> **Today's lecture:** basic, general methods for computing entropies and free energies from histograms taken in molecular simulation, with applications to phase equilibria.

Flat-histogram sampling

In the past two decades, so-called **flat histogram methods** have become major tools for computing free energies and phase equilibria. These methods are designed to construct an extended ensemble that generates a uniform or flat distribution in one or more parameters or reaction coordinates. By doing so, two goals are achieved:

- broad sampling and good statistics for the flat histogram parameters, which expands the range of conditions at which reweighting can be effectively performed

- computation of free energies or entropies along the flat histogram parameters, through a connection between them and the presence of a flat distribution

Such methods are often used in place of umbrella sampling approaches in Monte Carlo and molecular dynamics simulations because they automatically determine the weight functions, instead of requiring a specific form (e.g., a harmonic potential). Below we discuss a specific example in the grand-canonical ensemble; however, these approaches are general to any simulation ensemble and parameter of interest.

Example: Grand-Canonical Monte Carlo

Consider a grand canonical MC simulation with $\mu = \mu_1$, where both the energy and particle number fluctuate. The probability of seeing a particular configuration is given by

$$\wp(\mathbf{r}^N, N) \propto \frac{e^{-\beta U + \beta \mu_1 N}}{\Lambda(T)^{3N} N!}$$

If we integrate this expression over all configurations at a particular particle number, we can find the distribution in N:

$$\wp(N) \propto \int \wp(\mathbf{r}^N, N) d\mathbf{r}^N$$

$$= \int \frac{e^{-\beta U + \beta \mu_1 N}}{\Lambda(T)^{3N} N!} d\mathbf{r}^N$$

$$= e^{\beta \mu_1 N} \int \frac{e^{-\beta U}}{\Lambda(T)^{3N} N!} d\mathbf{r}^N$$

$$= e^{\beta \mu_1 N - \beta A(T,V,N)}$$

That is, the particle number distribution involves the Helmholtz free energy $A(T,V,N)$. We will actually drop the variables T,V since these do not fluctuate during the simulation:

$$\wp(N) \propto e^{\beta \mu_1 N - \beta A(N)}$$

Now consider that we have specified a temperature and chemical potential that places us at co-existence between a liquid and vapor phase. At coexistence, this distribution might look something like:
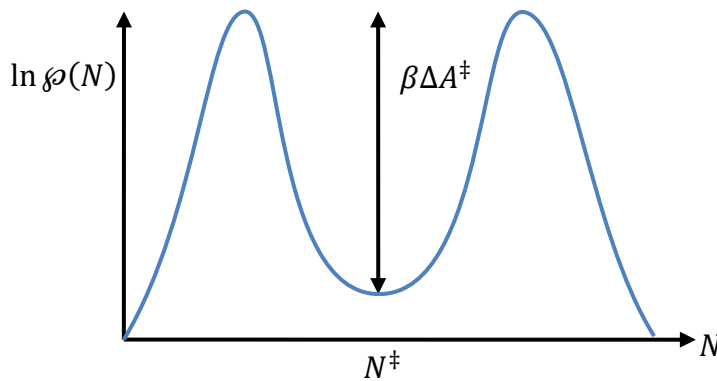
The two probability peaks correspond to the liquid and gas phase. Notice that there is an intermediate value of N with much lower probability. The probability that the system will visit this value of N is

$$\wp(N^{\ddagger}) \propto e^{-\beta \Delta A^{\ddagger}}$$

Small free energy differences can make the probability of intermediate values of N very, very tiny due to the exponential. Thus, even though the system is at coexistence per the bimodal distribution above, the implication is that fluctuations that traverse intermediate densities are very rare.

In a GCMC simulation, this means that we would rarely see the system interconvert between low and high particle numbers due to the low probability of visiting intermediate particle number states between them. We would have an extremely difficult time equilibrating a GCMC simulation at these conditions since it almost certainly would visit only one of the two phases during the simulation run, and not both.

We can use an extended ensemble to enhance the probability of intermediate states. One way to do this would be to modify the ensemble probabilities so that we would obtain a uniform

distribution or a **flat histogram** of particle numbers, between two predetermined limits N_{min} and N_{max}. Typically we pick $N_{\text{min}} = 1$ and N_{max} to be well beyond the particle number associated with the liquid density.

