

Metadynamics

Practical introduction with PLUMED
University of Padova

Thomas Löhr, 29/04/22

Vendruscolo Group, University of Cambridge, tl465@cam.ac.uk

Overview

Morning

1. Installing software
2. Alanine dipeptide (Ala2) without Metadynamics
3. Overview of Metadynamics
4. Accelerated sampling with Metadynamics

Afternoon

5. Postprocessing
6. Metainference (experimental restraints) + Metadynamics on a real-world system

Getting settled

These instructions assume you're using Linux or a Unix-like system like MacOS.

1. Install anaconda (if you haven't already): <https://www.anaconda.com/products/distribution>

2. Create a new environment for our simulations:

```
conda create --name metad-class  
conda activate metad-class
```

3. Install GROMACS (MD code) and PLUMED (Plugin for enhanced sampling), this will also install MPI, this means we can run multi-replica simulations:

```
conda install --strict-channel-priority -c plumed/label/  
masterclass-2022 -c conda-forge plumed gromacs
```

4. Test the installation:

```
gmx_mpi mdrun -h 2> /dev/null | grep -q plumed && echo ok  
plumed --has-mpi && echo ok
```

Getting settled

Installing python libraries for analysis:

```
conda install jupyter numpy matplotlib  
conda install -c conda-forge blas
```

Get the material (including these slides):

```
$ git clone github.com/tlhr/metad-course
```

Archive contents:

```
(plumed-class) $ ls -lh  
total 0  
drwxr-xr-x@  6 thomas  staff   192B  28 Apr 15:53 1-aladi-vanilla/  
drwxr-xr-x@ 18 thomas  staff   576B  28 Apr 15:54 2-aladi-metad/  
drwxr-xr-x  8 thomas  staff   256B  28 Apr 15:55 3-aladi-metad2d/  
drwxr-xr-x@ 29 thomas  staff   928B  28 Apr 18:25 4-egaw-mm/
```

Alanine dipeptide

No metadynamics (1-aladi-vanilla)

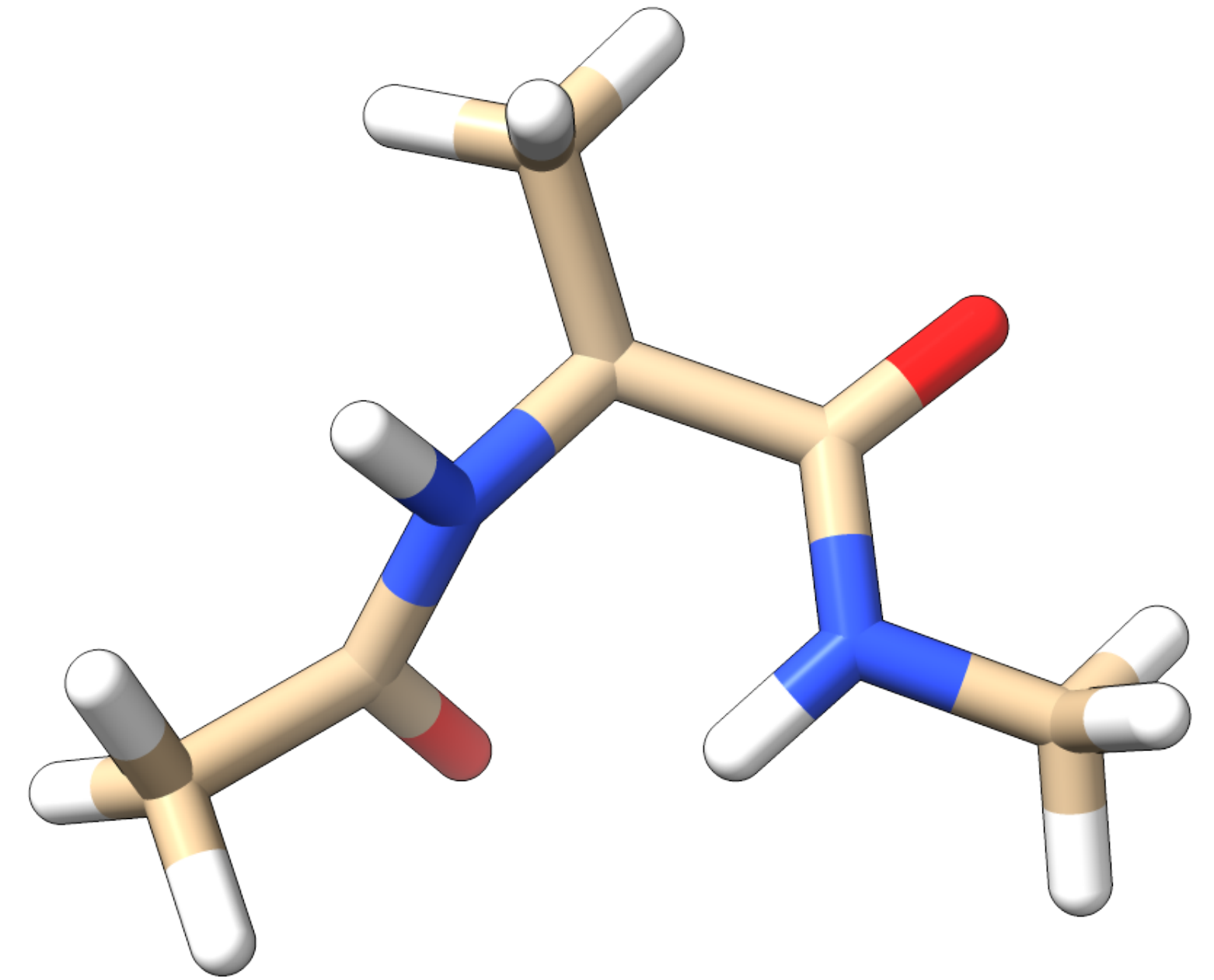
Running:

```
gmx_mpi mdrun -s topol.tpr -v -nstps 5000000
```

Output:

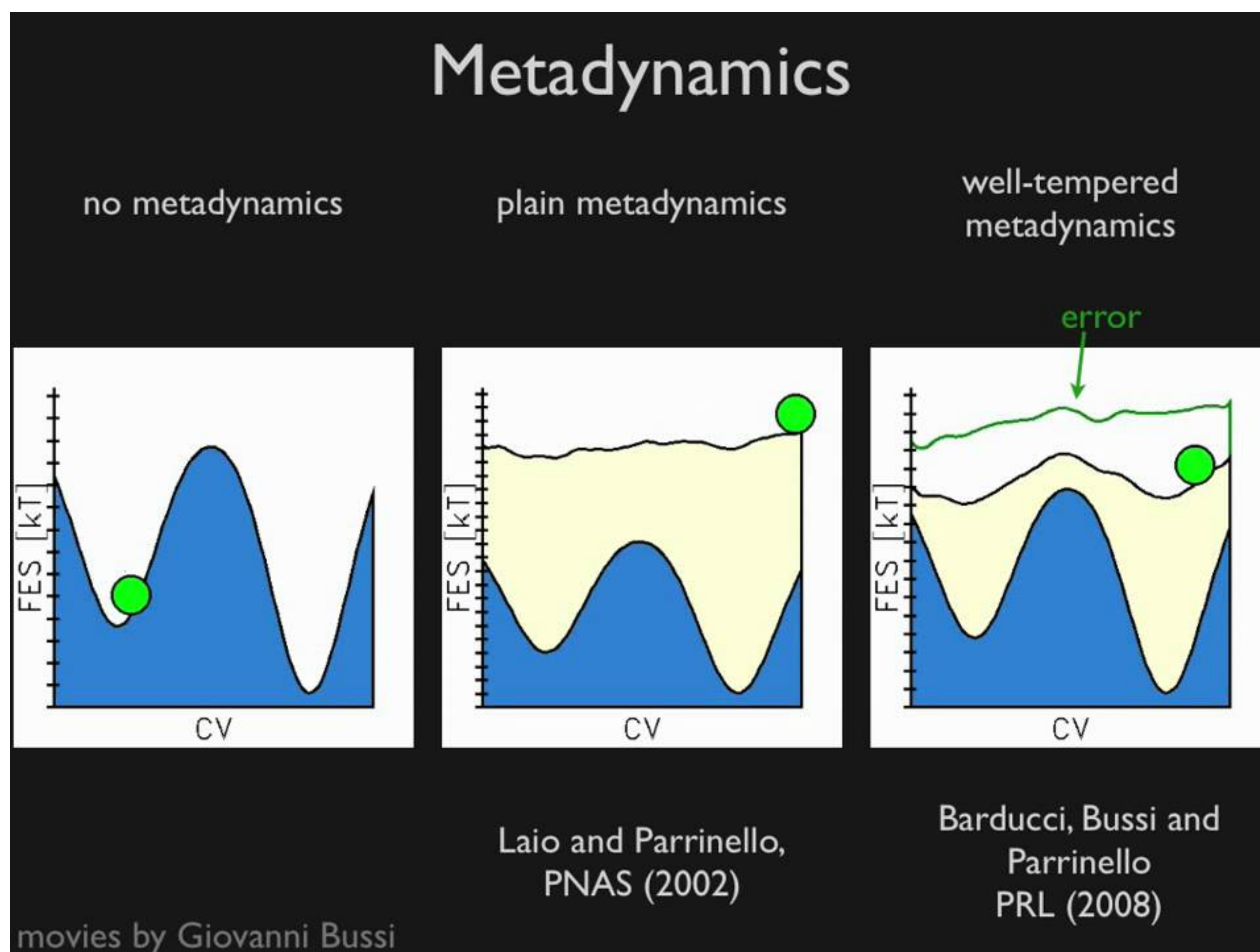
```
$ ls -lh
total 224
-rw-----@ 1 thomas  staff    1.5K  27 Apr  19:38 confout.gro
-rw-----@ 1 thomas  staff    1.7K  27 Apr  19:19 diala.pdb
-rw-----@ 1 thomas  staff   11K  27 Apr  19:38 ener.edr
-rw-----@ 1 thomas  staff   20K  27 Apr  19:38 md.log
-rw-----@ 1 thomas  staff   2.1K  27 Apr  19:38 state.cpt
-rw-----@ 1 thomas  staff   9.6K  27 Apr  19:13 topol.tpr
-rw-----@ 1 thomas  staff   34K  27 Apr  19:38 traj_comp.xtc
```

Final structure
Structure (included in tpr)
Energy over time
Logfile
Checkpoint
Run input file
Trajectory



Metadynamics

Conceptual overview



Update potential with Gaussian hill:

$$\dot{V}_{\text{MetaD}}(s, t) = \omega \exp \left[- \sum_{\alpha=1}^{N_{\text{CV}}} \frac{(s_{\alpha} - s_{\alpha}(t))^2}{2\sigma_{\alpha}^2} \right]$$

$$\dot{V}_{\text{WTMetaD}}(s, t) = \omega \exp \left[- \frac{V(s(t), t)}{k_B \Delta T} \right] \exp \left[- \sum_{\alpha=1}^{N_{\text{CV}}} \frac{(s_{\alpha} - s_{\alpha}(t))^2}{2\sigma_{\alpha}^2} \right]$$

