV` is the number of biased cv in that replica (not overall), `BIASEDCV` is the index of the biased cv in that replica (i.e. 1 for the first replica and so on); `LABEL` is a letter that identify the replica (usually is simply A for the first, B for the second and so on) this is useful if two replicas are biasing the same collective variable:

```
COLVAR.0:
#! FIELDS time cv1 cv2 cv3
#! ACTIVE 1 1 A
#! ..
...
COLVAR.1:
#! FIELDS time cv1 cv2 cv3
#! ACTIVE 1 2 B
#! ..
...
COLVAR.2:
#! FIELDS time cv1 cv2 cv3
#! ACTIVE 1 2 C
#! ..
...
COLVAR.3:
#! FIELDS time cv1 cv2 cv3
#! ACTIVE 0
#! ..
...
```

In the above case Replica 0 biases cv1; replicas 1 and 2 biases cv2 while replica 3 is a neutral (unbiased) replica. cv3 is unbiased in all the replicas.

The `ACTIVE` keyword must be the FIRST LINE in the `HILLS.#` files:

`HILLS.#:`

```
#! ACTIVE 1 1 A
#! FIELDS time cv1 sigma_cv1 height biasf
#! ..
...
```

The above notes hold for the `HILLS` files as well. In the folder `metagui` the script `check_for_metagui.sh` checks if the header of your file is compatible with `METAGUI`, but remember that this is not enough! Synchronisation of `COLVAR` and trajectory files is also needed. `HILLS` files can be written with a different frequency but times must be consistent.

**NOTE:** It is important to copy `HILLS` files in the `metagui` folder.

```
./check_for_metagui.sh ../COLVAR.0
```

will tell you that the `ACTIVE` keyword is missing, you need to modify all the header BEFORE proceeding with the tutorial!!

In the `metagui` folder there is a `metagui.input` file:

```
WHAM_EXE          wham_bemeta.x
BASINS_EXE         kinetic_basins.x
KT 2.4900
HILLS_FILE        HILLS.0
HILLS_FILE        HILLS.1
HILLS_FILE        HILLS.2
HILLS_FILE        HILLS.3
GRO_FILE          VIL.pdb
COLVAR_FILE COLVAR.0 ../traj0.xtc "psi-1"
```

```
COLVAR_FILE COLVAR.1 ../traj1.xtc "phi-2"
COLVAR_FILE COLVAR.2 ../traj2.xtc "psi-2"
COLVAR_FILE COLVAR.3 ../traj3.xtc "phi-3"
TRAJ_SKIP 10
CVGRID 1 -3.1415 3.1415 15 PERIODIC
CVGRID 2 -3.1415 3.1415 15 PERIODIC
CVGRID 3 -3.1415 3.1415 15 PERIODIC
CVGRID 4 -3.1415 3.1415 15 PERIODIC
ACTIVE 4 1 2 3 4
T_CLUSTER 0.
T_FILL 8000.
DELTA 4
GCORR 1
TR_N_EXP 5
```

where are defined the temperature in energy units, the place where to find COLVAR, HILLS and trajectory files. A reference gro or pdb file is needed to load the trajectories. The definition of the ranges and the number of bins for the available collective variables.

Now let's start with the analysis:

1. run VMD and load metagui
2. in metagui load the metagui.input file [belfast-8-mg1-fig](#)
3. In the left section of the interface "load all" the trajectories
4. Find the Microstates

In order to visualise the microstate it is convenient to align all the structures using the VMD RMSD Trajectory tool that can be found in Extensions->Analysis.

One or more microstates can be visualised by selecting them and clicking show.

You can sort the microstates using the column name tabs, for example by clicking on size the microstates will be ordered from the larger to the smaller. If you look at the largest one it is possible to observe that by using the four selected collective variables the backbone conformation of the peptide is well defined while the sidechains can populate different rotameric states.

The equilibrium time in the analysis panel should be such that by averaging over the two halves of the remind of the simulation the profiles are the same (i.e the profile averaged between  $T_{eq}$  and  $T_{eq}+(T_{tot}-T_{eq})/2$  should be the same of that averaged from  $T_{eq}+(T_{tot}-T_{eq})/2$  and  $T_{tot}$ ). By clicking on COMPUTE FREE ENERGIES, the program will first generate the 1D free energy profiles from the HILLS files and then run the WHAM analysis on the microstates. Once the analysis is done it is possible to visually check the convergence of the 1D profiles one by one by clicking on the K buttons next to the HILLS.# files. The BLUE and the RED profiles are the two profiles just defined, while the GREEN is the average of the two. Now it is possible for example to sort the microstates as a function of the free energy and save them by dumping the structures for further analysis.

If you look in the metagui folder you will see a lot of files, some of them can be very useful:

metagui/MICROSTATES: is the content of the microstates list table metagui/WHAM\_RUN/VG\_HILLS.#: are the opposite of the free energies calculated from the hills files metagui/WHAM\_RUN/\*.gnp: are gnuplot input files to plot the VG\_HILLS.# files (i.e. gnuplot -> load "convergence..") metagui/WHAM\_RUN/FES: is the result of the WHAM, for each cluster there is its free energy and the error estimate from WHAM

```
gnuplot> plot [0:40]'FES' u 2:3
```

plots the microstate error in the free energy estimate as a function of the microstates free energy. Finally in the folder metagui/FES there is script to integrate the multidimensional free energy contained in the MICROSTATES files to a 2D FES as a function of two of the used CV. To use it is enough to copy the MICROSTATES file in FES:

```
cp MICROSTATES FES/FES.4D
```

and edit the script to select the two columns of MICROSTATES on which show the integrated FES.

### 11.8.3.4 Multiple Walker Metadynamics

Multiple Walker metadynamics is the simplest way to parallelise a MetaD calculation: multiple simulation of the same system are run in parallel using metadynamics on the same set of collective variables. The deposited bias is shared among the replicas in such a way that the history dependent potential depends on the whole history.

We can use the same common input file defined above and then we can define four metadynamics bias in a similar way of what was done above for bias-exchange but now all the biases are defined on the same collective variables:

```
plumed.#.dat
INCLUDE FILE=plumed-common.dat

METAD ...
LABEL=mw
ARG=cv2,cv3
SIGMA=0.3,0.3
HEIGHT=0.2
PACE=100
BIASFACTOR=8
GRID_MIN=-pi,-pi
GRID_MAX=pi,pi
WALKERS_MPI
... METAD

PRINT ARG=cv1,cv2,cv3,cv4 STRIDE=1000 FILE=COLVAR
```

and the simulation can be run in a similar way without doing exchanges:

```
mpirun -np 4 gmx_mpi mdrun -s topol -plumed plumed -multi 4 >& log &
```

alternatively Multiple Walkers can be run as independent simulations sharing via the file system the biasing potential, this is useful because it provides a parallelisation that does not need a parallel code. In this case the walkers read with a given frequency the gaussians deposited by the others and add them to their own [METAD](#).

## 11.8.4 Reference

This tutorial is freely inspired to the work of Biarnes et al.

More materials can be found in

1. Marinelli, F., Pietrucci, F., Laio, A. & Piana, S. A kinetic model of trp-cage folding from multiple biased molecular dynamics simulations. *PLoS Comput. Biol.* 5, e1000452 (2009).
2. Biarnés, X., Pietrucci, F., Marinelli, F. & Laio, A. METAGUI. A VMD interface for analyzing metadynamics and molecular dynamics simulations. *Comput. Phys. Commun.* 183, 203–211 (2012).
3. Baftizadeh, F., Cossio, P., Pietrucci, F. & Laio, A. Protein folding and ligand-enzyme binding from bias-exchange metadynamics simulations. *Current Physical Chemistry* 2, 79–91 (2012).
4. Granata, D., Camilloni, C., Vendruscolo, M. & Laio, A. Characterization of the free-energy landscapes of proteins by NMR-guided metadynamics. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6817–6822 (2013).
5. Raiteri, P., Laio, A., Gervasio, F. L., Micheletti, C. & Parrinello, M. Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. *J. Phys. Chem. B* 110, 3533–3539 (2006).

## 11.9 Belfast tutorial: NMR restraints

### 11.9.1 Aims

This tutorial is about the use of experimental data, in particular NMR data, either as collective variables or as replica-averaged restraints in MD simulations. While the first is a just a simple extension of what we have been already doing in previous tutorials, the latter is an approach that can be used to increase the quality of a force-field in describing the properties of a specific system.

#### 11.9.1.1 Learning Outcomes

Once this tutorial is completed students will:

- know why and how to use experimental data to define a collective variable
- know why and how to use experimental data as replica-averaged restraints in MD simulations

### 11.9.2 Resources

The `tarball` for this project contains the following:

- system: the files use to generate the `topol?.tpr` files of the first and second example (the setup is for simulations in vacuum)
- first: an example on the use of chemical shifts as a collective variable
- second: an example on the use of chemical shifts as replica-averaged restraints
- third: an example on the use of RDCs (calculated with the theta-method) as replica-averaged restraints

### 11.9.3 Instructions

#### 11.9.3.1 Experimental data as Collective Variables

In the former tutorials it has been often discussed the possibility of measuring a distance with respect to a structure representing some kind of state for a system, i.e. [Belfast tutorial: Out of equilibrium dynamics](#). An alternative possibility is to use as a reference a set of experimental data that represent a state and measure the current deviation from the set. In plumed there are currently implemented the following NMR experimental observables: Chemical Shifts (only for proteins) [CS2BACKBONE](#), [NOE](#) distances, [JCOUPLING](#), [PRE](#) intensities, and Residual Dipolar couplings/pseudocontact shifts [RDC](#).

In the following we will write the [CS2BACKBONE](#) collective variable similar to the one used in Granata et al. (2013) (while the collective variable is still proportional to the square sum of the deviation of the calculated and experimental chemical shifts divided by a typical error, the exact definition is not the same. The sum is not done anymore with a flat bottom difference and the error used are not the one published, so the exact result of the scoring function can be different).

As a general rule, when using [CS2BACKBONE](#) or other experimental restraints it is better to increase the accuracy of the constraint algorithm due to the increased strain on the bonded structure. In the case of GROMACS it is safer to use `lincs-iter=2` and `lincs-order=6`.

```
prot: GROUP ATOMS=1-862
cs: CS2BACKBONE ATOMS=prot DATA=data NRES=56 CAMSHIFT
PRINT ARG=cs FILE=COLVAR STRIDE=100

ENDPLUMED
```

In this case the chemical shifts are those measured for the native state of the protein and can be used, together with other CVs and Bias-Exchange Metadynamics, to guide the system back and forth from the native structure. The experimental chemical shifts are in six files inside the "data/" folder (see first example in the resources tarball), one file for each nucleus. A 0 chemical shift is used where a chemical shift doesn't exist (i.e. CB of GLY) or where it has not been assigned. Additionally the data folder contains:

- camshift.db: this file is a parameter file for camshift, it is a standard file needed to calculate the chemical shifts from a structure
- template.pdb: this is a pdb file for the protein we are simulating (i.e. `editconf -f conf.gro -o template.pdb`) where atoms are ordered in the same way in which are included in the main code and again it is used to map the atom in plumed with those in almost.

This example can be executed as

```
gmx_mpi mdrun -s topol -plumed plumed
```

### 11.9.3.2 Replica-Averaged Restrained Simulations

NMR data, as all the equilibrium experimental data, are the result of a measure over an ensemble of structures and over time. In principle a "perfect" molecular dynamics simulations, that is a simulations with a perfect force-field and a perfect sampling can predict the outcome of an experiments in a quantitative way. Actually in most of the cases obtaining a qualitative agreement is already a fortunate outcome. In order to increase the accuracy of a force field in a system dependent manner it is possible to add to the force-field an additional term based on the agreement with a set of experimental data. This agreement is not enforced as a simple restraint because this would mean to ask the system to be always in agreement with all the experimental data at the same time, instead the restraint is applied over an AVERAGED COLLECTIVE VARIABLE where the average is performed over multiple independent simulations of the same system in the same conditions. In this way the is not a single replica that must be in agreement with the experimental data but they should be in agreement on average. It has been shown that this approach is equivalent to solving the problem of finding a modified version of the force field that will reproduce the provided set of experimental data without any additional assumption on the data themselves.

The second example included in the resources show how the amber force field can be improved in the case of protein domain GB3 using the native state chemical shifts a replica-averaged restraint. By the fact that replica-averaging needs the use of multiple replica simulated in parallel in the same conditions it is easily complemented with BIAS-EXCHANGE or MULTIPLE WALKER metadynamics to enhance the sampling.

```
prot: GROUP ATOMS=1-862
cs: CS2BACKBONE ATOMS=prot DATA=data NRES=56
enscs: ENSEMBLE ARG=(cs\hn_*), (cs\nh_*), (cs\ca_*), (cs\cb_*), (cs\co_*), (cs\ha_*)
stcs: STATS ARG=enscs.* SQDEVSUM PARARG=(cs\exphn_*), (cs\exph_*), (cs\expca_*), (cs\expcb_*), (cs\expco_*)
res: RESTRAINT ARG=stcs.sqdevsum AT=0. KAPPA=0. SLOPE=12

PRINT ARG=(cs\hn_*), (cs\nh_*), (cs\ca_*), (cs\cb_*), (cs\co_*), (cs\ha_*) FILE=CS STRIDE=1000
PRINT ARG=res.bias FILE=COLVAR STRIDE=10

ENDPLUMED
```

with respect to the case in which chemical shifts are used to define a standard collective variable, in this case `CS2BACKBONE` is a collective variable with multiple components, that are all the backcalculated chemical shifts, plus all the relative experimental values. The keyword function `ENSEMBLE` tells plumed to calculate the average of the arguments over the replicas (i.e. 4 replicas) and the function `STATS` compare the averaged back calculated chemical shifts with the experimental values and calculates the sum of the squared deviation. On this latter number it is possible to apply a linear `RESTRAINT` (because the variable is already a sum of squared differences) that is the new term we are adding to the underlying force field.

This example can be executed as

```
mpiexec -np 4 gmx_mpi mdrun -s topol -plumed plumed -multi 4
```

The third example show how `RDC` (calculated with the theta-methods) can be employed in the same way, in this case to describe the native state of Ubiquitin. In particular it is possible to observe how the RDC averaged restraint applied on the correlation between the calculated and experimental N-H RDCs result in the increase of the correlation of the RDCs for other bonds already on a very short time scale.

```
RDC ...
GYROM=-72.5388
SCALE=0.001060
ADDCOUPPLINGS
LABEL=nh
ATOMS1=20,21 COUPLING1=8.17
ATOMS2=37,38 COUPLING2=-8.271
ATOMS3=56,57 COUPLING3=-10.489
ATOMS4=76,77 COUPLING4=-9.871
#continue....
```

In this input the first four N-H RDCs are defined.

This example can be executed as

```
mpiexec -np 8 gmx_mpi mdrun -s topol -plumed plumed -multi 8
```

### 11.9.4 Reference

1. Granata, D., Camilloni, C., Vendruscolo, M. & Laio, A. Characterization of the free-energy landscapes of proteins by NMR-guided metadynamics. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6817–6822 (2013).
2. Cavalli, A., Camilloni, C. & Vendruscolo, M. Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle. *J. Chem. Phys.* 138, 094112 (2013).
3. Camilloni, C., Cavalli, A. & Vendruscolo, M. Replica-Averaged Metadynamics. *JCTC* 9, 5610–5617 (2013).
4. Roux, B. & Weare, J. On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method. *J. Chem. Phys.* 138, 084107 (2013).
5. Boomsma, W., Lindorff-Larsen, K. & Ferkinghoff-Borg, J. Combining Experiments and Simulations Using the Maximum Entropy Principle. *PLoS Comput. Biol.* 10, e1003406 (2014).
6. Camilloni, C. & Vendruscolo M. A Tensor-Free Method for the Structural and Dynamical Refinement of Proteins using Residual Dipolar Couplings. *J. PHYS. CHEM. B* 119, 653 (2015).

## 11.10 Belfast tutorial: Steinhardt Parameters

### 11.10.1 Aims

This tutorial is concerned with the collective variables that we use to study order/disorder transitions such as the transition between the solid and liquid phase. In this tutorial we will look at the transition from solid to liquid as this is easier to study using simulation. The CVs we will introduce can, and have, been used to study the transition from liquid to solid. More information can be found in the further reading section.

#### 11.10.1.1 Learning Outcomes

Once this tutorial is completed students will:

- Know of the existence of the Steinhardt Parameters and know how to create plumed input files for calculating them.
- Appreciate that the Steinhardt Parameter for a particular atom is a vector quantity and that you can thus measure local order in a material by taking dot products of the Steinhardt Parameter vectors of adjacent atoms. Students will know how to calculate such quantities using plumed.

### 11.10.2 Resources

The `tarball` for this project contains the following files:

- `in` : An input file for the simplemd code that forms part of plumed
- `input.xyz` : A configuration file for Lennard-Jones solid with an fcc solid structure

### 11.10.3 Instructions

#### 11.10.3.1 Simplemd

For this tutorial we will be using the MD code that is part of plumed - simplemd. This code allows us to do the simulations of `Lennard-Jones` that we require here but not much else. We will thus start this tutorial by doing some simulations with this code. You should have two files from the tarball, the first is called `input.xyz` and is basically an xyz file containing the positions of the atoms. The second meanwhile is called `in` and is the input to simplemd. If you open the file the contents should look something like this:

```
inputfile input.xyz
outputfile output.xyz
temperature 0.2
tstep 0.005
friction 1
forcecutoff 2.5
listcutoff 3.0
nstep 50000
nconfig 100 trajectory.xyz
nstat 10 energies.dat
```

Having run some molecular dynamics simulations in the past it should be pretty obvious what each line of the file does. One thing that might be a little strange is the units. Simplemd works with Lennard-Jones units so energy is in units of  $\epsilon$  and length is in units of  $\sigma$ . This means that temperature is in units of  $\frac{k_B T}{\epsilon}$ , which is why the temperature in the above file is apparently so low.

Use simplemd to run 50000 step calculations at 0.2, 0.8 and  $1.2 \frac{k_B T}{\epsilon}$ . (N.B. You will need an empty file called `plumed.dat` in order to run these calculations.) Visualise the trajectory output by each of your simulations using `VMD`. Please be aware that simplemd does not wrap the atoms into the cell box automatically. If you are at the tutorial we have resolved this problem by making it so that if you press `w` when you load all the atoms they will be wrapped into the box. At what temperatures did the simulations melt? What then is the melting point of the Lennard-Jones potential at this density?



### 11.10.3.2 Coordination Numbers

At the end of the previous section you were able to make very sophisticated judgements about whether or not the arrangement of atoms in your system was solid-like or liquid-like by simply looking at the configuration. The aim in the rest of this tutorial is to see if we can derive collective variables that are able to make an equally sophisticated judgement. For our first attempt let's use a CV which we have encountered elsewhere, the [COORDINATIONNUMBER](#).

Create a plumed input file that calculates the average [COORDINATIONNUMBER](#) of the atoms in your system and outputs it to a file every 100 steps. You will need to use the [COORDINATIONNUMBER](#) and [PRINT](#) actions. The switching function that tells plumed whether or not two atoms are bonded should be of type [RATIONAL](#) and should have its  $d_0$  parameter equal to 1.3 its  $r_0$  parameter equal to 0.2 and its  $n$  and  $m$  parameters set equal to 6 and 12 respectively.

Rerun the simplified simulations described at the end of the previous section. Is the average coordination number good at differentiating between solid and liquid configurations? Given your knowledge of physics/chemistry is this result surprising?

### 11.10.3.3 Steinhardt parameter

The solid and liquid phases of a material are both relatively dense so the result at the end of the last section - the fact that the coordination number is not particularly good at differentiating between them - should not be that much of a surprise. As you will have learnt early on in your scientific education when solids melt the atoms rearrange themselves in a much less ordered fashion. The bonds between them do not necessarily break it is just that, whereas in a solid the bonds were all at pretty well defined angles to each other, in a liquid the spread of bond angles is considerably larger. To detect the transition from solid to liquid what we need then is a coordinate that is able to recognise whether or not the geometry in the coordination spheres around each of the atoms in the system are the same or different. If these geometries are the same the system is in a solid-like configuration, whereas if they are different the system is liquid-like. The Steinhardt parameters [Q3](#), [Q4](#) and [Q6](#) are coordinates that allow us to examine the coordination spheres of atoms in precisely this way. The way in which these coordinates are able to do this is explained in the [video](#).

Repeat the calculations from the end of the previous section but edit the plumed.dat file so that the average [Q6](#) parameter is calculated and printed as well as the average [COORDINATIONNUMBER](#). In the Steinhardt parameter implementation in plumed the set of atoms in the coordination sphere of a particular atom are defined using a continuous switching function. For this calculation you should use a [RATIONAL](#) switching function with parameters  $d_0$  equal to 1.3,  $r_0$  equal to 0.2 and  $n$  and  $m$  set equal to 6 and 12 respectively. Is the average Q6 parameter able to differentiate between the solid and liquid states?

### 11.10.3.4 Local versus Global

At the end of the previous section you showed that the average Q6 parameter for a system of atoms is able to tell you whether or not that collection of atoms is in a solid-like or liquid-like configuration. In this section we will now ask the question - can the Steinhardt parameter always, unambiguously tell us whether particular tagged atoms are in a solid-like region of the material or whether they are in a liquid-like region of the material? This is an important question that might come up if we are looking at nucleation of a solid from the melt. Our question in this context reads - how do we unambiguously identify those atoms that are in the crystalline nucleus? A similar question would also come up when studying an interface between the solid and liquid phases. In this guise we would be asking about the extent of interface; namely, how far from the interface do we have to go before we can start thinking of the atoms as just another atom in the solid/liquid phase?

With these questions in mind our aim is to look at the distribution of values for the Q6 parameters of atoms in our system of Lennard-Jones have. If Q6 is able to unambiguously tell us whether or not an atom is in a solid-like pose or a liquid-like pose then there should be no overlap in the distributions obtained for the solid and liquid phases. If there is overlap, however, then we cannot use these coordinates for the applications described in the previous

paragraph. To calculate these distributions you will need to run two simulations - one at  $0.2 \frac{k_B T}{\epsilon}$  and one at  $1.2 \frac{k_B T}{\epsilon}$ . For these simulation you will need plumed.dat files that calculate and output the distribution of Steinhardt parameters. In writing the plumed input for these calculations you will need to use the [PRINT](#) command and the [Q6](#) command. In your Q6 instructions you will need to use the HISTOGRAM keyword - my recommendation would be to calculate a histogram consisting of 20 bins over a range starting at 0.0 and finishing at 1.0. Set the SMEAR parameter equal to 0.1. Do the distributions of Q6 parameters for the solid and liquid phase overlap? N.B. You can load the output from these simulations using GISMO and the all cvs button from the bar at the bottom.

#### 11.10.3.5 Local Steinhardt parameters

Hopefully you saw that the distribution of Q6 parameters for the solid and liquid phase overlap at the end of the previous section. Again this is not so surprising - as you go from solid to liquid the distribution of the geometries of the coordination spheres widens. The system is now able to arrange the atoms in the coordination spheres in a much wider variety of different poses. Importantly, however, the fact that an atom is in a liquid does not preclude it from having a very-ordered, solid-like coordination environment. As such in the liquid state we will find the occasional atom with a high value for the Q6 parameter. Consequently, an ordered coordination environment does not always differentiate solid-like atoms from liquid-like atoms. The major difference is the liquid the ordered atoms will always be isolated. That is to say in the liquid atoms with an ordered coordination will always be surrounded by atoms with a disordered coordination sphere. By contrast in the solid each ordered atom will be surrounded by further ordered atoms. This observation forms the basis of the local Steinhardt parameters and the locally averaged Steinhardt parameters that are explained in this [video](#).