To do this, we can add a weighting function to the ensemble probabilities that depends on N:

$$\wp^{\text{w}}(\mathbf{r}^N, N) \propto \frac{e^{-\beta U + \beta \mu_1 N + \eta(N)}}{\Lambda(T)^{3N} N!}$$

This is the function that we use in the determination of our acceptance criteria. Following through the detailed balance equation, we find that:
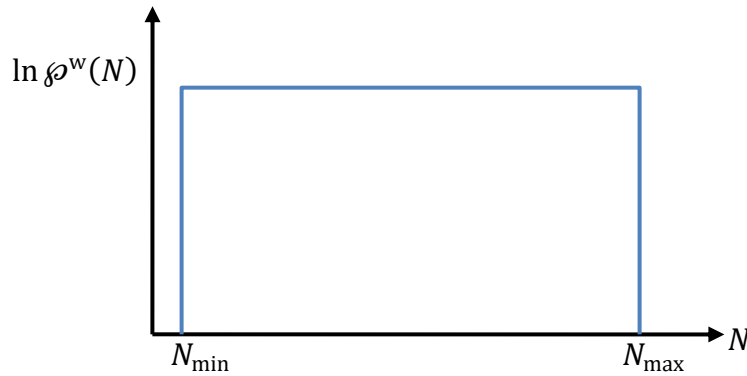
$$P_{12}^{\text{acc}} = \min\left[1, \frac{V}{N+1} e^{-\beta \Delta U + \beta \mu_1' + \Delta \eta}\right] \quad \text{for insertions}$$

$$P_{12}^{\text{acc}} = \min\left[1, \frac{N}{V} e^{-\beta \Delta U - \beta \mu_1' + \Delta \eta}\right] \quad \text{for deletions}$$

In each case, $\Delta \eta = \eta(N_2) - \eta(N_1)$.

Choice of weighting function and reweighting procedure

How do we pick the function $\eta(N)$? This function would be tabulated as an array in our simulation and would be defined in the range $[N_{\text{min}}, N_{\text{max}}]$. We want to pick this function so that our ultimate probability distribution in N looks flat:

We can figure out what we need to pick for $\eta(N)$ by examining the expected distribution in the weighted ensemble:

$$\wp^{\text{w}}(N) \propto \int \wp^{\text{w}}(\mathbf{r}^N, N) d\mathbf{r}^N$$

$$= e^{\beta \mu_1 N - \beta A(N) + \eta(N)}$$

We want this distribution to be flat, $\wp^w(N) = \text{const}$. This gives us a way to solve for $\eta(N)$:

$$\eta(N) = \beta A(N) - \beta \mu_1 N + \text{const}$$

Notice two important aspects of this result:

- We do not need to know the additive constant in $\eta(N)$, since the acceptance criterion depends only upon differences in this function at different values of N. Ultimately this is because additive shifts in $\eta(N)$ do not affect microstate probabilities, via the normalization condition. Typically we simply set $\eta(N_{\min}) = 0$ as a working assumption.

- To find $\eta(N)$ that gives a flat histogram, we need to know the Helmholtz free energy as a function of N. This may seem like a difficult task. In practice, however, this connection *provides* us with a way to determine $A(N)$: given a simulation with a trial function $\eta(N)$, if our biased simulation produces a uniform distribution in N, then we have computed the true $A(N)$. Various **flat histogram** techniques discussed below enable us to compute $A(N)$ based on this connection.

Let's say that we find some approximate $\eta(N)$ that gives rise to the distribution
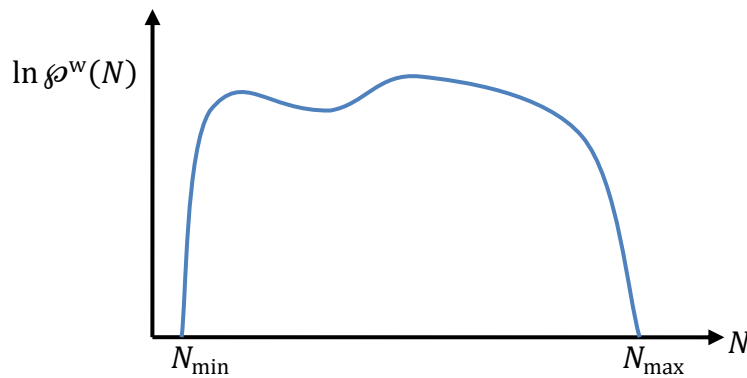
Even though this distribution is not totally flat, we still sample the intermediate values of N with much higher probability and our simulation will likely alternate between liquid and vapor densities with greater frequency than would be present in the unweighted ensemble. As a result, our simulation reaches equilibrium faster and we get good statistics at all particle numbers.

We measure the function $\wp^w(N)$ from a histogram. Now, we want to convert the measured $\wp^w(N)$ back to the unweighted $\wp(N)$, to compute the expected distribution (and averages) in the normal grand-canonical ensemble:

$$\wp(\mathbf{r}^N, N) \propto e^{-\eta(N)} \wp^w(\mathbf{r}^N, N)$$

Integrating over the particle positions,

$$\wp(N) \propto e^{-\eta(N)} \wp^{\mathrm{w}}(N)$$

Using this approach, we could **reweight** to different chemical potentials than the original chemical potential. In the unweighted ensemble,

$$\wp(N; \mu_2) \propto \wp(N; \mu_1) e^{\beta(\mu_2 - \mu_1)}$$

Making the above substitutions:

$$\wp(N; \mu_2) \propto \wp^{\mathrm{w}}(N; \mu_1) e^{-\eta(N) + \beta(\mu_2 - \mu_1)}$$

The constant of proportionality is given by the normalization condition. This equation enables the following general procedure:

- Specify a chemical potential μ_1 and approximate weighting function $\eta(N) \approx \beta A(N) - \beta \mu_1 N$.

