

Umbrella Sampling & Markov Chain Monte Carlo

Brandon N. Onusaitis

Department of Materials Science and Engineering

Northwestern University, Evanston, 60208, United States

December 27, 2023

Abstract

In this report, a derivation of a recursive relation is shown, which was adapted from Frenkel & Smit and Allen & Tildesley [1][2]. The recursive relation describes an algorithm that can be utilized to generate successive partition functions of a biased system. The partition functions are required to measure the Landau Free Energy of a system via Weighted Histogram Analysis Method (WHAM). Moreover, Metropolis Monte Carlo (MCMC) simulations were performed for 2D array of Ising spins under zero magnetic field ($B = 0$). Quality sampling was achieved by implementing Umbrella Sampling via a harmonic biasing potential. The system was sampled at temperatures above ($\beta < \beta_c$), around ($\beta \approx \beta_c$), and below ($\beta > \beta_c$) its critical temperature ($\beta_c \approx 0.44$). Preliminary free energy profiles were obtained along with histograms at various order parameter values (Magnetization per Site).

Contents

| | | |
|----------|-------------------|----------|
| 1 | Derivation | 2 |
| 2 | Methods | 6 |
| 3 | Results | 6 |
| 4 | Conclusion | 6 |

1 Derivation

The free energy of a system in the canonical ensemble depends on the partition function and the fixed temperature through the following relation:

$$F = -k_b T \log Q(N, V, T) \quad (1)$$

Since free energies are not absolute, only relative free energies can be measured in experiment or simulation. The free energy change between two states, denoted 1 and 2, is a function of their respective partition functions:

$$\Delta F = -k_b T \log \frac{\int dM^N e^{-\beta E_2(M^N)}}{\int dM^N e^{-\beta E_1(M^N)}} \quad (2)$$

where M^N is the dimension of the configuration space, and E_1 and E_2 are the energies of systems 1 and 2, respectively. Given the high-dimensional space of the partition functions and the Boltzmann sampling employed in traditional metropolis algorithm, there is a need to estimate the partition functions through biased sampling.

To relate the free energy to a measurable quantity in simulation, a relationship between the probability of finding a potential energy difference between system 2 and system 1 must be articulated. The probability, $P_2(\Delta E)$, of one such potential energy difference can be expressed as:

$$P_2(\Delta E) = \frac{\int dM^N e^{\beta E_2(M^N)} \delta(E_2(M^N) - (E_1(M^N) - \Delta E(M^N)))}{Q_2(N, V, T)} \quad (3)$$

where the delta function ensures the energy of system 2 is evaluated under the ΔE constraint. Specifically, the delta function yields the correct Boltzmann factor in terms of system 2's energy. Rewriting the Boltzmann factor for system 2's energy, we have:

$$e^{-\beta E_2(M^N)} = e^{-\beta(E_2(M^N) - \beta E_1(M^N))} e^{-\beta E_1(M^N)} = e^{-\beta \Delta E(M^N)} e^{-\beta E_1(M^N)} \quad (4)$$

By rewriting Equation (3) in terms of a ratio of partition functions, the expression becomes:

$$P_2(\Delta E) = \frac{\int dM^N e^{-\beta \Delta E(M^N)} e^{-\beta E_1(M^N)} \delta(E_2(M^N) - (E_1(M^N) - \Delta E(M^N)))}{Q_2(N, V, T)} \quad (5)$$

$$P_2(\Delta E) = \frac{e^{-\beta \Delta E(M^N)}}{Q_2(N, V, T)} \int dM^N e^{-\beta E_1(M^N)} \delta(E_2(M^N) - (E_1(M^N) - \Delta E(M^N))) \quad (6)$$

Based on our definition shown in Equation (3), the integral can be exchanged with the probability of system 1 multiplied by its partition function:

$$P_2(\Delta E) = \frac{e^{-\beta \Delta E(M^N)}}{Q_2(N, V, T)} Q_1(N, V, T) P_1(\Delta E) \quad (7)$$