Alanine dipeptide

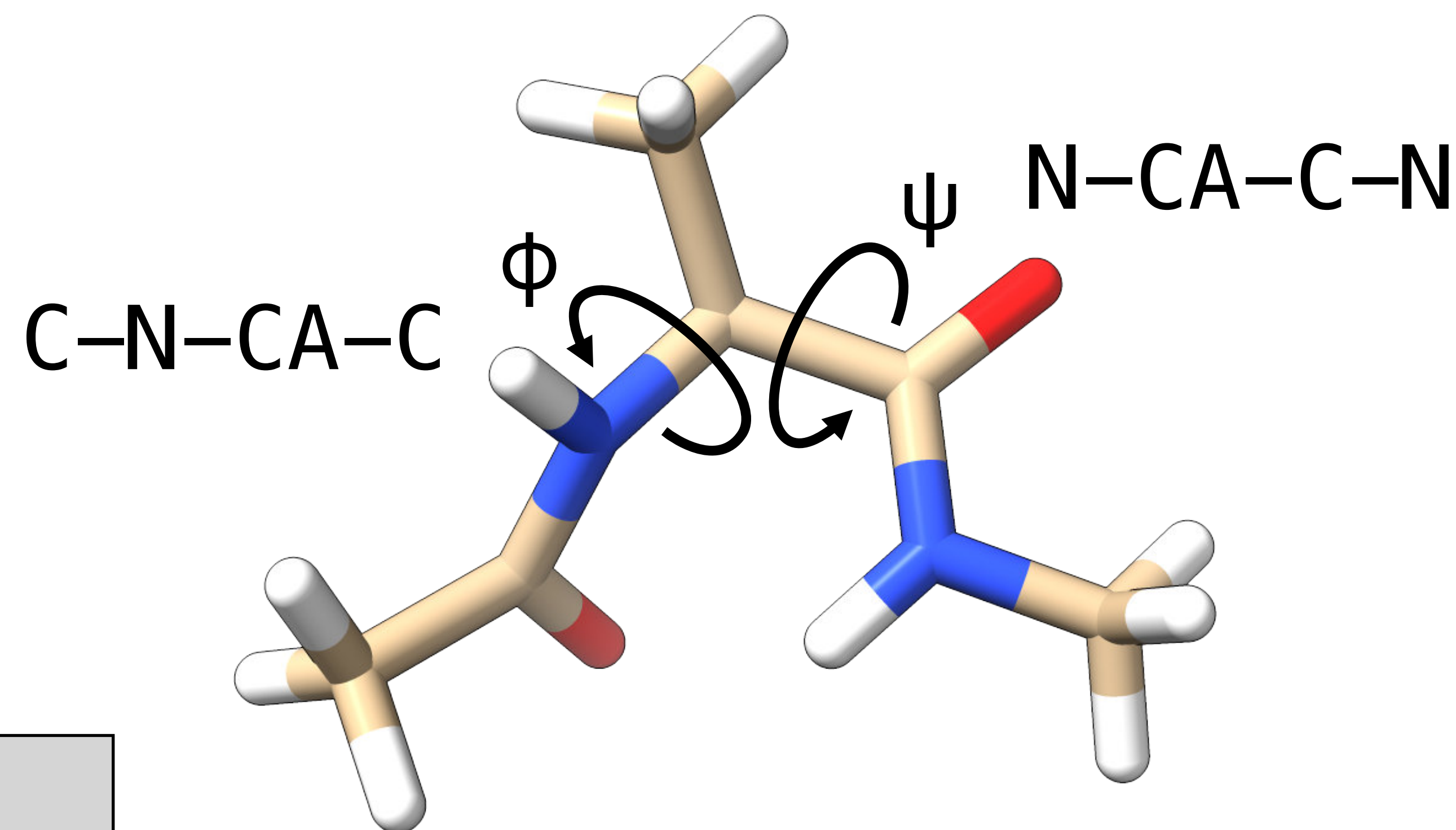
No metadynamics

PLUMED syntax:

```
label: ACTION F00=3.14 BAR=blah
```

PLUMED input (plumed.dat):

```
phi: TORSION ATOMS=????  
psi: TORSION ATOMS=????  
  
PRINT ARG=phi,psi FILE=COLVAR STRIDE=1
```



Analysis:

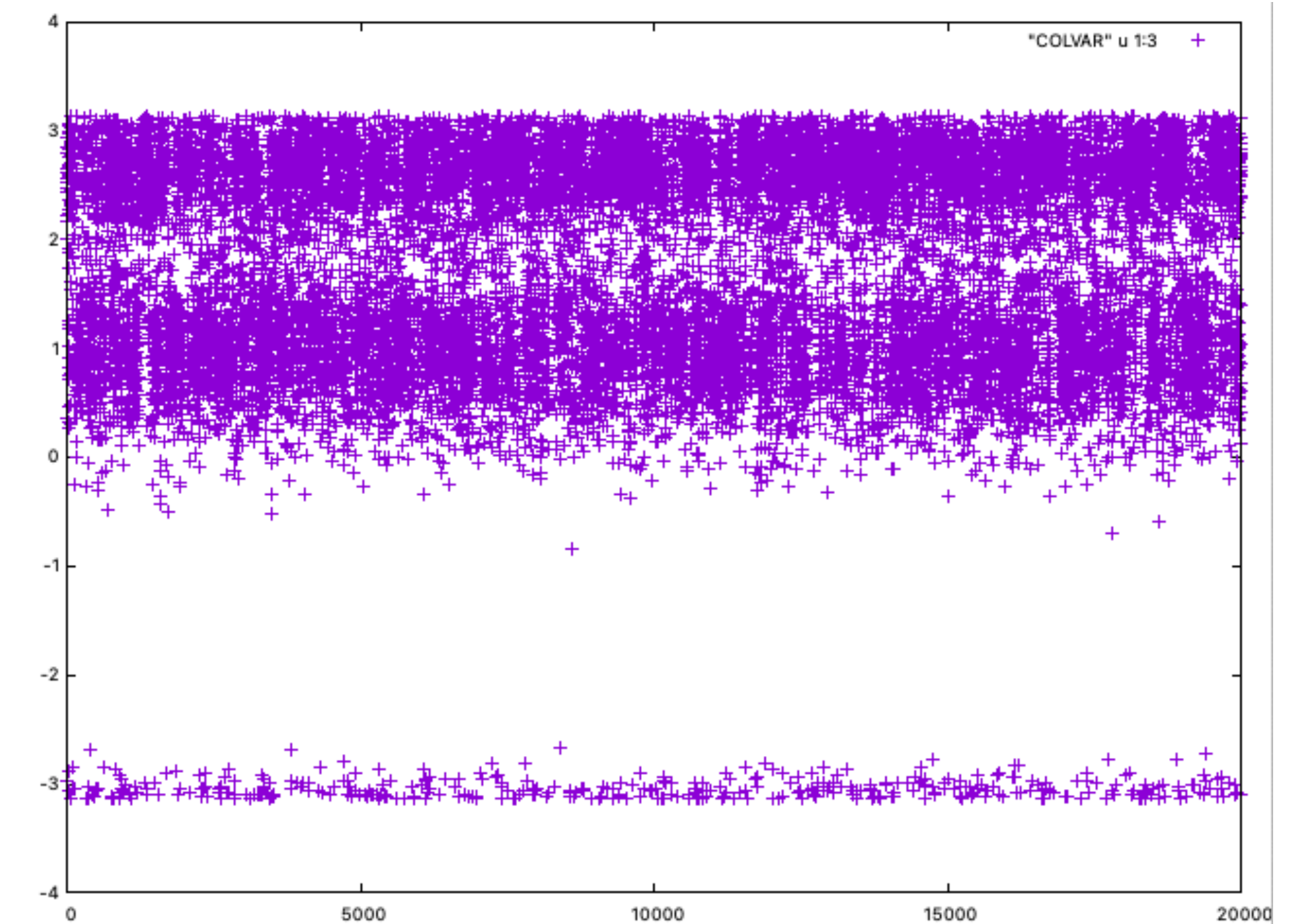
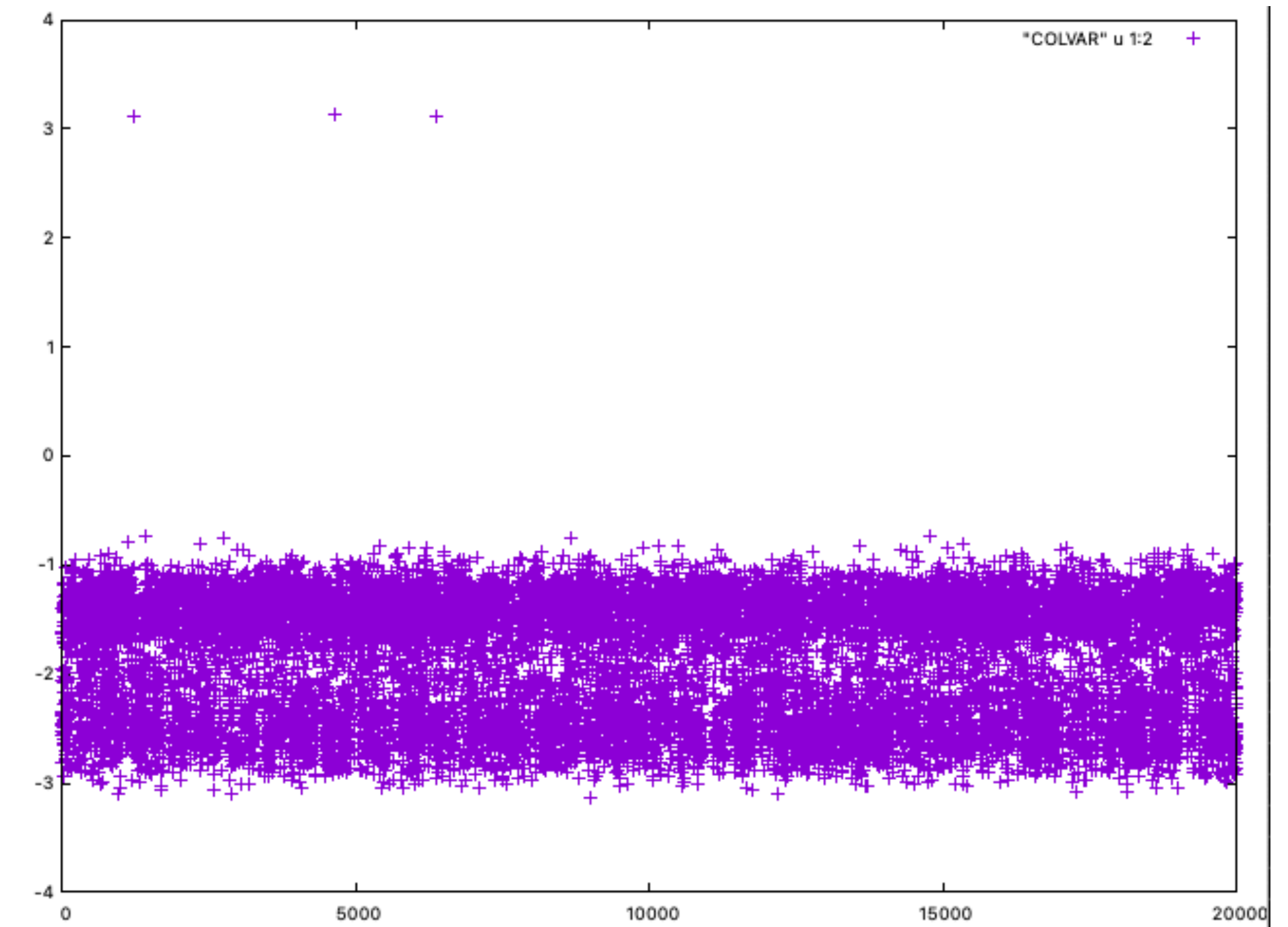
```
plumed driver --plumed plumed.dat --mf_xtc traj_comp.xtc
```