- Perform a simulation in the weighted ensemble and measure $\wp^{\mathrm{w}}(N; \mu_1)$ using a histogram.

- By using a weighted ensemble with a roughly flat distribution in N, we are accumulating good statistics in our histogram for a wide range of N values. Such would not be the case if we performed a traditional GCMC simulation, which has a narrowly peaked distribution of N.

- Use the reweighting equation to find the true $\wp(N; \mu_2)$ at any arbitrary chemical potential μ_2. The quality of the reweighted distribution is limited only by regions where the measured $\wp^{\mathrm{w}}(N; \mu_1)$ has few histogram entries.

Ultimately this approach enables us to find conditions of phase equilibrium to high accuracy: we tune the reweighting chemical potential μ_2 until the weight under the two peaks in the probability distribution is equal:
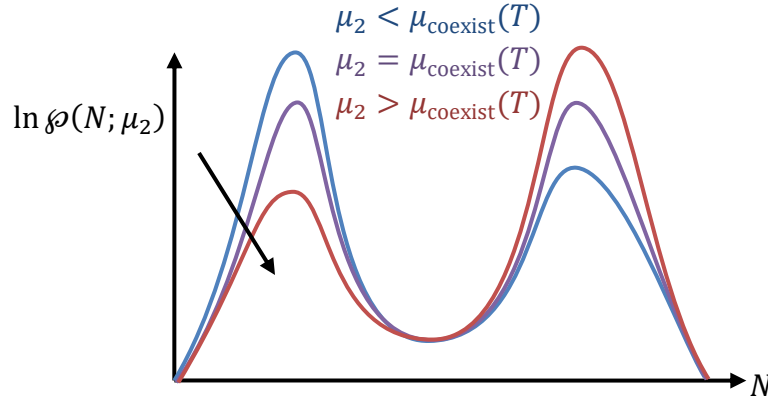
The multicanonical method

How do we find the optimal weighting function $\eta(N)$ such that a histogram of N in the weighted ensemble is flat? Berg and Neuhaus in 1992 devised an iterative solution to this problem called the **multicanonical method**. The basic idea is to perform a serial series of J simulations in which the weighting function is updated after each:

- Consider simulations number j and $j + 1$.

- We perform simulation j using weighting function $\eta_j(N)$ and measure $\wp_j^{\text{w}}(N)$.

- If $\wp_j^{\text{w}}(N)$ is flat, then $\eta_j(N)$ has converged and we are done.

- If it is not flat, we use deviations of $\wp_j^{\text{w}}(N)$ to update $\eta_j(N)$ to get $\eta_{j+1}(N)$.

- The process is repeated until convergence.

We can derive a simple update rule for extracting $\eta_{j+1}(N)$ from $\eta_j(N)$ by comparing the expected weighted ensemble distributions:

$$\wp_j^{\text{w}}(N) \propto \wp(N) e^{\eta_j(N)}$$

$$\wp_{j+1}^{\text{w}}(N) \propto \wp(N) e^{\eta_{j+1}(N)}$$

Dividing these two equations and taking the logarithm gives,

$$\ln \wp_{j+1}^{\text{w}}(N) - \ln \wp_j^{\text{w}}(N) = \eta_{j+1}(N) - \eta_j(N) + \text{const}$$

We demand that $\wp_{j+1}^{\text{w}}(N) = \text{const}$ so that we can find the optimal $\eta_{j+1}(N)$ at the next iteration of the weighting function. Applying this constraint and rearranging,

$$\eta_{j+1}(N) = \eta_j(N) - \ln \wp_j^{\mathrm{w}}(N) + \text{const}$$

This equation provides us with an update rule for determining the next weighting function upon each iteration. Note that,

- We can only know the weights to within an additive constant. This is not a problem, however, because only weight differences appear in the acceptance criterion. Typically we demand $\eta(N_{\min}) = 0$ with each of these updates.

- If $\wp_j^{\mathrm{w}}(N) = \text{const}$, then every value in the weights is shifted by the same amount. Since this effectively only serves to change the arbitrary additive constant, it therefore leaves the weights unchanged. This is a crucial feedback property of the method: if we measure a flat histogram, then we have converged to the desired $\eta(N)$.