Lets use plumed to calculate the distribution of [LOCAL\\_Q6](#) parameters in the solid and liquid phases. Repeat the calculations from the previous section but now use the HISTOGRAM keyword to calculate the distribution of [LOCAL\\_Q6](#) parameters. For the switching function in the [LOCAL\\_Q6](#) parameter instruction use a RATIONAL function with  $d_0$  equal to 1.3,  $r_0$  equal to 0.2 and  $n$  and  $m$  set equal to 6 and 12 respectively. For the HISTOGRAM use 20 bins starting at 0.0 and finishing at 1.0. Set the SMEAR parameter equal to 0.1. Do the distributions of [LOCAL\\_Q6](#) parameter for the solid and liquid phases overlap?

Lectner and Dellago have designed an alternative to the [LOCAL\\_Q6](#) parameter that is based on taking the [LOCAL\\_AVERAGE](#) of the Q6 parameter around a central atom. This quantity can be calculated using plumed. If you have time try to use the manual to work out how.

#### 11.10.4 Further Reading

There is a substantial literature on simulation of nucleation. Some papers are listed below but this list is far from exhaustive.

- F. Trudu, D. Donadio and M. Parrinello [Freezing of a Lennard-Jones Fluid: From Nucleation to Spinodal Regime](#), Phys. Rev. Lett. 97 105701 (2006)
- D. Quigley and P. M. Rodger [A metadynamics-based approach to sampling crystallization events](#), Mol. Simul. 2009
- W. Lechner and C. Dellago [Accurate determination of crystal structures based on averaged local bond order parameters](#) J. Chem. Phys 129 114707 (2008)

## 11.11 Cambridge tutorial

### Authors

Max Bonomi and Carlo Camilloni, part of the tutorial is based on other tutorials.

### Date

March 4, 2015

This document describes the PLUMED tutorial held for the Computational Biology MPhil, University of Cambridge, March 2015. The aim of this tutorial is to learn how to use PLUMED to run well-tempered and bias-exchange simulations and how to run replica-averaged metadynamics to incorporate experimental data into your simulation. Although the presented input files are correct, the users are invited to refer to the literature to understand how the parameters of enhanced sampling methods should be chosen in a real application.

Users are also encouraged to follow the links to the full PLUMED reference documentation and to wander around in the manual to discover the many available features and to do the other, more complete, tutorials. Here we are going to present only a very narrow selection of methods.

Users are expected to write PLUMED input files based on the instructions below.

### 11.11.1 Alanine dipeptide: our toy model

In this tutorial we will play with alanine dipeptide (see Fig. [cambridge-1-ala-fig](#)). This rather simple molecule is useful to make many benchmark that are around for data analysis and free energy methods. It is a rather nice example since it presents two metastable states separated by a high (free) energy barrier. Here metastable states are intended as states which have a relatively low free energy compared to adjacent conformation. It is conventional use to show the two states in terms of Ramachandran dihedral angles, which are denoted with  $\Phi$  and  $\Psi$  in Fig. [cambridge-1-transition-fig](#).

### 11.11.2 Exercise 1. Metadynamics

#### 11.11.2.1 Resources

The [tarball](#) for this project contains all the files needed to run this exercise.

#### 11.11.2.2 Summary of theory

In metadynamics, an external history-dependent bias potential is constructed in the space of a few selected degrees of freedom  $\vec{s}(q)$ , generally called collective variables (CVs) [41]. This potential is built as a sum of Gaussians deposited along the trajectory in the CVs space:

$$V(\vec{s}, t) = \sum_{k\tau < t} W(k\tau) \exp \left( - \sum_{i=1}^d \frac{(s_i - s_i(q(k\tau)))^2}{2\sigma_i^2} \right).$$

where  $\tau$  is the Gaussian deposition stride,  $\sigma_i$  the width of the Gaussian for the  $i$ th CV, and  $W(k\tau)$  the height of the Gaussian. The effect of the metadynamics bias potential is to push the system away from local minima into visiting

new regions of the phase space. Furthermore, in the long time limit, the bias potential converges to minus the free energy as a function of the CVs:

$$V(\vec{s}, t \rightarrow \infty) = -F(\vec{s}) + C.$$

In standard metadynamics, Gaussians of constant height are added for the entire course of a simulation. As a result, the system is eventually pushed to explore high free-energy regions and the estimate of the free energy calculated from the bias potential oscillates around the real value. In well-tempered metadynamics [43], the height of the Gaussian is decreased with simulation time according to:

$$W(k\tau) = W_0 \exp\left(-\frac{V(\vec{s}(q(k\tau)), k\tau)}{k_B \Delta T}\right),$$

where  $W_0$  is an initial Gaussian height,  $\Delta T$  an input parameter with the dimension of a temperature, and  $k_B$  the Boltzmann constant. With this rescaling of the Gaussian height, the bias potential smoothly converges in the long time limit, but it does not fully compensate the underlying free energy:

$$V(\vec{s}, t \rightarrow \infty) = -\frac{\Delta T}{T + \Delta T} F(\vec{s}) + C.$$

where  $T$  is the temperature of the system. In the long time limit, the CVs thus sample an ensemble at a temperature  $T + \Delta T$  which is higher than the system temperature  $T$ . The parameter  $\Delta T$  can be chosen to regulate the extent of free-energy exploration:  $\Delta T = 0$  corresponds to standard molecular dynamics,  $\Delta T \rightarrow \infty$  to standard metadynamics. In well-tempered metadynamics literature and in PLUMED, you will often encounter the term "biasfactor" which is the ratio between the temperature of the CVs ( $T + \Delta T$ ) and the system temperature ( $T$ ):

$$\gamma = \frac{T + \Delta T}{T}.$$

The biasfactor should thus be carefully chosen in order for the relevant free-energy barriers to be crossed efficiently in the time scale of the simulation.

Additional information can be found in the several review papers on metadynamics [59] [60] [61].

### 11.11.2.3 Setup, run, and analyse a well-tempered metadynamics simulation

In this exercise, we will run a well-tempered metadynamics simulation on alanine dipeptide in vacuum, using as CVs the two backbone dihedral angles phi and psi.

In order to run this simulation we need to prepare the PLUMED input file (plumed.dat) as follows.

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate metadynamics in phi and psi
# depositing a Gaussian every 500 time steps,
# with height equal to 1.2 kJoule/mol,
# and width 0.35 rad for both CVs.
# Well-tempered metadynamics is activated,
# and the biasfactor is set to 6.0
#
metad: METAD ARG=phi,psi PACE=500 HEIGHT=1.2 SIGMA=0.35,0.35 FILE=HILLS BIASFACTOR=6.0

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR
```

(see [TORSION](#), [METAD](#), and [PRINT](#)).

The syntax for the command [METAD](#) is simple. The directive is followed by a keyword ARG followed by the labels of the CVs on which the metadynamics potential will act. The keyword PACE determines the stride of Gaussian deposition in number of time steps, while the keyword HEIGHT specifies the height of the Gaussian in kJoule/mol. To run well-tempered metadynamics, we need two additional keywords BIASFACTOR and TEMP, which specify how fast the bias potential is decreasing with time and the temperature of the simulation, respectively. For each CVs, one has to specify the width of the Gaussian by using the keyword SIGMA. Gaussian will be written to the file indicated by the keyword FILE.

Once the PLUMED input file is prepared, one has to run Gromacs with the option to activate PLUMED and read the input file:

```
gmx_mpi mdrun -plumed
```

During the metadynamics simulation, PLUMED will create two files, named COLVAR and HILLS. The COLVAR file contains all the information specified by the PRINT command, in this case the value of the CVs every 10 steps of simulation, along with the current value of the metadynamics bias potential. The HILLS file contains a list of the Gaussians deposited along the simulation. If we give a look at the header of this file, we can find relevant information about its content. The last column contains the value of the biasfactor used, while the last but one the height of the Gaussian, which is rescaled during the simulation following the well-tempered recipe.

```
#! FIELDS time phi psi sigma_phi sigma_psi height biasf
#! SET multivariate false
#! SET min_phi -pi
#! SET max_phi pi
#! SET min_psi -pi
#! SET max_psi pi
1.0000 -1.3100 0.0525 0.35 0.35 1.4400 6
2.0000 -1.4054 1.9742 0.35 0.35 1.4400 6
3.0000 -1.9997 2.5177 0.35 0.35 1.4302 6
4.0000 -2.2256 2.1929 0.35 0.35 1.3622 6
```

If we carefully look at the height column, we will notice that in the beginning the value reported is higher than the initial height specified in the input file, which should be 1.2 kJoule/mol. In fact, this column reports the height of the Gaussian rescaled by the pre-factor that in well-tempered metadynamics relates the bias potential to the free energy. In this way, when we will use [sum\\_hills](#), the sum of the Gaussians deposited will directly provide the free-energy, without further rescaling needed.

We can plot the time evolution of the CVs along with the height of the Gaussian.

The system is initialized in one of the local minimum where it starts accumulating bias. As the simulation progresses and the bias added grows, the Gaussian height is progressively reduced. After a while ( $t=0.8$  ns), the system is able to escape the local minimum and explore a new region of the phase space. As soon as this happens, the Gaussian height is restored to the initial value and starts to decrease again. In the long time, the Gaussian height becomes smaller and smaller while the system diffuses in the entire CVs space.

#### 11.11.2.4 Calculate free-energies and monitor convergence

One can estimate the free energy as a function of the metadynamics CVs directly from the metadynamics bias potential. In order to do so, the utility [sum\\_hills](#) should be used to sum the Gaussians deposited during the simulation and stored in the HILLS file.

To calculate the two-dimensional free energy as a function of phi and psi, it is sufficient to use the following command line:

```
plumed sum_hills --hills HILLS
```

The command above generates a file called `fes.dat` in which the free-energy surface as function of  $\phi$  and  $\psi$  is calculated on a regular grid. One can modify the default name for the free energy file, as well as the boundaries and bin size of the grid, by using the following options of `sum_hills` :

```
--outfile - specify the outputfile for sumhills
--min - the lower bounds for the grid
--max - the upper bounds for the grid
--bin - the number of bins for the grid
--spacing - grid spacing, alternative to the number of bins
```

It is also possible to calculate one-dimensional free energies from the two-dimensional metadynamics simulation. For example, if one is interested in the free energy as a function of the  $\phi$  dihedral alone, the following command line should be used:

```
plumed sum_hills --hills HILLS --idw phi --kt 2.5
```

The result should look like this:

To assess the convergence of a metadynamics simulation, one can calculate the estimate of the free energy as a function of simulation time. At convergence, the reconstructed profiles should be similar. The option `--stride` should be used to give an estimate of the free energy every  $N$  Gaussians deposited, and the option `--mintozero` can be used to align the profiles by setting the global minimum to zero. If we use the following command line:

```
plumed sum_hills --hills HILLS --idw phi --kt 2.5 --stride 500 --mintozero
```

one free energy is calculated every 500 Gaussians deposited, and the global minimum is set to zero in all profiles. Try to visualize the different estimates of the free energy as a function of time.

To assess the convergence of the simulation more quantitatively, we can calculate the free-energy difference between the two local minima in the one-dimensional free energy along  $\phi$  as a function of simulation time. We can use the bash script `analyze_FES.sh` to integrate the multiple free-energy profiles in the two basins defined by the following intervals in  $\phi$  space: basin A,  $-3 < \phi < -1$ , basin B,  $0.5 < \phi < 1.5$ .

```
./analyze_FES.sh NFES -3.0 -1.0 0.5 1.5 KBT
```

where `NFES` is the number of profiles (free-energy estimates at different times of the simulation) generated by the option `--stride` of `sum_hills`, and `KBT` is the temperature in energy units (in this case  $KBT=2.5$ ).

This analysis, along with the observation of the diffusive behavior in the CVs space, suggest that the simulation is converged.

### 11.11.3 Exercise 2. Bias-Exchange Metadynamics

#### 11.11.3.1 Resources

The `tarball` for this project contains all the files needed to run this exercise.

## 11.11.3.2 Summary of theory

In well-tempered metadynamics the free-energy landscape of the system is reconstructed by gradually filling the local minima with gaussian hills. The dimensionality of the landscape is equal to the number of CVs which are biased, and typically a number of CVs smaller than four is employed. The reason for this is that qualitatively, if the CVs are not correlated among them, the simulation time required to fill the free-energy landscape grows exponentially with the number of CVs. This limitation can be severe when studying complex transformations or reactions in which more than say three relevant CVs can be identified.

Alternative solutions employ metadynamics together with other enhanced sampling techniques (i.e. Parallel Tempering or more generally Hamiltonian Replica Exchange). In particular in Bias-Exchange metadynamics the problem of sampling a multi-dimensional free-energy surface is recast in that of sampling many one-dimensional free-energies. In this approach a relatively large number  $N$  of CVs is chosen to describe the possible transformations of the system (e.g., to study the conformations of a peptide one may consider all the dihedral angles between amino acids). Then,  $N$  metadynamics simulations (replicas) are run on the same system at the same temperature, biasing a different CV in each replica. Normally, in these conditions, each bias profile would converge very slowly to the equilibrium free-energy, due to hysteresis.

In order to tackle this problem, in the bias-exchange approach every fixed number of steps (say 10,000) an exchange is attempted between a randomly selected pair of replicas  $a$  and  $b$ . The exchanges allow the replicas to gain from the sampling of the other replicas and enable the system to explore the conformational space along a large number of different directions.

The probability to accept the exchange is given by a Metropolis rule:

$$\min \left( 1, \exp \left[ \beta (V_G^a(x^a, t) + V_G^b(x^b, t) - V_G^a(x^b, t) - V_G^b(x^a, t)) \right] \right)$$

where  $x^a$  and  $x^b$  are the coordinates of replicas  $a$  and  $b$  and  $V_G^{a(b)}(x, t)$  is the metadynamics potential acting on the replica  $a$  ( $b$ ). Each trajectory evolves through the high dimensional free energy landscape in the space of the CVs sequentially biased by different metadynamics potentials acting on one CV at each time. The results of the simulation are  $N$  one-dimensional projections of the free energy.

## 11.11.3.3 Setup, run, and analyse a well-tempered bias-exchange metadynamics simulation

In the following example, a bias-exchange simulation is performed on Alanine di-peptide. A typical input file for BE-WTMetaD is the following:

- a common input file in which all the collective variables are defined:

```
plumed-common.dat:

# read a pdb file to get topology info
MOLINFO STRUCTURE=ALAD.pdb

# tell PLUMED to generate random exchange trials
RANDOM_EXCHANGES

# set up four variables for Phi, Psi, Theta, Xi dihedral angles
phi: TORSION ATOMS=@phi-2
psi: TORSION ATOMS=@psi-2
theta: TORSION ATOMS=6,5,7,9
xi: TORSION ATOMS=16,15,17,19

PRINT ARG=phi,psi,theta,xi STRIDE=250 FILE=COLVAR
```

(see [MOLINFO](#) and [RANDOM\\_EXCHANGES](#)).

- additional input files that [INCLUDE](#) the common input and define the [METAD](#) along the selected CVs:

```
plumed.dat.0:
```

```
INCLUDE FILE=plumed-common.dat
METAD ARG=phi HEIGHT=1.2 SIGMA=0.25 PACE=100 GRID_MIN=-pi GRID_MAX=pi BIASFACTOR=10.0
ENDPLUMED
```

```
plumed.dat.1:
```

```
INCLUDE FILE=plumed-common.dat
METAD ARG=psi HEIGHT=1.2 SIGMA=0.25 PACE=100 GRID_MIN=-pi GRID_MAX=pi BIASFACTOR=10.0
ENDPLUMED
```

```
plumed.dat.2:
```

```
INCLUDE FILE=plumed-common.dat
METAD ARG=theta HEIGHT=1.2 SIGMA=0.50 PACE=100 GRID_MIN=-pi GRID_MAX=pi BIASFACTOR=10.0
ENDPLUMED
```

```
plumed.dat.3:
```

```
INCLUDE FILE=plumed-common.dat
METAD ARG=xi HEIGHT=1.2 SIGMA=0.50 PACE=100 GRID_MIN=-pi GRID_MAX=pi BIASFACTOR=10.0
ENDPLUMED
```

The, in this case, two replicas start from the same GROMACS topology file replicated twice: `topol0.tpr`, `topol1.tpr`, `topol2.tpr` and `topol3.tpr`. Finally, GROMACS is launched as a parallel run on 4 cores, with one replica per core, with the command

```
mpirun -np 4 gmx_mpi mdrun -s topol -plumed plumed -multi 4 -replex 10000 >& log &
```

where `-replex 10000` indicates that every 10000 molecular-dynamics steps exchanges are attempted (as printed in the GROMACS `*.log` files).

In order to have an idea of how Bias-Exchange Metadynamics works we can compare the sampling of the four replicas along the first two collective variables with respect to a free energy surface obtained by sampling in two dimensions.

The main features are that all the replicas can sample all the relevant free energy minima even if their collective variable is not meant to sample them, only the replicas with a collective variable meant to pass a barrier can sample that transition, so high energy regions are sampled only locally. This can seem a weakness in the case of alanine dipeptide but is the real strength of BE-MetaD, indeed by not knowing anything of a system it is possible to choose as many collective variables as possible and even if most of them are not particularly usefull a posteriori, they all gain from the good ones and nothing is lost.

#### 11.11.3.4 Calculate free-energies and monitor convergence

As for the former case, one can estimate the free energy as a function of the metadynamics CVs directly from the metadynamics bias potential. This approach can only be used to recover one-dimensional free-energy profiles:

```
plumed sum_hills --hills HILLS.0 --mintozero
```



The command above generates a file called `fes.dat` in which the free-energy surface as function of  $\phi$  is calculated on a regular grid.

It is also possible to calculate one-dimensional free energies from the two-dimensional metadynamics simulation. For example, if one is interested in the free energy as a function of the  $\phi$  dihedral alone, the following command line should be used:

The result should look like this (compared with the one obtained before):

As above we can assess the convergence of a metadynamics simulation by calculating the estimate of the free energy as a function of simulation time. At convergence, the reconstructed profiles should be similar. The option `--stride` should be used to give an estimate of the free energy every  $N$  Gaussians deposited, and the option `--mintozero` can be used to align the profiles by setting the global minimum to zero. If we use the following command line:

```
plumed sum_hills --hills HILLS.0 --stride 500 --mintozero
```

Try to visualize the different estimates of the free energy as a function of time.

To assess the convergence of the simulation more quantitatively, we can calculate the free-energy difference between the two local minima in the one-dimensional free energy along  $\phi$  as a function of simulation time. We can use the bash script `analyze_FES.sh` to integrate the multiple free-energy profiles in the two basins defined by the following intervals in  $\phi$  space: basin A,  $-3 < \phi < -1$ , basin B,  $0.5 < \phi < 1.5$ .

```
./analyze_FES.sh NFES -3.0 -1.0 0.5 1.5 KBT
```

where `NFES` is the number of profiles (free-energy estimates at different times of the simulation) generated by the option `--stride` of `sum_hills`, and `KBT` is the temperature in energy units (in this case  $KBT=2.5$ ).

In addition to check for the convergence of the one dimensional profiles, it is important to verify the rate of exchange among the replicas. In fact if the rate is too low or if it is not enough homogeneous then the replicas can still suffer from histeris problems (ie with one replica trapped somewhere in the conformational space). Generally speaking a good probability of exchange is between 10 and 30%.

It is possible to check the exchange statistics and the detailed diffusion of the replicas among the different biases using a script provided:

```
./demux_m.pl md0.log
```

which will create two files, called `replica_temp.xvg` and `replica_index.xvg`. We can plot the first file, which reports the bias index of each configuration at a given time of the simulation. This file is extremely useful, because it allows us monitoring the replicas diffusion in bias.

### 11.11.4 Exercise 3. Replica-Average Metadynamics

#### 11.11.4.1 Resources

The [tarball](#) for this project contains all the files needed to run this exercise.

#### 11.11.4.2 Summary of theory

Equilibrium experimental data are the result of a measure over an ensemble of structures and over time. In principle a "perfect" molecular dynamics simulations, that is a simulations with a perfect force-field and a perfect sampling can predict the outcome of an experiment in a quantitative way. Actually in most of the cases obtaining a qualitative agreement is already a fortunate outcome. In order to increase the accuracy of a force field in a system dependent manner it is possible to add to the force-field an additional term based on the agreement with a set of experimental data. This agreement is not enforced as a simple restraint because this would mean to ask the system to be always in agreement with all the experimental data at the same time, instead the restraint is applied over an AVERAGED COLLECTIVE VARIABLE where the average is performed over multiple identical simulations (replicas). In this way there is not a single replica that must be in agreement with the experimental data but they should be in agreement on average. It has been shown that this approach is equivalent to solve the problem of finding a modified version of the force field that reproduces the provided set of experimental data without any additional assumption on the data themselves [65] [66] [67] [68] .

#### 11.11.4.3 The system: Chignolin

In this exercise we will model the equilibrium ensemble of Chignolin by combining an implicit solvent force-field (Amber99sb) with synthetic experimental data derived from an experimentally determined ensemble of structures.

The experimental data are the following distances among CA carbons that have been averaged over the whole ensemble.

#RES	RES	DISTANCE (nm)
1	3	0.656214
1	4	0.963703
1	5	1.197160
1	6	1.219970
1	7	1.206760
1	8	1.060110
1	9	0.858911
1	10	0.872988
2	4	0.656913
2	5	0.931405
2	6	0.953046
2	7	0.901493
2	8	0.809623
2	9	0.716444
2	10	0.882070
3	5	0.588322
3	6	0.660023
3	7	0.704165
3	8	0.720063
3	9	0.799961
3	10	0.981838
4	6	0.550374
4	7	0.572299
4	8	0.764540
4	9	0.939752
4	10	1.186040
5	7	0.613088
5	8	0.847734
5	9	1.084080
5	10	1.296630
6	8	0.591571
6	9	0.888388
6	10	1.102920
7	9	0.701717
7	10	0.987368
8	10	0.654310

#### 11.11.4.4 Setup, run and analysis

Here we will give partial information on how to setup the simulations. Users should refer to the manual and the literature cited on how to complete the information provided and thus successfully perform the exercise.

Step 1: Add the collective variables that describe the distances averaged over the ensemble to the template input files provided.

For each of the above experimental data one should define:

```
#this is the distance between CA atoms 5 and 33 belonging to residues 1 and 3, respectively.
d1: DISTANCE ATOMS=5,33
# this is the averaging of the distance among the replicas
ed1: ENSEMBLE ARG=d1
# this is the restraint applied on the averaged distance to match the experimental one
RESTRAINT ARG=ed1.d1 KAPPA=100000 AT=0.656214
```

(see [DISTANCE](#), [ENSEMBLE](#), and [RESTRAINT](#)).

Step 2: Setup the Bias-Exchange simulations

We will run Bias-Exchange simulations using four CVs, two of them have already been chosen (see `plumed-common.dat`) while the others should be selected by you. Additionally Metadynamics parameters [METAD](#) for the two selected cvs should be added in the `plumed.dat.0` and `plumed.dat.1` files. In particular while the PACE, HEIGHT and BIASFACTOR can be kept the same of those defined for the preselected CVs, the SIGMA should be half of the fluctuations of the chosen CV in an unbiased run ( $>1$  ns).

Step 3: Run

The simulation should be run until convergence of the four monodimensional free energies, this will typically take more than 200 ns per replica.

Step 4: Analysis

The user should check the diffusion of the replicas among the biases (see above) and send the converged free energy profiles.

Step 5: Only for the brave!

The same simulation without experimental restraints could be performed in such a way to compare the free energy profiles obtained with and without the inclusion of experimental data.

## 11.12 Cineca tutorial

### Authors

Max Bonomi, stealing most of the material from other tutorials. Alejandro Gil Ley is acknowledged for beta-testing this tutorial.

### Date

November 20, 2015

This document describes the PLUMED tutorial held at CINECA, November 2015. The aim of this tutorial is to learn how to use PLUMED to analyze molecular dynamics (MD) simulations on-the-fly, to analyze existing trajectories, and to perform enhanced sampling. Although the presented input files are correct, the users are invited to **refer to the literature to understand how the parameters of enhanced sampling methods should be chosen in a real application**.

Users are also encouraged to follow the links to the full PLUMED reference documentation and to wander around in the manual to discover the many available features and to do the other, more complete, tutorials. Here we are going to present only a very narrow selection of methods.

All the tests here are performed on a toy system, alanine dipeptide, simulated in vacuum using the AMBER99SB force field. Simulations are carried out with GROMACS 5.0.4, which is here assumed to be already patched with PLUMED 2.2 and properly installed. However, these examples could be easily converted to other MD software.

### 11.12.1 Resources

All the GROMACS input files and analysis scripts are provided in this [tarball](#). Users are expected to write PLUMED input files based on the instructions below.

We can start by downloading the tarball and uncompressing it:

```
> tar xzvf cineca.tar.gz
```

#### Warning

Exercises should be run inside the newly-created cineca directory, by creating new subdirectories, one per exercise.

### 11.12.2 Alanine dipeptide: our toy model

In this tutorial we will play with alanine dipeptide (see Fig. [cineca-1-ala-fig](#)). This rather simple molecule is useful to benchmark data analysis and free-energy methods. This system is a nice example because it presents two metastable states separated by a high free-energy barrier. It is conventional use to characterize the two states in terms of Ramachandran dihedral angles, which are denoted with  $\Phi$  and  $\Psi$  in Fig. [cineca-1-transition-fig](#).

### 11.12.3 Monitoring collective variables

The main goal of PLUMED is to compute collective variables, which are complex descriptors of a system than can be used to analyze a conformational change or a chemical reaction. This can be done either on-the-fly, that is during MD, or a posteriori, using PLUMED as a post-processing tool. In both cases one should create an input file with a specific PLUMED syntax. A sample input file is below:

```
# compute distance between atoms 1 and 10
d: DISTANCE ATOMS=1,10
# create a virtual atom in the center between atoms 20 and 30
center: CENTER ATOMS=20,30
# compute torsional angle between atoms 1,10,20 and center
phi: TORSION ATOMS=1,10,20,center
# print d and phi every 10 step
PRINT ARG=d,phi STRIDE=10
```

(see [DISTANCE](#), [CENTER](#), [TORSION](#), and [PRINT](#))

PLUMED works using kJ/nm/ps as energy/length/time units. This can be personalized using [UNITS](#). Notice that variables should be given a name (in the example above, `d`, and `phi`), which is then used to refer to these variables. Lists of atoms should be provided as comma separated numbers, with no space. Virtual atoms can be created and assigned a name for later use. You can find more information on the PLUMED syntax at [Getting started](#) page of the manual. The complete documentation for all the supported collective variables can be found at the [Collective variables](#) page.

## 11.12.3.1 Exercise 1: on-the-fly analysis

Here we will run a plain MD simulation on alanine dipeptide and compute two torsional angles on-the-fly. GROMACS needs a .tpr file, which is a binary file containing initial positions as well as force-field parameters. For this tutorial, it is sufficient to use the .tpr files provided in the SETUP directory of the package:

- topolA.tpr - setup in vacuum, initialized in state A
- topolB.tpr - setup in vacuum, initialized in state B

However, we also provide .gro, .mdp, and .top files, that can be modified and used to generate a new .tpr file.

GROMACS can be run (interactively) using the following command:

```
> gmx_mpi mdrun -s ../SETUP/topolA.tpr -nsteps 10000 -x traj.xtc
```

The nsteps flags can be used to change the number of timesteps and topolA.tpr is the name of the tpr file. While running, GROMACS will produce a md.log file, with log information, and a traj.xtc file, with a binary trajectory. The trajectory can be visualized with VMD using:

```
> vmd confout.gro traj.xtc
```

To activate PLUMED during a GROMACS MD simulation, you need to add the -plumed flag

```
> gmx_mpi mdrun -s ../SETUP/topolA.tpr -nsteps 10000 -plumed plumed.dat -x traj.xtc
```

Here plumed.dat is the name of the PLUMED input file. Notice that PLUMED will write information in the md.log that could be useful to verify if the simulation has been set up properly.

In this exercise, we will run a plain MD simulation and monitor the  $\Phi$  and  $\Psi$  dihedral angles on-the-fly. In order to do so, you need to prepare the following PLUMED input file:

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
PRINT ARG=phi,psi STRIDE=100 FILE=COLVAR
```

(see [TORSION](#) and [PRINT](#))

PLUMED is going to compute the collective variables only when necessary, that is, in this case, every 100 steps. This is not very relevant for simple variables such as torsional angles, but provides a significant speed-up when using expensive collective variables.

PLUMED will write a textual file named COLVAR containing three columns: simulation time,  $\Phi$  and  $\Psi$ . Results can be plotted using gnuplot:

```
> gnuplot
# this shows phi as a function of time
gnuplot> plot "COLVAR" u 2
# this shows psi as a function of time
gnuplot> plot "COLVAR" u 3
# this shows psi as a function of phi
gnuplot> plot "COLVAR" u 2:3
```

Now try to do the same using the two different initial configurations that we provided (topolA.tpr and topolB.tpr). Results from 200ps (100000 steps) trajectories in vacuum are shown in Figure [cineca-ala-traj](#).

Notice that the result depends heavily on the starting structure. For the simulation in vacuum, the two free-energy minima are separated by a large barrier so that the system cannot cross it in such short simulation time. Also notice that the two clouds are well separated, indicating that these two collective variables are appropriate to properly distinguish the two minima.

As a final comment, consider that if you run twice the same calculation in the same directory, you might overwrite the output files. GROMACS takes automatic backup of these files, and PLUMED does it as well. In case you are restarting a simulation, you can add the keyword [RESTART](#) at the beginning of the PLUMED input file. This will tell PLUMED to *append* files instead of taking a backup copy.

### 11.12.3.2 Exercise 2: analysis with the driver tool

Imagine you have already run a simulation, with or without PLUMED. You might want to compute the collective variables a posteriori, i.e. from the trajectory file. You can do this by using the PLUMED executable on the command line. Type

```
> plumed driver --help
```

to have an idea of the possible options. See [driver](#) for the full documentation.

Here we will use the driver to compute  $\Phi$  and  $\Psi$  on the trajectory previously generated. Let's assume the trajectory is named `traj.xtc`. You should prepare a PLUMED input file named `analysis.dat` as:

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
PRINT ARG=phi,psi FILE=COLVAR_ANALYSIS
```

(see [TORSION](#) and [PRINT](#)) Notice that typically when using the driver we do not provide a STRIDE keyword to PRINT. This implies "print at every step" which, analyzing a trajectory, means "print for all the available snapshots". Then, you can use the following command:

```
> plumed driver --mf_xtc traj.xtc --plumed analysis.dat
```

Notice that PLUMED has no way to know the value of physical time from the trajectory. If you want physical time to be printed in the output file you should give more information to the driver, e.g.:

```
> plumed driver --mf_xtc traj.xtc --plumed analysis.dat --timestep 0.002 --trajectory-stride 1000
```

(see [driver](#))

In this case we inform the driver that the `traj.xtc` file was produced in a run with a timestep of 0.002 ps and saving a snapshot every 1000 timesteps.

You might want to analyze a different collective variable, such as the gyration radius. The gyration radius tells how extended is a molecule in space. You can do it with the following PLUMED input file

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

heavy: GROUP ATOMS=1,5,6,7,9,11,15,16,17,19
gyr: GYRATION ATOMS=heavy

# the same could have been achieved with
# gyr: GYRATION ATOMS=1,5,6,7,9,11,15,16,17,19

PRINT ARG=phi,psi,gyr FILE=analyze
```

(see [TORSION](#), [GYRATION](#), [GROUP](#), and [PRINT](#))

Now try to compute the time serie of the gyration radius.

### 11.12.4 Biasing collective variables

We have seen that PLUMED can be used to compute collective variables. However, PLUMED is most often used to add forces on the collective variables. To this aim, we have implemented a variety of possible biases acting on collective variables. A bias works in a manner conceptually similar to the `PRINT` command, taking as argument one or more collective variables. However, here the STRIDE is usually omitted (that is equivalent to setting it to 1), which means that forces are applied at every timestep. Starting from PLUMED 2.2 you can change the STRIDE also for bias potentials, but that's another story. In the following we will see how to build an adaptive bias potential with metadynamics and how to apply harmonic restraints. The complete documentation for all the biasing methods available in PLUMED can be found at the [Bias](#) page.

#### 11.12.4.1 Metadynamics

##### 11.12.4.1.1 Summary of theory

In metadynamics, an external history-dependent bias potential is constructed in the space of a few selected degrees of freedom  $\vec{s}(q)$ , generally called collective variables (CVs) [41]. This potential is built as a sum of Gaussians deposited along the trajectory in the CVs space:

$$V(\vec{s}, t) = \sum_{k\tau < t} W(k\tau) \exp \left( - \sum_{i=1}^d \frac{(s_i - s_i(q(k\tau)))^2}{2\sigma_i^2} \right).$$

where  $\tau$  is the Gaussian deposition stride,  $\sigma_i$  the width of the Gaussian for the  $i$ th CV, and  $W(k\tau)$  the height of the Gaussian. The effect of the metadynamics bias potential is to push the system away from local minima into visiting new regions of the phase space. Furthermore, in the long time limit, the bias potential converges to minus the free energy as a function of the CVs:

$$V(\vec{s}, t \rightarrow \infty) = -F(\vec{s}) + C.$$

In standard metadynamics, Gaussians of constant height are added for the entire course of a simulation. As a result, the system is eventually pushed to explore high free-energy regions and the estimate of the free energy calculated from the bias potential oscillates around the real value. In well-tempered metadynamics [43], the height of the Gaussian is decreased with simulation time according to:

$$W(k\tau) = W_0 \exp \left( - \frac{V(\vec{s}(q(k\tau)), k\tau)}{k_B \Delta T} \right),$$

where  $W_0$  is an initial Gaussian height,  $\Delta T$  an input parameter with the dimension of a temperature, and  $k_B$  the Boltzmann constant. With this rescaling of the Gaussian height, the bias potential smoothly converges in the long time limit, but it does not fully compensate the underlying free energy:

$$V(\vec{s}, t \rightarrow \infty) = - \frac{\Delta T}{T + \Delta T} F(\vec{s}) + C.$$

where  $T$  is the temperature of the system. In the long time limit, the CVs thus sample an ensemble at a temperature  $T + \Delta T$  which is higher than the system temperature  $T$ . The parameter  $\Delta T$  can be chosen to regulate the extent of free-energy exploration:  $\Delta T = 0$  corresponds to standard MD,  $\Delta T \rightarrow \infty$  to standard metadynamics. In well-tempered metadynamics literature and in PLUMED, you will often encounter the term "biasfactor" which is the ratio between the temperature of the CVs ( $T + \Delta T$ ) and the system temperature ( $T$ ):

$$\gamma = \frac{T + \Delta T}{T}.$$

The biasfactor should thus be carefully chosen in order for the relevant free-energy barriers to be crossed efficiently in the time scale of the simulation.

Additional information can be found in the several review papers on metadynamics [59] [60] [61].

If you do not know exactly where you would like your collective variables to go, and just know (or suspect) that some variables have large free-energy barriers that hinder some conformational rearrangement or some chemical reaction, you can bias them using metadynamics. In this way, a time dependent, adaptive potential will be constructed that tends to disfavor visited configurations in the collective-variable space. The bias is usually built as a sum of Gaussian deposited in the already visited states.

### 11.12.4.1.2 Exercise 3

In this exercise, we will run a well-tempered metadynamics simulation on alanine dipeptide in vacuum, using as CV the backbone dihedral angle phi. In order to run this simulation we need to prepare the PLUMED input file (plumed.dat) as follows.

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate well-tempered metadynamics in phi depositing
# a Gaussian every 500 time steps, with initial height equal
# to 1.2 kJoule/mol, biasfactor equal to 10.0, and width to 0.35 rad

METAD ...
LABEL=metad
ARG=phi
PACE=500
HEIGHT=1.2
SIGMA=0.35
FILE=HILLS
BIASFACTOR=10.0
TEMP=300.0
GRID_MIN=-pi
GRID_MAX=pi
GRID_SPACING=0.1
... METAD

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR
```

(see [TORSION](#), [METAD](#), and [PRINT](#)).

The syntax for the command [METAD](#) is simple. The directive is followed by a keyword ARG followed by the labels of the CVs on which the metadynamics potential will act. The keyword PACE determines the stride of Gaussian deposition in number of time steps, while the keyword HEIGHT specifies the height of the Gaussian in kJoule/mol. For each CVs, one has to specified the width of the Gaussian by using the keyword SIGMA. Gaussian will be written to the file indicated by the keyword FILE.

The bias potential will be stored on a grid, whose boundaries are specified by the keywords GRID\_MIN and GRID\_MAX. Notice that you should provide either the number of bins for every collective variable (GRID\_BIN) or the desired grid spacing (GRID\_SPACING). In case you provide both PLUMED will use the most conservative choice (highest number of bins) for each dimension. In case you do not provide any information about bin size (neither GRID\_BIN nor GRID\_SPACING) and if Gaussian width is fixed PLUMED will use 1/5 of the Gaussian width as grid spacing. This default choice should be reasonable for most applications.

Once the PLUMED input file is prepared, one has to run GROMACS with the option to activate PLUMED and read the input file:



```
gmx_mpi mdrun -s ../SETUP/topolA.tpr -plumed plumed.dat -nsteps 5000000 -x traj.xtc
```

During the metadynamics simulation, PLUMED will create two files, named COLVAR and HILLS. The COLVAR file contains all the information specified by the PRINT command, in this case the value of the CVs every 10 steps of simulation, along with the current value of the metadynamics bias potential. We can use COLVAR to visualize the behavior of the CV during the simulation:

By inspecting Figure [cineca-metad-phi-fig](#), we can see that the system is initialized in one of the two metastable states of alanine dipeptide. After a while ( $t=0.1$  ns), the system is pushed by the metadynamics bias potential to visit the other local minimum. As the simulation continues, the bias potential fills the underlying free-energy landscape, and the system is able to diffuse in the entire phase space.

The HILLS file contains a list of the Gaussians deposited along the simulation. If we give a look at the header of this file, we can find relevant information about its content:

```
#! FIELDS time phi psi sigma_phi sigma_psi height biasf
#! SET multivariate false
#! SET min_phi -pi
#! SET max_phi pi
#! SET min_psi -pi
#! SET max_psi pi
```

The line starting with FIELDS tells us what is displayed in the various columns of the HILLS file: the time of the simulation, the value of phi and psi, the width of the Gaussian in phi and psi, the height of the Gaussian, and the biasfactor. We can use the HILLS file to visualize the decrease of the Gaussian height during the simulation, according to the well-tempered recipe:

If we look carefully at the scale of the y-axis, we will notice that in the beginning the value of the Gaussian height is higher than the initial height specified in the input file, which should be 1.2 kJoule/mol. In fact, this column reports the height of the Gaussian rescaled by the pre-factor that in well-tempered metadynamics relates the bias potential to the free energy. In this way, when we will use [sum\\_hills](#), the sum of the Gaussians deposited will directly provide the free-energy, without further rescaling needed (see below).

One can estimate the free energy as a function of the metadynamics CVs directly from the metadynamics bias potential. In order to do so, the utility [sum\\_hills](#) should be used to sum the Gaussians deposited during the simulation and stored in the HILLS file.

To calculate the free energy as a function of phi, it is sufficient to use the following command line:

```
plumed sum_hills --hills HILLS
```

The command above generates a file called fes.dat in which the free-energy surface as function of phi is calculated on a regular grid. One can modify the default name for the free energy file, as well as the boundaries and bin size of the grid, by using the following options of [sum\\_hills](#) :

```
--outfile - specify the outputfile for sumhills
--min - the lower bounds for the grid
--max - the upper bounds for the grid
--bin - the number of bins for the grid
--spacing - grid spacing, alternative to the number of bins
```

The result should look like this:

To assess the convergence of a metadynamics simulation, one can calculate the estimate of the free energy as a function of simulation time. At convergence, the reconstructed profiles should be similar. The option `--stride` should be used to give an estimate of the free energy every N Gaussians deposited, and the option `--mintozero` can be used to align the profiles by setting the global minimum to zero. If we use the following command line:

```
plumed sum_hills --hills HILLS --stride 100 --mintozero
```

one free energy is calculated every 100 Gaussians deposited, and the global minimum is set to zero in all profiles. The resulting plot should look like the following:

To assess the convergence of the simulation more quantitatively, we can calculate the free-energy difference between the two local minima of the free energy along  $\phi$  as a function of simulation time. We can use following script to integrate the multiple free-energy profiles in the two basins defined by the following regions in  $\phi$  space: basin A,  $-3 < \phi < -1$ , basin B,  $0.5 < \phi < 1.5$ .

```
# number of free-energy profiles
nfes= # put here the number of profiles
# minimum of basin A
minA=-3
# maximum of basin A
maxA=1
# minimum of basin B
minB=0.5
# maximum of basin B
maxB=1.5
# temperature in energy units
kbt=2.5

for((i=0;i<nfes;i++))
do
  # calculate free-energy of basin A
  A='awk 'BEGIN{tot=0.0}{if($1=="#!" && $1>min && $1<max)tot+=exp(-$2/kbt)}END{print -kbt*log(tot)}' min=${minA}
  # and basin B
  B='awk 'BEGIN{tot=0.0}{if($1=="#!" && $1>min && $1<max)tot+=exp(-$2/kbt)}END{print -kbt*log(tot)}' min=${minB}
  # calculate difference
  Delta=$(echo "${A} - ${B}" | bc -l)
  # print it
  echo $i $Delta
done
```

Please notice that `nfes` should be set to the number of profiles (free-energy estimates at different times of the simulation) generated by the option `--stride` of [sum\\_hills](#).

This analysis, along with the observation of the diffusive behavior in the CVs space, suggest that the simulation is converged.

#### 11.12.4.1.3 Exercise 4

In this exercise, we will run a well-tempered metadynamics simulation on alanine dipeptide in vacuum, using as CV the backbone dihedral angle  $\psi$ . In order to run this simulation we need to prepare the PLUMED input file (`plumed.dat`) as follows.

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate well-tempered metadynamics in psi depositing
# a Gaussian every 500 time steps, with initial height equal
# to 1.2 kJoule/mol, biasfactor equal to 10.0, and width to 0.35 rad

METAD ...
LABEL=metad
ARG=psi
PACE=500
HEIGHT=1.2
SIGMA=0.35
FILE=HILLS
BIASFACTOR=10.0
```

```

TEMP=300.0
GRID_MIN=-pi
GRID_MAX=pi
GRID_SPACING=0.1
... METAD

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR

```

(see [TORSION](#), [METAD](#), and [PRINT](#)).

Once the PLUMED input file is prepared, one has to run GROMACS with the option to activate PLUMED and read the input file. We will submit this job using the PBS queue system, as done for the previous exercise.

As we did in the previous exercise, we can use COLVAR to visualize the behavior of the CV during the simulation. Here we will plot at the same time the evolution of the metadynamics CV psi and of the other dihedral phi:

By inspecting Figure [cineca-metad-psi-phi-fig](#), we notice that something different happened compared to the previous exercise. At first the behavior of psi looks diffusive in the entire CV space. However, around  $t=1$  ns, psi seems trapped in a region of the CV space in which it was previously diffusing without problems. The reason is that the non-biased CV phi after a while has jumped into a different local minima. Since phi is not directly biased, one has to wait for this (slow) degree of freedom to equilibrate before the free energy along psi can converge. Try to repeat the analysis done in the previous exercise (calculate the estimate of the free energy as a function of time and monitor the free-energy difference between basins) to assess the convergence of this metadynamics simulation.

#### 11.12.4.2 Restraints

##### 11.12.4.2.1 Biased sampling theory

A system at temperature  $T$  samples conformations from the canonical ensemble:

$$P(q) \propto e^{-\frac{U(q)}{k_B T}}$$

. Here  $q$  are the microscopic coordinates and  $k_B$  is the Boltzmann constant. Since  $q$  is a highly dimensional vector, it is often convenient to analyze it in terms of a few collective variables (see [Belfast tutorial: Analyzing CVs](#), [Belfast tutorial: Adaptive variables I](#), and [Belfast tutorial: Adaptive variables II](#)). The probability distribution for a CV  $s$  is

$$P(s) \propto \int dq e^{-\frac{U(q)}{k_B T}} \delta(s - s(q))$$

This probability can be expressed in energy units as a free energy landscape  $F(s)$ :

$$F(s) = -k_B T \log P(s)$$

. Now we would like to modify the potential by adding a term that depends on the CV only. That is, instead of using  $U(q)$ , we use  $U(q) + V(s(q))$ . There are several reasons why one would like to introduce this potential. One is to avoid that the system samples some un-desired portion of the conformational space. As an example, imagine that you want to study dissociation of a complex of two molecules. If you perform a very long simulation you will be able to see association and dissociation. However, the typical time required for association will depend on the size of the simulation box. It could be thus convenient to limit the exploration to conformations where the distance between the two molecules is lower than a given threshold. This could be done by adding a bias potential on the distance between the two molecules. Another example is the simulation of a portion of a large molecule taken out from its initial context. The fragment alone could be unstable, and one might want to add additional potentials to keep the fragment in place. This could be done by adding a bias potential on some measure of the distance from the experimental structure (e.g. on root-mean-square deviation).

Whatever CV we decide to bias, it is very important to recognize which is the effect of this bias and, if necessary, remove it a posteriori. The biased distribution of the CV will be

$$P'(s) \propto \int dq e^{-\frac{U(q)+V(s(q))}{k_B T}} \delta(s - s(q)) \propto e^{-\frac{V(s(q))}{k_B T}} P(s)$$

and the biased free energy landscape

$$F'(s) = -k_B T \log P'(s) = F(s) + V(s) + C$$

Thus, the effect of a bias potential on the free energy is additive. Also notice the presence of an undetermined constant  $C$ . This constant is irrelevant for what concerns free-energy differences and barriers, but will be important later when we will learn the weighted-histogram method. Obviously the last equation can be inverted so as to obtain the original, unbiased free-energy landscape from the biased one just subtracting the bias potential

$$F(s) = F'(s) - V(s) + C$$

Additionally, one might be interested in recovering the distribution of an arbitrary observable. E.g., one could add a bias on the distance between two molecules and be willing to compute the unbiased distribution of some torsional angle. In this case there is no straightforward relationship that can be used, and one has to go back to the relationship between the microscopic probabilities:

$$P(q) \propto P'(q) e^{\frac{V(s(q))}{k_B T}}$$

The consequence of this expression is that one can obtain any kind of unbiased information from a biased simulation just by weighting every sampled conformation with a weight

$$w \propto e^{\frac{V(s(q))}{k_B T}}$$

That is, frames that have been explored in spite of a high (disfavoring) bias potential  $V$  will be counted more than frames that has been explored with a less disfavoring bias potential.

#### 11.12.4.2.2 Umbrella sampling theory

Often in interesting cases the free-energy landscape has several local minima. If these minima have free-energy differences that are on the order of a few times  $k_B T$  they might all be relevant. However, if they are separated by a high saddle point in the free-energy landscape (i.e. a low probability region) than the transition between one and the other will take a lot of time and these minima will correspond to metastable states. The transition between one minimum and the other could require a time scale which is out of reach for MD. In these situations, one could take inspiration from catalysis and try to favor in a controlled manner the conformations corresponding to the transition state.

Imagine that you know since the beginning the shape of the free-energy landscape  $F(s)$  as a function of one CV  $s$ . If you perform a MD simulation using a bias potential which is exactly equal to  $-F(s)$ , the biased free-energy landscape will be flat and barrierless. This potential acts as an "umbrella" that helps you to safely cross the transition state in spite of its high free energy.

It is however difficult to have an a priori guess of the free-energy landscape. We will see later how adaptive techniques such as metadynamics ([Belfast tutorial: Metadynamics](#)) can be used to this aim. Because of this reason, umbrella sampling is often used in a slightly different manner.

