
Funnel hopping Monte Carlo: An efficient method to overcome broken ergodicity [1]

Matthias Vogler

December 11, 2020

1 MARKOV CHAIN MONTE CARLO

The goal of Markov chain Monte Carlo (MCMC) is to generate samples of a target distribution. In our case the target distribution will always be the Boltzmann distribution

$$P(\mathbf{r}) = \frac{1}{Z(T)} \exp\left(\frac{-E(\mathbf{r})}{k_B T}\right) \quad (1.1)$$

We start with a initial configuration \mathbf{r} , we then propose a new configuration \mathbf{r}' and accept the new configuration with probability α given by the Metropolis-Hastings criterion [2]

$$\alpha = \min\left(1, \frac{P(\mathbf{r}') g(r|r')}{P(\mathbf{r}) g(r'|r)}\right), \quad (1.2)$$

where $P(\mathbf{r})$ is our target distribution and $g(r|r')$ the probability to propose a move from r to r' . The heart of MCMC lies in proposing the right moves. The displacements should be sufficiently small or the moves will always be rejected and a convergent solution will not be found. On the other hand, proposing moves that are too small will result again in almost no progress. Thus there is a inherent trade of between the proposed moves and the acceptance rate. Another problem that arises in MCMC when trying to sample a configuration space to calculate a quantity of a system (e.g. heat capacity) are high energy barriers between different regions of minima, especially at low temperatures of interest. This leads to never sampling the whole configuration space and to the namesake phenomenon of broken ergodicity which the paper of Finkler and Goedecker tries to overcome with their novel method of *Funnel Hopping Monte Carlo*.

1.1 Energy landscapes and Funnels

Energy landscapes can best be characterized by their local minima. These minima can be grouped together with the higher lying minima into so-called *funnels*. Different funnels are usaly seperated by high energy barriers which can't be overcome in the computation

time available. This concept is best visualized with energy disconnectivity graphs as seen in figure 1.1.

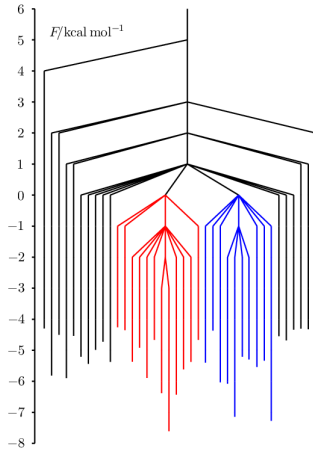


Figure 1.1: Energy disconnectivity graph for met-enkephalin at 298 K. The low-energy region of the graph contains two funnels, highlighted in red and blue. Source: [3]

1.2 Already available methods: Paralell Tempering

Paralell Tempering (PT) is a MC method that runs many copies of the system in paralell at different temperatures. The idea is then to make configurations at high temperatures available to the simulations at low temperatures. This is useful since at high temperatures it is much more likely to overcome higher energy barriers. This information is the passed down to the low temperature instances. The method however has the drawback of being computationally expensive since many instances have to be run in paralell.

2 FUNNEL HOPPING MONTE CARLO

The basic idea of Funnel Hopping Monte Carlo (FHMC) is to use precomputed knowledge about the local minima in the energy landscape. This includes the location of the minima which can easily be extracted with global optimization methods such as minima hopping. To construct the FHMC method, there are a lot of preliminary steps one has to take. This will be explained in the following sections.

