

Umbrella Sampling + WHAM

Henry Bloom, Arnav Brahmāsandra

February 28, 2025

1 Summary

Molecular systems often exhibit high free energy barriers that impede adequate sampling of configurational space in standard simulations. Umbrella sampling addresses this by introducing biasing potentials to enhance sampling in targeted regions, stitching the biased distributions together using WHAM. WHAM provides a systematic framework to reconstruct unbiased free energy profiles from biased simulations.

2 Umbrella Sampling

Consider some system with a potential $U(r^N)$. Traditional sampling methods may suffer from poor ergodicity due to certain energy landscapes, particularly if the configuration space has many high-energy barriers that may trap the trajectory in a local minimum. Introduced by Torrie and Valleau in 1977, umbrella sampling uses bias potentials to encourage sampling of potentially inaccessible states. The biased distributions are then used to reconstruct an unbiased distribution and calculate quantities of interest, such as the free energy of the system.

2.1 Reaction Coordinate

We will introduce the notion of a reaction coordinate $q = f(r^N)$, an abstract one-dimensional parameter that unambiguously measures progress along a reaction pathway or transition between states. Some examples of a reaction coordinate, both geometric and non-geometric, include bond distance, bond angle, torsion, dihedral angle, bond order, Hydrogen bonds, and RMSD. The free energy of a system is often considered as a function of reaction coordinate to demonstrate in some schematic form the potential energy profile associated to the reaction.

2.2 Bias Potentials

Umbrella sampling introduces a window-specific soft restraint biasing potential $W_i(q, s)$, such that the total potential for window i becomes $U_i^{\text{biased}}(r^N) =$

$U(r^N) + W_i(f(r^N))$. Often, the restraint is chosen to be a harmonic potential of the form $W_i(f(r^N), s) = \frac{1}{2}\kappa(f(r^N) - s_i)^2$ centered around a chosen s_i .

2.3 Simulations

In umbrella sampling, run a series of n simulations with different biasing potentials, collecting a biased distribution of values of q . Essentially, each simulation can be thought of as generating a histogram of values of q primarily located within a "window" corresponding to the bias potential chosen. The aim following the computation of these biased distributions is to first unbiased them, and then to use the unbiased distributions to calculate the free energy of the system.

3 WHAM

The weighted histogram analysis method (WHAM) is the method for unbiasing the biased distributions that is used in conjunction with umbrella sampling.

$$\begin{aligned}\tilde{P}(q, s_k) &= e^{\beta A_k} \int d^N r^N e^{-\beta U(r^N)} e^{-\beta W(f(r^N), s^{(k)})} \delta(f(r^N) - q) \\ e^{-\beta A_k} &= \int d^N r^N e^{-\beta U(r^N)} e^{-\beta W(f(r^N), s^{(k)})} = e^{-\beta A_0} \left\langle e^{-\beta W(f(r^N), s^{(k)})} \right\rangle \\ e^{-\beta A_0} &= \int d^N r^N e^{-\beta U(r^N)} = Z\end{aligned}$$

We can use this process to calculate the unbiased potential from the k -th window:

$$P_k(q) = e^{-\beta(A_k - A_0)} e^{\beta W(q, s^{(k)})} \tilde{P}(q, s^{(k)})$$

We want to use the k unbiased distribution estimates from the simulations to construct the global distribution $P(q)$ while minimizing statistical error. We can set this up as follows:

$$\begin{aligned}P(q) &= \sum_{k=1}^n C_k(q) P_k(q) \\ \text{subject to } &\sum_{k=1}^n C_k(q) = 1\end{aligned}$$

The aim is to optimize the statistical error of $P(q)$ while