Alanine dipeptide

No metadynamics

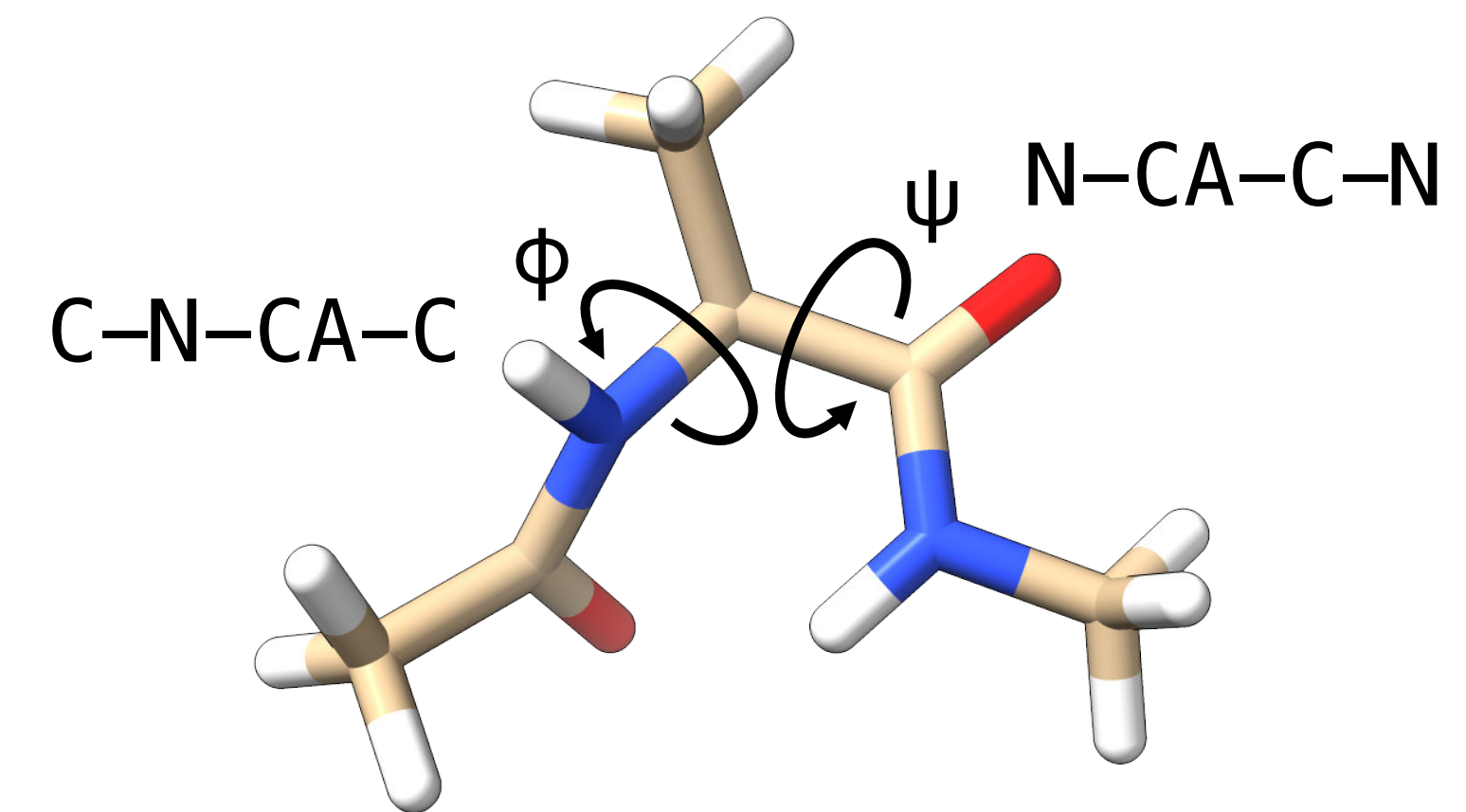
Plotting CV over time:

```
gnuplot
plot "COLVAR" u 1:2 w l
replot "COLVAR" u 1:3 w l
```



Alanine dipeptide

Metadynamics (2-aladi-metad)



PLUMED input (plumed.dat):

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

metad: METAD ...
      ARG=phi
      PACE=500 HEIGHT=1.2 BIASFACTOR=8
      # Choose sigma based on fluctuations in unbiased run (1/2 or 1/3)
      SIGMA=????
      FILE=HILLS GRID_MIN=-pi GRID_MAX=pi
      ...

PRINT ARG=phi,psi,metad.bias FILE=COLVAR STRIDE=100
```

Running:

```
gmx_mpi mdrun -s topol.tpr -v -nsteps 5000000 -plumed plumed.dat
```

Detour: Parallel-bias metadynamics

Biasing many collective variables at once

Problem: Depositing a high-dimensional Gaussian falls victim to the curse of dimensionality

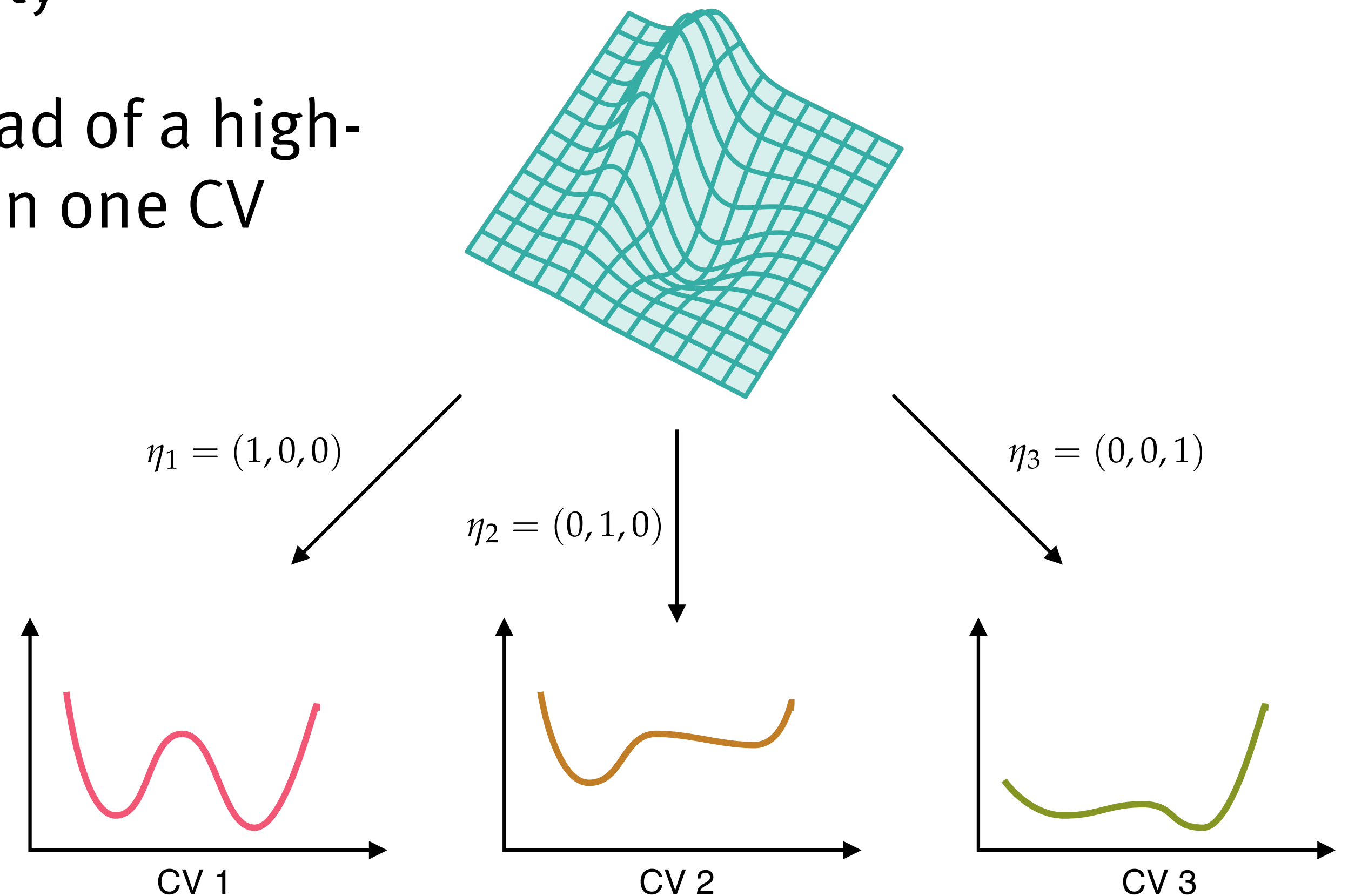
Idea: deposit many 1D-Gaussians instead of a high-dimensional Gaussian, scale the bias on one CV based on the bias on the others

$$V_{\text{PB}}(\mathbf{S}, t) = -\frac{1}{\beta} \log \left[\sum_{S \in \mathcal{S}} \exp(-\beta V_G(S, t)) \right]$$

$$V_G(S_i, t) = k_B \Delta T_i \log \left[1 + \frac{\omega_i N(S_i, \boldsymbol{\eta} = \mathbf{e}_i, t)}{k_B \Delta T_i} \right]$$

$$\omega_i(t) = \omega_i \exp \left(-\frac{V_G(S_i, t)}{k_B \Delta T_i} \right) P(\boldsymbol{\eta}_i \mid \mathbf{R})$$

$$P(\boldsymbol{\eta}_i \mid \mathbf{R}) = \frac{\exp(-\beta V_G(S_i, t))}{\sum_i \exp(-\beta V_G(S_i, t))}$$



Alanine dipeptide

Metadynamics postprocessing

Plotting CV over time:

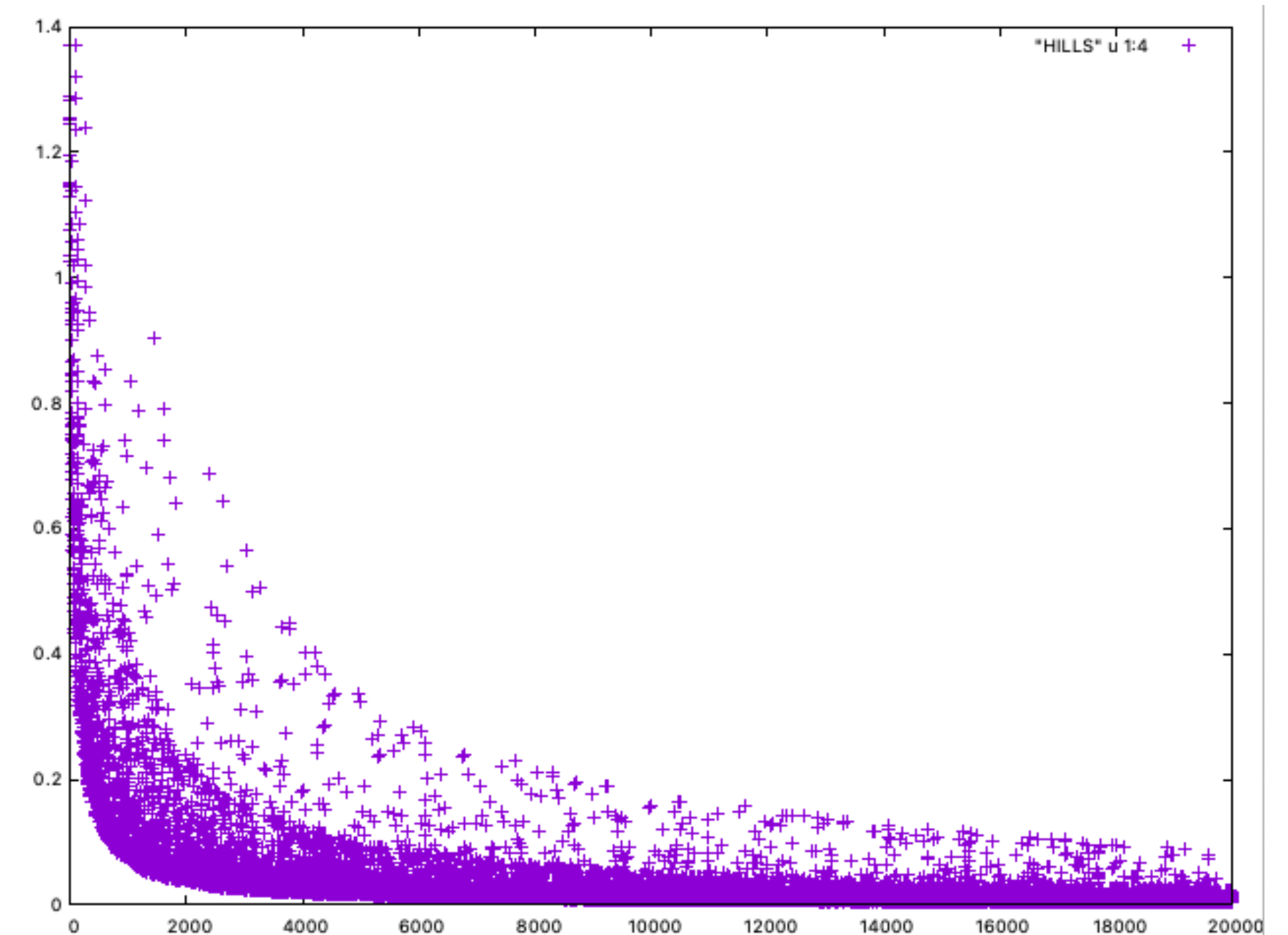
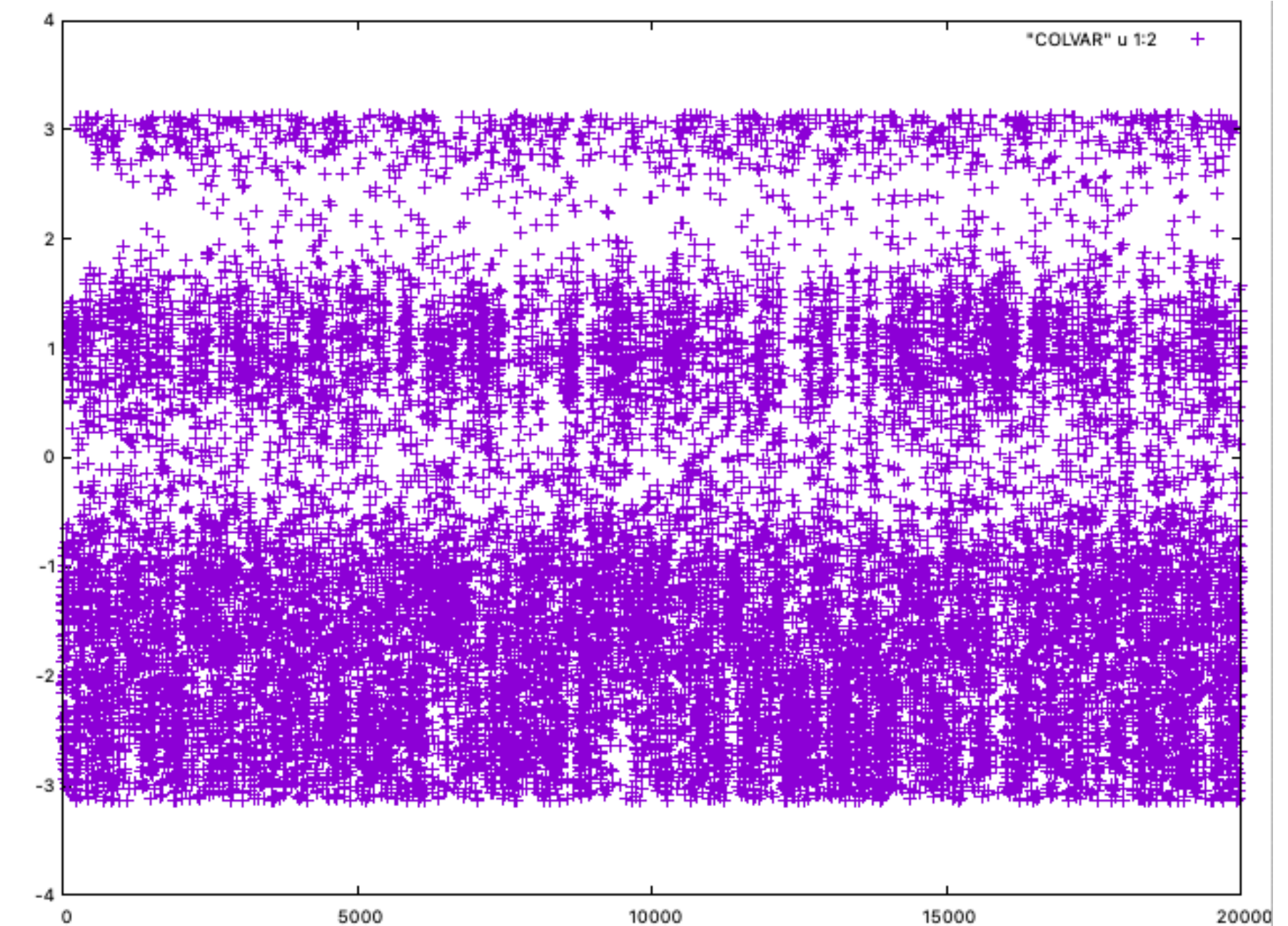
```
gnuplot
plot "COLVAR" u 1:2
replot "COLVAR" u 1:3
```

HILLS file content:

```
#! FIELDS time phi sigma_phi height biasf
#! SET multivariate false
#! SET kerneltype stretched-gaussian
#! SET min_phi -pi
#! SET max_phi pi
```

Plot the gaussian height over time:

```
gnuplot
plot "HILLS" u 1:4
```



Alanine dipeptide

Metadynamics postprocessing

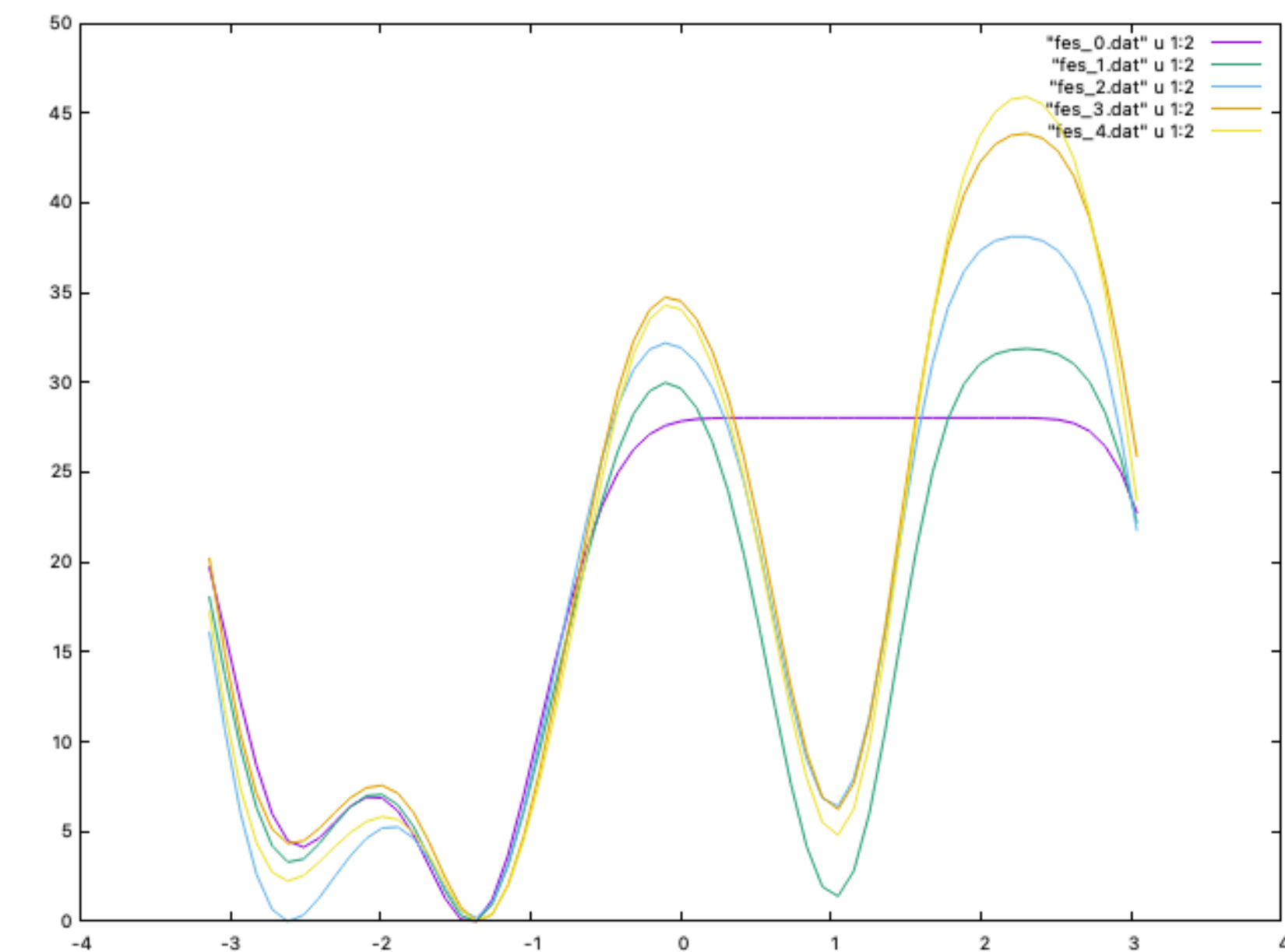
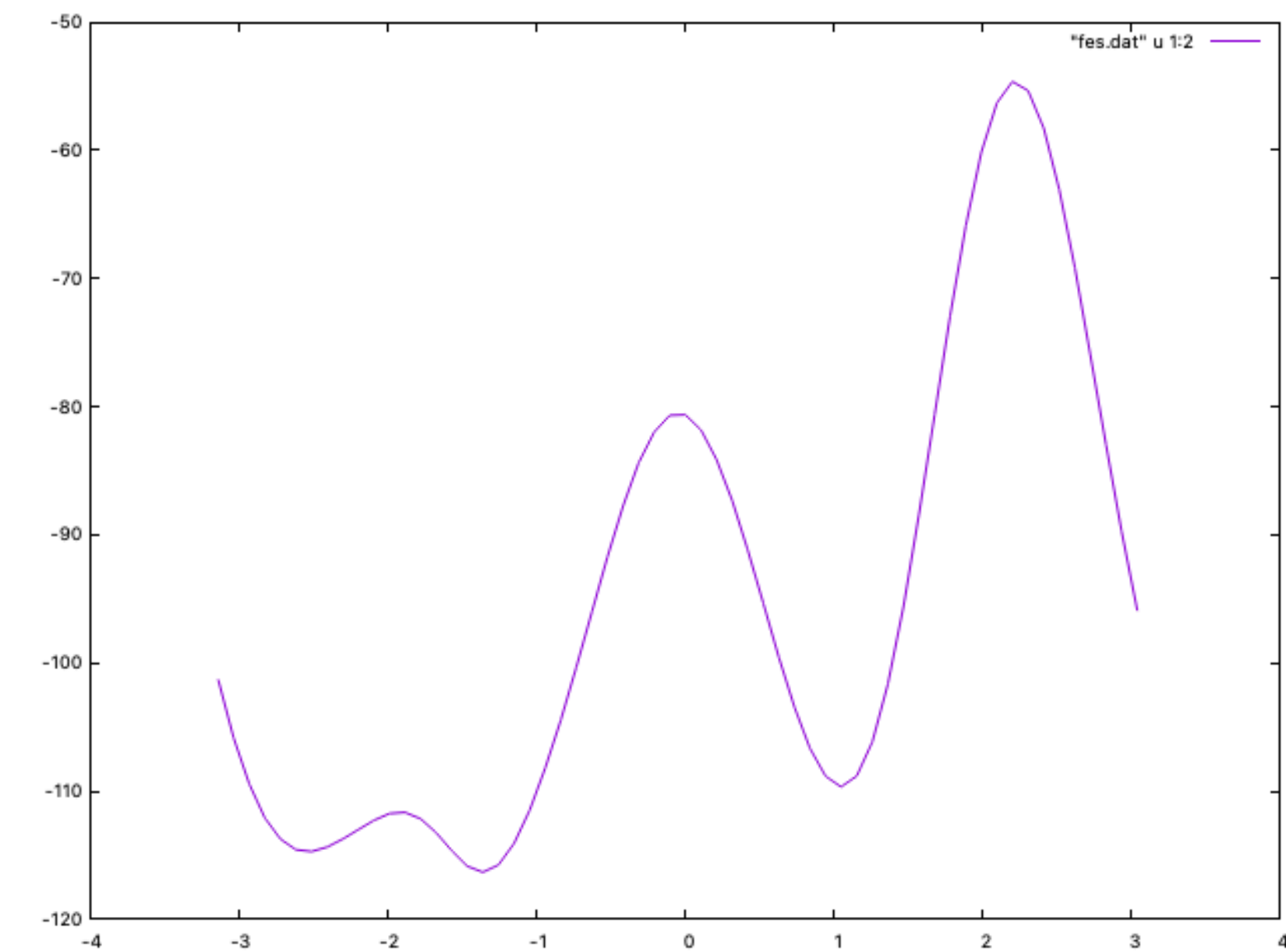
Estimate the free energy surface:

```
plumed sum_hills --hills HILLS
```

Estimate the free energy surface over time:

```
plumed sum_hills --hills HILLS --stride 100 -mintozero  
# Plot the first few fes.dat files
```

Is it converged?



Alanine dipeptide

Metadynamics postprocessing

Un-biasing the simulation:

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

metad: METAD ...
      ARG=phi
      PACE=10000000 HEIGHT=0.0 BIASFACTOR=8
      SIGMA=???? # Use your value here!
      FILE=HILLS GRID_MIN=-pi GRID_MAX=pi
      RESTART=YES
...

PRINT ARG=phi,psi,metad.bias FILE=COLVAR_RW STRIDE=1
```

Run on the trajectory:

```
plumed driver --plumed plumed-rw.dat --mf_xtc traj_comp.xtc --kt 2.494339
```


Alanine dipeptide

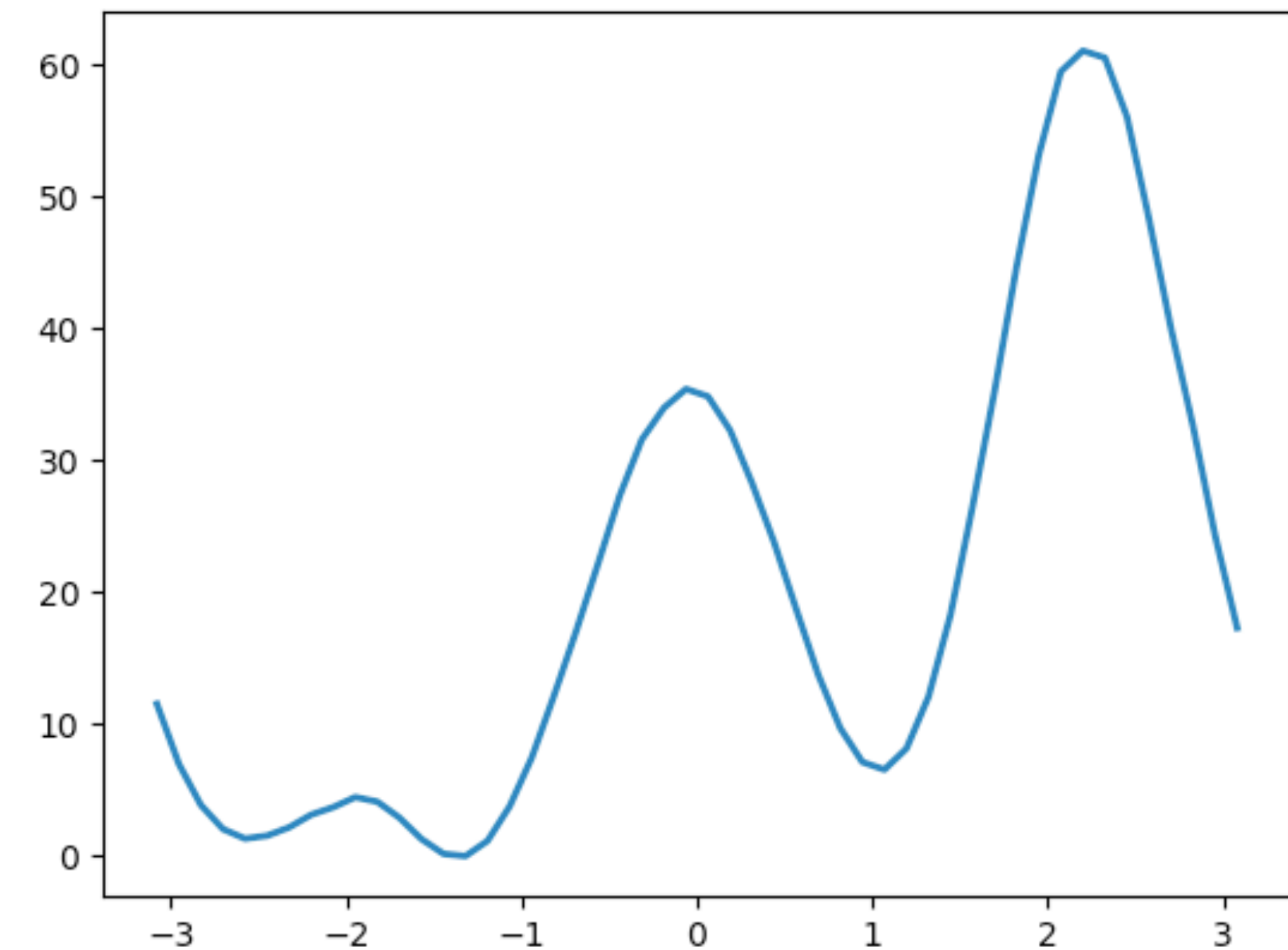
Metadynamics postprocessing

Calculating weights:

```
import numpy as np
```

```
bias = np.loadtxt("COLVAR_RW")[:, -1]  
weights = np.exp(bias / 2.494339)  
weights /= weights.sum()
```

$$w \propto \exp \left[\frac{V_{\text{MetaD}}(s)}{k_B T} \right]$$



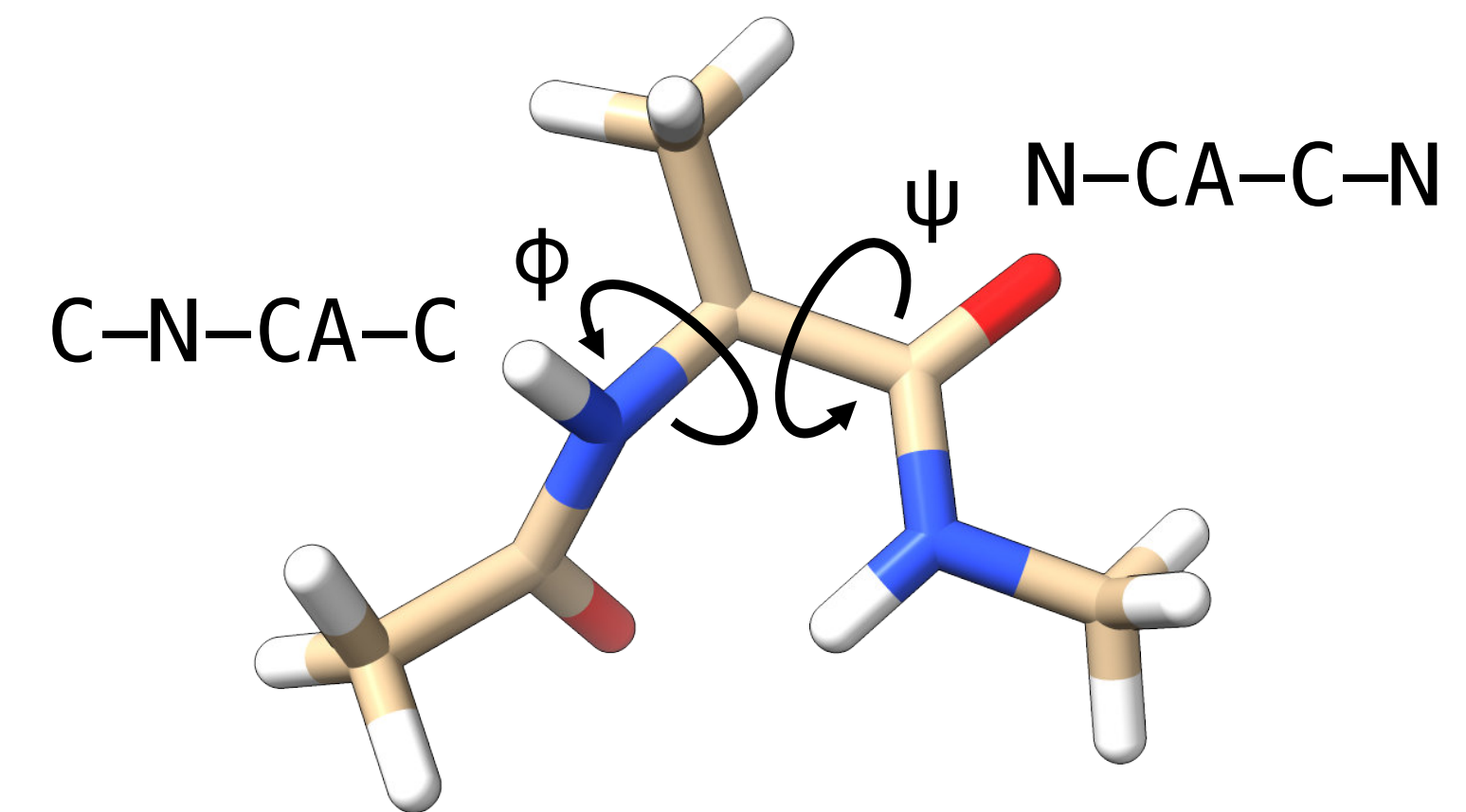
Creating a 1D free energy surface:

```
import matplotlib.pyplot as plt
```

```
phi, psi = np.loadtxt("COLVAR_RW")[:, [1, 2]].T  
hist, edges = np.histogram(phi, bins=50, range=(-np.pi, np.pi),  
                           density=True, weights=weights)  
x = 0.5 * (edges[:-1] + edges[1:])  
F = -2.494339 * np.log(hist)  
F -= F.min()  
plt.plot(x, F); plt.show()
```