This formulation facilitates the connection between the probability of potential energy differences and partition functions, which provides a means to leverage biased sampling in monte carlo (MC) simulations. It is important to clarify, that $P_2(\Delta E)$ indicates the probability of finding an energy difference between system 2 and 1 while Boltzmann sampling system 2—the same can be said for $P_1(\Delta E)$. Moreover, taking the natural log of Equation (7) and using the definition shown in Equation (2) we get an expression relating the probabilities and the free energy:

$$\log P_2(\Delta E) = \log \frac{Q_1(N, V, T)}{Q_2(N, V, T)} - \beta \Delta E + \log P_1(\Delta E) \quad (8)$$

$$\log P_2(\Delta E) = \beta \log(\Delta F) - \beta \Delta E + \log P_1(\Delta E) \quad (9)$$

In principle, the free energy can be determined by employing Equation (9) through Metropolis MC simulations with Boltzmann sampling for both systems 1 and 2. That is, the disparity separating each systems' histogram is directly related to the free energy difference between them. However, this method falls short when attempting to capture the free energies between systems separated by a large energy barrier—a scenario where Boltzmann proves inadequate. As a remedy, the established practice is conducting MC simulation at intermediate energy separations through a biased potential. Subsequently, the histograms are integrated, and the probabilities are written as a linear combination that sums to the unbiased distribution function value. While I won't include the derivation of this method, I will present the final outcome, as described in Frenkel and Smit [1].

$$P_0^{est} = \frac{\sum_i^S H_i(M)}{\sum_{i=1}^S C_i \frac{Q_0(N, V, T)}{Q_i(N, V, T)} e^{-\beta V_{bias, i}}} \quad (10)$$

$$Q_i(N, V, T) = \int dM^N e^{-\beta V_{bias, i}} \frac{\sum_{j=1}^n H_j(M)}{\sum_{k=1}^n C_k \frac{e^{-\beta V_{bias, k}}}{Q_k(N, V, T)}} \quad (11)$$

Equation (10) gives the target probability distribution, which is crucial to estimating the Landau free energy. Together, these equations constitute the self-consistent histogram method, also known as WHAM. Solving Equations (10) and (11) involves an iterative process, usually through fixed-point direct iteration method.³ However, in this work a recursion relation

will be derived, as discussed in Frenkel and Smit. The derivation of the recursion relation demands two key assumptions: the use of an infinite square well biasing potential and a window only overlaps with its adjacent window functions. These assumptions will be addressed subsequent to our derivation of the recursion relation. Let our window function be defined as,

$$V_{bias,i} = \begin{cases} \infty & M_i < M_i^{min}, \\ 0 & M_i^{min} \leq M_i \leq M_i^{max}, \\ \infty & M_i \geq M_i^{max}. \end{cases} \quad (12)$$

where M_i is the magnetization (order parameter) over which the biasing potential samples. Ideally, the biasing potential, $V_{bias,i}$ is chosen such that it overlaps exclusively with the adjacent windows, as formulated in Equation (13) and shown in Figure 1.

$$\begin{cases} M_{i-2}^{max} < M_i^{min} < M_{i-1}^{max} \\ M_{i+2}^{min} > M_i^{max} > M_{i+1}^{min} \end{cases} \quad (13)$$

The recursive relation can be derived for the biasing function in Equation (12) and by evaluating Equation (11). In Equation (11), H_j is histogram j and C_k is the number of points falling within histogram j . Thus,

$$P_j = \frac{H_j}{C_k} \quad \forall j, k : \delta_{jk} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \quad (14)$$

substituting Equation (14) into Equation (11) yields,

$$Q_i(N, V, T) = \int dM e^{-\beta V_{bias,i}} \frac{\sum_i^n P_i(M)}{\sum_i^n \frac{e^{-\beta V_{bias,i}}}{Q_i}} \quad (15)$$

where P_i is the i th simulation's probability density. When evaluating the partition function Q_i , we perform integration over all space. However, the Boltzmann factor of the biasing function alters the bound, which renders the integrand zero for all magnetization values outside the interval $M \in [M_i^{min}, M_i^{max}]$.