2.1 Smart darting

The main idea of smart darting [4] is to construct special moves out of the knowledge of the minima. Given a set of M local minima $\{\mathbf{R}_i\}_{i=1,2,\dots,M}$, we can define *darting vectors* \mathbf{D}_{ij} as

$$\mathbf{D}_{ij} = \mathbf{R}_j - \mathbf{R}_i. \quad (2.1)$$

Around each minima epsilon regions are placed. Configurations can now lie within a minima with the criterion

$$\|\mathbf{R}_i - \mathbf{r}_i\| < \epsilon \quad (2.2)$$

with ϵ chosen such that the regions don't overlap. We can then replace a certain fraction of the MC moves with darting moves. If a darting move is to be executed, first we check if

the current configuration \mathbf{r} lies within a ϵ region of a minima \mathbf{R}_i . If true, a darting vector starting at i will randomly be chosen and a new configuration is proposed with

$$\mathbf{r}' = \mathbf{r} + \mathbf{D}_{ij} \quad (2.3)$$

and accepted or rejected with the standard Metropolis criterion (eq. 1.2). This method works very well in some cases. But it has some key problems in certain systems. If we have systems with invariances such as rotation and translations as in clusters, the epsilon regions become hard to define. The minima don't correspond to a single point in configuration space, but rather a hypersurface. This is further exacerbated in systems with indistinguishable atoms which are invariant under permutations. It then becomes impossible to define these regions with the method described earlier. It also assumes spherical epsilon regions, where as in reality regions of low energy are often ellipsoidal. It is therefore important to have a method which makes no assumptions about the shape of the high probability regions.

2.2 Removing degrees of freedom: Eckart space

To decide if a configuration is inside the high probability regions, we have to be able to identify similar configurations. In $3N$ dimensional configuration space it is non trivial to identify similar configurations which are rotationally invariant. To do this, we first select a reference configuration \mathbf{R} in a local minima. We then try to identify the optimal rotation \mathcal{R} and permutation \mathcal{P} of \mathbf{r} such that the root mean squared deviation (RMSD) is minimal. We define the RMSD as

$$\text{RMSD}(\mathbf{r}, \mathbf{R}) = \sqrt{\frac{\sum_{i=1}^N ||\vec{r}_i - \vec{R}_i||^2}{N}} \quad (2.4)$$

and the displacement \mathbf{d} as

$$\vec{d}_i = \vec{r}_i - \vec{R}_i. \quad (2.5)$$

The RMSD is then minimal if the modified *Eckart conditions* are met [5]:

$$\sum_{i=1}^N \vec{d}_i = \vec{0} \quad (2.6)$$

$$\sum_{i=1}^N \vec{d}_i \times \vec{R}_i = \vec{0} \quad (2.7)$$

We can now define vectors, to which all displacement vectors obtained with minimal RMSD alignment are orthogonal to:

$$V^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}, V^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, V^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, V^4 = \begin{pmatrix} 0 \\ \vec{R}_{1,z} \\ -\vec{R}_{1,y} \\ 0 \\ \vec{R}_{2,z} \\ -\vec{R}_{2,y} \\ 0 \\ \vdots \end{pmatrix}, V^5 = \begin{pmatrix} -\vec{R}_{1,z} \\ 0 \\ \vec{R}_{1,x} \\ -\vec{R}_{2,z} \\ 0 \\ \vec{R}_{2,x} \\ \vec{R}_{3,z} \\ \vdots \end{pmatrix}, V^6 = \begin{pmatrix} \vec{R}_{1,y} \\ -\vec{R}_{1,x} \\ 0 \\ \vec{R}_{2,y} \\ -\vec{R}_{2,x} \\ 0 \\ \vec{R}_{3,y} \\ \vdots \end{pmatrix} \quad (2.8)$$

This gives as the possibility to construct $3N - 6$ new basis vectors B^i that are orthogonal to the vectors V^j . Using B^i as our new basis removes 6 degrees of freedom from our system and fixes rotation and translation. The $3N$ dimensional vector \mathbf{d} is then transformed via

$$\mathbf{d}' = \mathbf{d} \cdot \mathbf{B}^i. \quad (2.9)$$

To get the total minimum of the RMSD in systems of indistinguishable atoms, one has to minimize with respect to permutations as well. There however, one encounters a causality dilemma. The optimal rotation depends on the permutation where as the optimal permutation depends on the optimal rotation. The paper however suggests a clever method using the Hungarian algorithm [6] in conjunction with a quaternion method [7].