Alanine dipeptide

2D Metadynamics (3-aldi-metad2d)



PLUMED input (plumed.dat):

```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

metad: METAD ...
      ARG=phi,psi
      PACE=500 HEIGHT=1.2 BIASFACTOR=14
      # Choose sigma based on fluctuations in unbiased run (1/2 or 1/3)
      SIGMA=????,????
      FILE=HILLS GRID_MIN=-pi,-pi GRID_MAX=pi,pi
      ...

PRINT ARG=phi,psi,metad.bias FILE=COLVAR STRIDE=100
```

Running:

```
gmx_mpi mdrun -s topol.tpr -v -nsteps 5000000 -plumed plumed.dat
```

Alanine dipeptide

2D Metadynamics postprocessing

PLUMED input (plumed.dat):

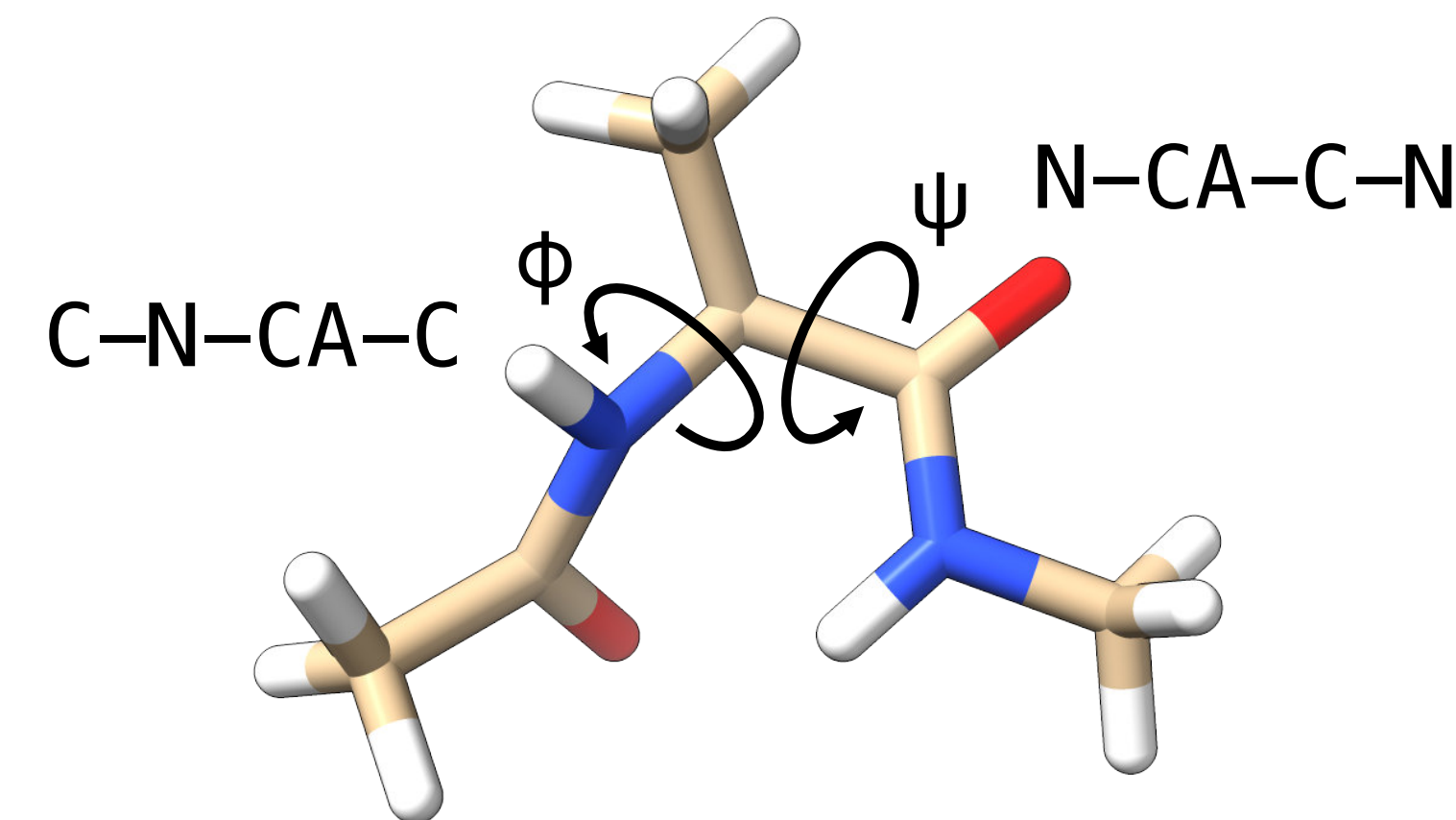
```
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

metad: METAD ...
  ARG=phi,psi
  PACE=10000000 HEIGHT=0.0 BIASFACTOR=14
  # Choose previous values
  SIGMA=????,????
  FILE=HILLS GRID_MIN=-pi,-pi GRID_MAX=pi,-pi
  RESTART=YES
...

PRINT ARG=phi,psi,metad.bias FILE=COLVAR STRIDE=100
```

Running:

```
plumed driver --plumed plumed-rw.dat --mf_xtc traj_comp.xtc --kt 2.494339
```



Alanine dipeptide

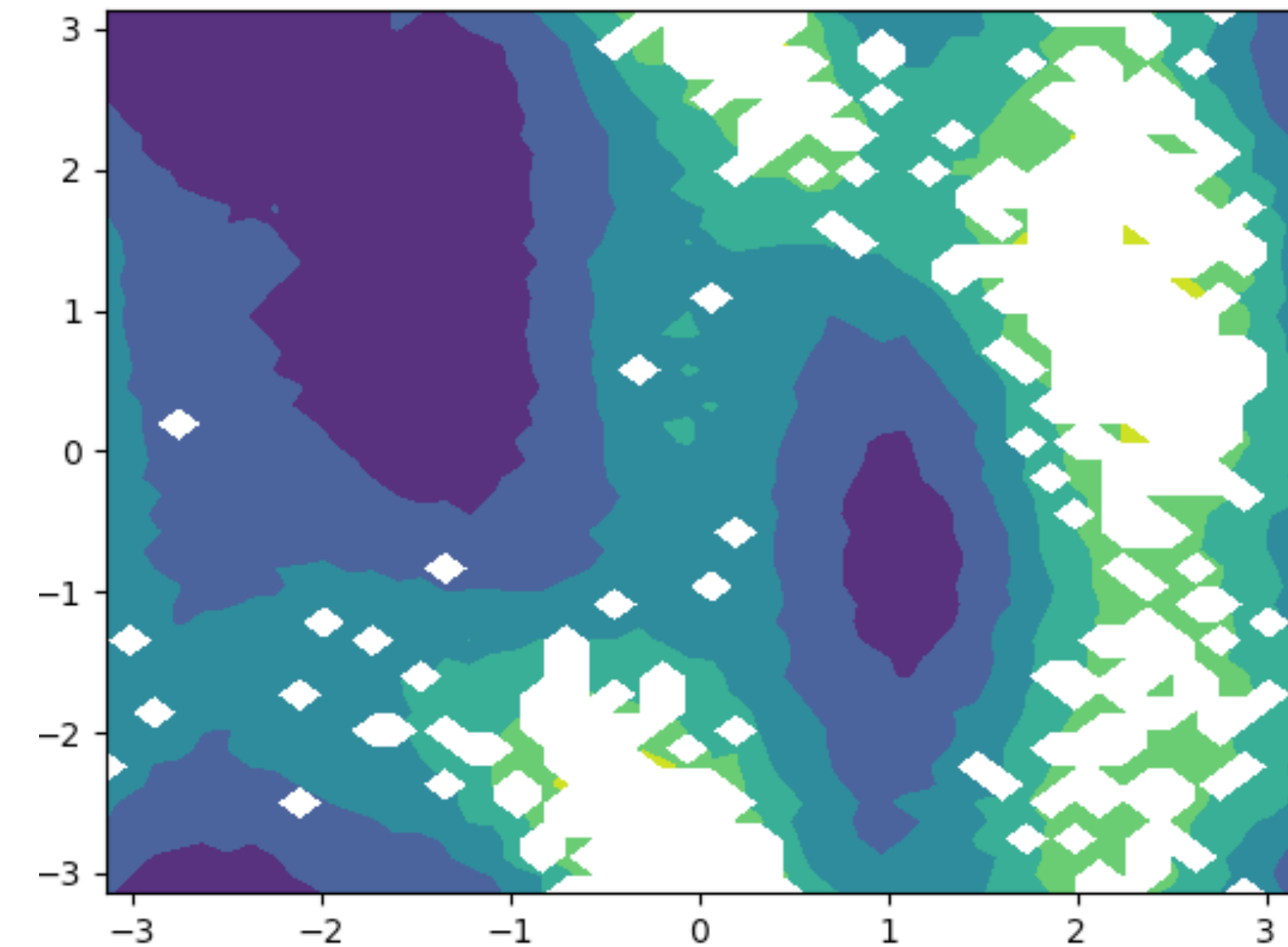
2D Metadynamics postprocessing

Calculate weights as before $w \propto \exp \left[\frac{V_{\text{MetaD}}(s)}{k_B T} \right]$

Create a 2D free energy surface:

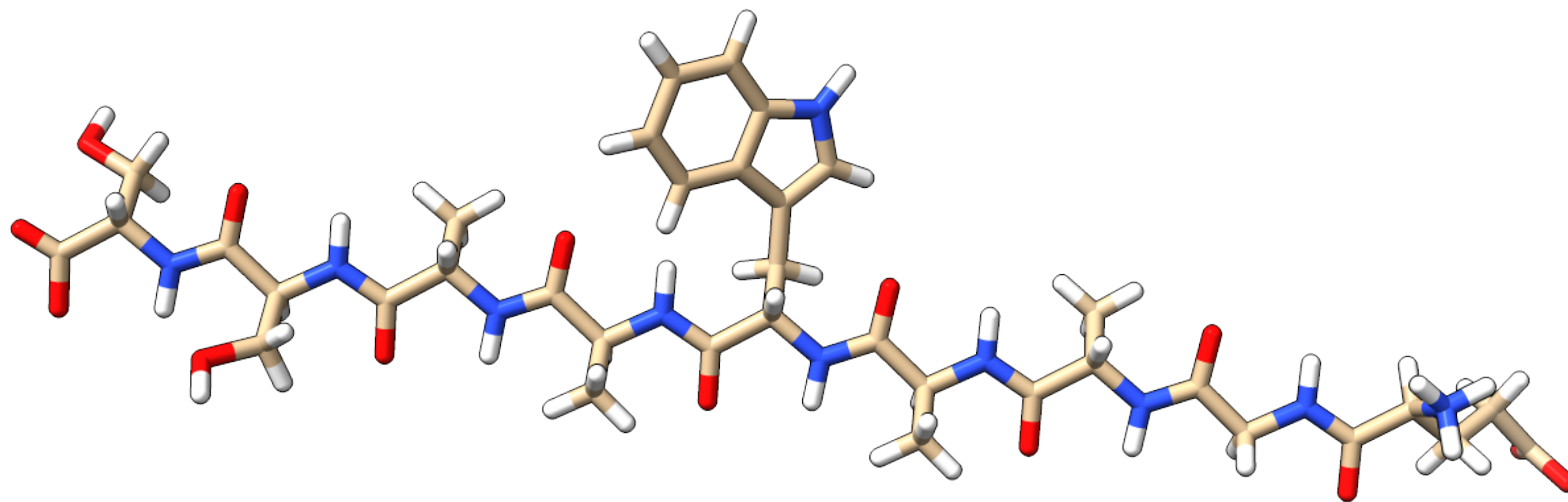
```
import matplotlib.pyplot as plt

phi, psi = np.loadtxt("COLVAR_RW")[:, [1, 2]].T
hist, edges = np.histogram2d(
    phi, psi, bins=50, density=True, weights=weights,
    range=(-np.pi, np.pi), (-np.pi, np.pi))
xx, yy = np.mgrid[-np.pi:np.pi:50j, -np.pi:np.pi:50j]
F = -2.494339 * np.log(hist)
F -= F.min()
plt.contourf(xx, yy, F); plt.show()
```



Real system: EGAWAASS peptide

Metadynamics + experimental restraints (4-egaw-mm)



Enhance sampling with metadynamics,
and add experimental restraints (from
NMR) to improve accuracy

Real system: EGAAWAASS peptide

Metadynamics + experimental restraints

To apply restraints we need a *forward model* mapping from our structure (XYZ coordinates) to our experimental observable (typically a scalar):

$$f : \text{Structure} \rightarrow \mathbb{R}$$

Use the *Karplus* relation to go from dihedral angles to J-coupling:

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

Then apply a restraint in form of a harmonic potential:

$$V(X, \sigma \mid d) \propto \frac{(f(X) - d)^2}{\sigma^2}$$

Real system: EGAAWAASS peptide

Metainference

$$V_{\text{MI}}(\mathbf{X}, \boldsymbol{\sigma} \mid \mathbf{d}) = \sum_{r=1}^{N_r} \left\{ V_{\text{FF}}(X_r) + k_B T \sum_{i=1}^{N_d} \left[\frac{(\langle f_i(\mathbf{X}) \rangle - d_i)^2}{2\sigma_{r,i}^2} + \frac{1}{2} \ln(2\pi\sigma_{r,i}^2) + \frac{1}{2} \ln \left(\frac{\sigma_{r,i}^2}{2} \right) \right] \right\}$$

$\mathbf{X} = [X_r]$ Ensemble of conformations

$\boldsymbol{\sigma} = [\sigma_{r,i}]$ Errors

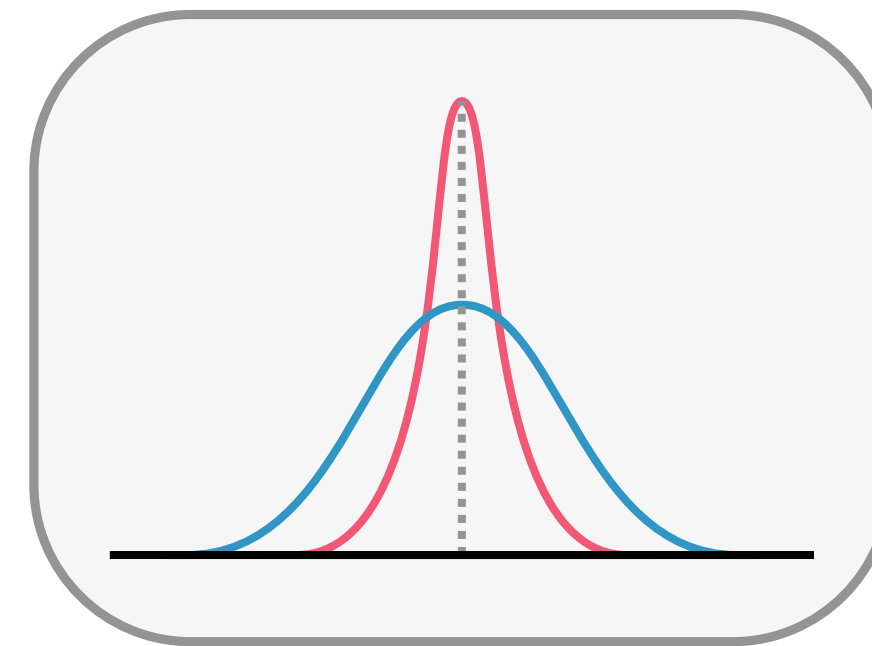
$\mathbf{d} = [d_i]$ Datapoints

$f(\mathbf{X})$ Forward model

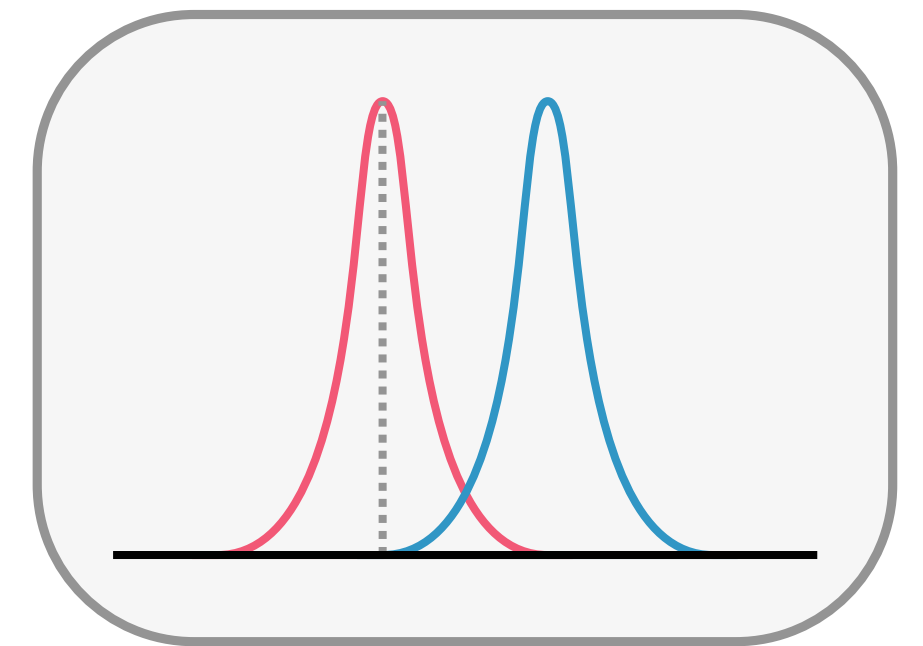
N_d Number of datapoints

N_r Number of replicas

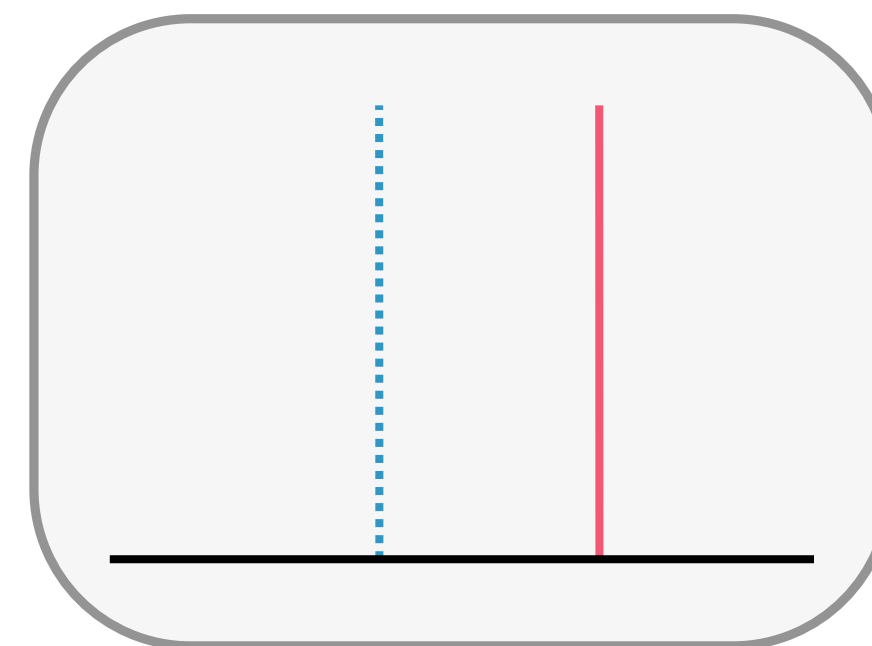
Random error



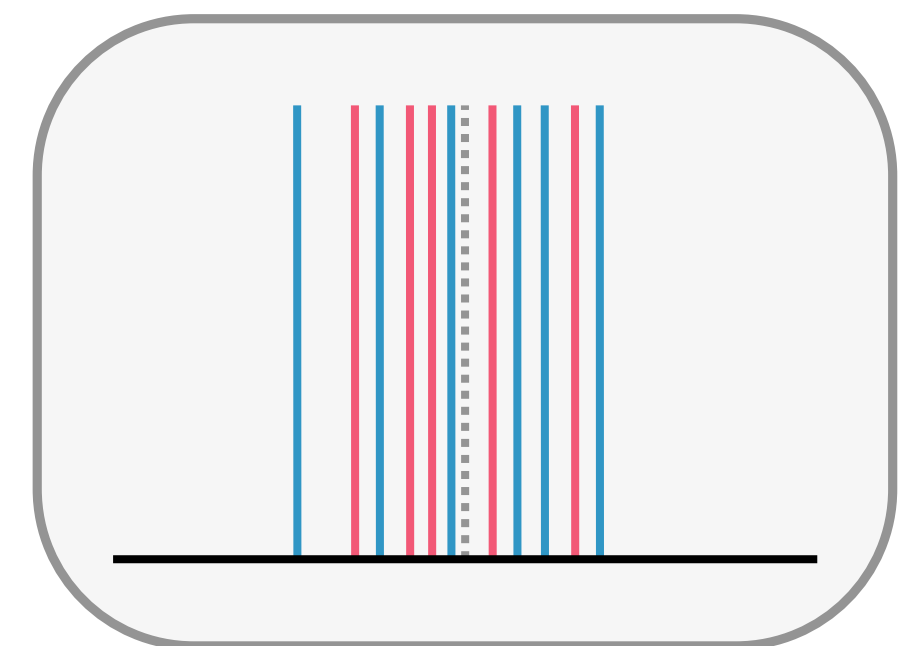
Systematic error



Forward model error



Ensemble error



Real system: EGAAWAASS peptide

Metadynamics + experimental restraints

Forward model in PLUMED (H^α–N, based on Ψ dihedrals):

```
JCOUPLING ...  
  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  
...
```