$$Q_i(N, V, T) = \int_{-\infty}^{+\infty} dM e^{-\beta V_{bias,i}} \frac{\sum_i^n P_i(M)}{\sum_i^n \frac{e^{-\beta V_{bias,i}}}{Q_i}} = \int_{M_i^{min}}^{M_i^{max}} dM e^{-\beta V_{bias,i}} \frac{\sum_i^n P_i(M)}{\sum_i^n \frac{e^{-\beta V_{bias,i}}}{Q_i}} \quad (16)$$

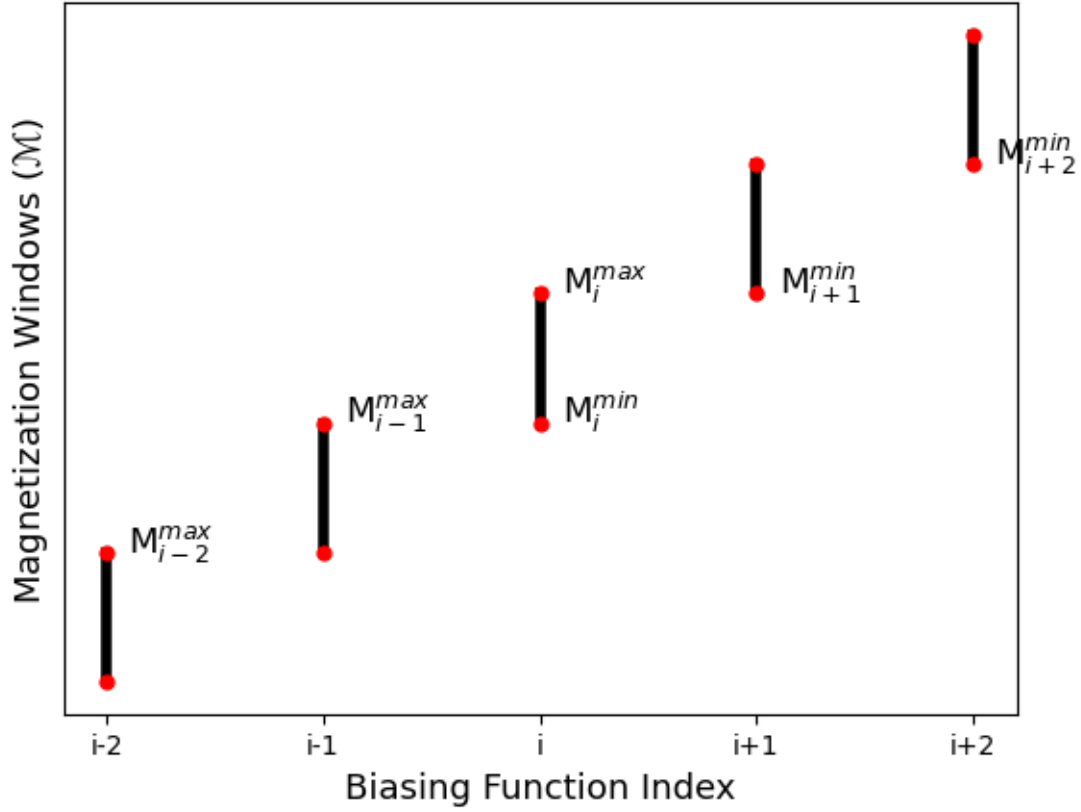


Figure 1: Window function with magnetization range on the ordinate and window function index on the abscissa. This illustrates the ideal scenario where only adjacent windows overlap over a specified range of magnetizations.

