

Chapter 1

Rare events

1.1 Path-metadynamics

In this exercise, we will perform metadynamics [1] and path-metadynamics (PMD) [2, 3, 4] simulations of a system that can undergo a transition between two free energy minima. These two stable states are separated by a transition state barrier, so that the transition process is a *rare event* for “normal” (i.e. unbiased) MD simulations. PMD is able to simultaneously and efficiently converge the transition path (which can be either an average transition path or a minimum free energy path (MFEP)) and the free energy profile of the process along this path. This is achieved by performing metadynamics on an adaptive path connecting two known stable states. In order to understand the principles and advantages of PMD, we will first perform standard metadynamics and get familiar with its parameters and functionalities. Subsequently, we will add the features corresponding to the adaptive path. To ease computation and visualization we will employ a simple model system. This exercise is implemented in PLUMED [5], a popular open-source library that can be interfaced with many well-known MD engines, which means that you can potentially use it for your own research! Here, to keep the practicum general, we will use PLUMED’s own internal engine: *PLUMED pesmd*.

1.1.1 Introduction

In order to overcome the timescale limitations of MD and be able to observe relevant molecular transitions, the community has developed a plethora of enhanced sampling methods. Among them, we can identify the so-called biasing methods, which work by exerting an external potential (or force) on a few key descriptive degrees of freedom of the transition. We refer to these descriptors as collective variables (CVs), a concept similar to an order parameter, or a reaction coordinate. The choice and design of CVs is a task of great importance. In the following, we will assume that the CVs are known. We will consider a molecular system whose free energy surface (FES) is fully described by a set of N CVs, $\{z_i(\vec{q})\}$ with $i = 1 \dots N$, which are all functions of the system’s particle positions $\vec{q}(t)$. We will also assume that these particle positions $\vec{q}(t)$ evolve in time with a canonical equilibrium distribution at temperature T under a potential $V(\vec{q})$.

1.1.2 Metadynamics

We are interested in exploring the space spanned by the CVs, $\{z_i\}$, and in quantifying the free energy. In order to do so, in metadynamics we exert a history-dependent repulsive potential by summing Gaussian kernels over time [1]:

$$V_{\text{bias}}(\vec{z}, t) = \sum_{\kappa\tau < t} H(\kappa\tau) \exp\left(-\sum_{i=1}^N \frac{(z_i - z_i(q(\kappa\tau)))^2}{2W_i^2}\right), \quad (1.1)$$

with Gaussian height H , widths along each CV W_i , and depositing frequency τ . This bias drives the system away from already visited configurations and into new regions of CV-space. More importantly, in the long-time limit, the bias potential converges to the minus free energy as function of the CVs: $V_{\text{bias}}(\vec{z}, t \rightarrow \infty) = -F(\vec{z})$.

As a well-established sampling technique, metadynamics also has several extensions to accelerate convergence and improve computational performance. Consult Refs. [6] and [7] to learn more about these developments, like the multiple-walker and well-tempered versions.

Metadynamics can handle a few CVs simultaneously in a trivial manner, which spares us from having to find a single perfect CV. Instead, an appropriate set of CVs allows us to converge an insightful multidimensional free energy landscape, in which (meta)stable states and connecting MFEPs can be identified. In practice however, especially when investigating complex transitions, the number of CVs is limited to $N \approx 3$, because of the exponential growth of computational cost with CV dimensionality.

1.1.3 Path-metadynamics

In order to tackle complex transitions that require many CVs, path-based methods were developed. In these schemes, we introduce a path-CV. That is, a parametrized curve that connects two FES basins — the known stable states A and B — in the space spanned by the CVs: $\vec{s}(\sigma) : \mathbb{R} \rightarrow \mathbb{R}^N$, where the parameter $\sigma(\vec{z})|_s : \mathbb{R}^N \rightarrow [0, 1]$ yields the progress along the path from A to B , such that $\vec{s}(0) \in A$ and $\vec{s}(1) \in B$. This progress parameter is a well-defined reaction coordinate and can in principle be connected to the committor value. Since the transition path curve is not known a priori, the path-CV must be adaptive. If we assume that, in the vicinity of the path, the iso-committor planes S_σ are perpendicular to $\vec{s}(\sigma)$, and that the configurational probability is a good indicator of the transition flux density, then we can converge the average transition path by iteratively adapting a guess path to the cumulative average density: $\vec{s}_g(\sigma_g) = \langle \vec{z} \rangle_{\sigma_g}$ [3].