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

Real system: EGAAWAASS peptide

Metadynamics + experimental restraints

Ensemble-averaged statistics:

```
ENSEMBLE ...  
    ARG=(jhan\.j-.*),pb.bias REWEIGHT  
    LABEL=ens  
... ENSEMBLE  
  
STATS ...  
    ARG=(ens\.jhan\.j-.* ) PARARG=(jhan\.exp-.* )  
    LABEL=jhanst  
... STATS  
  
PRINT ARG=jhanst.* STRIDE=2000 FILE=../STATS.J
```

Weighted instantaneous ensemble average

$$w(s_r, t) = \exp\left(\frac{V(s_r, t)}{k_B T}\right) \left[\sum_{j=1}^{N_{\text{replica}}} \exp\left(\frac{V(s_j, t)}{k_B T}\right) \right]^{-1}$$

Statistics (e.g. correlation of forward model with experimental value)

Print to file

Real system: EGAAWAASS peptide

Metadynamics + experimental restraints

Ensemble-averaged restraints with error-accounting:

```
METAINFERENCE ...  
  ARG=(jhan\.j-.*),(jhahn\.j-.*),pb.bias  
  PARARG=(jhan\.exp-.*),(jhahn\.exp-.*)  
  NOISETYPE=MGAUSS REWEIGHT OPTSIGMA MEAN=SEM AVERAGING=200  
  SIGMA0=5.0 SIGMA_MIN=0.0001 SIGMA_MAX=15.0 DSIGMA=0.1  
  WRITE_STRIDE=10000  
  LABEL=byj  
...
```

$$V_{\text{MI}}(\mathbf{X}, \sigma \mid \mathbf{d}) = \sum_{r=1}^{N_r} \left\{ V_{\text{FF}}(X_r) + k_B T \sum_{i=1}^{N_d} \left[\frac{(\langle f_i(\mathbf{X}) \rangle - d_i)^2}{2\sigma_{r,i}^2} + \frac{1}{2} \ln(2\pi\sigma_{r,i}^2) + \frac{1}{2} \ln\left(\frac{\sigma_{r,i}^2}{2}\right) \right] \right\}$$

$$\sigma_{r,i}^2 = (\sigma_{r,i}^{\text{B}})^2 + (\sigma_i^{\text{SEM}})^2$$

Sampled by Monte Carlo

Estimated based on windowed average (AVERAGING)

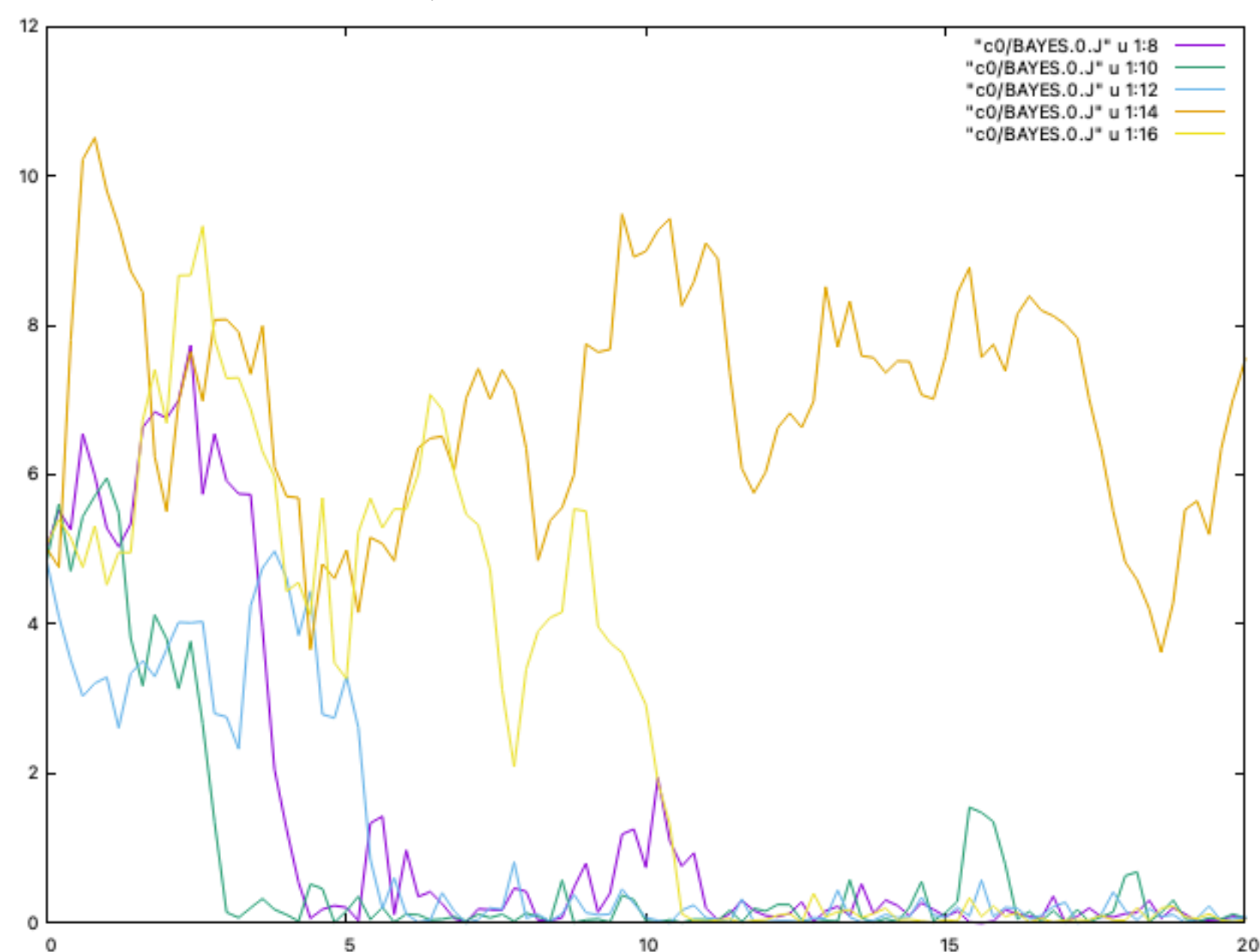
Real system: EGAAWAASS peptide

Metadynamics + experimental restraints

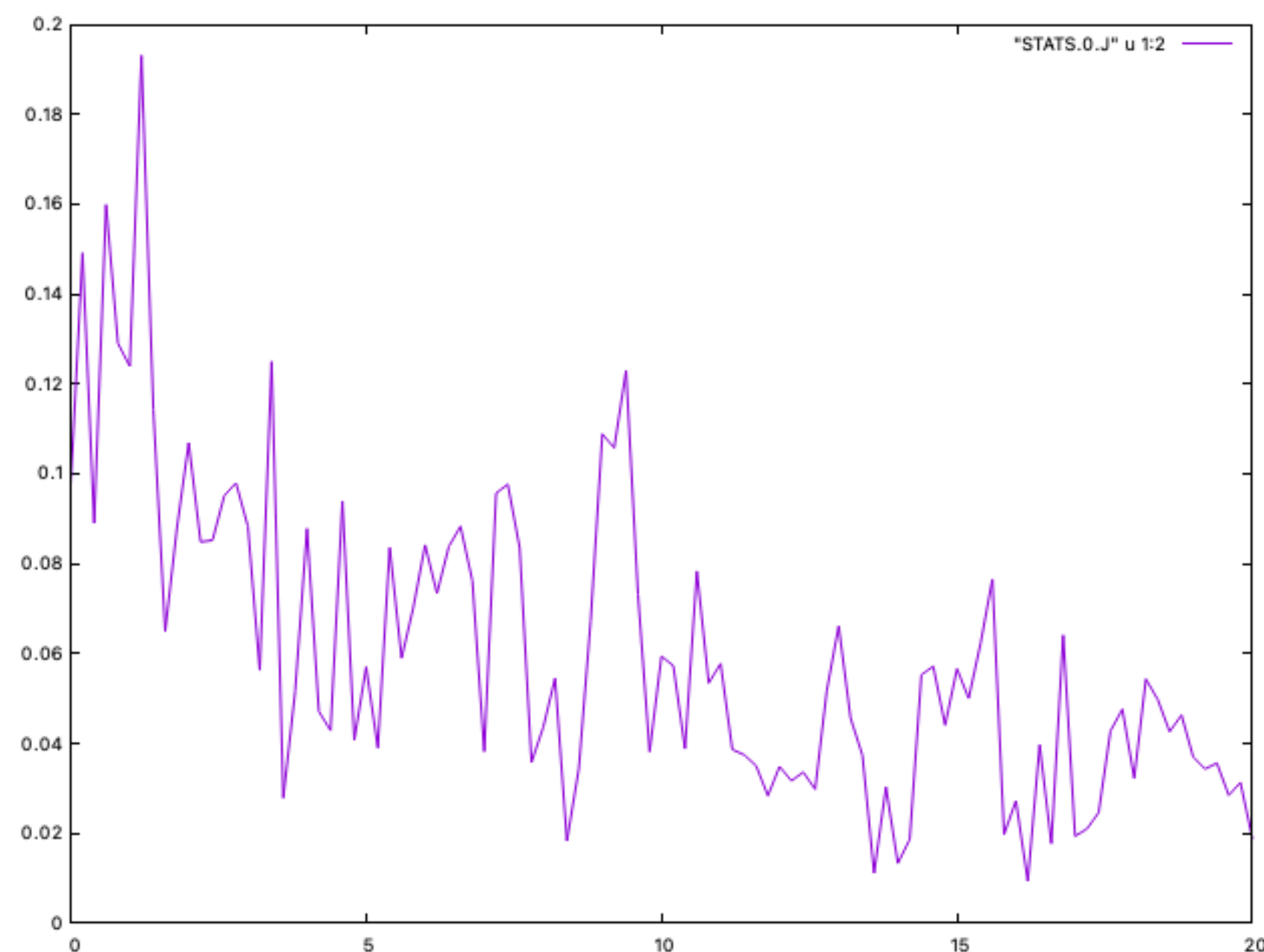
Running in multi-replica mode (requires at least 4 cores):

```
mpirun -np 4 gmx_mpi mdrun -s topol -plumed ../plumed.dat  
-multidir c? -v -nsteps 10000
```

Per-datapoint sigma over time



Squared deviation over time



Useful resources

- GROMACS
 - Simple tutorial: <http://www.mdtutorials.com/gmx/lysozyme/>
 - User guide: <https://manual.gromacs.org/documentation/current/user-guide/index.html>
- PLUMED
 - Brief intro: https://www.plumed.org/doc-v2.8/user-doc/html/master-i_s_d_d-1.html
 - Metadynamics: https://www.plumed.org/doc-v2.8/user-doc/html/master-i_s_d_d-2.html
 - Metainference: <https://www.plumed.org/doc-v2.8/user-doc/html/isdb-1.html>
 - Reference: <https://www.plumed.org/doc-v2.8/user-doc/html/index.html>
- Python
 - Standard library: <https://docs.python.org/3/>
 - Numpy: <https://numpy.org/doc/stable/>

Relevant papers / books

- Metadynamics
 - How-to guide — Bussi & Branduardi
 - Well-tempered metadynamics — Barducci 2008
 - Parallel-bias metadynamics — Pfaendtner & Bonomi 2015
 - Adaptive gaussians — Branduardi 2012
- Forward models
 - J-couplings — Karplus 1963
 - 3-J coupling A , B , C parameters — Pérez 2001
- Metainference
 - Original method — Bonomi & Camilloni 2015
 - With metadynamics — Bonomi & Camilloni 2016
 - Practical guide on the EGAAWAASS peptide