Entropic Sampling Monte Carlo

P. B. Sunil Kumar
Department on Physics
Indian Institute of Technology Madras
Chennai 600036
www.physics.iitm.ac.in/~sunil

Typically in equilibrium statistical mechanics you are equipped with some model expression for the energy of the system of interest: $E(\sigma)$, where σ is the microscopic state of the system.

Mostly the system is in contact with a heat bath at constant temperature T

So we calculate the partition function of the system and then from that the associated free energy, in the Helmholtz, Gibbs or the grand canonical ensemble.

In general this is a formidable task, and some approximation methods like mean-field theory or computer simulation techniques must be employed.

Entropic sampling is one technique that can be employed in a Monte Carlo simulation to compute the free energy of the system.

A reminder on importance sampling Monte Carlo

The canonical partition function is given by $$Z = \sum_{\sigma} e^{-\beta E(\sigma)}$$

The probability for a closed system to be in a macrostate σ is $$P(\sigma) = \frac{e^{-\beta E(\sigma)}}{Z}$$

In numerical simulations,
the partition function is approximated by $$Z = \sum_{i=1}^{M} e^{-\beta E(\sigma_i)}$$

and the average of any operator by $$< O >= \sum_{i=1}^{M} O(\sigma_i) \frac{e^{-\beta E(\sigma_i)}}{Z}$$

The accuracy of these estimates depend on the "representative" subset M that is used

In importance sampling the idea is to pick a representative set of conformations which are not completely random, but biased towards conformations which are significantly populated in equilibrium.

A particularly important situation is when $P(\sigma_i) \propto e^{-\beta E(\sigma_i)}$

The average of the operator is now

$$< O >= \frac{\sum_\sigma O(\sigma) e^{-\beta E(\sigma)}}{\sum_\sigma e^{-\beta E(\sigma)}} = \frac{\sum_\sigma O(\sigma) e^{-\beta E(\sigma)} \frac{P(\sigma)}{P(\sigma)}}{\sum_\sigma e^{-\beta E(\sigma)} \frac{P(\sigma)}{P(\sigma)}} = \frac{\sum_i^M O(\sigma_i) e^{-\beta E(\sigma_i)} / P(\sigma_i)}{\sum_i^M e^{-\beta E(\sigma_i)} / P(\sigma_i)} = \frac{\sum_i^M O(\sigma_i)}{M}$$

How do we generate $P(\sigma_i) \propto e^{-\beta E(\sigma_i)}$

The most common way is to generate this as the stationary distribution of a Markov process

From the master equation for probability distributions,

$$P^{n+1}(\sigma) = P^n(\sigma) + \sum_{\sigma'}(P^n(\sigma')W_{\sigma'\sigma} - P^n(\sigma)W_{\sigma\sigma'})$$

we see that the stationarity is ensured by the detailed balance condition.

The Metropolis algorithm, in which P(σ) is proportional to the canonical distribution will give a stationary distribution if

$$W_{\sigma\sigma'} = min(1, e^{-\beta(E(\sigma')-E(\sigma))})$$

- The problem of calculating Z is not solved
- The sampling of high energy states are poor
- Contribution to averages from high energy states are not always negligible
- Estimates of free energy away from the equilibrium (or a minimum in energy) will be bad
- Low probability to cross barriers

Entropic Sampling

The canonical partition function can be written as $Z = \sum_{\sigma} e^{-\beta E(\sigma)} = \sum_{E} \rho(E) e^{-\beta E}$

where, ρ(E) is the density of states with $\rho(E) = e^{-S(E)/k_B}$

If we can compute ρ(E), we can obtain Z(T) from it Laplace transform.

The sampling in configuration space an now be replaced with that in energy space

$$P(E) = \frac{e^{(S(E)/k_B - \beta E)}}{Z}$$

Remember we cannot use the Metropolis importance sampling scheme to get ρ(E) as the statistics at high E is bad

Let us say that instead of the canonical distribution we sample with $P(E) \propto e^{(S(E)/k_B - \phi(E))}$

The transition rates will now be $W_{\sigma\sigma'} = min[1, e^{-(\phi(E(\sigma')) - \phi(E(\sigma)))}]$

Note that ϕ(E)=βE will be the usual Boltzmann sampling and an umbrella potential will be the umbrella sampling

Entropic sampling is obtained when we set $\phi(E) = S(E)/k_B$

which means that the distribution is now P(E)=constant and we simply execute a random walk in energy space !

The trouble is that we do not know what S(E) is !!