Zero-entry bins

When we measure $\wp_j^{\mathrm{w}}(N)$, we use a histogram of counts $c_j(N)$. However, this histogram can lead to a zero probability in bins that have no counts. This would make the update equation above ill-defined due to the logarithm term. A statistical analysis by Smith and Bruce [J. Phys. A, 1995] shows that the following update equation is a more reliable statistical estimator and avoids this problem:

$$\eta_{j+1}(N) = \eta_j(N) - \ln\big[c_j(N) + 1\big] + \text{const}$$

Convergence

Several iterations can be required for convergence. The following shows the results for four iterations using the Lennard-Jones system at $V = 125$ and $\mu' = -3.7$.
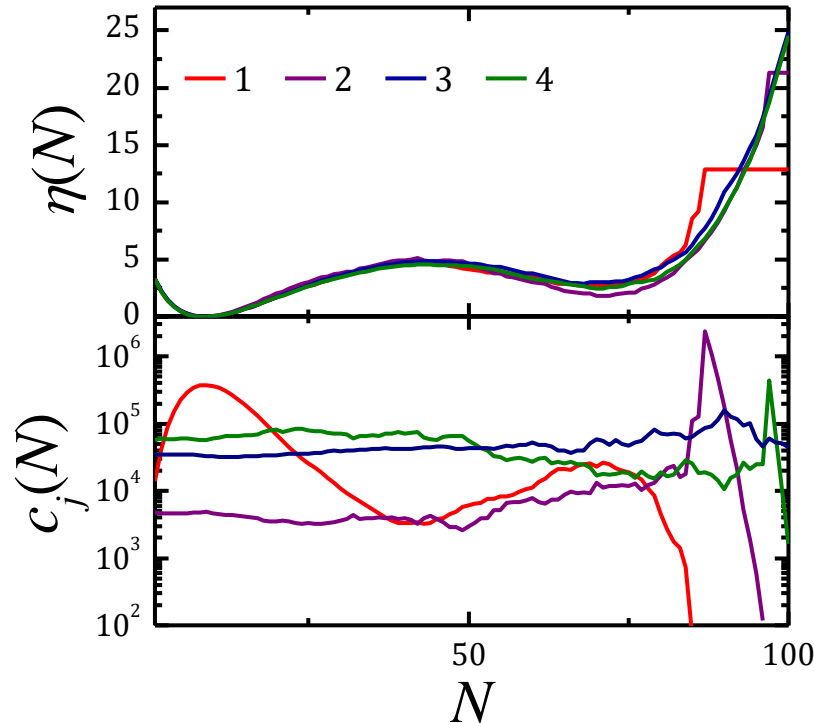
Note that, even if the $\eta(N)$ is not fully converged, a "flat enough" histogram can be used to compute the unweighted average via

$$\wp(N; \mu_2) \propto \wp^w(N; \mu_1) e^{-\eta(N) + \beta(\mu_2 - \mu_1)}$$

Error analysis

The multicanonical approach provides a way to estimate $\eta(N)$ that gives a fairly flat histogram. In principle this method should give $\eta(N) = \beta A(N) - \beta \mu_1 N + \text{const}$ for a perfectly flat histogram. Thus, we might be able to estimate from the final iteration

$$A(N) = \mu_1 N + k_B T \eta_J(N) + \text{const}$$

In practice, the multicanonical method is not able to resolve $\eta(N)$ to sufficient statistical accuracy so that we can determine $A(N)$ to high quality. Each iteration introduces statistical error into this estimate since the histograms employed have a finite number of counts.

Wang-Landau sampling

The **Wang-Landau method** [Wang and Landau, PRL 86, 2050 (2001); PRE 64, 056161] is a recent approach to estimating $\eta(N)$ to high statistical accuracy. It overcomes some of the statistical problems with the multicanonical method and has been rapidly adopted as one of the major tools in flat histogram algorithms. Our discussion here entails a slightly different presentation than the original formulation, for clarity in connecting the approach to the current example.

In this approach, we do not iterate over different simulations to find the optimal $\eta(N)$ that gives a flat histogram. Rather, we modify this function *at every MC step* in a way that enforces a flat distribution and allows it to converge to its optimal value. After every MC move, we update our running estimate for $\eta(N)$ via

$$\eta(N) \leftarrow \eta(N) - g$$

Here, g is termed the **modification factor**. It is a number that we choose so as to gauge the rate at which the weighting function is updated. Typically, simulations begin with $g = 1$.

Think about what the behavior of this update would be:

- Values of N that are oversampled will have their weights decreased, on average, more than values of N that are undersampled, since there will be more updates to them

- When all particle numbers are sampled with equal probability, on average, we will update $\eta(N)$ uniformly across all N. This results in a net additive shift in the weights, which ultimately doesn't affect the microstate probabilities. Thus, if we are sampling a uniform distribution in N, the weights are not effectively modified. In any case, we always demand that $\eta(N_{\mathrm{min}}) = 0$ by shifting the curve.