Imagine that you do not know the exact height of the free-energy barrier but you have an idea of where the barrier is located. You could try to just favor the sampling of the transition state by adding a harmonic restraint on the CV, e.g. in the form

$$V(s) = \frac{k}{2}(s - s_0)^2$$

. The sampled distribution will be

$$P'(q) \propto P(q) e^{-\frac{k(s(q)-s_0)^2}{2k_B T}}$$

For large values of  $k$ , only points close to  $s_0$  will be explored. It is thus clear how one can force the system to explore only a predefined region of the space adding such a restraint. By combining simulations performed with different values of  $s_0$ , one could obtain a continuous set of simulations going from one minimum to the other crossing the transition state. In the next section we will see how to combine the information from these simulations.

If you want to just bring a collective variables to a specific value, you can use a simple restraint. Let's imagine that we want to force the  $\Phi$  angle to visit a region close to  $\Phi = \pi/2$ . We can do it adding a restraint in  $\Phi$ , with the following input

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
res: RESTRAINT ARG=phi AT=0.5pi KAPPA=5
PRINT ARG=phi,psi,res.bias
```

(see [TORSION](#), [RESTRAINT](#), and [PRINT](#)).

Notice that here we are printing a quantity named `res.bias`. We do this because [RESTRAINT](#) does not define a single value (that here would be theoretically named `res`) but a structure with several components. All biasing methods (including [METAD](#)) do so, as well as many collective variables (see e.g. [DISTANCE](#) used with `COMPONENTS` keyword). Printing the bias allows one to know how much a given snapshot was penalized. Also notice that PLUMED understands numbers in the form `{number}pi`. This is convenient when using torsions, since they are expressed in radians.

Now you can plot your trajectory with `gnuplot` and see the effect of `KAPPA`. You can also try different values of `KAPPA`. The stiffer the restraint, the less the collective variable will fluctuate. However, notice that a too large `kappa` could make the MD integrator unstable.

#### 11.12.4.3 Using multiple replicas

##### Warning

Notice that multireplica simulations with PLUMED are fully supported with GROMACS, but only partly supported with other MD engines.

Some free-energy methods are intrinsically parallel and requires running several simultaneous simulations. This can be done with GROMACS using the multi replica framework. That is, if you have 4 `.tpr` files named `topol0.tpr`, `topol1.tpr`, `topol2.tpr`, `topol3.tpr` you can run 4 simultaneous simulations.

```
> mpirun -np 4 gmx_mpi mdrun -s topol.tpr -plumed plumed.dat -multi 4 -nsteps 500000
```

Each of the 4 replicas will open a different `topol` file, and GROMACS will take care of adding the replica number before the `.tpr` suffix. PLUMED deals with the extra number in a slightly different way. In this case, for example, PLUMED first look for a file named `plumed.X.dat`, where `X` is the number of the replica. In case the file is not found, then PLUMED looks for `plumed.dat`. If also this is not found, PLUMED will complain. As a consequence, if all the replicas should use the same input file it is sufficient to put a single `plumed.dat` file, but one has also the flexibility of using separate files named `plumed.0.dat`, `plumed.1.dat` etc.

Also notice that providing the flag `-replex` one can instruct GROMACS to perform a replica exchange simulation. Namely, from time to time GROMACS will try to swap coordinates among neighboring replicas and accept or reject the exchange with a Monte Carlo procedure which also takes into account the bias potentials acting on the replicas, even if different bias potentials are used in different replicas. That is, PLUMED allows to easily implement many forms of Hamiltonian replica exchange.

#### 11.12.4.4 Using multiple restraints with replica exchange

##### 11.12.4.4.1 Weighted histogram analysis method theory

Let's now consider multiple simulations performed with restraints located in different positions. In particular, we will consider the  $i$ -th bias potential as  $V_i$ . The probability to observe a given value of the collective variable  $s$  is:

$$P_i(s) = \frac{P(s)e^{-\frac{V_i(s)}{k_B T}}}{\int ds' P(s')e^{-\frac{V_i(s')}{k_B T}}} = \frac{P(s)e^{-\frac{V_i(s)}{k_B T}}}{Z_i}$$

where

$$Z_i = \sum_q e^{-(U(q)+V_i(q))}$$

The likelihood for the observation of a sequence of snapshots  $q_i(t)$  (where  $i$  is the index of the trajectory and  $t$  is time) is just the product of the probability of each of the snapshots. We use here the minus-logarithm of the likelihood (so that the product is converted to a sum) that can be written as

$$\mathcal{L} = - \sum_i \int dt \log P_i(s_i(t)) = \sum_i \int dt \left( -\log P(s_i(t)) + \frac{V_i(s_i(t))}{k_B T} + \log Z_i \right)$$

One can then maximize the likelihood by setting  $\frac{\delta \mathcal{L}}{\delta P(\mathbf{s})} = 0$ . After some boring algebra the following expression can be obtained

$$0 = \sum_i \int dt \left( -\frac{\delta_{\mathbf{s}_i(t), \mathbf{s}}}{P(\mathbf{s})} + \frac{e^{-\frac{V_i(\mathbf{s})}{k_B T}}}{Z_i} \right)$$

In this equation we aim at finding  $P(s)$ . However, also the list of normalization factors  $Z_i$  is unknown, and they should be found selfconsistently. Thus one can find the solution as

$$P(\mathbf{s}) \propto \frac{N(\mathbf{s})}{\sum_i \int dt \frac{e^{-\frac{V_i(\mathbf{s})}{k_B T}}}{Z_i}}$$

where  $Z$  is selfconsistently determined as

$$Z_i \propto \int ds' P(\mathbf{s}') e^{-\frac{V_i(\mathbf{s}')}{k_B T}}$$

These are the WHAM equations that are traditionally solved to derive the unbiased probability  $P(s)$  by the combination of multiple restrained simulations. To make a slightly more general implementation, one can compute the weights that should be assigned to each snapshot, that turn out to be:

$$w_i(t) \propto \frac{1}{\sum_j \int dt \frac{e^{-\beta V_j(\mathbf{s}_i(t))}}{Z_j}}$$

The normalization factors can in turn be found from the weights as

$$Z_i \propto \frac{\sum_j \int dt e^{-\beta V_i(\mathbf{s}_j(t))} w_j(t)}{\sum_j \int dt w_j(t)}$$

This allows to straightforwardly compute averages related to other, non-biased degrees of freedom, and it is thus a bit more flexible. It is sufficient to precompute this factors  $w$  and use them to weight every single frame in the trajectory.

## 11.12.4.4.2 Exercise 5

In this exercise we will run multiple restraint simulations and learn how to reweight and combine data with WHAM to obtain free-energy profiles. We start with running in a replica-exchange scheme 32 simulations with a restraint on  $\phi$  in different positions, ranging from -3 to 3. We will instruct GROMACS to attempt an exchange between different simulations every 1000 steps.

First we need to create the 32 PLUMED input files and .tpr files, using the following script:

```
nrep=32
dx='echo "6.0 / ( $nrep - 1 )" | bc -l`

for((i=0;i<nrep;i++))
do
# center of the restraint
AT='echo "$i * $dx - 3.0" | bc -l`

cat >plumed.${i}.dat << EOF
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Impose an umbrella potential on phi
# with a spring constant of 200 kJoule/mol
# and centered in phi=AT
#
restraint-phi: RESTRAINT ARG=phi KAPPA=200.0 AT=$AT
# monitor the two variables and the bias potential
PRINT STRIDE=100 ARG=phi,psi,restraint-phi.bias FILE=COLVAR
EOF

# we initialize some replicas in A and some in B:
if((i%2==0)); then
cp ../SETUP/topolA.tpr topol${i}.tpr
else
cp ../SETUP/topolB.tpr topol${i}.tpr
fi
done
```

And then run GROMACS with the following command:

```
mpirun -np 32 gmx_mpi mdrun -plumed plumed.dat -s topol.tpr -multi 32 -replex 1000 -nsteps 500000 -x traj.xtc
```

To be able to combine data from all the simulations, it is necessary to have an overlap between statistics collected in two adjacent umbrellas.

Have a look at the plots of ( $\phi$ , $\psi$ ) for the different simulations to understand what is happening. To visualize them all at once as in Fig. [cineca-usrem-phi-all](#), you can prepare the following script:

```
awk 'BEGIN{print ""; ORS=""; print "p \"COLVAR.0\" u 2:3 notitle,"; for(i=1;i<=30;i++) print "\"COLVAR.\"i\" u 2:3 notitle,"; print ""}
```

and then open gnuplot and load it:

```
gnuplot> load "plot.plt"
```

An often misunderstood fact about WHAM is that data of the different trajectories can be mixed and it is not necessary to keep track of which restraint was used to produce every single frame. Let's get the concatenated trajectory

```
trjcat_mpi -cat -f traj*.xtc -o alltraj.xtc
```

Now we should compute the value of each of the bias potentials on the entire (concatenated) trajectory.

```
nrep=32
dx='echo "6.0 / ( $nrep - 1 )" | bc -l`

for i in `seq 0 $(( $nrep - 1 ))`
do
# center of the restraint
AT='echo "$i * $dx - 3.0" | bc -l`

cat >plumed.dat << EOF
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
restraint-phi: RESTRAINT ARG=phi KAPPA=200.0 AT=$AT

# monitor the two variables and the bias potential
PRINT STRIDE=100 ARG=phi,psi,restraint-phi.bias FILE=ALLCOLVAR.$i
EOF

plumed driver --mf_xtc alltraj.xtc --trajectory-stride=10 --plumed plumed.dat

done
```

It is very important that this script is consistent with the one used to generate the multiple simulations above. Now, single files named ALLCOLVAR.XX will contain on the fourth column the value of the bias centered in a given position, computed on the entire concatenated trajectory.

Next step is to compute the weights self-consistently solving the WHAM equations, using the python script "wham.py" contained in the SCRIPTS directory. We will run this script as follows:

```
bash ../SCRIPTS/wham.sh ALLCOLVAR.*
```

This script will produce several files. Let's visualize "phi\_fes.dat", which contains the free energy as a function of phi, and compare this with the result previously obtained with metadynamics.

#### 11.12.4.4.3 Exercise 6

In the previous exercise, we use multiple restraint simulations to calculate the free energy as a function of the dihedral phi. The resulting free energy was in excellent agreement with our previous metadynamics simulation. In this exercise we will repeat the same procedure for the dihedral psi. At the end of the steps defined above, we can plot the free energy "psi\_fes.dat" and compare it with the profile calculated from a metadynamics simulations using both phi and psi as CVs.

We can easily spot from the plot above that something went wrong in this multiple restraint simulations, despite we used the very same approach we adopted for the phi dihedral. The problem here is that psi is a "bad" collective variable, and the system is not able to equilibrate the missing slow degree of freedom phi in the short time scale of the umbrella simulation (1 ns). In the metadynamics exercise in which we biased only psi, we detect problems by observing the behavior of the CV as a function of simulation time. How can we detect problems in multiple restraint simulations? This is slightly more complicated, but running this kind of simulation in a replica-exchange scheme offers a convenient way to detect problems.

The first thing we need to do is to demux the replica-exchange trajectories and reconstruct the continuous trajectories of the replicas across the different restraint potentials. In order to do so, we can use the following script:

```
demux.pl md0.log
trjcat_mpi -f traj?.xtc traj??.xtc -demux replica_index.xvg
```

This commands will generate 32 continuous trajectories, named XX\_trajout.xtc. We will use the driver to calculate the value of the CVs phi and psi on these trajectories.



```
nrep=32

for i in `seq 0 $(( $nrep - 1 ))`
do

cat >plumed.dat << EOF
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

# monitor the two variables
PRINT STRIDE=100 ARG=phi,psi FILE=COLVARDEMUX.$i
EOF

plumed driver --mf_xtc ${i}_trajout.xtc --trajectory-stride=10 --plumed plumed.dat

done
```

The COLVARDEMUX.XX files will contain the value of the CVs on the demuxed trajectory. If we visualize these files (see previous exercise for instructions) we will notice that replicas sample the CVs space differently. In order for each umbrella to equilibrate the slow degrees of freedom phi, the continuous replicas must be ergodic and thus sample the same distribution in phi and psi.

## 11.13 Using Hamiltonian replica exchange with GROMACS

When patching GROMACS with PLUMED, it is also possible to perform Hamiltonian replica exchange with different topologies. Although this feature is provided together with PLUMED, it is actually a new feature for GROMACS itself that can be enabled using the `-hrex` flag of `mdrun`. This implementation is very close to the one used to produce the data in this paper [69]. In case you find it useful, please cite this paper.

### Warning

This feature is currently used by several groups and should be robust enough. However, be sure that you understand its limitations and to perform all the tests discussed in this page before using it in production.

In this short tutorial you will learn how to do two things:

- Generate scaled topologies with `plumed partial_tempering`. This is actually not directly related to the `-hrex` flag.
- Run GROMACS with replica exchange and multiple topologies.

### 11.13.1 Generate scaled topologies

Plumed comes with a `partial_tempering` command line tool that can be used to generate scaled topologies. Notice that you might want to generate these topologies by yourself. This step is totally independent from the use of the `-hrex` flag.

The `partial_tempering` tool can be invoked as follows:

```
plumed partial_tempering scale < processed.top
```

where `scale` is the Hamiltonian scaling factor and `processed.top` is a post-processed topology file (i.e. produced with `grompp -pp`) where each "hot" atom has a "\_" appended to the atom type, e.g. change this

```
1          HC      1      ACE      HH31      1      0.1123      1.008      ; qtot 0.1123
```

to this:

```
1          HC_      1      ACE      HH31      1      0.1123      1.008      ; qtot 0.1123
```

Notice that the section that should be edited is the [atoms] section for all the molecules that you wish to affect (typically only for the solute, but you may also want to change solvent parameters). If you select all the atoms of the solute, you will be able to prepare topologies such as those needed in a REST2 simulation [70].

Also remember to first produce the processed.top file with grompp -pp. Editing a normal topol.top file will not work, because it does not contain all the parameters. The processed.top file should not have any "#include" statement.

```
# produce a processed topology
grompp -pp
# choose the "hot" atoms
vi processed.top
# generate the actual topology
plumed partial_tempering $scale < processed.top > topol$i.top
```

Warnings:

- It's not very robust and there might be force-field dependent issues!
- It certainly does not work with charmm cmap

Suggested tests:

1. Compare partial\_tempering with scale=1.0 to non-scaled force field. E.g.

```
grompp -o topol-unscaled.tpr
grompp -pp
vi processed.top # choose the "hot" atoms appending "_". You can choose whatever.
plumed partial_tempering 1.0 < processed.top > topol-scaled.top # scale with factor 1
grompp -p topol-scaled.top -o topol-scaled.tpr
# Then do a rerun on a trajectory
mdrun -s topol-unscaled.tpr -rerun rerun.trr
mdrun -s topol-scaled.tpr -rerun rerun.trr
# and compare the resulting energy files. they should be identical
```

2. Compare partial\_tempering with scale=0.5 to non-scaled force field. Repeat the same procedure but using "plumed partial\_tempering 0.5". Choose all the atoms in all the relevant [atoms] sections (e.g. solute, solvent and ions). In the two resulting energy files you should see: long range electrostatics, LJ, and dihedral energy is *half* in the scaled case all other terms (bonds/bends) are identical.

### 11.13.2 Run GROMACS

If GROMACS has been patched with PLUMED it should accept the -hrex option in mdrun. Please double check this (mdrun -h should list this possibility). Notice that not all the versions of GROMACS allow this feature. First of all prepare separate topologies for each replicas using plumed partial\_tempering tool as shown above (or using some other tool). Then run a normal replica exchange with gromacs adding the flag "-hrex" on the command line.

A complete run script could be adapted from the following

```

# five replicas
nrep=5
# "effective" temperature range
tmin=300
tmax=1000

# build geometric progression
list=$(
  awk -v n=$nrep \
    -v tmin=$tmin \
    -v tmax=$tmax \
    'BEGIN{for(i=0;i<n;i++){
      t=tmin*exp(i*log(tmax/tmin)/(n-1));
      printf(t); if(i<n-1)printf(",");
    }
  }'
)

# clean directory
rm -fr \#*
rm -fr topol*

for((i=0;i<nrep;i++))
do

# choose lambda as T[0]/T[i]
# remember that high temperature is equivalent to low lambda
lambda=$(echo $list | awk 'BEGIN{FS=",";}{print $1/$'${(i+1)}';}')
# process topology
# (if you are curious, try "diff topol0.top topol1.top" to see the changes)
plumed partial_tempering $lambda < processed.top > topol$i.top
# prepare tpr file
# -maxwarn is often needed because box could be charged
grompp_mpi_d -maxwarn 1 -o topol$i.tpr -f grompp$i.mdp -p topol$i.top
done
mpirun -np $nrep gmx_mpi mdrun_d -v -plumed plumed.dat -multi $nrep -replex 100 -nsteps 15000000 -hrex

```

Notice that total cell could be charged. This happens whenever the scaled portion of the system is not neutral. There should be no problem in this. When used with pbc, GROMACS will add a compensating background.

Suggested check:

- Try with several identical force fields (hardcode the same lambda for all replicas in the script above) and different seed/starting point. Acceptance should be 1.0

Notice that when you run with GPUs acceptance could be different from 1.0. The reason is that to compute the acceptance GROMACS is sending the coordinates to the neighboring replicas which then recompute energy. If all the tpr files are identical, one would expect energy to be identically to the originally computed one. However, calculations made with GPUs are typically not reproducible to machine precision. For a large system, even an error in the total energy for the last significant digit could be on the order of a kJ. This results in practice in a lower acceptance. The error induced on the final ensemble is expected to be very small.

Warnings:

- Topologies should have the same number of atoms, same masses and same constraint topology. Some of these differences (e.g. masses) are not explicitly checked and might lead to unnoticed errors in the final results.
- Choose neighbor list update (nstlist) that divides replex. Notice that running with GPUs GROMACS is going to change nstlist automatically, be sure that it still divides replex.
- Option -hrex requires also option -plumed. If you do not care about plumed, just provide an empty plumed.dat file.
- It should work correctly if replicas have different force-field, temperature, lambda, pressure, in any combination. However, not all these combinations have been tried in practice, so please first test the code on a system for which you know the result.

## 11.14 Julich tutorial: Developing CVs in plumed

### 11.14.1 Aims

The aim of this tutorial is to introduce users to the tutorials in the developer manual of PLUMED. We will learn a little bit about the structure of the code and how this structure can be visualised using the tree diagrams in the developer manual. You will then learn how to implement a new collective variable in a way that uses as much of the functionality that is already within the code and that thus avoids the duplication of code. Finally, you will learn how to write regression tests and how these can be used for continuous integration.

### 11.14.2 Learning Outcomes

By the end of this session you will know how to:

- access the developer manual for PLUMED.
- find the tutorial information for implementing new collective variables, multicolvars, functions, biases and analysis routines.
  - find the tree diagrams showing the class structure in the PLUMED developer manual.
- exploit the functionality within `PLMD::multicolvar` in order to write a reasonably complex collective variable quickly.
- write regression tests for PLUMED and understand how these are used for continuous integration.

To do this module you must understand the basics of object-oriented programming. Information on object oriented programming and how it is used within plumed can be found [here](#).

### 11.14.3 Resources

The `tarball` for this project contains the following directories:

- `rt-coord` : a directory containing input files for doing a regtest on an already existing collective variable.
- `rt-second-shell` : a directory containing input files for doing a regtest on the collective variable you will implement within this tutorial

At the start of the exercise you should move these two directories to the `plumed2/regtest/multicolvar` directory of your PLUMED source. You should see many other directories called `rt...` within the same directory. If you change into any of these directories and issue the following set of commands:

```
source ../../../../sourceme.sh
make
```

Then you will have run one of the regression tests of PLUMED. If the test runs without a hitch you should get an output something like this:

```
../../../../scripts/run
Thu Aug 20 14:33:00 IST 2015
Running regtest in /Users/gareth/Projects/CVception/plumed2/regtest/multicolvar/rt22
cp: ../tmp is a directory (not copied).
++ Test type: driver
++ Arguments: --plumed plumed.dat --trajectory-stride 10 --timestep 0.005 --ixyz trajectory.xyz --dump-forces
++ Processors: 0
/Users/gareth/Projects/CVception/plumed2/regtest/multicolvar/rt22/tmp
Run driver
Done. Here is the error file:
```

### 11.14.4 Introduction

PLUMED has two manuals: a manual that is for users of the code and a manual that is for developers of the code. If you are interested in implementing new features in the code your first point of call should thus be the developer manual, which can be found [here](#) . Alternatively, you can get to the front page of the developer manual by clicking the USER/DEVELOPER logo in the top right hand corner of any page of the user manual.

One of PLUMED's nicest features for the developer is that all the code and documentation for any new PLUMED command all appears together in a single file. When you come to implement a new feature it is thus relatively unlikely that you will have to modify any of the files you downloaded. Adding a new feature is simply a matter of adding one further cpp file containing the new method and the documentation for this method. We are able to achieve this by exploiting abstract base classes and polymorphism. All classes that calculate collective variables, print these variables or calculate biases thus inherit from a single base class called Action. You can read about the Action class [here](#) . Notice also that this page also shows how all the various classes within the code inherit from this single base class. It is perhaps worth spending a little while browsing through the various branches of this tree and understanding how the classes at each level become increasingly specialised and thus fit for particular purposes.

Lets say you want to implement a new collective variable in PLUMED. One way to start this task would be to write a new class that inherits from the `PLMD::Colvar` base class.

If you click on the box in the tree for `PLMD::Colvar` and follow various links on the various subject pages you will eventually get to the following [page](#) , which will give you a step-by-step set of instructions for implementing a new collective variable. If you look through the manual you can find similar pages that provide you with instructions for implementing new analysis methods, functions and so on. Our suggestion when you implement something new would thus be to find some similar functionality in the code, look at how it is implemented and to look in the developer manual at the descriptions of the classes that have been inherited. This process is what we will try and take you through in this short tutorial.

### 11.14.5 Instructions

#### 11.14.5.1 Calculating a reasonably complex collective variable

In this first exercise I would like you to look through the manual and work out how to use the functionality that is already available in PLUMED to calculate the following collective variable:

$$s = \sum_{i=1}^{108} 1 - \frac{1 - \left(\frac{c_i}{4}\right)^6}{1 - \left(\frac{c_i}{4}\right)^{12}}$$

where  $c_i$  is equal to the coordination number of atom  $i$ , i.e.:

$$c_i = \sum_{j \neq i} \frac{1 - \left(\frac{r_{ij}}{1}\right)^6}{1 - \left(\frac{r_{ij}}{1}\right)^{12}}$$

So  $r_{ij}$  is the distance between atom  $i$  and atom  $j$ . This collective variable measures the number of coordination numbers that are more than 4 so the summations in the above expressions run over all 108 atoms in this particular system.

Write an input file that computes  $s$  and outputs its value to a file called colvar and that computes both the analytical and numerical derivatives of  $s$  and outputs this information to a file called deriv. You are going to run this calculation in the rt-coord directory that you downloaded with the tarball for this exercise so you should create this input file within that directory. You will need to output the quantities in the colvar and deriv files with only four decimal places. When you are content with your input run the calculation by typing make. If you have done everything correctly you should see the output that was discussed above.

In setting up your input files you may find it useful to watch our [introductory](#) video and [this video](#) on some of the features that are available in the code. Obviously, the user manual will be indispensable as well.

## 11.14.5.2 Implementing a new collective variable

In this second exercise I would like you to implement the following collective variable:

$$s = \sum_{i=1}^{108} \int_{57}^{59} \exp\left(-\frac{(x_i - x)^2}{2}\right) dx$$

where  $x_i$  is number of atoms in the second coordination sphere of the  $i$ th atom, i.e.:

$$s = \sum_{i \neq j} \int_2^4 \exp\left(-\frac{(r_{ij} - r)^2}{2}\right) dr$$

So  $r_{ij}$  is the distance between atom  $i$  and atom  $j$ . This collective variable measure the number of atoms that have between 57 and 59 atoms in their second coordination sphere. There are a number of observations we can make here that will help you enormously:

- Notice that CV is similar to the coordination number CV you calculated in the first example. For both CVs there is a sum over a quantity that is calculated for different sets of atoms. You should thus look at what was done by the routines that calculate the first CV and see what you can reuse.
- Notice that with the first CV we can calculate the number of coordination numbers that are within a certain range as well as the number of coordination numbers that are less than a certain threshold.
- You can set the link cell cutoff equal to