Lee's algorithm for entropic sampling

J. Lee New Monte Carlo algorithm: entropic sampling Phys. Rev. Lett. 71 211 (1993).

The idea here is to start with a $\phi(E)$ and take it to $S(E)/k_B$ iteratively

$$\phi^k(E) \rightarrow S(E)/k_B \ for \ k \rightarrow \infty$$

First we choose a suitable energy range [Emin,Emax], which we subdivide into N_{bin} bins of equal size, which serves as a discretization of E. E_i is thus the energy of the i'th bin

1. As a start we set $\Phi^{(0)}=\Phi^{(0)}(E_i)=0$ and histogram $H_i = 0$ for all i.

2. Now, carry out a series of N_{MC} Monte Carlo sweeps, where a histogram Hi is collected.
Hi is thus the number of micro-states visited falling into the i'th energy bin.
In the first iteration step all trial moves will be accepted.

3. At the end of the N_{MC} Monte Carlo sweeps, Φ is updated:
$\Phi_i^{(k+1)} = \Phi_i^{(k)} + (1 - \delta_{0,Hi}) \cdot \ln(H_i)$
So, $\ln(H_i)$ is added to $\Phi^{(k)}$ unless the i'th bin was not visited. Notice, that the bins which have been visited now is less favored in the Metropolis algorithm.

4. Reset the histogram, $H_i = 0$, and repeat step 2,3 above. This iterative procedure can go on until a k_{max} where a reasonable flat histogram H_i has been obtained ($P^{(kmax)}(E) \simeq$ constant).

The obtained $\Phi_i^{(kmax)}$ is now an estimate for $S(E_i)/k_B$, which can be used for many purposes, e.g. for numerical evaluation of the internal energy at any temperature T.

Once the construction of the function $\phi(E)$ is completed we carry out a "production" run to collect an "entropic ensemble" of macrostates.

The canonical average at any temperature can then be calculated by re-weighting

$$< O >= \frac{\sum_{i=1}^{M} O(\sigma_i) e^{(-\beta E(\sigma_i)+\phi(E(\sigma_i))}}{\sum_{i=1}^{M} e^{(-\beta E(\sigma_i)+\phi(E(\sigma_i))}}$$

Thus, by one sampling run we have the averages at all temperatures .
Since the sampling is through a random walk in energy space, we sample all rare events .

In general, Lee's procedure is safe, i.e. it is converging, but slowly.

The practical usage has thus been limited.

Wang-Landau method

- This is a variation of the Lee's algorithm.
- The basis is still the entropic sampling idea.
- The major difference is that Φ_i is updated after every MCS, making the stochastic process highly non-Markovian.
- Φi is updated by a convergence factor f, which is adjusted iteratively.
- The energy histogram is updated as well, but not directly used in the update of Φ_i as in Lee's algorithm.
- A flatness criteria on the energy histograms is used to update f.

1. Set $\Phi^{(0)} = \Phi^{(0)}(E_i) = 0$ and $H_i = 0$ for all i, like in the Lee algorithm, and the convergence factor is initialized, e.g. $f_i = 0.5$.

2. Carry out a series of N_k Monte Carlo sweeps. For each update (k), E_i is identified and Φ_i updated: $\Phi_i^{(k+1)} = \Phi i^{(k)} + f_i$

3. After the N_k MCS the histogram i updated:
the current energy E_i is identified and $H_i^{(l+1)} = H_i^{(l)} + 1$

4. Return to 2 and perform these iterations N_l times.

5. Check, if the histogram $\{H_i\}$ is flat.
If NO: return to 2 and continue with another $N_k \times N_l$ updates.
If YES: Update convergence factor $f = 0.5 \times f$ and start again with the histogram $H_i = 0$ for all i.
If f gets below some criteria, e.g. $f_{min} \sim 10^{-8}$, stop, otherwise return to 2.

Few points to note about Wang Landau method

As a flatness criteria, Wang and Landau used the relation $H_i > \eta \sum_{i=1}^{N_{bin}} H_i/N_{bin}$ for all i, where flatness parameter $\eta \simeq 0.8 - 0.9$.

The algorithm does not have the same strong theoretical foundation as Lee's algorithm

E.g. continuous update of $\phi(E)$ implies that the Metropolis algorithm has completely lost its original meaning as a Markov process transition probability to ensure a particular stationary probability distribution.

It has just become an importance sampling machine which ensures that the already visited energy bins are penalized.

However, in the limit of $f \to 0$ the algorithm starts to look like Lee's algorithm and we are on safe theoretical ground.

We don't know which convergence criteria are necessary to acquire a particular accuracy in the estimate of the density of states.

But IT WORKS very efficiently!