The key advantage of PMD is that the metadynamics biasing can be performed on the 1D progress parameter along the adaptive path-CV, instead of the ND CV-space. PMD is able to converge a path, and the free energy along it, with a sublinear rise in computational cost with respect to the growth in CV dimensionality (instead of the exponential relation when not using a path-CV) [4].

Numerically, the adaptive path-CV is implemented as a set of M ordered nodes (coordinates in CV-space) $\vec{s}_g(\sigma_g, t) \rightarrow \vec{s}_j^{t_i}$ with $j = 1, \dots, M$, where t_i is the discrete time of path updates. The projection of any point \vec{z} onto the path is done considering the closest two nodes, and σ is obtained by interpolation. This approach requires equidistant nodes, which are imposed by a reparametrization step after each path update. The update step for the path nodes is done by:

$$\begin{aligned} \vec{s}_j^{t_{i+1}} &= \vec{s}_j^{t_i} + \frac{\sum_k w_k \cdot (\vec{z}_k - \vec{s}^{t_i}(\sigma(\vec{z}_k)))}{\sum_k \xi^{t_i-k} w_{j,k}}, \text{ with} \\ w_k &= \max \left[0, \left(1 - \frac{\|\vec{s}_j^{t_i} - \vec{s}^{t_i}(\sigma(\vec{z}_k))\|}{\|\vec{s}_j^{t_i} - \vec{s}_{j+1}^{t_i}\|} \right) \right], \end{aligned} \quad (1.2)$$

where k is the current MD step and w is the weight of the adjustment, which is non-zero only for the two nodes closest to the average CV density. We also introduce the fade factor $\xi = \exp(-\ln(2)/\tau_{1/2})$, with $\tau_{1/2}$ as a half-life: the number of MD steps after which the distances measured from the path weight only 50% of its original value. A short half-life increases path flexibility by allowing it to rapidly “forget” a bad initial guess, while a long half-life restricts the path fluctuations and leads eventually to optimal convergence. The first and last nodes, located at the stable states A and B respectively, remain fixed. Extra trailing nodes can be placed at both ends of the path to better capture the free energy valleys. Wall potentials can be used to keep the sampling near the relevant $[0, 1]$ σ -interval.

Additionally, a harmonic restraint set on $\|\vec{z}_k - \vec{s}^{t_i}(\sigma(\vec{z}_k))\|$ can help in converging to a specific transition path in scenarios with multiple or ill-defined transition channels. We refer to this restraint as a “tube” potential. In the limit of an infinitesimally narrow tube potential, the path update step follows the local free energy gradient and PMD converges to the MFEP instead of the average transition path. This control

over the behavior of the algorithm is an advantageous feature, as switching from a density-driven to a gradient-driven path optimization leads either to the average transition path or the MFEP.

Metadynamics, the method of choice to sample the path-CV, drives the system back and forth along the path, providing sufficient sampling of the CV density to localize the transition path. At the same time, metadynamics continuously self-corrects the free energy by overwriting the previously deposited Gaussians with new ones. This leads to a converged free energy along the localized path: $V_{\text{bias}}(\sigma, t \rightarrow \infty) = -F(\sigma)$. Most of the algorithmic extensions developed for metadynamics (well-tempered, transition-tempered, multiple walkers, etc.) can be straightforwardly combined with the adaptive path-CV. Moreover, other biasing techniques (umbrella sampling, constrained MD, steered MD, etc.) can be also be used to sample along the path, or even in the direction perpendicular to it.