```
std::numeric_limits<double>::max()
```

- Lastly, [this tool](#) will prove very useful.

Write an input file that computes  $s$  and outputs its value to a file called colvar and that computes both the analytical and numerical derivatives of  $s$  and outputs this information to a file called deriv. When outputting numbers to colvar and deriv you should output numbers to four and three decimal places respectively. You are going to run this calculation in the rt-second-shell directory that you downloaded with the tarball for this exercise so you should create this input file within that directory. When you are content with your input run the calculation by typing make. If you have done everything correctly you should see the output that was discussed above.

If you have sufficient time try use [doxygen within PLUMED](#) to write the documentation for your new collective variable.

### 11.14.6 Final thoughts

The aim of this tutorial is not so much to implement the CV. I do not think it is a particularly interesting or useful CV. Instead the hope is that by doing it you will get some idea of how to implement things within PLUMED. Based on my experience there are three key things I wish I could go back and tell my younger self about programming, which I would urge you to learn now:

- **Write less code** The more code you write the more bugs you generate. To be totally clear this is not me trying to say I am better at coding than you are - the same statement holds true if I replace each you in the above with an I - it holds true for everyone. In addition, there are lots of smart people out in the world writing code (and again I am not talking about the developers of plumed). Their code is properly tested and faster than anything that you will write so spend your time learning how it works so that you can re-use it.
- **Test your code** Notice that I set up directories that you could use to test your code for this tutorial. Testing should always be part of your development workflow and coming up with ways to test your code is hard, which is why we often don't do it enough. Also don't develop stuff in a vacuum get in touch with people who are using the code. You need to test your code on systems that people actually want to simulate and not just on dummy problems. In addition, you should learn to use profiling tools such as valgrind and instruments on the mac so that you can find memory leaks and bottlenecks in your code. When you start doing this properly you will realise that you cannot possibly properly maintain all the code that you have written and you will see clearly that you that you need to write less code.
- **Write documentation for your code** Undocumented software is useless. You need to explain how to use your code and to fill your manual with examples of how the code can be used to solve specific and interesting problems. That is to say that you need to provide examples of calculations that users actually want to perform. To be clear here writing a manual that tells users how something like the shake algorithm works in general terms is actually pretty useless. After all a user can go off and find out this information from a textbook, you could thus replace your description with a link to a textbook. What users want from you is an example that is similar to the calculation they actually want to set up. If I were providing documentation for an implementation of shake in an MD code I would thus explain in detail how to set up shake to constrain all the angles and bonds in all the water molecules in my system as it is a fair bet that this is how it will be being used by many users.

I hope you have seen from this tutorial that PLUMED tries to help you with all of these aims. We have a developer manual where we try to explain how to use the code that is already there. In addition, you can quite quickly generate nicely formatted user documentation for new CVs and it is easy to add new regression tests that allow you to test new features. Notice also that whenever commits are made on the origin git repository we use a website called [travis-ci](#) to test that the code works across a range of platforms.

## 11.15 Moving from PLUMED 1 to PLUMED 2

Syntax in PLUMED 2 has been completely redesigned based on our experience using PLUMED 1, hopefully making it clearer, more flexible, and less error prone. The main difference is that whereas in PLUMED 1 lines could be inserted in any order, in PLUMED 2 the order of the lines matters. This is due to a major change in the internal architecture of PLUMED. In version 2, commands (or "actions") are executed in the order they are found in the input file. Because of this, you must e.g. first compute a collective variable and then print it later. More information can be found in the Section about [Getting started](#).

Other important changes are in the way groups and units are used, as discussed below. Finally, many features appear under a different name in the new version.

### 11.15.1 New syntax

We know that changing the input syntax requires a lot of work from the user side to update their input files. However, we believe that the new syntax is easier to read and that it allows for more flexibility. As an example, something that in PLUMED 1.3 was:

```
HILLS HEIGHT 0.4 W_STRIDE 600
WELLTEMPERED SIMTEMP 300 BIASFACTOR 15
RGYR LIST <all>
all->
1 5 6 7 9 11 15 16 17 19
all<-
TORSION LIST 5 7 9 15 SIGMA 0.1
TORSION LIST 7 9 15 17
PRINT W_STRIDE 100
```

in PLUMED 2.x becomes:

```
all:  GROUP      ATOMS=1,5,6,7,9,11,15,16,17,19
rg:   GYRATION   ATOMS=all
t1:   TORSION    ATOMS=5,7,9,15
METAD ...
  LABEL=meta
  ARG=t1 SIGMA=0.1 HEIGHT=0.4 PACE=600
  BIASFACTOR=15 TEMP=300
... METAD
t2:   TORSION    ATOMS=7,9,15,17
PRINT ARG=t1,t2,meta.bias STRIDE=100 FILE=COLVAR
```

Giving all quantities an explicit name makes the input easier to interpret. Additionally, all the parameters related to **METAD** are placed on a single line (actually, this is done here exploiting continuation lines). Also notice that one can customize the name of the COLVAR file. By specifying to **PRINT** which collective variables should be printed, one can easily decide what to print exactly and using which stride. By repeating the **PRINT** line one can also monitor very expensive variables with a larger stride, just putting the result on a separate file.

You might have noticed that ideas that were very difficult to implement in PLUMED 1.3 now become immediately available. As an example, one can now apply concurrently several **METAD** potentials [47] .

### 11.15.2 Groups

In PLUMED 1 groups (lists) were used for two tasks:

- To provide centers of masses to collective variables such as distances, angles, etc. This is now done by defining virtual atoms using either **CENTER** or **COM**
- To provide lists of atoms to collective variables such as coordination, gyration radius, etc. This is now done directly in the line that defines the collective variable.

If you would still like to use groups you can use the **GROUP** commands. Whenever the label for a **GROUP** action appears in the input it is replaced by the list of atoms that were specified in the **GROUP**.

A restraint on the distance between centers of mass in PLUMED 1 was something like:

```
DISTANCE LIST <g1> <g2>
g1->
17 20 22 30
g1<-
g2->
LOOP 37 40
g2<-
UMBRELLA CV 1 KAPPA 200 AT 1.0
PRINT W_STRIDE 100
```



The same in PLUMED 2.x reads:

```
g1: COM ATOMS=17,20,22,30
g2: COM ATOMS=37-40
d:  DISTANCE ATOMS=g1,g2
r:  RESTRAINT ARG=d KAPPA=200 AT=1.0
PRINT STRIDE=100 FILE=COLVAR ARG=d,r.*
```

Notice that virtual atoms are very powerful tools in PLUMED 2. Actually, they can be used in any collective variable where normal atoms can be used, just by calling them by name. This allows to straightforwardly define variables such as coordination between centers of mass, which would have required an ad hoc implementation in PLUMED 1.

In the example above you can also appreciate the advantage of calling collective variables by name. It is obvious here that **RESTRAINT** is acting on distance *d*, whereas in PLUMED 1 one had to keep track of the number of the collective variables. This was easy for a single collective variable, but could become cumbersome for complex input files.

### 11.15.3 Names in output files

Another advantage of having names is that when PLUMED produces an output file it can insert explicit names in the file. Consider for example this PLUMED 1 input

```
DISTANCE LIST 1 2
ANGLE     LIST 3 4 5
PRINT W_STRIDE 100
```

The first line of the COLVAR file was then

```
#!  FIELDS time cv1 cv2
```

The equivalent input file in PLUMED 2 is

```
d: DISTANCE ATOMS=1,2
a: ANGLE     ATOMS=3,4,5
PRINT ARG=d,a FILE=COLVAR STRIDE=100
```

The first line of the COLVAR file now is

```
#!  FIELDS time d a
```

This makes it easy to remember what's the meaning of each column of the COLVAR file without the need to always go back to the PLUMED input to check in which order the variables were declared.

### 11.15.4 Units

In PLUMED 1 the input file of PLUMED was expected to be written using the same units of the MD code. This choice was made to allow users to adopt the same units they were used to. However, we realized later that this choice was not allowing the PLUMED input files to be ported between different MD code. Let's say that one was using this keyword with GROMACS (kJ/mol - nm)

```
UMBRELLA CV 1 KAPPA 200 AT 1.0
```

Let's assume the variable CV 1 was a distance. The same keyword in NAMD (kcal/mol - Å) should have been converted to

```
UMBRELLA CV 1 KAPPA .4780 AT 10.0
```

The conversion of AT is straightforward (1nm=10Å), but the conversion on KAPPA is more error prone. This is because KAPPA is measured in units of energy divided by distance squared. Notice that a different factor should have been used if the CV was an angle.

Learning from this, we designed PLUMED 2 in a way that units in the PLUMED input are independent of the MD code. Technically, this is achieved by doing the conversion when coordinates and forces are passed back and forth. In this way, the same PLUMED input could be used with GROMACS and NAMD.

We decided to use as standard units in PLUMED kJ/mol, nm, and ps. Perhaps this is because most of the developers are using GROMACS, which also adopts these units. However, we still allow the personalization of units by means of the [UNITS](#) keyword. Since this keyword is included directly in the PLUMED input file, even when using personalized units the input files remains perfectly portable across different MD engines.

### 11.15.5 Directives

What follows is a list of all the documented directives of PLUMED 1 together with their plumed 2 equivalents. Be aware that the input syntaxes for these directives are not totally equivalent. You should read the documentation for the PLUMED 2 Action.

HILLS	<a href="#">METAD</a>
WELLTEMPERED	<a href="#">METAD</a> with BIASFACTOR
GRID	<a href="#">METAD</a> with GRID_MIN, GRID_MAX, and GRID_BIN
WRITE_GRID	<a href="#">METAD</a> with GRID_WFILE, GRID_WSTRIDE
READ_GRID	<a href="#">METAD</a> with GRID_RFILE
MULTIPLE_WALKERS	<a href="#">METAD</a> with options WALKERS_ID, WALKERS_N, WALKERS_DIR, and WALKERS_RSTRIDE
NOHILLS	not needed (collective variables are not biased by default)
INTERVAL	<a href="#">METAD</a> with INTERVAL
INVERT	<b>currently missing</b>
PTMETAD	not needed (replica exchange detected from MD engine)
BIASXMD	not needed (replica exchange detected from MD engine); one should anyway use <a href="#">RANDOM_EXCHANGES</a> to get the normal behavior
UMBRELLA	<a href="#">RESTRAINT</a>
STEER	<a href="#">MOVINGRESTRAINT</a>
STEERPLAN	<a href="#">MOVINGRESTRAINT</a>
ABMD	<a href="#">ABMD</a>
UWALL	<a href="#">UPPER_WALLS</a>
LWALL	<a href="#">LOWER_WALLS</a>
EXTERNAL	<a href="#">EXTERNAL</a>

COMMITMENT	<a href="#">COMMITTOR</a>
PROJ_GRAD	<a href="#">DUMPPROJECTIONS</a>
DAFED	a similar method using a Langevin thermostat, with <a href="#">EXTENDED_LAGRANGIAN</a>
DISTANCE	<a href="#">DISTANCE</a> - POINT_FROM_AXIS and PROJ_ON_AXIS can be reproduced with <a href="#">DISTANCE</a> and <a href="#">MATHEVAL</a>
POSITION	<a href="#">POSITION</a>
MINDIST	<a href="#">DISTANCES</a> with keyword MIN
ANGLE	<a href="#">ANGLE</a>
TORSION	<a href="#">TORSION</a>
COORD	<a href="#">COORDINATION</a>
HBOND	<b>currently missing</b> , can be emulated with <a href="#">COORDINATION</a>
WATERBRIDGE	<a href="#">BRIDGE</a>
RGYR	<a href="#">GYRATION</a>
DIPOLE	<a href="#">DIPOLE</a>
DIHCOR	<a href="#">DIHCOR</a>
ALPHABETA	<a href="#">ALPHABETA</a>
ALPHARMSD	<a href="#">ALPHARMSD</a>
ANTIBETARMSD	<a href="#">ANTIBETARMSD</a>
PARABETARMSD	<a href="#">PARABETARMSD</a>
ELSTPOT	currently missing, but a related quantity can be obtained with <a href="#">DHENERGY</a>
PUCKERING	<a href="#">PUCKERING</a>
S_PATH	<a href="#">PATHMSD</a> , s component
Z_PATH	<a href="#">PATHMSD</a> , z component
TARGETED	<a href="#">RMSD</a>
ENERGY	<a href="#">ENERGY</a>
HELIX	<b>currently missing</b>
PCA	<a href="#">PCARMSD</a>
SPRINT	<a href="#">SPRINT</a>
RDF	<a href="#">DISTANCES</a> , used in combination with HISTOGRAM / BETWEEN keyword
ADF	<a href="#">ANGLES</a> , used in combination with HISTOGRAM / BETWEEN keyword
POLY	<a href="#">COMBINE</a>
FUNCTION	<a href="#">MATHEVAL</a>
ALIGN_ATOMS	<a href="#">WHOLEMOLECULES</a>

## 11.16 Munster tutorial

### Authors

Max Bonomi and Giovanni Bussi, stealing a lot of material from other tutorials. Richard Cunha is acknowledged for beta-testing this tutorial.

### Date

March 11, 2015

This document describes the PLUMED tutorial held in Munster, March 2015. The aim of this tutorial is to learn how to use PLUMED to analyze molecular dynamics simulations on the fly, to analyze existing trajectories, and to perform enhanced sampling. Although the presented input files are correct, the users are invited to **refer to the literature to understand how the parameters of enhanced sampling methods should be chosen in a real application.**

Users are also encouraged to follow the links to the full PLUMED reference documentation and to wander around in the manual to discover the many available features and to do the other, more complete, tutorials. Here we are going to present only a very narrow selection of methods.

We here use PLUMED 2.1 syntax and we explicitly note if some syntax is expected to change in PLUMED 2.2, which will be released later in 2015. All the tests here are performed on a toy system, alanine dipeptide, simulated using the AMBER99SB force field. We provide both a setup that includes explicit water, which is more realistic but slower, and a setup in gas phase, which is much faster. Simulations are made using GROMACS 4.6.7, which is here assumed to be already patched with PLUMED and properly installed. However, these examples could be easily converted to other MD software.

All the gromacs input files and analys scripts are provided in this [tarball](#) .

Users are expected to write PLUMED input files based on the instructions below.

### 11.16.1 Alanine dipeptide: our toy model

In this tutorial we will play with alanine dipeptide (see Fig. [munster-1-ala-fig](#)). This rather simple molecule is useful to make benchmark that are around for data analysis and free energy methods. It is a nice example since it presents two metastable states separated by a high (free) energy barrier. Here metastable states are intended as states which have a relatively low free energy compared to adjacent conformations. It is conventional use to show the two states in terms of Ramachandran dihedral angles, which are denoted with  $\Phi$  and  $\Psi$  in Fig. [munster-1-transition-fig](#) .

### 11.16.2 Monitoring collective variables

The main goal of PLUMED is to compute collective variables, which are complex descriptors than can be used to analyze a conformational change or a chemical reaction. This can be done either on the fly, that is during molecular dynamics, or a posteriori, using PLUMED as a post-processing tool. In both cases one should create an input file with a specific PLUMED syntax. A sample input file is below:

```
# compute distance between atoms 1 and 10
d: DISTANCE ATOMS=1,10
# create a virtual atom in the center between atoms 20 and 30
center: CENTER ATOMS=20,30
# compute torsional angle between atoms 1,10,20 and center
phi: TORSION ATOMS=1,10,20,center
# compute some function of previously computed variables
d2: MATHEVAL ARG=phi FUNC=cos(x) PERIODIC=NO
# print both of them every 10 step
PRINT ARG=d,phi,d2 STRIDE=10
```

(see [DISTANCE](#), [CENTER](#), [TORSION](#), [MATHEVAL](#), and [PRINT](#))

PLUMED works using kJ/nm/ps as energy/length/time units. This can be personalized using [UNITS](#). Notice that variables should be given a name (in the example above, `d`, `phi`, and `d2`), which is then used to refer to these variables. Lists of atoms should be provided as comma separated numbers, with no space. Virtual atoms can be created and assigned a name for later use. You can find more information on the PLUMED syntax at [Getting started](#) page of the manual. The complete documentation for all the supported collective variables can be found at the [Collective variables](#) page.

## 11.16.2.1 Analyze on the fly

Here we will run a plain MD on alanine dipeptide and compute two torsional angles on the fly. GROMACS needs a .tpr file, which is a binary file containing initial positions as well as force-field parameters. We also provide .gro, .mdp, and .top files, that can be modified and used to generate a new .tpr file. For this tutorial, it is sufficient to use the provided .tpr files. You will find several tpr files, namely:

- topolAwat.tpr - setup in water, initialized in state A
- topolBwat.tpr - setup in water, initialized in state B
- topolA.tpr - setup in vacuum, initialized in state A
- topolB.tpr - setup in vacuum, initialized in state B

Gromacs md can be run using on the command line:

```
> gmx_mpi mdrun -s topolA.tpr -nsteps 10000
```

The nsteps flags can be used to change the number of timesteps and topolA.tpr is the name of the tpr file. While running, gromacs will produce an md.log file, with log information, and a traj.xtc file, with a binary trajectory. The trajectory can be visualized with VMD using a command such as

```
> vmd confout.gro tra.xtc
```

To run a simulation with gromacs+plumed you just need to add a -plumed flag

```
> gmx_mpi mdrun -s topolA.tpr -nsteps 10000 -plumed plumed.dat
```

Here plumed.dat is the name of the plumed input file. Notice that PLUMED will write information in the md.log that could be useful to verify if the simulation has been set up properly.

## 11.16.2.1.1 Exercise 0

In this exercise, we will run a plain molecular dynamics simulation and monitor the  $\Phi$  and  $\Psi$  dihedral angles on the fly. Using the following PLUMED input file you can monitor  $\Phi$  and  $\Psi$  angles during the MD simulation

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
PRINT ARG=phi,psi STRIDE=100 FILE=colvar
```

(see [TORSION](#) and [PRINT](#))

Notice that PLUMED is going to compute the collective variables only when necessary, that is, in this case, every 100 steps. This is not very relevant for simple variables such as torsional angles, but provides a significant speedup when using expensive collective variables.

PLUMED will write a textual file named colvar containing three columns: physical time,  $\Phi$  and  $\Psi$ . Results can be plotted using gnuplot:

```
> gnuplot
# this shows phi as a function of time
gnuplot> plot "colvar" u 2
# this shows psi as a function of time
gnuplot> plot "colvar" u 3
# this shows psi as a function of phi
gnuplot> plot "colvar" u 2:3
```

Now try to do the same using the two different initial configurations that we provided (topolA.tpr and topolB.tpr). You can try both setup (water and vacuum). Results from 200ps (100000 steps) trajectories in vacuum are shown in Figure [munster-ala-traj](#).

Notice that the result depends heavily on the starting structure. For the simulation in vacuum, the two free-energy minima are separated by a large barrier and, in such a short simulation, the system cannot cross it. In water the barrier is smaller and you might see some crossing. Also notice that the two clouds are well separated, indicating that these two collective variables are good enough to properly distinguish among the two minima.

As a final comment, notice that if you run twice the same calculation in the same directory, you might overwrite the resulting files. GROMACS takes automatic backup of the output files, and PLUMED does it as well. In case you are restarting a simulation, you can add the keyword [RESTART](#) at the beginning of the PLUMED input file. This will tell PLUMED to *append* files instead of taking a backup copy.

### 11.16.2.2 Analyze using the driver

Imagine you already made a simulation, with or without PLUMED. You might want to compute the collective variables a posteriori, from the trajectory file. You can do this by using the plumed executable on the command line. Type

```
> plumed driver --help
```

to have an idea of the possible options. See [driver](#) for the full documentation.

Here we will use the driver to compute  $\Phi$  and  $\Psi$  on the already generated trajectory. Let's assume the trajectory is named `traj.xtc`. You should prepare an PLUMED input file named `analysis.dat` as:

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
PRINT ARG=phi,psi FILE=analysis
```

(see [TORSION](#) and [PRINT](#)) Notice that typically when using the driver we do not provide a STRIDE keyword to PRINT. This implies "print at every step" which, analyzing a trajectory, means "print for all the available snapshots". Then, you can use the following command:

```
> plumed driver --mf_xtc traj.xtc --plumed analysis.dat
```

Notice that PLUMED has no way to know the value of physical time from the trajectory. If you want physical time to be printed in the `analysis` file you should give more information to the driver, e.g.:

```
> plumed driver --mf_xtc traj.xtc --plumed analysis.dat --timestep 0.002 --trajectory-stride 1000
```

(see [driver](#))

In this case we inform the driver that the `traj.xtc` file was produced in a run with a timestep of 0.002 ps and saving a snapshot every 1000 timesteps.

You might want to analyze a different collective variable, such as the gyration radius. The gyration radius tells how extended is the molecules in space. You can do it with the following plumed input file

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

heavy: GROUP ATOMS=1,5,6,7,9,11,15,16,17,19
gyr: GYRATION ATOMS=heavy

# the same could have been achieved with
# gyr: GYRATION ATOMS=1,5,6,7,9,11,15,16,17,19

PRINT ARG=phi,psi,gyr FILE=analyze
```

(see [TORSION](#), [GYRATION](#), [GROUP](#), and [PRINT](#))

Now try to compute the time series of the gyration radius.

## 11.16.2.3 Periodic boundaries and explicit water

In case you are running the simulation in water, you might see that at some point this variable shows some crazy jump. The reason is that the trajectory contains coordinates where molecules are broken across periodic-boundary conditions. This happens with GROMACS and some other MD code. These codes typically have tools to process trajectories and restore whole molecules. This trick is ok for the a-posteriori analysis we are trying now, but cannot be used when one needs to compute a collective variable on-the-fly or, as we will see later, one wants to add a bias to that collective variable. For this reason, we implemented a workaround in PLUMED, that is the molecule should be made whole using the [WHOLEMOLECULES](#) command. What this command is doing is making a loop over all the atoms (in the order they are provided) and set the coordinates of each of them in the periodic image which is as close as possible to the coordinates of the preceeding atom. In most cases it is sufficient to list all atoms of a molecule in order. Look in the [WHOLEMOLECULES](#) page to get more information. Here this will be enough:

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

# notice the 1-22 syntax, a shortcut for a list 1,2,3,...,22
WHOLEMOLECULES ENTITY0=1-22

heavy: GROUP ATOMS=1,5,6,7,9,11,15,16,17,19
gyr: GYRATION ATOMS=heavy

PRINT ARG=phi,psi,gyr FILE=analyze
```

(see [TORSION](#), [WHOLEMOLECULES](#), [GROUP](#), [GYRATION](#), and [PRINT](#))

This is a very important issue that should be kept in mind when using PLUMED. Notice that starting with version 2.2 PLUMED will make molecules used in [GYRATION](#) (as well as in other variables) whole automatically, so that this extra command will not be necessary.

Notice that you can instruct PLUMED to dump on a file not only the collective variables (as we are doing with [PRINT](#)) but also the atomic positions. This is a very good way to understand what [WHOLEMOLECULES](#) is actually doing. Try the following input

```
MOLINFO STRUCTURE=../TOPO/reference.pdb
DUMPATOMS FILE=test1.gro ATOMS=1-22
WHOLEMOLECULES ENTITY0=1-22
DUMPATOMS FILE=test2.gro ATOMS=1-22
```

(see [MOLINFO](#), [DUMPATOMS](#), and [WHOLEMOLECULES](#)).

[DUMPATOMS](#) writes on a gro file the coordinates of the alanine dipeptide atoms. Here PLUMED will produce two files, one with coordinates *before* the application of [WHOLEMOLECULES](#), and one with coordinates *after* the application of [WHOLEMOLECULES](#). You can load both trajectories in VMD to see the difference. The [MOLINFO](#) command is here used to provide atom names to PLUMED so that the resulting gro file looks nicer in VMD.

Notice that PLUMED has several commands that manipulate atomic coordinates. One example is [WHOLEMOLECULES](#), that fixes problems with periodic boundary conditions. [COM](#) and [CENTER](#) add new atoms, and [FIT\\_TO\\_TEMPLATE](#) can actually move atoms from their original position to align them on a template. [DUMPATOMS](#) is this very useful to check what these commands are doing and for using the PLUMED [driver](#) to manipulate MD trajectories.

### 11.16.2.4 Other analysis tools

PLUMED also allows you to make some analysis on the collective variables you are calculating. For example, you can compute a histogram with an input like this one

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
heavy: GROUP ATOMS=1,5,6,7,9,11,15,16,17,19
gyr: GYRATION ATOMS=heavy
PRINT ARG=phi,psi,gyr FILE=analyze
HISTOGRAM ...
  ARG=gyr
  USE_ALL_DATA
  KERNEL=discrete
  GRID_MIN=0
  GRID_MAX=1
  GRID_BIN=50
  GRID_WFILE=histogram
... HISTOGRAM
```

(see [TORSION](#), [WHOLEMOLECULES](#), [GROUP](#), [GYRATION](#), [PRINT](#), and [HISTOGRAM](#))

An histogram with 50 bins will be performed on the gyration radius. Try to compute the histogram for the  $\Phi$  and  $\Psi$  angles.