2.3 Funnel Hopping Monte Carlo

With the procedures described above, we can finally get to the actual method at hand. First we assign each point in configuration space to its nearest minimum. For each minimum \mathbf{R}_i one defines a probability distribution $q_i(\mathbf{r})$ which lives in the $3N - 6$ dimensional fixed frame coordinate space. To propose a Funnel Hopping move one determines the nearest minimum \mathbf{R}_i to the current configuration and randomly chooses another minimum j . Then one draws a configuration from the corresponding q_j . Determining j can be done completely random or via a transition matrix T with T_{ij} being the probability to choose j while being in minimum i . This can help with proposing moves that connect minima with similar energy. The move is then accepted or rejected over a modified Metropolis criterion:

$$\alpha(\mathbf{r} \rightarrow \mathbf{r}') = \min \left(1, \exp \left(-\frac{E(\mathbf{r}') - E(\mathbf{r})}{k_B T} \right) \frac{q_i(\mathbf{r})}{q_j(\mathbf{r}')} \frac{T_{ji}}{T_{ij}} \frac{h_{\alpha i}}{h_{\alpha j}} \right) \quad (2.10)$$

The distributions q_i are still undefined, but the main role they play is that they are used to generate the proposed configuration. The paper proposes to define these distributions over Gaussian mixtures which are defined as follows:

$$q_i^{g.m.}(r) = \sum_{k=1}^m a_i^k \mathcal{N}_i^k(r) \quad \left| \quad \sum_{k=1}^m a_i^k = 1 \text{ and } a_i^k \geq 0 \forall k \right. \quad (2.11)$$

with \mathcal{N}_i^k are normalized multivariate Gaussians with means μ_i^k and covariance Σ_i^k . The parameters a_i^k , μ_i^k and Σ_i^k are then fitted to the Boltzmann distribution.

REFERENCES

- [1] J. A. Finkler and S. Goedecker, “Funnel hopping monte carlo: An efficient method to overcome broken ergodicity,” *The Journal of Chemical Physics*, vol. 152, no. 16, p. 164106, 2020. DOI: 10.1063/5.0004106. eprint: <https://doi.org/10.1063/5.0004106>. [Online]. Available: <https://doi.org/10.1063/5.0004106>.
- [2] W. K. Hastings, “Monte carlo sampling methods using markov chains and their applications,” *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970, ISSN: 00063444. [Online]. Available: <http://www.jstor.org/stable/2334940>.
- [3] D. Wales, “Energy landscapes and properties of biomolecules.,” *Physical biology*, vol. 2 4, S86–93, 2005.
- [4] I. Andricioaei, J. E. Straub, and A. F. Voter, “Smart darting monte carlo,” *The Journal of Chemical Physics*, vol. 114, no. 16, pp. 6994–7000, 2001. DOI: 10.1063/1.1358861. eprint: <https://doi.org/10.1063/1.1358861>. [Online]. Available: <https://doi.org/10.1063/1.1358861>.
- [5] K. N. Kudin and A. Y. Dymarsky, “Eckart axis conditions and the minimization of the root-mean-square deviation: Two closely related problems,” *The Journal of Chemical Physics*, vol. 122, no. 22, p. 224105, 2005. DOI: 10.1063/1.1929739. eprint: <https://doi.org/10.1063/1.1929739>. [Online]. Available: <https://doi.org/10.1063/1.1929739>.
- [6] H. W. Kuhn, “The hungarian method for the assignment problem,” *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955. DOI: 10.1002/nav.3800020109. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/nav.3800020109>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/nav.3800020109>.
- [7] E. A. Coutsiias, C. Seok, and K. A. Dill, “Using quaternions to calculate rmsd,” *Journal of Computational Chemistry*, vol. 25, no. 15, pp. 1849–1857, 2004. DOI: <https://doi.org/10.1002/jcc.20110>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.20110>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20110>.