Thus, the Wang-Landau method enforces a kind of feedback loop between the sampled distribution of N and the determination of the weighting function.

Modification factor schedule
Still, one must address two important issues:

- Modifying the weighting factor at every step breaks detailed balance, because it changes the state probabilities with time.

- We cannot resolve $\eta(N)$ to differences less than the modification factor g.

To resolve the issues, Wang and Landau proposed that $g \to 0$ over the course of the simulation run. They suggested the following procedure:

1. Set $\eta(N) = 0$ and $g = 1$ initially.

2. Perform a simulation, updating $\eta(N)$ at every simulation step using the modification factor. Start a new histogram of N and collect observations throughout the run.

3. When the histogram is "flat enough", scale down the value of the modification factor according to:

$$g \leftarrow \frac{1}{2}g$$

4. Re-zero the histogram counts and return to step 2. Continue until g is very small. Typically, we stop when $g < 10^{-6} - 10^{-8}$.

The histogram can be determined to be flat enough using the 80% rule: when the number of counts in the least-visited histogram bin is no less than 80% of the average number of counts over all bins, the histogram can be considered "flat".

In this way, the simulation proceeds in stages of decreasing values of the modification factor. Initial stages help to rapidly build a good estimate of $\eta(N)$, while later stages refine this calculation to increasing precision and satisfy detailed balance asymptotically.

Reweighting of results
At the end of the simulation, we have computed $\eta(N)$ to such high accuracy that we can use it *directly* to perform reweighting. That is, we assume that $\wp^w(N) = $ const:

$$\wp(N; \mu_2) \propto e^{-\eta(N) + \beta(\mu_2 - \mu_1)}$$

In fact, $\eta(N)$ provides a high-quality estimate of the free energy:

$$A(N) = \mu_1 N + k_B T \eta(N) + \text{const}$$

Choice of initial state
Since the Wang-Landau method automatically and dynamically determines the weighting function, it is typical to choose the initial state of the simulation so that the only factor appearing in the probability distribution is the weighting function itself. In this case, we would choose $\mu_1 = 0$ so that:

$$A(N) = k_B T \eta(N) + \text{const}$$

and the reweighting equation becomes

$$\wp(N; \mu_2) \propto e^{-\eta(N) + \beta \mu_2}$$

Transition matrix methods
The most recent methods to have emerged in biased simulations are those based on **transition matrix estimators.** These have been shown to be very easy to implement and to provide very high quality estimates of free energies, perhaps better than the Wang-Landau approach. Errington and coworkers have pioneered the application of these methods to fluid phase equilibria [Erringron, JCP 118, 9915 (2003)]. Transition matrix approaches are applied to Monte Carlo simulations because they rely on the detailed balance equation.

The idea of transition matrix estimators is that we measure **macrostate transition probabilities** and use these to compute underlying free energies. A macrostate transition probability is simply the probability associated with transitions of the system between different values of some macroscopic parameter. In this example, we measure the conditional probability that a system will make a transition between one value N_1 to another value N_2, given that it initially has N_1 particles:

$$\Pi(N_1 \rightarrow N_2)$$

Notice that, in a grand canonical simulation, we only make incremental steps in N. Thus, $\Pi(N_1 \rightarrow N_2) = 0$ if N_2 is not one of $(N_1 - 1, N_1, N_1 + 1)$.

Definition of the macroscopic transition probability

We can relate the macroscopic transition probability to the microscopic ones that we presented in the initial discussion of MC simulations:

$$\Pi(N_1 \rightarrow N_2) = \frac{\sum_{m \in \{N_1\}} \sum_{n \in \{N_2\}} \wp_m \pi_{mn}}{\sum_{m \in \{N_1\}} \wp_m}$$

Here, the sums over microstates m and n are performed for all states with N_1 and N_2 particles, respectively.

If the microstate transition probabilities obey detailed balance:

$$\wp_m \pi_{mn} = \wp_n \pi_{nm}$$

then a simple summation of this equation over all m and n shows that the macrostate transition probabilities also obey detailed balance:

$$\boxed{\wp(N_1)\Pi(N_1 \rightarrow N_2) = \wp(N_2)\Pi(N_2 \rightarrow N_1)}$$

Estimating free energies from transition matrices

The equation above allows us to estimate free energies from macroscopic transition probabilities. We rearrange it according to

$$\ln \frac{\Pi(N_1 \rightarrow N_2)}{\Pi(N_2 \rightarrow N_1)} = \ln \frac{\wp(N_2)}{\wp(N_1)}$$
$$= \beta\mu(N_2 - N_1) - \beta[A(N_2) - A(N_1)]$$

Solving for the free energy difference,

$$A(N_2) - A(N_1) = \mu(N_2 - N_1) + k_B T \ln \frac{\Pi(N_2 \rightarrow N_1)}{\Pi(N_1 \rightarrow N_2)}$$