PLUMED can do much more than a histogram, more information on analysis can be found at the page [Analysis](#)

Notice that the plumed driver can also be used directly from VMD taking advantage of the PLUMED collective variable tool developed by Toni Giorgino (<http://multiscalelab.org/utilities/PlumedGUI>). Just open a recent version of VMD and go to Extensions/Analysis/Collective Variable Analysis (PLUMED). This graphical interface can also be used to quickly build PLUMED input files based on template lines.

## 11.16.3 Biasing collective variables

We have seen that PLUMED can be used to compute collective variables. However, PLUMED is most often used to add forces on the collective variables. To this aim, we have implemented a variety of possible biases acting on collective variables. A bias works in a manner conceptually similar to the [PRINT](#) command, taking as argument one or more collective variables. However, here the STRIDE is usually omitted (that is equivalent to setting it to 1), which means that forces are applied at every timestep. In PLUMED 2.2 you will be able to change the STRIDE also for bias potentials, but that's another story. In the following we will see how to apply harmonic restraints and how to build an adaptive bias potential with metadynamics. The complete documentation for all the biasing methods available in PLUMED can be found at the [Bias](#) page.

### 11.16.3.1 Metadynamics

#### 11.16.3.1.1 Summary of theory

In metadynamics, an external history-dependent bias potential is constructed in the space of a few selected degrees of freedom  $\vec{s}(q)$ , generally called collective variables (CVs) [41]. This potential is built as a sum of Gaussians deposited along the trajectory in the CVs space:

$$V(\vec{s}, t) = \sum_{k\tau < t} W(k\tau) \exp \left( - \sum_{i=1}^d \frac{(s_i - s_i(q(k\tau)))^2}{2\sigma_i^2} \right).$$



where  $\tau$  is the Gaussian deposition stride,  $\sigma_i$  the width of the Gaussian for the  $i$ th CV, and  $W(k\tau)$  the height of the Gaussian. The effect of the metadynamics bias potential is to push the system away from local minima into visiting new regions of the phase space. Furthermore, in the long time limit, the bias potential converges to minus the free energy as a function of the CVs:

$$V(\vec{s}, t \rightarrow \infty) = -F(\vec{s}) + C.$$

In standard metadynamics, Gaussians of constant height are added for the entire course of a simulation. As a result, the system is eventually pushed to explore high free-energy regions and the estimate of the free energy calculated from the bias potential oscillates around the real value. In well-tempered metadynamics [43], the height of the Gaussian is decreased with simulation time according to:

$$W(k\tau) = W_0 \exp\left(-\frac{V(\vec{s}(q(k\tau)), k\tau)}{k_B \Delta T}\right),$$

where  $W_0$  is an initial Gaussian height,  $\Delta T$  an input parameter with the dimension of a temperature, and  $k_B$  the Boltzmann constant. With this rescaling of the Gaussian height, the bias potential smoothly converges in the long time limit, but it does not fully compensate the underlying free energy:

$$V(\vec{s}, t \rightarrow \infty) = -\frac{\Delta T}{T + \Delta T} F(\vec{s}) + C.$$

where  $T$  is the temperature of the system. In the long time limit, the CVs thus sample an ensemble at a temperature  $T + \Delta T$  which is higher than the system temperature  $T$ . The parameter  $\Delta T$  can be chosen to regulate the extent of free-energy exploration:  $\Delta T = 0$  corresponds to standard molecular dynamics,  $\Delta T \rightarrow \infty$  to standard metadynamics. In well-tempered metadynamics literature and in PLUMED, you will often encounter the term "biasfactor" which is the ratio between the temperature of the CVs ( $T + \Delta T$ ) and the system temperature ( $T$ ):

$$\gamma = \frac{T + \Delta T}{T}.$$

The biasfactor should thus be carefully chosen in order for the relevant free-energy barriers to be crossed efficiently in the time scale of the simulation.

Additional information can be found in the several review papers on metadynamics [59] [60] [61].

If you do not know exactly where you would like your collective variables to go, and just know (or suspect) that some variables have large free-energy barriers that hinder some conformational rearrangement or some chemical reaction, you can bias them using metadynamics. In this way, a time dependent, adaptive potential will be constructed that tends to disfavor visited configurations in the collective-variable space. The bias is usually built as a sum of Gaussian deposited in the already visited states.

## 11.16.3.1.2 Exercise 1

Now run a metadynamics simulation with the following input

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
METAD ARG=phi,psi HEIGHT=1.0 BIASFACTOR=10 SIGMA=0.35,0.35 PACE=100 GRID_MIN=-pi,-pi GRID_MAX=pi,pi
```

(see [TORSION](#) and [METAD](#)) Thus, a single METAD line will contain all the metadynamics related options, such as Gaussian height (HEIGHT, here in kJ/mol), stride (PACE, here in number of time steps), bias factor (BIASFACTOR, here indicates that we are going to effectively boost the temperature of the collective variables by a factor 10), and width (SIGMA, an array with same size as the number of collective variables).

There are two additional keywords that are optional, namely GRID\_MIN and GRID\_MAX. These keywords sets the range of the collective variables and tell PLUMED to keep the bias potential stored on a grid. This affects speed but, in principle, not the accuracy of the calculation. You can try to remove those keywords and see the difference.

Now, run a metadynamics simulations and check the explored collective variable space. Results from a 200ps (100000 steps) trajectory in vacuum are shown in Figure [munster-ala-traj-metad](#).

As you can see, exploration is greatly enhanced. Notice that the explored ensemble can be tuned using the biasfactor  $\gamma$ . Larger  $\gamma$  implies that the system will explore states with higher free energy. As a rule of thumb, if you expect a barrier of the order of  $\Delta G^*$ , a reasonable choice for the biasfactor is  $\gamma \approx \frac{\Delta G}{2k_B T}$ .

Finally, notice that [METAD](#) potential depends on the previously visited trajectories. As such, when you restart a previous simulation, it should read the previously deposited HILLS file. This is automatically triggered by the [RESTART](#) keyword.

## 11.16.3.1.3 Exercise 2

In this exercise, we will run a well-tempered metadynamics simulation on alanine dipeptide in vacuum, using as CV the backbone dihedral angle phi. In order to run this simulation we need to prepare the PLUMED input file (plumed.dat) as follows.

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate well-tempered metadynamics in phi depositing
# a Gaussian every 500 time steps, with initial height equal
# to 1.2 kJoule/mol, biasfactor equal to 10.0, and width to 0.35 rad

METAD ...
LABEL=metad
ARG=phi
PACE=500
HEIGHT=1.2
SIGMA=0.35
FILE=HILLS
BIASFACTOR=10.0
TEMP=300.0
GRID_MIN=-pi
GRID_MAX=pi
GRID_SPACING=0.1
... METAD

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR
```

(see [TORSION](#), [METAD](#), and [PRINT](#)).

The syntax for the command [METAD](#) is simple. The directive is followed by a keyword ARG followed by the labels of the CVs on which the metadynamics potential will act. The keyword PACE determines the stride of Gaussian deposition in number of time steps, while the keyword HEIGHT specifies the height of the Gaussian in kJoule/mol. For each CVs, one has to specify the width of the Gaussian by using the keyword SIGMA. Gaussian will be written to the file indicated by the keyword FILE.

The bias potential will be stored on a grid, whose boundaries are specified by the keywords GRID\_MIN and GRID\_MAX. Notice that you should provide either the number of bins for every collective variable (GRID\_BIN) or the desired grid spacing (GRID\_SPACING). In case you provide both PLUMED will use the most conservative choice (highest number of bins) for each dimension. In case you do not provide any information about bin size (neither GRID\_BIN nor GRID\_SPACING) and if Gaussian width is fixed PLUMED will use 1/5 of the Gaussian width as grid spacing. This default choice should be reasonable for most applications.

Once the PLUMED input file is prepared, one has to run Gromacs with the option to activate PLUMED and read the input file:

```
gmx_mpi mdrun -s ../TOPO/topolA.tpr -plumed plumed.dat -nsteps 5000000
```

During the metadynamics simulation, PLUMED will create two files, named COLVAR and HILLS. The COLVAR file contains all the information specified by the PRINT command, in this case the value of the CVs every 10 steps of simulation, along with the current value of the metadynamics bias potential. We can use COLVAR to visualize the behavior of the CV during the simulation:

By inspecting Figure [munster-metad-phi-fig](#), we can see that the system is initialized in one of the two metastable states of alanine dipeptide. After a while ( $t=0.1$  ns), the system is pushed by the metadynamics bias potential to visit the other local minimum. As the simulation continues, the bias potential fills the underlying free-energy landscape, and the system is able to diffuse in the entire phase space.

The HILLS file contains a list of the Gaussians deposited along the simulation. If we give a look at the header of this file, we can find relevant information about its content:

```
#! FIELDS time phi psi sigma_phi sigma_psi height biasf
#! SET multivariate false
#! SET min_phi -pi
#! SET max_phi pi
#! SET min_psi -pi
#! SET max_psi pi
```

The line starting with FIELDS tells us what is displayed in the various columns of the HILLS file: the time of the simulation, the value of phi and psi, the width of the Gaussian in phi and psi, the height of the Gaussian, and the biasfactor. We can use the HILLS file to visualize the decrease of the Gaussian height during the simulation, according to the well-tempered recipe:

If we look carefully at the scale of the y-axis, we will notice that in the beginning the value of the Gaussian height is higher than the initial height specified in the input file, which should be 1.2 kJoule/mol. In fact, this column reports the height of the Gaussian rescaled by the pre-factor that in well-tempered metadynamics relates the bias potential to the free energy. In this way, when we will use [sum\\_hills](#), the sum of the Gaussians deposited will directly provide the free-energy, without further rescaling needed (see below).

One can estimate the free energy as a function of the metadynamics CVs directly from the metadynamics bias potential. In order to do so, the utility [sum\\_hills](#) should be used to sum the Gaussians deposited during the simulation and stored in the HILLS file.

To calculate the free energy as a function of phi, it is sufficient to use the following command line:

```
plumed sum_hills --hills HILLS
```

The command above generates a file called `fes.dat` in which the free-energy surface as function of  $\phi$  is calculated on a regular grid. One can modify the default name for the free energy file, as well as the boundaries and bin size of the grid, by using the following options of `sum_hills` :

```
--outfile - specify the outputfile for sumhills
--min - the lower bounds for the grid
--max - the upper bounds for the grid
--bin - the number of bins for the grid
--spacing - grid spacing, alternative to the number of bins
```

The result should look like this:

To assess the convergence of a metadynamics simulation, one can calculate the estimate of the free energy as a function of simulation time. At convergence, the reconstructed profiles should be similar. The option `--stride` should be used to give an estimate of the free energy every  $N$  Gaussians deposited, and the option `--mintozero` can be used to align the profiles by setting the global minimum to zero. If we use the following command line:

```
plumed sum_hills --hills HILLS --stride 100 --mintozero
```

one free energy is calculated every 100 Gaussians deposited, and the global minimum is set to zero in all profiles. The resulting plot should look like the following:

To assess the convergence of the simulation more quantitatively, we can calculate the free-energy difference between the two local minima of the free energy along  $\phi$  as a function of simulation time. We can use following script to integrate the multiple free-energy profiles in the two basins defined by the following intervals in  $\phi$  space: basin A,  $-3 < \phi < -1$ , basin B,  $0.5 < \phi < 1.5$ .

```
# number of free-energy profiles
nfes= # put here the number of profiles
# minimum of basin A
minA=-3
# maximum of basin A
maxA=1
# minimum of basin B
minB=0.5
# maximum of basin B
maxB=1.5
# temperature in energy units
kbt=2.5

for((i=0;i<nfes;i++))
do
  # calculate free-energy of basin A
  A='awk 'BEGIN{tot=0.0}{if($1=="#!" && $1>min && $1<max)tot+=exp(-$2/kbt)}END{print -kbt*log(tot)}' min=${minA}
  # and basin B
  B='awk 'BEGIN{tot=0.0}{if($1=="#!" && $1>min && $1<max)tot+=exp(-$2/kbt)}END{print -kbt*log(tot)}' min=${minB}
  # calculate difference
  Delta=$(echo "${A} - ${B}" | bc -l)
  # print it
  echo $i $Delta
done
```

notice that `nfes` should be set to the number of profiles (free-energy estimates at different times of the simulation) generated by the option `--stride` of `sum_hills`.

This analysis, along with the observation of the diffusive behavior in the CVs space, suggest that the simulation is converged.

## 11.16.3.1.4 Exercise 3

In this exercise, we will run a well-tempered metadynamics simulation on alanine dipeptide in vacuum, using as CV the backbone dihedral angle psi. In order to run this simulation we need to prepare the PLUMED input file (plumed.dat) as follows.

```
# set up two variables for Phi and Psi dihedral angles
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Activate well-tempered metadynamics in psi depositing
# a Gaussian every 500 time steps, with initial height equal
# to 1.2 kJoule/mol, biasfactor equal to 10.0, and width to 0.35 rad

METAD ...
LABEL=metad
ARG=psi
PACE=500
HEIGHT=1.2
SIGMA=0.35
FILE=HILLS
BIASFACTOR=10.0
TEMP=300.0
GRID_MIN=-pi
GRID_MAX=pi
GRID_SPACING=0.1
... METAD

# monitor the two variables and the metadynamics bias potential
PRINT STRIDE=10 ARG=phi,psi,metad.bias FILE=COLVAR
```

(see [TORSION](#), [METAD](#), and [PRINT](#)).

Once the PLUMED input file is prepared, one has to run Gromacs with the option to activate PLUMED and read the input file:

```
gmx_mpi mdrun -s ../TOPO/topolA.tpr -plumed plumed.dat -nsteps 5000000
```

As we did in the previous exercise, we can use COLVAR to visualize the behavior of the CV during the simulation. Here we will plot at the same time the evolution of the metadynamics CV psi and of the other dihedral phi:

By inspecting Figure [munster-metad-psi-phi-fig](#), we notice that something different happened compared to the previous exercise. At first the behavior of psi looks diffusive in the entire CV space. However, around  $t=1$  ns, psi seems trapped in a region of the CV space in which it was previously diffusing without problems. The reason is that the non-biased CV phi after a while has jumped into a different local minima. Since phi is not directly biased, one has to wait for this (slow) degree of freedom to equilibrate before the free energy along psi can converge. Try to repeat the analysis done in the previous exercise (calculate the estimate of the free energy as a function of time and monitor the free-energy difference between basins) to assess the convergence of this metadynamics simulation.

## 11.16.3.2 Restraints

## 11.16.3.2.1 Biased sampling theory

A system at temperature  $T$  samples conformations from the canonical ensemble:

$$P(q) \propto e^{-\frac{U(q)}{k_B T}}$$

. Here  $q$  are the microscopic coordinates and  $k_B$  is the Boltzmann constant. Since  $q$  is a highly dimensional vector, it is often convenient to analyze it in terms of a few collective variables (see [Belfast tutorial: Analyzing CVs](#), [Belfast tutorial: Adaptive variables I](#), and [Belfast tutorial: Adaptive variables II](#)). The probability distribution for a CV  $s$  is

$$P(s) \propto \int dq e^{-\frac{U(q)}{k_B T}} \delta(s - s(q))$$

This probability can be expressed in energy units as a free energy landscape  $F(s)$ :

$$F(s) = -k_B T \log P(s)$$

Now we would like to modify the potential by adding a term that depends on the CV only. That is, instead of using  $U(q)$ , we use  $U(q) + V(s(q))$ . There are several reasons why one would like to introduce this potential. One is to avoid that the system samples some un-desired portion of the conformational space. As an example, imagine that you want to study dissociation of a complex of two molecules. If you perform a very long simulation you will be able to see association and dissociation. However, the typical time required for association will depend on the size of the simulation box. It could be thus convenient to limit the exploration to conformations where the distance between the two molecules is lower than a given threshold. This could be done by adding a bias potential on the distance between the two molecules. Another example is the simulation of a portion of a large molecule taken out from its initial context. The fragment alone could be unstable, and one might want to add additional potentials to keep the fragment in place. This could be done by adding a bias potential on some measure of the distance from the experimental structure (e.g. on root-mean-square deviation).

Whatever CV we decide to bias, it is very important to recognize which is the effect of this bias and, if necessary, remove it a posteriori. The biased distribution of the CV will be

$$P'(s) \propto \int dq e^{-\frac{U(q) + V(s(q))}{k_B T}} \delta(s - s(q)) \propto e^{-\frac{V(s(q))}{k_B T}} P(s)$$

and the biased free energy landscape

$$F'(s) = -k_B T \log P'(s) = F(s) + V(s) + C$$

Thus, the effect of a bias potential on the free energy is additive. Also notice the presence of an undetermined constant  $C$ . This constant is irrelevant for what concerns free-energy differences and barriers, but will be important later when we will learn the weighted-histogram method. Obviously the last equation can be inverted so as to obtain the original, unbiased free-energy landscape from the biased one just subtracting the bias potential

$$F(s) = F'(s) - V(s) + C$$

Additionally, one might be interested in recovering the distribution of an arbitrary observable. E.g., one could add a bias on the distance between two molecules and be willing to compute the unbiased distribution of some torsional angle. In this case there is no straightforward relationship that can be used, and one has to go back to the relationship between the microscopic probabilities:

$$P(q) \propto P'(q) e^{\frac{V(s(q))}{k_B T}}$$

The consequence of this expression is that one can obtain any kind of unbiased information from a biased simulation just by weighting every sampled conformation with a weight

$$w \propto e^{\frac{V(s(q))}{k_B T}}$$

That is, frames that have been explored in spite of a high (disfavoring) bias potential  $V$  will be counted more than frames that has been explored with a less disfavoring bias potential.

## 11.16.3.2.2 Umbrella sampling theory

Often in interesting cases the free-energy landscape has several local minima. If these minima have free-energy differences that are on the order of a few times  $k_B T$  they might all be relevant. However, if they are separated by a high saddle point in the free-energy landscape (i.e. a low probability region) than the transition between one and the other will take a lot of time and these minima will correspond to metastable states. The transition between one minimum and the other could require a time scale which is out of reach for molecular dynamics. In these situations, one could take inspiration from catalysis and try to favor in a controlled manner the conformations corresponding to the transition state.

Imagine that you know since the beginning the shape of the free-energy landscape  $F(s)$  as a function of one CV  $s$ . If you perform a molecular dynamics simulation using a bias potential which is exactly equal to  $-F(s)$ , the biased free-energy landscape will be flat and barrierless. This potential acts as an "umbrella" that helps you to safely cross the transition state in spite of its high free energy.

It is however difficult to have an a priori guess of the free-energy landscape. We will see later how adaptive techniques such as metadynamics ([Belfast tutorial: Metadynamics](#)) can be used to this aim. Because of this reason, umbrella sampling is often used in a slightly different manner.

Imagine that you do not know the exact height of the free-energy barrier but you have an idea of where the barrier is located. You could try to just favor the sampling of the transition state by adding a harmonic restraint on the CV, e.g. in the form

$$V(s) = \frac{k}{2}(s - s_0)^2$$

. The sampled distribution will be

$$P'(q) \propto P(q) e^{-\frac{k(s(q)-s_0)^2}{2k_B T}}$$

For large values of  $k$ , only points close to  $s_0$  will be explored. It is thus clear how one can force the system to explore only a predefined region of the space adding such a restraint. By combining simulations performed with different values of  $s_0$ , one could obtain a continuous set of simulations going from one minimum to the other crossing the transition state. In the next section we will see how to combine the information from these simulations.

If you want to just bring a collective variables to a specific value, you can use a simple restraint. Let's imagine that we want to force the  $\Phi$  angle to visit a region close to  $\Phi = \pi/2$ . We can do it adding a restraint in  $\Phi$ , with the following input

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
res: RESTRAINT ARG=phi AT=0.5pi KAPPA=5
PRINT ARG=phi,psi,res.bias
```

(see [TORSION](#), [RESTRAINT](#), and [PRINT](#)).

Notice that here we are printing a quantity named `res.bias`. We do this because [RESTRAINT](#) does not define a single value (that here would be theoretically named `res`) but a structure with several components. All biasing methods (including [METAD](#)) do so, as well as many collective variables (see e.g. [DISTANCE](#) used with `COMP↔ONENTS` keyword). Printing the bias allows one to know how much a given snapshot was penalized. Also notice that PLUMED understands numbers in the form `{number}pi`. This is convenient when using torsions, since they are expressed in radians.

Now you can plot your trajectory with gnuplot and see the effect of KAPPA. You can also try different values of KAPPA. The stiffer the restraint, the less the collective variable will fluctuate. However, notice that a too large kappa could make the MD integrator unstable.

### 11.16.3.3 Moving restraints

A restraint can also be modified as a function of time. For example, if you want to bring the system from one minimum to the other, you can use a moving restraint on  $\Phi$ :

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
# notice that a long line can be splitted with this syntax
MOVINGRESTRAINT ...
# also notice that a LABEL keyword can be used and is equivalent
# to adding the name at the beginning of the line with colon, as we did so far
  LABEL=res
  ARG=phi
  STEP0=0 AT0=-0.5pi KAPPA0=5
  STEP1=10000 AT0=0.5pi
...
PRINT ARG=phi,psi,res.work,res.phi_cntr FILE=colvar
```

(see [TORSION](#), [MOVINGRESTRAINT](#), and [PRINT](#)).

Notice that here we are plotting a few new components, namely `work` and `phi_cntr`. The former gives the work performed in pulling the restraint, and the latter the position of the restraint. Notice that if pulling is slow enough one can compute free energy profile from the work. You can plot the putative free-energy landscape with

```
> gnuplot
# column 5 is res.phi_cntr
# column 4 is res.work
gnuplot> p "colvar" u 5:4
```

### 11.16.3.4 Using multiple replicas

#### Warning

Notice that multireplica simulations with PLUMED are fully supported with GROMACS, but only partly supported with other MD engines.