The integral on the right-hand side of Equation (16) can be decomposed into the sum of integrals over three continuous intervals, collectively covering the entire domain, $M \in [M_i^{min}, M_i^{max}]$.

$$\int_{M_i^{min}}^{M_i^{max}} dM e^{-\beta V_{bias,i}} \frac{\sum_i^n P_i(M)}{\sum_i^n \frac{e^{-\beta V_{bias,i}}}{Q_i}} = \int_{M_i^{min}}^{M_{i-1}^{max}} dM + \int_{M_{i-1}^{max}}^{M_{i+1}^{min}} dM + \int_{M_{i+1}^{min}}^{M_i^{max}} dM e^{-\beta V_{bias,i}} \frac{\sum_i^n P_i(M)}{\sum_i^n \frac{e^{-\beta V_{bias,i}}}{Q_i}} \quad (17)$$

Equation (17) is much simpler to evaluate since many terms will cancel out due to the three sub intervals. For example, in the numerator, the summation of probabilities across each window ensures only two terms survive in each integral. This is a consequence of the probability being zero outside the domain encompassed by the integration bounds. Therefore,

$$\int_{M_{i-1}^{max}}^{M_{i+1}^{min}} dM \sum_i^n P_i(M) = \int_{M_{i-1}^{max}}^{M_{i+1}^{min}} dM P_1 + \dots + P_{i-1} + P_i + P_{i+1} + \dots + P_n = \int_{M_{i-1}^{max}}^{M_{i+1}^{min}} dM P_{i-1} + P_{i+1} \quad (18)$$

where this same argument holds true for the remaining two integrals. The subsequent task is to evaluate the denominator of each integral in Equation (17). By expanding the sum of each biased probability and evaluating each window function both within and outside its bounds, it becomes evident only two non-trivial terms persist:

$$\left\{ \begin{array}{l} \int_{M_{i-1}^{max}}^{M_{i+1}^{min}} dM \frac{1}{\sum_i^n \frac{e^{-\beta V_{bias,i}}}{Q_i}} = \\ \int_{M_{i-1}^{max}}^{M_{i+1}^{min}} dM \frac{1}{\frac{e^{-\beta V_{bias,1}}}{Q_1} + \dots + \frac{e^{-\beta V_{bias,i-1}}}{Q_{i-1}} + \frac{e^{-\beta V_{bias,i}}}{Q_i} + \frac{e^{-\beta V_{bias,i+1}}}{Q_{i+1}} + \dots + \frac{e^{-\beta V_{bias,n}}}{Q_n}} = \\ \int_{M_{i-1}^{max}}^{M_{i+1}^{min}} dM \frac{1}{0 + \dots + \frac{e^{-\beta 0}}{Q_{i-1}} + 0 + \frac{e^{-\beta 0}}{Q_{i+1}} + \dots + 0} \end{array} \right. \quad (19)$$

Substituting the results from Equations (18) and (19) into Equation (17), the following expression is obtained:

$$Q_i(N, V, T) = \int_{M_i^{min}}^{M_{i-1}^{max}} dM \frac{Q_{i-1} Q_i [P_i(M) + P_{i-1}(M)]}{Q_{i-1} + Q_i} + \int_{M_{i-1}^{max}}^{M_{i+1}^{min}} dM \frac{Q_{i+1} Q_{i-1} [P_{i+1}(M) + P_{i-1}(M)]}{Q_{i-1} + Q_{i+1}} + \int_{M_{i+1}^{min}}^{M_i^{max}} dM \frac{Q_i Q_{i+1} [P_i(M) + P_{i+1}(M)]}{Q_i + Q_{i+1}} \quad (20)$$

2 Methods

3 Results

4 Conclusion

References

¹D. Frenkel and B. Smit, “Understanding molecular simulation”, (2001).

²M. P. Allen and D. J. Tildesley, “Computer simulation of liquids”, (2017).