Looking at neighboring particle numbers,

$$A(N+1) - A(N) = \mu + k_B T \ln \frac{\Pi(N+1 \to N)}{\Pi(N \to N+1)}$$

Using this equation, we could map out an entire $A(N)$ curve by computing free energy differences at each particle number. To do that, we would need to measure the relative probabilities of seeing transitions between N_1 and N_2. We can do this using a histogram:

$$c_{N_1,N_2} = \text{counts of observations of transitions from } N_1 \text{ to } N_2$$

Then,

$$\Pi(N_1 \to N_2) = \frac{c_{N_1,N_2}}{\sum_i c_{N_1,N_i}}$$

We can actually do better than this. Instead of tallying counts, we can tally the actual acceptance probabilities computed for use in the Metropolis criterion:

$$c_{N_1,N_2} = \text{sums of observations of } P^{\text{acc}}_{N_1,N_2}$$

Computing the weighting function from transition probabilities

The transition probabilities provide us with a way to periodically update an estimate for $\eta(N)$:

$$\eta(N) = \beta A(N) - \beta\mu N + \text{const}$$

where $A(N)$ is determined using the above procedure.

Note that, to determine $A(N)$ using the transition probabilities, we need to be in the unweighted ensemble (since that is what we used to derive the relationship). If we add a weighting function, then it would seem that we need to take this into account. However, the advantage of this approach is that we do not need to take this into account if we sum the acceptance probabilities in the *unweighted* ensemble in our transition matrix, and not in the weighted ensemble. That is, we compute two acceptance probabilities.

For particle additions:

$$P^{\text{acc}}_{12} = \min\left[1, \frac{V}{N+1} e^{-\beta\Delta U + \beta\mu' + \Delta\eta}\right] \quad \to \text{used in acceptance of moves}$$

$$P^{\text{acc}}_{12} = \min\left[1, \frac{V}{N+1} e^{-\beta\Delta U + \beta\mu'}\right] \quad \to \text{used in updating sums in transition probabilities}$$

Notice that we don't include the weight factor in the update of the transition probability matrix. This enables us to compute transition probabilities as if we were still in the unweighted ensemble,

even though we are actually performing a simulation using a weighting function. A similar case exists for particle deletions.

Stratification

When a flat histogram simulation is performed, the system experiences large fluctuations in the flat histogram variable. As such, the correlation time for that variable can be quite large. In other words, it can take the system a very long period of time to explore the complete range of interest in the flat histogram. In our example, this means that the system has a long time scale for traversing values of N between $N_{\min}$ and $N_{\max}$. The time it takes for the system to perform a complete walk of particle numbers between the two limits is called the **tunneling time.**

The tunneling time in flat histogram simulations can grow to be very long for a wide range $N_{\min}$ to $N_{\max}$. If the system performs a random walk in N, we expect

$$\tau_{\text{tunnel}} \propto (N_{\max} - N_{\min})^2$$

In reality, the walk is not completely random due to correlations, even if the system has a perfectly flat distribution in N. Typically the tunneling time grows with an exponent greater than 2.

One way around this problem is to use **stratification** to determine the underlying free energy or entropy function that governs a flat histogram variable. In stratification, we break the entire range of N into smaller, overlapping sub-ranges or **windows.** Then we perform a separate simulation in each. At the end of the run, we know that the underlying free energy estimates should be the same to within an unknown additive constant. We can shift these constants so as to obtain overlap.

Consider that we perform multiple grand canonical simulations for a number of windows each with different $N_{\min}, N_{\max}$. The simulations would reject moves that took their particle number outside of the specified sub-range. From each, we would then obtain a high-quality estimate of $A(N)$:

$$A(N) = \mu N + k_B T \eta(N) + \text{const}$$

Then, we patch the $A(N)$ together by shifting to form a master curve:
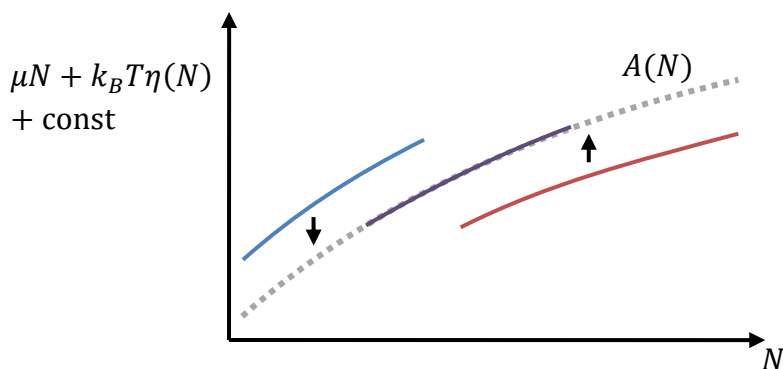
However, the flat histogram approach makes this shifting much easier than the earlier un-weighted case in which we used Ferrenberg-Swendsen reweighting. Here, because each value of N is sampled with the same frequency, we expect the same statistical error in our estimates for $A(N)$ at each value of N. This means that we can use a simple least-squares procedure to optimally shift each curve so as to obtain overlap.