Some free-energy methods are intrinsically parallel and requires running several simultaneous simulations. This can be done with gromacs using the multi replica framework. That is, if you have 4 tpr files named `topol0.tpr`, `topol1.tpr`, `topol2.tpr`, `topol3.tpr` you can run 4 simultaneous simulations.

```
> mpirun -np 4 gmx_mpi mdrun -s topol.tpr -plumed plumed.dat -multi 4 -nsteps 500000
```

Each of the 4 replicas will open a different `topol` file, and GROMACS will take care of adding the replica number before the `.tpr` suffix. PLUMED deals with the extra number in a slightly different way. In this case, for example, PLUMED first look for a file named `plumed.dat.X`, where `X` is the number of the replica. In case the file is not found, then PLUMED looks for `plumed.dat`. If also this is not found, PLUMED will complain. As a consequence, if all the replicas should use the same input file it is sufficient to put a single `plumed.dat` file, but one has also the flexibility of using separate files named `plumed.dat.0`, `plumed.dat.1` etc. Finally, notice that the way PLUMED adds suffixes will change in version 2.2, and names will be `plumed.0.dat` etc.

Also notice that providing the flag `-replex` one can instruct gromacs to perform a replica exchange simulation. Namely, from time to time gromacs will try to swap coordinates among neighboring replicas and accept or reject the exchange with a Monte Carlo procedure which also takes into account the bias potentials acting on the replicas, even if different bias potentials are used in different replicas. That is, PLUMED allows to easily implement many forms of Hamiltonian replica exchange.



## 11.16.3.5 Using multiple restraints with replica exchange

## 11.16.3.5.1 Weighted histogram analysis method theory

Let's now consider multiple simulations performed with restraints located in different positions. In particular, we will consider the  $i$ -th bias potential as  $V_i$ . The probability to observe a given value of the collective variable  $s$  is:

$$P_i(s) = \frac{P(s)e^{-\frac{V_i(s)}{k_B T}}}{\int ds' P(s')e^{-\frac{V_i(s')}{k_B T}}} = \frac{P(s)e^{-\frac{V_i(s)}{k_B T}}}{Z_i}$$

where

$$Z_i = \sum_q e^{-(U(q)+V_i(q))}$$

The likelihood for the observation of a sequence of snapshots  $q_i(t)$  (where  $i$  is the index of the trajectory and  $t$  is time) is just the product of the probability of each of the snapshots. We use here the minus-logarithm of the likelihood (so that the product is converted to a sum) that can be written as

$$\mathcal{L} = -\sum_i \int dt \log P_i(s_i(t)) = \sum_i \int dt \left( -\log P(s_i(t)) + \frac{V_i(s_i(t))}{k_B T} + \log Z_i \right)$$

One can then maximize the likelihood by setting  $\frac{\delta \mathcal{L}}{\delta P(\mathbf{s})} = 0$ . After some boring algebra the following expression can be obtained

$$0 = \sum_i \int dt \left( -\frac{\delta_{\mathbf{s}_i(t), \mathbf{s}}}{P(\mathbf{s})} + \frac{e^{-\frac{V_i(\mathbf{s})}{k_B T}}}{Z_i} \right)$$

In this equation we aim at finding  $P(s)$ . However, also the list of normalization factors  $Z_i$  is unknown, and they should be found selfconsistently. Thus one can find the solution as

$$P(\mathbf{s}) \propto \frac{N(\mathbf{s})}{\sum_i \int dt \frac{e^{-\frac{V_i(\mathbf{s})}{k_B T}}}{Z_i}}$$

where  $Z$  is selfconsistently determined as

$$Z_i \propto \int ds' P(\mathbf{s}') e^{-\frac{V_i(\mathbf{s}')}{k_B T}}$$

These are the WHAM equations that are traditionally solved to derive the unbiased probability  $P(s)$  by the combination of multiple restrained simulations. To make a slightly more general implementation, one can compute the weights that should be assigned to each snapshot, that turn out to be:

$$w_i(t) \propto \frac{1}{\sum_j \int dt \frac{e^{-\beta V_j(\mathbf{s}_i(t))}}{Z_j}}$$

The normalization factors can in turn be found from the weights as

$$Z_i \propto \frac{\sum_j \int dt e^{-\beta V_i(\mathbf{s}_j(t))} w_j(t)}{\sum_j \int dt w_j(t)}$$

This allows to straightforwardly compute averages related to other, non-biased degrees of freedom, and it is thus a bit more flexible. It is sufficient to precompute this factors  $w$  and use them to weight every single frame in the trajectory.

## 11.16.3.5.2 Exercise 4

In this exercise we will run multiple restraint simulations and learn how to reweight and combine data with WHAM to obtain free-energy profiles. We start with running in a replica-exchange scheme 32 simulations with a restraint on phi in different positions, ranging from -3 to 3. We will instruct gromacs to attempt an exchange between different simulations every 1000 steps.

```
nrep=32
dx='echo "6.0 / ( $nrep - 1 )" | bc -l`

for ((i=0;i<nrep;i++))
do
# center of the restraint
AT='echo "$i * $dx - 3.0" | bc -l`

cat >plumed.dat.$i << EOF
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
#
# Impose an umbrella potential on phi
# with a spring constant of 200 kJoule/mol
# and centered in phi=AT
#
restraint-phi: RESTRAINT ARG=phi KAPPA=200.0 AT=$AT
# monitor the two variables and the bias potential
PRINT STRIDE=100 ARG=phi,psi,restraint-phi.bias FILE=COLVAR
EOF

# we initialize some replicas in A and some in B:
if((i%2==0)); then
cp ../TOPO/topolA.tpr topol$i.tpr
else
cp ../TOPO/topolB.tpr topol$i.tpr
fi
done

# run REM
mpirun -np $nrep gmx_mpi mdrun -plumed plumed.dat -s topol.tpr -multi $nrep -replex 1000 -nsteps 500000
```

To be able to combine data from all the simulations, it is necessary to have an overlap between statistics collected in two adjacent umbrellas.

Have a look at the plot of (phi,psi) for the different simulations to understand what is happening.

An often misunderstood fact about WHAM is that data of the different trajectories can be mixed and it is not necessary to keep track of which restraint was used to produce every single frame. Let's get the concatenated trajectory

```
trjcat_mpi -cat -f traj*.xtc -o alltraj.xtc
```

Now we should compute the value of each of the bias potentials on the entire (concatenated) trajectory.

```
nrep=32
dx='echo "6.0 / ( $nrep - 1 )" | bc -l`

for i in `seq 0 $(( $nrep - 1 ))`
do
# center of the restraint
AT='echo "$i * $dx - 3.0" | bc -l`

cat >plumed.dat << EOF
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17
restraint-phi: RESTRAINT ARG=phi KAPPA=200.0 AT=$AT

# monitor the two variables and the bias potential
PRINT STRIDE=100 ARG=phi,psi,restraint-phi.bias FILE=ALLCOLVAR.$i
EOF

plumed driver --mf_xtc alltraj.xtc --trajectory-stride=10 --plumed plumed.dat

done
```

It is very important that this script is consistent with the one used to generate the multiple simulations above. Now, single files named ALLCOLVAR.XX will contain on the fourth column the value of the bias centered in a given position, computed on the entire concatenated trajectory.

Next step is to compute the weights self-consistently solving the WHAM equations, using the python script "wham.py" contained in the SCRIPTS directory. To use this code:

```
../SCRIPTS/wham.sh ALLCOLVAR.*
```

This script will produce several files. Let's visualize "phi\_fes.dat", which contains the free energy as a function of phi, and compare this with the result previously obtained with metadynamics.

### 11.16.3.5.3 Exercise 5

In the previous exercise, we use multiple restraint simulations to calculate the free energy as a function of the dihedral phi. The resulting free energy was in excellent agreement with our previous metadynamics simulation. In this exercise we will repeat the same procedure for the dihedral psi. At the end of the steps defined above, we can plot the free energy "psi\_fes.dat" and compare it with the reference profile calculated from a metadynamics simulations using both phi and psi as CVs.

We can easily spot from the plot above that something went wrong in this multiple restraint simulations, despite we used the very same approach we adopted for the phi dihedral. The problem here is that psi is a "bad" collective variable, and the system is not able to equilibrate the missing slow degree of freedom phi in the short time scale of the umbrella simulation (1 ns). In the metadynamics exercise in which we biased only psi, we detect problems by observing the behavior of the CV as a function of simulation time. How can we detect problems in multiple restraint simulations? This is slightly more complicated, but running this kind of simulation in a replica-exchange scheme offers a convenient way to detect problems.

The first thing we need to do is to demux the replica-exchange trajectories and reconstruct the continuous trajectories of the replicas across the different restraint potentials. In order to do so, we can use the following script:

```
demux.pl md0.log
trjcat_mpi -f traj*.xtc -demux replica_index.xvg
```

This commands will generate 32 continuous trajectories, named XX\_trajout.xtc. We will use the driver to calculate the value of the CVs phi and psi on these trajectories.

```
nrep=32

for i in `seq 0 $(( $nrep - 1 ))`
do

cat >plumed.dat << EOF
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

# monitor the two variables
PRINT STRIDE=100 ARG=phi,psi FILE=COLVARDEMUX.$i
EOF

plumed driver --mf_xtc ${i}_trajout.xtc --trajectory-stride=10 --plumed plumed.dat

done
```

The COLVARDEMUX.XX files will contain the value of the CVs on the demuxed trajectory. If we visualize these files we will notice that replicas sample the CVs space differently. In order for each umbrella to equilibrate the slow degrees of freedom phi, the continuous replicas must be ergodic and thus sample the same distribution in phi and psi.



## Chapter 12

# Index of Actions

The following page contains an alphabetically ordered list of all the Actions and command line tools that are available in PLUMED 2. For lists of Actions classified in accordance with the particular tasks that are being performed see:

- [Collective variables](#) tells you about the ways that you can calculate functions of the positions of the atoms.
- [Analysis](#) tells you about the various forms of analysis you can run on trajectories using PLUMED.
- [Bias](#) tells you about the methods that you can use to bias molecular dynamics simulations with PLUMED.

### 12.1 Full list of actions

<a href="#">ABMD</a>	BIAS	Adds a ratchet-and-pawl like restraint on one or more variables.
<a href="#">ALIGNED_MATRIX</a>	MATRIX	Adjacency matrix in which two molecule are adjacent if they are within a certain cutoff and if they have the same orientation.
<a href="#">ALPHABETA</a>	COLVAR	Measures a distance including pbc between the instantaneous values of a set of torsional angles and set of reference values.
<a href="#">ALPHARMSD</a>	COLVAR	Probe the alpha helical content of a protein structure.
<a href="#">ANGLES</a>	MCOLVAR	Calculate functions of the distribution of angles .
<a href="#">ANGLE</a>	COLVAR	Calculate an angle.
<a href="#">ANTIBETARMSD</a>	COLVAR	Probe the antiparallel beta sheet content of your protein structure.
<a href="#">AROUND</a>	VOLUMES	This quantity can be used to calculate functions of the distribution of collectivevariables for the atoms that lie in a particular, user-specified part of of the cell.
<a href="#">AVERAGE</a>	GRIDCALC	Calculate the ensemble average of a collective variable
<a href="#">BIASVALUE</a>	BIAS	Takes the value of one variable and use it as a bias
<a href="#">BRIDGE</a>	MCOLVAR	Calculate the number of atoms that bridge two parts of a structure
<a href="#">CAVITY</a>	VOLUMES	This quantity can be used to calculate functions of the distribution of collectivevariables for the atoms that lie in a box defined by the positions of four atoms.
<a href="#">CELL</a>	COLVAR	Calculate the components of the simulation cell
<a href="#">CENTER_OF_MULTICOLVAR</a>	VATOM	Calculate a a weighted average position based on the value of some multicolvar.

<a href="#">CENTER</a>	VATOM	Calculate the center for a group of atoms, with arbitrary weights.
<a href="#">CLASSICAL_MDS</a>	DIMRED	Create a low-dimensional projection of a trajectory using the classical multidimensional scaling algorithm.
<a href="#">CLUSTER_DIAMETER</a>	CONCOMP	Print out the diameter of one of the connected components
<a href="#">CLUSTER_DISTRIBUTION</a>	CONCOMP	Calculate functions of the distribution of properties in your connected components.
<a href="#">CLUSTER_NATOMS</a>	CONCOMP	Gives the number of atoms in the connected component
<a href="#">CLUSTER_PROPERTIES</a>	CONCOMP	Calculate properties of the distribution of some quantities that are part of a connected component
<a href="#">CLUSTER_WITHSURFACE</a>	MATRIXF	Take a connected component that was found using a clustering algorithm and create a new cluster that contains those atoms that are in the cluster together with those atoms that are within a certain cutoff of the cluster.
<a href="#">COLUMNSUMS</a>	MATRIXF	Sum the columns of a contact matrix
<a href="#">COMBINE</a>	FUNCTION	Calculate a polynomial combination of a set of other variables.
<a href="#">COMMITTOR</a>	PRINTANALYSIS	Does a committor analysis.
<a href="#">COM</a>	VATOM	Calculate the center of mass for a group of atoms.
<a href="#">CONSTANT</a>	COLVAR	Return one or more constant quantities with or without derivatives.
<a href="#">CONTACTMAP</a>	COLVAR	Calculate the distances between a number of pairs of atoms and transform each distance by a switching function. The transformed distance can be compared with a reference value in order to calculate the squared distance between two contact maps. Each distance can also be weighted for a given value. CONTACTMAP can be used together with <a href="#">FUNCPATHMSD</a> to define a path in the contactmap space.
<a href="#">CONTACT_MATRIX</a>	MATRIX	Adjacency matrix in which two atoms are adjacent if they are within a certain cutoff.
<a href="#">CONVERT_TO_FES</a>	GRIDANALYSIS	Convert a histogram, $H(x)$ , to a free energy surface using $F(x) = -k_B T \ln H(x)$ .
<a href="#">COORDINATIONNUMBER</a>	MCOLVAR	Calculate the coordination numbers of atoms so that you can then calculate functions of the distribution of coordination numbers such as the minimum, the number less than a certain quantity and so on.
<a href="#">COORDINATION</a>	COLVAR	Calculate coordination numbers.
<a href="#">CS2BACKBONE</a>	COLVAR	This collective variable calculates the backbone chemical shifts for a protein.
<a href="#">DEBUG</a>	GENERIC	Set some debug options.
<a href="#">DENSITY</a>	MCOLVAR	Calculate functions of the density of atoms as a function of the box. This allows one to calculate the number of atoms in half the box.
<a href="#">DFSCLUSTERING</a>	MATRIXF	Find the connected components of the matrix using the depth first search clustering algorithm.
<a href="#">DHENERGY</a>	COLVAR	Calculate Debye-Huckel interaction energy among GROUPA and GROUPB.
<a href="#">DIHCOR</a>	COLVAR	Measures the degree of similarity between dihedral angles.

DIPOLE	COLVAR	Calculate the dipole moment for a group of atoms.
DISTANCE_FROM_CONTOUR	COLVAR	Calculate the perpendicular distance from a Willard-Chandler dividing surface.
DISTANCES	MCOLVAR	Calculate the distances between one or many pairs of atoms. You can then calculate functions of the distribution of distances such as the minimum, the number less than a certain quantity and so on.
DISTANCE	COLVAR	Calculate the distance between a pair of atoms.
driver-float	TOOLS	Equivalent to <a href="#">driver</a> , but using single precision reals.
driver	TOOLS	driver is a tool that allows one to use plumed to post-process an existing trajectory.
DRMSD	DCOLVAR	Calculate the distance RMSD with respect to a reference structure.
DUMPATOMS	PRINTANALYSIS	Dump selected atoms on a file.
DUMPCUBE	GRIDANALYSIS	Output a three dimensional grid using the Gaussian cube file format.
DUMPPERIVATIVES	PRINTANALYSIS	Dump the derivatives with respect to the input parameters for one or more objects (generally CVs, functions or biases).
DUMPFORCES	PRINTANALYSIS	Dump the force acting on one of a values in a file.
DUMPGRID	GRIDANALYSIS	Output the function on the grid to a file with the PLUMED grid format.
DUMPMASSCHARGE	PRINTANALYSIS	Dump masses and charges on a selected file.
DUMPMULTICOLVAR	PRINTANALYSIS	Dump atom positions and multicolvar on a file.
DUMPPROJECTIONS	PRINTANALYSIS	Dump the derivatives with respect to the input parameters for one or more objects (generally CVs, functions or biases).
EFFECTIVE_ENERGY_DRIFT	GENERIC	Print the effective energy drift described in Ref [57]
ENERGY	COLVAR	Calculate the total energy of the simulation box.
ENSEMBLE	FUNCTION	Calculates the replica averaging of a collective variable over multiple replicas.
ERMSD	COLVAR	Calculate eRMSD with respect to a reference structure.
EXTENDED_LAGRANGIAN	BIAS	Add extended Lagrangian.
EXTERNAL	BIAS	Calculate a restraint that is defined on a grid that is read during start up
FAKE	COLVAR	This is a fake colvar container used by cltools or various other actions and just support input and period definition
FCCUBIC	MCOLVAR	Measure how similar the environment around atoms is to that found in a FCC structure.
FIND_CONTOUR_SURFACE	GRIDANALYSIS	Find an isocontour by searching along either the x, y or z direction.
FIND_CONTOUR	GRIDANALYSIS	Find an isocontour in a smooth function.
FIND_SPHERICAL_CONTOUR	GRIDANALYSIS	Find an isocontour in a three dimensional grid by searching over a Fibonacci sphere.
FIT_TO_TEMPLATE	GENERIC	This action is used to align a molecule to a template.
FIXEDATOM	VATOM	Add a virtual atom in a fixed position.
FLUSH	GENERIC	This command instructs plumed to flush all the open files with a user specified frequency. Notice that all files are flushed anyway every 10000 steps.

FOURIER_TRANSFORM	GRIDANALYSIS	Compute the Discrete Fourier Transform (DFT) by means of FFTW of data stored on a 2D grid.
FRET	COLVAR	Calculate the FRET efficiency between a pair of atoms. The efficiency is calculated using the Forster relation:
FUNCPATHMSD	FUNCTION	This function calculates path collective variables.
FUNCSUMHILLS	FUNCTION	This function is intended to be called by the command line tool <code>sum_hills</code> and it is meant to integrate a HILLS file or an HILLS file interpreted as a histogram in a variety of ways. Therefore it is not expected that you use this during your dynamics (it will crash!)
gentemplate	TOOLS	gentemplate is a tool that you can use to construct template inputs for the various actions
GHOST	VATOM	Calculate the absolute position of a ghost atom with fixed coordinates in the local reference frame formed by three atoms. The computed ghost atom is stored as a virtual atom that can be accessed in an atom list through the label for the GHOST action that creates it.
GPROPERTYMAP	COLVAR	Property maps but with a more flexible framework for the distance metric being used.
GRADIENT	MCOLVARF	Calculate the gradient of the average value of a multi-colvar value
GROUP	GENERIC	Define a group of atoms so that a particular list of atoms can be referenced with a single label in definitions of CVs or virtual atoms.
GYRATION	COLVAR	Calculate the radius of gyration, or other properties related to it.
HBOND_MATRIX	MATRIX	Adjacency matrix in which two atoms are adjacent if there is a hydrogen bond between them.
HISTOGRAM	GRIDCALC	Accumulate the average probability density along a few CVs from a trajectory.
INCLUDE	GENERIC	Includes an external input file, similar to "#include" in C preprocessor.
INCYLINDER	VOLUMES	This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a particular, user-specified part of the cell.
info	TOOLS	This tool allows you to obtain information about your plumed version
INPLANEDISTANCES	MCOLVAR	Calculate distances in the plane perpendicular to an axis
INSPHERE	VOLUMES	This quantity can be used to calculate functions of the distribution of collective variables for the atoms that lie in a particular, user-specified part of the cell.
INTERMOLECULARTORSIONS	MCOLVARF	Calculate torsions between axis of adjacent molecules
INTERPOLATE_GRID	GRIDANALYSIS	Interpolate a smooth function stored on a grid onto a grid with a smaller grid spacing.
JCOUPLING	COLVAR	Calculates $^3J$ coupling constants for a dihedral angle.
kt	TOOLS	Print out the value of $k_B T$ at a particular temperature
LOAD	GENERIC	Loads a library, possibly defining new actions.
LOCAL_AVERAGE	MCOLVARF	Calculate averages over spherical regions centered on atoms
LOCALENSEMBLE	FUNCTION	Calculates the average over multiple arguments.



LOCAL_Q3	MCOLVARF	Calculate the local degree of order around an atoms by taking the average dot product between the $q_3$ vector on the central atom and the $q_3$ vector on the atoms in the first coordination sphere.
LOCAL_Q4	MCOLVARF	Calculate the local degree of order around an atoms by taking the average dot product between the $q_4$ vector on the central atom and the $q_4$ vector on the atoms in the first coordination sphere.
LOCAL_Q6	MCOLVARF	Calculate the local degree of order around an atoms by taking the average dot product between the $q_6$ vector on the central atom and the $q_6$ vector on the atoms in the first coordination sphere.
LOWER_WALLS	BIAS	Defines a wall for the value of one or more collective variables, which limits the region of the phase space accessible during the simulation.
manual	TOOLS	manual is a tool that you can use to construct the manual page for a particular action
MATHEVAL	FUNCTION	Calculate a combination of variables using a matheval expression.
METAD	BIAS	Used to performed MetaDynamics on one or more collective variables.
METAINFERENCE	BIAS	Calculate the Metainference Score for a set of back calculated experimental data.
MFILTER_BETWEEN	MFILTERS	This action can be used to filter the colvar values calculated by a multicolvarso that one can compute the mean and so on for only those multicolvars within a certain range.
MFILTER_LESS	MFILTERS	This action can be used to filter the distribution of colvar values in a multicolvarso that one can compute the mean and so on for only those multicolvars less than a tolerance.
MFILTER_MORE	MFILTERS	This action can be used to filter the distribution of colvar values in a multicolvarso that one can compute the mean and so on for only those multicolvars more than a tolerance.
MOLECULES	MCOLVAR	Calculate the vectors connecting a pair of atoms in order to represent the orientation of a molecule.
MOLINFO	TOPOLOGY	This command is used to provide information on the molecules that are present in your system.
MOVINGRESTRAINT	BIAS	Add a time-dependent, harmonic restraint on one or more variables.
MTRANSFORM_BETWEEN	MTRANSFORMS	This action can be used to transform the colvar values calculated by a multicolvar using a <a href="#">histogrambead</a>
MTRANSFORM_LESS	MTRANSFORMS	This action can be used to transform the colvar values calculated by a multicolvar using a <a href="#">switchingfunction</a>
MTRANSFORM_MORE	MTRANSFORMS	This action can be used to transform the colvar values calculated by a multicolvar using one minus a <a href="#">switchingfunction</a>
MULTICOLVARDENS	GRIDCALC	Evaluate the average value of a multicolvar on a grid.
MULTI-RMSD	DCOLVAR	Calculate the RMSD distance moved by a number of separated domains from their positions in a reference structure.
NLINKS	MCOLVARF	Calculate number of pairs of atoms/molecules that are "linked"

NOE	COLVAR	Calculates NOE intensities as sums of $1/r^6$ , also averaging over multiple equivalent atoms or ambiguous N↔OE.
OUTPUT_CLUSTER	CONCOMP	Output the indices of the atoms in one of the clusters identified by a clustering object
PARABETARMSD	COLVAR	Probe the parallel beta sheet content of your protein structure.
PATHMSD	COLVAR	This Colvar calculates path collective variables.
PATH	COLVAR	Path collective variables with a more flexible framework for the distance metric being used.
PBMETAD	BIAS	Used to performed Parallel Bias MetaDynamics.
PCARMSD	DCOLVAR	Calculate the PCA components ( see [21] and [22] ) for a number of provided eigenvectors and an average structure. Performs optimal alignment at every step and reports the rmsd so you know if you are far or close from the average structure. It takes the average structure and eigenvectors in form of a pdb. Note that beta and occupancy values in the pdb are neglected and all the weights are placed to 1 (differently from the RMSD colvar for example)
PCA	DIMRED	Perform principal component analysis (PCA) using either the positions of the atoms a large number of collective variables as input.
PCAVARS	COLVAR	Projection on principal component eigenvectors or other high dimensional linear subspace
PIECEWISE	FUNCTION	Compute a piecewise straight line through its arguments that passes through a set of ordered control points.
PLANES	MCOLVAR	Calculate the plane perpendicular to two vectors in order to represent the orientation of a planar molecule.
POSITION	COLVAR	Calculate the components of the position of an atom.
PRE	COLVAR	Calculates the Paramagnetic Resonance Enhancement intensity ratio between two atoms. The reference atom for the spin label is added with SPINLABEL, the affected atom(s) are given as numbered GROUPA1, GROUPA2, ... The additional parameters needed for the calculation are given as INEPT, the inepttime, TAUC the correlation time, OMEGA, the larmor frequency and RTWO for the relaxation time.
PRINT	PRINTANALYSIS	Print quantities to a file.
PROPERTYMAP	COLVAR	Calculate generic property maps.
PUCKERING	COLVAR	Calculate sugar pseudorotation coordinates.
Q3	MCOLVAR	Calculate 3rd order Steinhardt parameters.
Q4	MCOLVAR	Calculate 4th order Steinhardt parameters.
Q6	MCOLVAR	Calculate 6th order Steinhardt parameters.
RANDOM_EXCHANGES	GENERIC	Set random pattern for exchanges.
RDC	COLVAR	Calculates the (Residual) Dipolar Coupling between two atoms.
READ	GENERIC	Read quantities from a colvar file.
RESET_CELL	GENERIC	This action is used to rotate the full cell
RESTART	GENERIC	Activate restart.
RESTRAINT	BIAS	Adds harmonic and/or linear restraints on one or more variables.