1.1.4 Model system

In this exercise we will sample the well-known Müller-Brown (MB) potential energy surface (PES). You can find more about it in Ref. [8]. We add harmonic walls to limit the sampling within the interesting region of the PES.

1.1.5 Using PLUMED

To install PLUMED, go to the folder `plumed` in the folder for this exercise. In the script `install`, change the prefix to the installation folder of your choice, and then run the script. If you have issues with the installation you can check: https://plumed.github.io/doc-v2.3/user-doc/html/_installation.html. You will also need gnuplot to visualize the results of this exercise. You can obtain it using `apt-get` or `brew install` in Unix systems. Plotting script examples are provided (all starting with `plot_` or `anim_`). Make sure to change the corresponding variables according to your choices of the metadynamics and path-CV parameters.

- To run the PLUMED `pesmd` on the MB PES, use the command `plumed pesmd < path/to/input`. The `input` file is located in the **Exercise** folder. Review the PLUMED `pesmd` documentation: <https://plumed.github.io/doc-v2.4/user-doc/html/pesmd.html>.
- To perform metadynamics, you can review the keywords on: https://plumed.github.io/doc-v2.3/user-doc/html/_m_e_t_a_d.html.
- To use the adaptive path-CV, you can review the keywords in the header of the file `PathCV.cpp` in the folder `plumed` in the folder for this exercise. To see the html documentation on the path-CV in a browser, open `PMD_exercise/plumed/plumed-2.3.0/user-doc/html/_p_a_t_h_c_v.html` in a browser.
- To construct a FES from the metadynamics hills, use the `plumed sum_hills` functionality. Consult its features on: https://plumed.github.io/doc-v2.3/user-doc/html/sum_hills.html. We recommend to use `--stride 1` and `--mintozero`.

1.1.6 Questions

1. Use PLUMED `pesmd` to run regular MD on the MB PES. What do you observe? How would you change the behaviour of the system without using an enhanced sampling method (just modifying the MD run parameters)?
2. Use PLUMED `pesmd` to run metadynamics on the MB PES. Try to obtain a good FES reflecting the two minima. Vary the width of the Gaussians for each CV in a systematic manner, what effect do you observe? Vary the height and deposition frequency of the Gaussians, what trends do you observe?

3. Run metadynamics along a fixed path-CV on the MB PES. We recommend using 20 transition nodes, with 20 trailing nodes at each end (60 nodes in total). Use harmonic walls to keep the sampling within a relevant $[-0.2, 1.2]$ σ -range. Now there's only one dimension for the Gaussian width, how to do interpret it? Does the diffusion of the system along σ reflect anything particular? What could be the reason? What happens if you add and strengthen a tube potential? What is the final path and free energy that you obtain?
4. Run metadynamics along an adaptive path-CV on the MB PES. For simplification, start with the same pace for the metadynamics Gaussians and the path update. For simplification, you can keep only 20 transition nodes, and not use trailing nodes. Use harmonic walls to keep the sampling within a relevant $[-0.2, 1.2]$ σ -range. What do you expect if you vary the ratio between the metadynamics and the path update frequency? Vary the tube potential strength and the half-life parameter, what trend do you observe?

Bibliography

- [1] Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566.
- [2] Branduardi, D.; Gervasio, F.L.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 054103.
- [3] Díaz Leines, G.; Ensing, B. *Phys. Rev. Lett.* **2012**, *109*, 020601.
- [4] Pérez de Alba Ortíz, A.; Tiwari, A.; Puthenkalathil, R.C.; Ensing, B. *J. Chem. Phys.* **2018**, *149*, 072320.
- [5] Tribello, G.A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comp. Phys. Comm.* **2014**, *185*, 604–613.
- [6] Raiteri, P.; Laio, A.; Gervasio, F.L.; Micheletti, C.; Parrinello, M. *J. Phys. Chem. B* **2006**, *110*, 3533–3539.
- [7] Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100*, 020603.
- [8] Müller, K.; Brown, L.D. *Theor. Chim. Acta* **1997**, *53*, 75–93.