Other parameters

In all of the above examples, we presented flat histogram calculations in the context of a grand canonical simulation. There, we computed a weighting function $\eta(N)$ that had a relation with the N-dependence of the underlying Helmholtz free energy $A(N)$.

All of these methods presented can be used to compute any arbitrary free energy or entropy function, including potentials of mean force like the one presented for umbrella sampling. In general,

> To compute the free energy along a given reaction coordinate, we need to bias the simulation to perform flat-histogram sampling along that coordinate.

The table at the end of this section summarizes the microstate probabilities and reweighting expression that one might use in a number of different flat-histogram sampling. Keep in mind that the microstate probabilities are used to determine acceptance criteria in Monte Carlo simulations.

Keep in mind that continuous coordinates, like the energy or a distance-based reaction coordinate, require us to discretize our weighting function, histograms, and reweighting procedure.

Microscopic order parameters and metadynamics

For example, let's imagine that we want to measure a potential of mean force along some microscopic reaction coordinate, $F(z)$. This could be the free energy of a molecule approaching a

surface, as we discussed in umbrella sampling. We can define an effective or weighted Hamiltonian or potential that includes a biasing term:

$$U^w(\mathbf{r}^N) = U(\mathbf{r}^N) - k_B T \, \eta(z)$$

where U is the normal potential and $\eta(z)$ is a dimensionless function that captures the bias. The weighted probability distribution is

$$\wp^w(\mathbf{r}^N) \propto \wp(\mathbf{r}^N) e^{\eta(z)}$$

If we integrate over all positions except for the reaction coordinate z, we find:

$$\wp^w(z) \propto e^{-\beta F(z)} e^{\eta(z)}$$

Thus if we want flat-histogram like sampling in our weighted ensemble such that $\wp^w(z) = const$, we need:

$$\eta(z) = \beta F(z)$$

Metadynamics is a method that uses this result to determine the PMF. It was introduced by Laio and Parrinello in 2002 and is similar in spirit to the Wang-Landau approach in that it adaptively estimates the underlying bias potential (which is an estimate of the PMF). A major difference with the WL approach is that metadynamics does not accumulate counts of observations in a discrete biasing potential, but estimates the bias as a sum of Gaussians. That is, the approach periodically drops down Gaussians in the biasing potential centered at the current value of the reaction coordinate. In other words,

$$\eta(z) = \sum_{t \in \{t_1, t_2, \dots\}} \omega \exp\left(-\frac{\left(z - z(t)\right)^2}{2\sigma^2} \right)$$

Conceptually, the approach looks like the following, taken from "Free-energy calculations with metadynamics: Theory and practice" by Bussi and Branduardi (2015):
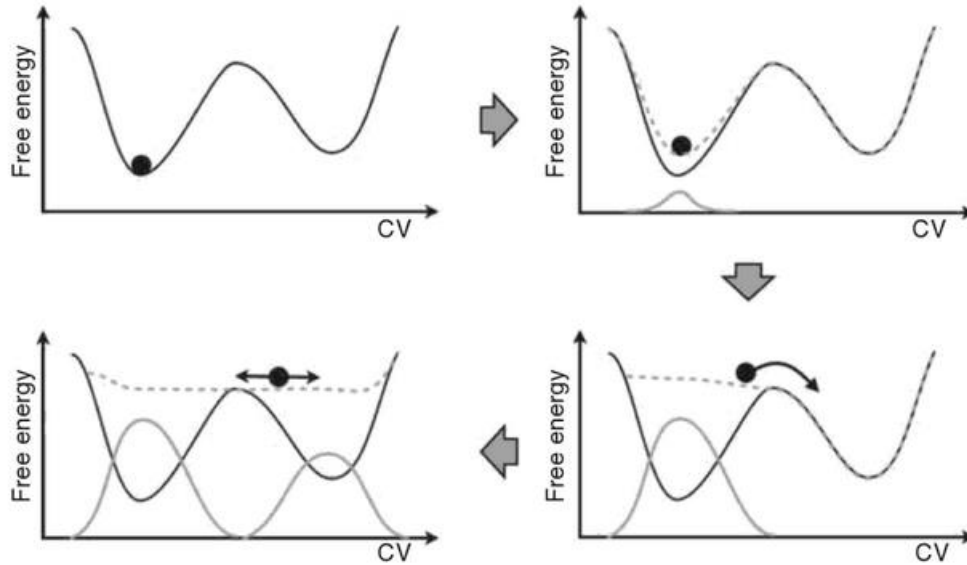
Note here that there are several parameters to choose/tune in the approach:

- the frequency of adding Gaussians

- the width of the Gaussians σ

- the height or magnitude of the Gaussians ω

In the "well-tempered metadynamics" method, the Gaussian height is systematically reduced over time, just as in the Wang Landau approach.