REWEIGHT_BIAS	REWEIGHTING	Calculate weights for ensemble averages that negate the effect the bias has on the region of phase space explored
REWEIGHT_METAD	REWEIGHTING	Calculate the weights configurations should contribute to the histogram in a simulation in which a metadynamics bias acts upon the system.
REWEIGHT_TEMP	REWEIGHTING	Calculate weights for ensemble averages allow for the computing of ensemble averages at temperatures lower/higher than that used in your original simulation.
RMSD	DCOLVAR	Calculate the RMSD with respect to a reference structure.
ROWSUMS	MATRIXF	Sum the rows of a adjacency matrix.
SIMPLECUBIC	MCOLVAR	Calculate whether or not the coordination spheres of atoms are arranged as they would be in a simplecubic structure.
simplemd	TOOLS	simplemd allows one to do molecular dynamics on systems of Lennard-Jones atoms.
SMAC_MATRIX	MATRIX	Adjacency matrix in which two molecules are adjacent if they are within a certain cutoff and if the angle between them is within certain ranges.
SMAC	MCOLVARF	Calculate a variant on the SMAC collective variable discussed in [26]
SORT	FUNCTION	This function can be used to sort colvars according to their magnitudes.
SPRINT	MATRIXF	Calculate SPRINT topological variables from an adjacency matrix.
STATS	FUNCTION	Calculates statistical properties of a set of collective variables with respect to a set of reference values.↔ In particular it calculates and store as components the sum of the squared deviations, the correlation, the slope and the intercept of a linear fit.
sum_hills	TOOLS	sum_hills is a tool that allows one to use plumed to post-process an existing hills/colvar file
TARGET	DCOLVAR	This function measures the pythagorean distance from a particular structure measured in the space defined by some set of collective variables.
TEMPLATE	COLVAR	This file provides a template for if you want to introduce a new CV.
TETRAHEDRALPORE	VOLUMES	This quantity can be used to calculate functions of the distribution of collective variables for the atoms lie that lie in a box defined by the positions of four atoms at the corners of a tetrahedron.
TETRAHEDRAL	MCOLVAR	Calculate the degree to which the environment about ions has a tetrahedral order.
TIME	GENERIC	retrieve the time of the simulation to be used elsewhere
TORSIONS	MCOLVAR	Calculate whether or not a set of torsional angles are within a particular range.
TORSION	COLVAR	Calculate a torsional angle.
UNITS	GENERIC	This command sets the internal units for the code. A new unit can be set by either specifying how to convert from the plumed default unit into that new unit or by using the shortcuts described below. This directive MUST appear at the BEGINNING of the plumed.dat file. The same units must be used throughout the plumed.dat file.

UPDATE_IF	PRINTANALY↔ SIS	Conditional update of other actions.
UPPER_WALLS	BIAS	Defines a wall for the value of one or more collective variables, which limits the region of the phase space accessible during the simulation.
UWALLS	MCOLVARB	Add <a href="#">UPPER_WALLS</a> restraints on all the multicolvar values
VOLUME	COLVAR	Calculate the volume of the simulation box.
WHOLEMOLECULES	GENERIC	This action is used to rebuild molecules that can become split by the periodic boundary conditions.
WRAPAROUND	GENERIC	Rebuild periodic boundary conditions around chosen atoms.
XDISTANCES	MCOLVAR	Calculate the x components of the vectors connecting one or many pairs of atoms. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.
XYDISTANCES	MCOLVAR	Calculate distance between a pair of atoms neglecting the z-component. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.
XZDISTANCES	MCOLVAR	Calculate distance between a pair of atoms neglecting the y-component. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.
YDISTANCES	MCOLVAR	Calculate the y components of the vectors connecting one or many pairs of atoms. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.
YZDISTANCES	MCOLVAR	Calculate distance between a pair of atoms neglecting the x-component. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.
ZDISTANCES	MCOLVAR	Calculate the z components of the vectors connecting one or many pairs of atoms. You can then calculate functions of the distribution of values such as the minimum, the number less than a certain quantity and so on.

## Chapter 13

# Bug List

### Page [CENTER\\_OF\\_MULTICOLVAR](#)

The virial contribution for this type of virtual atom is not currently evaluated so do not use in bias functions unless the volume of the cell is fixed

### Page [ENERGY](#)

Acceptance for replica exchange when [ENERGY](#) is biased is computed correctly only if all the replicas have the same potential energy function. This is for instance not true when using GROMACS with lambda replica exchange or with plumed-hrex branch.

### Page [lammmps-6Apr13](#)

The provided patch does not pass the virial correctly to PLUMED (see <https://github.com/plumed/plumed2/issues/377>). If you need to perform constant pressure simulations you should download a LAMMPS version newer than November 2018, where the mentioned bug has been fixed, that natively supports PLUMED 2.4. If you need to use LAMMPS with fixed virial with PLUMED 2.3, you can compile LAMMPS with PLUMED 2.4 using runtime binding and then set the variable PLUMED\_KERNEL to load PLUMED 2.3.

### Page [MOLINFO](#)

At the moment the HA1 atoms in a GLY residues are treated as if they are the CB atoms. This may or may not be true - GLY is problematic for secondary structure residues as it is achiral.

If you use WHOLEMOLECULES RESIDUES=1-10 for a 18 amino acid protein ( 18 amino acids + 2 terminal groups = 20 residues ) the code will fail as it will not be able to interpret terminal residue 1.

### Page [namd-2.8](#)

NAMD does not currently take into account virial contributions from PLUMED. Please use constant volume simulations only.

### Page [namd-2.9](#)

NAMD does not currently take into account virial contributions from PLUMED. Please use constant volume simulations only.

### Page [PCAVARS](#)

It is not possible to use the [DRMSD](#) metric with this variable. You can get around this by listing the set of distances you wish to calculate for your DRMSD in the plumed file explicitly and using the EUCLIDEAN metric. MAHALONOBIS and NORM-EUCLIDEAN also do not work with this variable but using these options makes little sense when projecting on a linear subspace.

### Page [PUCKERING](#)

The 6-membered ring implementation distributed with this version of PLUMED leads to qx and qy values that have an opposite sign with respect to those originally defined in [19]. The bug will be fixed in a later version but is here kept to preserve reproducibility.



# Bibliography

- [1] Gareth A. Tribello, Massimiliano Bonomi, Davide Branduardi, Carlo Camilloni, and Giovanni Bussi. Plumed 2: New feathers for an old bird. *Comput. Phys. Commun.*, 185(2):604–613, 2014. [1](#)
- [2] Massimiliano Bonomi, Davide Branduardi, Giovanni Bussi, Carlo Camilloni, Davide Provasi, Paolo Raiteri, Davide Donadio, Fabrizio Marinelli, Fabio Pietrucci, Ricardo A Broglia, and Michele Parrinello. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications*, 180(10):1961–1972, 2009. [1](#)
- [3] Pratyush Tiwary and Michele Parrinello. A time-independent free energy estimator for metadynamics. *The Journal of Physical Chemistry B*, 119(3):736–742, Jan 2015. [14](#), [337](#), [341](#)
- [4] F. Pietrucci and A. Laio. A collective variable for the efficient exploration of protein beta-structures with metadynamics: application to sh3 and gb1. *J. Chem. Theory Comput.*, 5(9):2197–2201, 2009. [64](#), [67](#), [96](#)
- [5] R. B. Best, G. Hummer, and W. A. Eaton. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. U.S.A.*, 110(44):17874–17879, 2013. [73](#)
- [6] KJ Kohlhoff, Paul Robustelli, Andrea Cavalli, Xavier Salvatella, and Michele Vendruscolo. Fast and accurate predictions of protein NMR chemical shifts from interatomic distances. *J. Am. Chem. Soc.*, 131(39):13894–13895, 2009. [75](#)
- [7] Paul Robustelli, Kai Kohlhoff, Andrea Cavalli, and Michele Vendruscolo. Using NMR chemical shifts as structural restraints in molecular dynamics simulations of proteins. *Structure*, 18(8):923–933, 2010. [76](#)
- [8] Daniele Granata, Carlo Camilloni, Michele Vendruscolo, and Alessandro Laio. Characterization of the free-energy landscapes of proteins by NMR-guided metadynamics. *Proc. Natl. Acad. Sci. U.S.A.*, 110(17):6817–6822, 2013. [76](#), [77](#)
- [9] Carlo Camilloni, Paul Robustelli, Alfonso De Simone, Andrea Cavalli, and Michele Vendruscolo. Characterization of the Conformational Equilibrium between the Two Major Substates of RNase A Using NMR Chemical Shifts. *J. Am. Chem. Soc.*, 134(9):3968–3971, 2012. [76](#), [78](#)
- [10] Carlo Camilloni, Andrea Cavalli, and Michele Vendruscolo. Assessment of the Use of NMR Chemical Shifts as Replica-Averaged Structural Restraints in Molecular Dynamics Simulations to Characterize the Dynamics of Proteins. *J. Phys. Chem. B*, 117(6):1838–1843, 2013. [76](#), [78](#)
- [11] Trang N. Do, Paolo Carloni, Gabriele Varani, and Giovanni Bussi. Rna/peptide binding driven by electrostatics—insight from bidirectional pulling simulations. *Journal of Chemical Theory and Computation*, 9(3):1720–1730, 2013. [78](#)
- [12] C. Bartels and M. Karplus. Probability Distributions for Complex Systems: Adaptive Umbrella Sampling of the Potential Energy. *J. Phys. Chem. B*, 102(5):865–880, 1998. [85](#)
- [13] M. Bonomi and M. Parrinello. Enhanced sampling in the well-tempered ensemble. *Phys. Rev. Lett.*, 104:190601, 2010. [85](#), [432](#)
- [14] Sandro Bottaro, Francesco Di Palma, and Giovanni Bussi. The role of nucleobase interactions in rna structure and dynamics. *Nucleic acids research*, 21(42):13306–13314, 2014. [86](#)
- [15] Vojtech Spiwok and Blanka Králová. Metadynamics in the conformational space nonlinearly dimensionally reduced by Isomap. *Journal of Chemical Physics*, 135(22):224504, December 2011. [89](#), [107](#)

- [16] Jiří Vymětal and Jiří Vondrášek. Gyration- and Inertia-Tensor-Based Collective Coordinates for Metadynamics. Application on the Conformational Behavior of Polyalanine Peptides and Trp-Cage Folding. *J. Phys. Chem. A*, page 110930112611005, 2011. [91](#)
- [17] Davide Branduardi, Francesco Luigi Gervasio, and Michele Parrinello. From A to B in free energy space. *J. Chem. Phys.*, 126(5):054103, Feb 2007. [98](#), [99](#), [127](#)
- [18] Ming Huang, Timothy J Giese, Tai-Sung Lee, and Darrin M York. Improvement of dna and rna sugar pucker profiles from semiempirical quantum methods. *Journal of chemical theory and computation*, 10(4):1538–1545, 2014. [108](#)
- [19] D t Cremer and JA Pople. General definition of ring puckering coordinates. *Journal of the American Chemical Society*, 97(6):1354–1358, 1975. [109](#), [509](#)
- [20] Xevi Biarnés, Albert Ardevol, Antoni Planas, Carme Rovira, Alessandro Laio, and Michele Parrinello. The conformational free energy landscape of  $\beta$ -d-glucopyranose. implications for substrate preactivation in  $\beta$ -glucoside hydrolases. *Journal of the American Chemical Society*, 129(35):10686–10693, 2007. [109](#)
- [21] L Sutto, M D Abramo, and F L Gervasio. Comparing the efficiency of biased and unbiased molecular dynamics in reconstructing the free energy landscape of met-enkephalin. *J. Chem. Theory Comput.*, 6(12):3640–3646, Dec 2010. [115](#), [119](#), [506](#)
- [22] Vojtech Spiwok, Petra Lipovová, and Blanka Králová. Metadynamics in essential coordinates: free energy simulation of conformational changes. *The journal of physical chemistry B*, 111(12):3073–6, Mar 2007. [115](#), [119](#), [506](#)
- [23] S. K. Kearsley. On the orthogonal transformation used for structural comparison. *Acta Cryst. A*, 45:208–210, 1989. [121](#)
- [24] Andrea Pérez-Villa, Maria Darvas, and Giovanni Bussi. Atp dependent ns3 helicase interaction with rna: insights from molecular simulations. *Nucleic Acids Research*, 43(18):8725, 2015. [133](#)
- [25] Wolfgang Lechner and Christoph Dellago. Accurate determination of crystal structures based on averaged local bond order parameters. *The Journal of Chemical Physics*, 129(11):–, 2008. [140](#), [227](#), [230](#)
- [26] Federico Giberti, Matteo Salvalaglio, Marco Mazzotti, and Michele Parrinello. Insight into the nucleation of urea crystals from the melt. *Chemical Engineering Science*, 121:51 – 59, 2015. 2013 Danckwerts Special Issue on Molecular Modelling in Chemical Engineering. [140](#), [244](#), [507](#)
- [27] Gareth A. Tribello, Jérôme Cuny, Hagai Eshet, and Michele Parrinello. Exploring the free energy surfaces of clusters using reconnaissance metadynamics. *J. Chem. Phys.*, 135(11):114109, 2011. [141](#)
- [28] Stefano Angioletti-Uberti, Michele Ceriotti, Peter D. Lee, and Mike W. Finnis. Solid-liquid interface free energy through metadynamics simulations. *Phys. Rev. B*, 81:125416, Mar 2010. [154](#)
- [29] Bingqing Cheng, Gareth A. Tribello, and Michele Ceriotti. Solid-liquid interfacial free energy out of equilibrium. *Phys. Rev. B*, 92:180102, Nov 2015. [154](#)
- [30] Federico Giberti, Gareth A. Tribello, and Michele Parrinello. Transient polymorphism in nacl. *Journal of Chemical Theory and Computation*, 9(2526-2530):null, 2013. [224](#)
- [31] Gareth A. Tribello, Federico Giberti, Gabriele C. Sosso, Matteo Salvalaglio, and Michele Parrinello. Analyzing and driving cluster formation in atomistic simulations. *Journal of Chemical Theory and Computation*, 13(3):1317–1327, 2017. [260](#), [261](#), [263](#), [271](#), [278](#), [280](#), [281](#), [284](#)
- [32] Gabriele C. Sosso, Gareth A. Tribello, Andrea Zen, Philipp Pedevilla, and Angelos Michaelides. Ice formation on kaolinite: Insights from molecular dynamics simulations. *The Journal of Chemical Physics*, 145(21):211927, 2016. [267](#)
- [33] Pratyush Tiwary and Michele Parrinello. A time-independent free energy estimator for metadynamics. *The Journal of Physical Chemistry B*, 119(3):736–742, 2015. PMID: 25046020. [301](#)
- [34] Adam P. Willard and David Chandler. Instantaneous liquid interfaces. *The Journal of Physical Chemistry B*, 114(5):1954–1958, 2010. [314](#), [315](#)



- [35] M. Marchi and P. Ballone. Adiabatic bias molecular dynamics: A method to navigate the conformational space of complex molecular systems. *J. Chem. Phys.*, 110(8):3697–3702, 1999. [327](#)
- [36] D. Provasi and M. Filizola. Putative active states of a prototypic g-protein-coupled receptor from biased molecular dynamics. *Biophys. J.*, 98:2347–2355, 2010. [327](#)
- [37] C. Camilloni, R. A. Broglia, and G. Tiana. Hierarchy of folding and unfolding events of protein g, ci2, and acbp from explicit-solvent simulations. *J. Chem. Phys.*, 134:045105, 2011. [327](#)
- [38] M. Iannuzzi, A. Laio, and M. Parrinello. Efficient exploration of reactive potential energy surfaces using car-parrinello molecular dynamics. *Phys. Rev. Lett.*, 90:238302, 2003. [331](#)
- [39] L. Maragliano and E. Vanden-Eijnden. A temperature-accelerated method for sampling free energy and determining reaction pathways in rare events simulations. *Chem. Phys. Lett.*, 426:168–175, 2006. [331](#)
- [40] Jerry B. Abrams and Mark E. Tuckerman. Efficient and Direct Generation of Multidimensional Free Energy Surfaces via Adiabatic Dynamics without Coordinate Transformations. *J. Phys. Chem. B*, 112(49):15742–15757, DEC 11 2008. [331](#)
- [41] A. Laio and M. Parrinello. Escaping free energy minima. *Proc. Natl. Acad. Sci. USA*, 99:12562–12566, 2002. [336](#), [346](#), [425](#), [451](#), [463](#), [488](#)
- [42] V. Babin, C. Roland, and C. Sagui. Adaptively biased molecular dynamics for free energy calculations. *J. Chem. Phys.*, 128:134101, 2008. [336](#)
- [43] A. Barducci, G. Bussi, and M. Parrinello. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.*, 100(2):020603, Jan 2008. [336](#), [347](#), [425](#), [452](#), [463](#), [489](#)
- [44] D. Branduardi, G. Bussi, and M. PARRINELLO. Metadynamics with adaptive Gaussians. *J. Chem. Theory Comput.*, 8(7):2247–2254, 2012. [337](#), [407](#)
- [45] Fahimeh Baftizadeh, Pilar Cossio, Fabio Pietrucci, and Alessandro Laio. Protein folding and ligand-enzyme binding from bias-exchange metadynamics simulations. *Curr Phys Chem*, 2:79–91, 2012. [337](#), [347](#)
- [46] P. Raiteri, A. Laio, F.L. Gervasio, C. Micheletti, and M. Parrinello. Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. *J. Phys. Chem. B*, 110:3533–3539, 2006. [337](#), [340](#), [347](#)
- [47] Alejandro Gil-Ley and Giovanni Bussi. Enhanced conformational sampling using replica exchange with collective-variable tempering. *Journal of chemical theory and computation*, 11(3):1077–1085, 2015. [338](#), [480](#)
- [48] Pratyush Tiwary and Michele Parrinello. From metadynamics to dynamics. *Phys. Rev. Lett.*, 111:230602, Dec 2013. [341](#)
- [49] Andrew White, James Dama, and Gregory A Voth. Designing free energy surfaces that match experimental data with metadynamics. *J. Chem. Theory Comput.*, 11(6):2451–2460, 2015. [341](#)
- [50] Fabrizio Marinelli and José D Faraldo-Gómez. Ensemble-biased metadynamics: A molecular simulation method to sample experimental distributions. *Biophys. J.*, 108(12):2779–2782, 2015. [341](#)
- [51] Alejandro Gil-Ley and Giovanni Bussi. Empirical corrections to the amber rna force field with target metadynamics, 2016. [341](#)
- [52] James F Dama, Michele Parrinello, and Gregory A Voth. Well-tempered metadynamics converges asymptotically. *Phys. Rev. Lett.*, 112(24):240602, 2014. [341](#)
- [53] H. Grubmüller, B. A. Heymann, and P. Tavan. *Science*, 271:997–999, 1996. [344](#)
- [54] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, 1997. [344](#)
- [55] Jim Pfaendtner and Massimiliano Bonomi. Efficient sampling of high-dimensional free-energy landscapes with parallel bias metadynamics. *Journal of Chemical Theory and Computation*, 11(11):5062–5067, 2015. [346](#)
- [56] Stefano Piana and Alessandro Laio. A bias-exchange approach to protein folding. *J. Phys. Chem. B*, 111(17):4553–9, 2007. [377](#)

- [57] Marco Jacopo Ferrarotti, Sandro Bottaro, Andrea Perez-Villa, and Giovanni Bussi. Accurate multiple time step in biased molecular simulations. *J. Chem. Theory Comput.*, 11(1):139–146, 2015. [389](#), [503](#)
- [58] G.M. Torrie and J.P. Valleau. Nonphysical sampling distributions in monte carlo free energy estimation: Umbrella sampling. *J. Comput. Phys.*, 23:187–199, 1977. [415](#)
- [59] Alessandro Laio and Francesco Luigi Gervasio. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.*, 71:126601, 2008. [425](#), [452](#), [464](#), [489](#)
- [60] Alessandro Barducci, Massimiliano Bonomi, and Michele Parrinello. Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(5):826–843, 2011. [425](#), [452](#), [464](#), [489](#)
- [61] Ludovico Sutto, Simone Marsili, and Francesco Luigi Gervasio. New advances in metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(5):771–779, 2012. [425](#), [452](#), [464](#), [489](#)
- [62] Yuji Sugita and Yuko Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.*, 314(1–2):141–151, November 1999. [432](#)
- [63] Giovanni Bussi, Francesco Luigi Gervasio, Alessandro Laio, and Michele Parrinello. Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. *J. Am. Chem. Soc.*, 128(41):13435–41, 2006. [432](#)
- [64] Michael Deighan, Massimiliano Bonomi, and Jim Pfendtner. Efficient simulation of explicitly solvated proteins in the well-tempered ensemble. *Journal of Chemical Theory and Computation*, 8(7):2189–2192, 2012. [432](#), [438](#)
- [65] Andrea Cavalli, Carlo Camilloni, and Michele Vendruscolo. Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle. *J. Chem. Phys.*, 138(9):094112, March 2013. [458](#)
- [66] Carlo Camilloni, Andrea Cavalli, and Michele Vendruscolo. Replica-Averaged Metadynamics. *J. Chem. Theory Comput.*, 9(12):5610–5617, December 2013. [458](#)
- [67] Carlo Camilloni and Michele Vendruscolo. Statistical mechanics of the denatured state of a protein using replica-averaged metadynamics. *J. Am. Chem. Soc.*, 136(25):8982–8991, June 2014. [458](#)
- [68] Wouter Boomsma, Kresten Lindorff-Larsen, and Jesper Ferkinghoff-Borg. Combining Experiments and Simulations Using the Maximum Entropy Principle. *PLoS Comput. Biol.*, 10(2):e1003406, February 2014. [458](#)
- [69] Giovanni Bussi. Hamiltonian replica-exchange in gromacs: a flexible implementation. *Mol. Phys.*, 2013. DOI: 10.1080/00268976.2013.824126. [473](#)
- [70] Lingle Wang, Richard A Friesner, and BJ Berne. Replica exchange with solute scaling: A more efficient version of replica exchange with solute tempering (rest2). *The Journal of Physical Chemistry B*, 115(30):9431–9438, 2011. [474](#)