The method has been extended to "expanded ensembles"

The expanded ensemble is usually characterized by some reaction coordinate D, which groups states of a system to sub-states with different values of reaction coordinate.

We can then determine the multi-dimensional density of state ρ(E,D) by constructing the histogram, now not only in E but also in D

Each sub-state of the expanded ensemble is weighted by the probability density g(D), where D is a reaction coordinate

For example: if we want to look at the free energy of a polymer as a function of its radius of gyration

1. Set up a usual Metropolis MC simulation for the polymer
2. After N MC moves, sample the energy E and radius of gyration R of the polymer and construct the histogram H(E,R).
3. Update ϕ(E,R) until the two dimensional histogram H(E,R) is flat.
4. The two dimensional density of state is now obtained as exp(ϕ(E,R)) from which $g(R) = \sum_E \exp(\phi(E,R))$

Wang-Landau sampling along a reaction coordinate - an example from polymer Physics

How to calculate the free energy of a polymer as a function of radius of gyration

Here the reaction coordinate is the radius of gyration Rg and the density of states ρ(E) is now replaced by probability $\rho(E,R_g)$.

A histogram in (E,R_g) space $H(E,R_g)$ should be used to compute this probability, from which we can compute the free energy $F(R_g, T)$.

To simplify the numerical process we compute a reduced probability density g(Rg) and sample, instead of histogram in energy space H(E), a histogram H(Rg).

Our aim is to get the equilibrium distribution of Rg of a PE, from which the free energy F(Rg) can be calculated as, $$F(R_g) = -k_B T \ln g(R_g)$$

In the simulations we will use the variable $\phi(R_g) = -k_B T \ln g(R_g)$ and write the transition probability between two states as $$W_{\sigma\sigma'} = min[1, e^{-\beta((\phi(R_g)+E(R_g))-(\phi(R'_g)+E'(R_g)))}]$$

The procedure is now

0. Choose a range of R_g values and bin it to intervals labelled by i

1. Set $\Phi^{(0)} = \Phi^{(0)}(R_g) = 0$ and $H(Rg) = 0$ for all i, the convergence factor is initialized, e.g. $f_i = 0.5$.

2. Carry out a series of N_k Monte Carlo sweeps. For each update (k), R_g is identified and Φ_i is updated: $\Phi_i^{(k+1)} = \Phi i^{(k)} + f_i$

3. After the N_k MCS the histogram i updated:
the current energy E_i is identified and $H_i^{(l+1)} = H_i^{(l)} + 1$

4. Return to 2 and perform these iterations N_l times.

5. Check, if the histogram $\{H_i\}$ is flat.
If NO: return to 2 and continue with another $N_k \times N_l$ updates.
If YES: Update convergence factor $f = 0.5 \times f$ and start again with the histogram $H_i = 0$ for all i.
If f gets below some criteria, e.g. $f_{min} \sim 10^{-8}$, stop, otherwise return to 2.

Umbrella sampling to compute the free energy of a polymer

We saw that the distribution function and the free energy are related through

$$F(R_g) = -k_B T \ln P(R_g)$$

P(R_g) is given by $$P(R_g) =< \delta(R - R_g) > -\frac{\int e^{-\beta E}\delta(R - R_g)dR}{\int e^{-\beta E}dR}$$

We also saw that the idea of umbrella sampling is to use a bias potential.

Here we will use a harmonic bias potential $V(R_g) = C(R_g - R_0)^2/2$

By varying the value of R_0 we can sample the entire range of R_g values, including regions where the canonical probability is small.

The biased sampling probability is then $$P_b(R_g) =< \delta(R - R_g) >_b = \frac{\int e^{-\beta(E+V)}\delta(R - R_g)dR}{\int e^{-\beta(E+V)}dR}$$

The unbiased probability is then given by

$$P(R_g) = < \delta(R - R_g) >= \frac{\int e^{-\beta(E+V)}\delta(R - R_g)e^{\beta V}dR}{\int e^{-\beta(E+V)}e^{\beta V}dR} = \frac{\left\langle e^{\beta V}\delta(R - R_g)\right\rangle_b}{\left\langle e^{\beta V}\right\rangle b}$$

$$= \frac{e^{\beta V}\left\langle \delta(R - R_g)\right\rangle_b}{\left\langle e^{\beta V}\right\rangle b} = \frac{e^{\beta V}P_b(R_g)}{\left\langle e^{\beta V}\right\rangle b}$$

The free energy is then given by

$$\beta F(R_g) = -\beta V - \ln P_b(R_g) + \ln \left\langle e^{\beta V}\right\rangle_b$$

Thank you for your attention