Metadynamics is widely used, more so than the Wang Landau approach, because it can be easily coupled with molecular dynamics. Since $\eta(z)$ is the sum of continuous functions, we can easily obtain derivatives of it. So using our weighted potential, we can derive modified equations of motion. For an atom i:

$$\mathbf{f}_i^{\mathrm{w}} = -\frac{dU^{\mathrm{w}}}{d\mathbf{r}_i}$$

$$= -\frac{dU}{d\mathbf{r}_i} - k_B T \frac{d\eta(z)}{dz} \frac{dz}{d\mathbf{r}_i}$$

We see that the bias effectively adds an additional force that performs the flat-histogram sampling. The last term, $dz/d\mathbf{r}_i$, is a purely geometric one and depends on how the reaction coordinate z is defined in terms of the atomic coordinates. For example, if z is the z-axis center of mass of a collection of atoms:

$$z = \frac{\sum_j m_j z_j}{\sum_j m_j}$$

then we have

$$\frac{dz}{d\mathbf{r}_i} = \frac{m_i}{\sum_j m_j}$$

Common flat-histogram ensembles

	initial ensemble	specified	flat histogram variables	microstate probabilities	ideal weighting function
1	canonical	T_1	U	$\wp^{\mathrm{w}}(\mathbf{r}^N) \propto e^{-\beta_1 U + \eta(U)}$	$\eta(U) = \beta_1 U - S(U)$
2	grand canonical	T_1, μ_1	U, N	$\wp^{\mathrm{w}}(\mathbf{r}^N, N) \propto e^{-\beta_1 U + \beta_1 \mu_1 N + \eta(U,N)}$	$\eta(U,N) = \beta_1 U - \beta_1 \mu_1 N - S(U,N)$
3	isothermal-isobaric	T_1, P_1	U, V	$\wp^{\mathrm{w}}(\mathbf{r}^N, V) \propto e^{-\beta_1 U - \beta_1 P_1 V + \eta(U,V)}$	$\eta(U,V) = \beta_1 U + \beta_1 P_1 V - S(U,V)$
4	grand canonical	T_1, μ_1	N	$\wp^{\mathrm{w}}(\mathbf{r}^N, N) \propto e^{-\beta_1 U + \beta_1 \mu_1 N + \eta(N)}$	$\eta(N) = -\beta_1 \mu_1 N + \beta_1 A(N; T_1)$
5	isothermal-isobaric	T_1, P_1	V	$\wp^{\mathrm{w}}(\mathbf{r}^N, V) \propto e^{-\beta_1 U - \beta_1 P_1 V + \eta(V)}$	$\eta(V) = \beta_1 P_1 V + \beta A(V; T_1)$
6	reaction coord.	T_1	ξ	$\wp^{\mathrm{w}}(\mathbf{r}^N) \propto e^{-\beta_1 U + \eta(\xi)}$	$\eta(\xi) = \beta F(\xi; T_1)$

	initial ensemble	reweightable conditions	reweighting expression
1	canonical	T_2	$\wp(U; T_2) \propto \wp^{\mathrm{w}}(U) e^{-(\beta_2 - \beta_1)U - \eta(U)}$
2	grand canonical	T_2, μ_2	$\wp(U, N; T_2, \mu_2) \propto \wp^{\mathrm{w}}(U, N) e^{-(\beta_2 - \beta_1)U + (\beta_2 \mu_2 - \beta_1 \mu_1)N - \eta(U,N)}$
3	isothermal-isobaric	T_2, P_2	$\wp(U, V; T_2, P_2) \propto \wp^{\mathrm{w}}(U, V) e^{-(\beta_2 - \beta_1)U - (\beta_2 P_2 - \beta_1 P_1)V - \eta(U,V)}$
4	grand canonical	$\mu_2 \; (T_2 = T_1)$	$\wp(N; \mu_2) \propto \wp^{\mathrm{w}}(N) e^{(\beta_2 \mu_2 - \beta_1 \mu_1)N - \eta(N)}$
5	isothermal-isobaric	$P_2 \; (T_2 = T_1)$	$\wp(V; P_2) \propto \wp^{\mathrm{w}}(V) e^{-(\beta_2 P_2 - \beta_1 P_1)V - \eta(V)}$
6	reaction coord.	none	N/A

* S denotes the dimensionless configurational entropy; A denotes the configurational free energy

** All chemical potentials are relative by a factor of $-3 k_B T \ln \Lambda(T)$

 last modified 11/30/2022