Chem 860. Lecture 13

Basic Monte Carlo

March 3, 2009

As we discussed in the beginning of the course, one can obtain equilibrium static properties from stochastic simulations in which configurations are NOT related by equation of motion. The requirement (in the most basic scheme) is that these configurations follow pre-defined distribution - which often is the Boltzmann distribution. These stochastic simulation methods are often referred to as "Monte Carlo" methods due to the obvious reason. Several major advantages of Monte Carlo methods compared to Molecular Dynamics, which will soon be clear, are:

- Force evaluation is NOT required
- Easy to implement for different ensembles
- Flexible moves (e.g., identity exchange) and biases (e.g., multi-canonical MC, parallel tempering)

1 Metropolis Monte Carlo

1.1 Motivation for Importance Sampling

To compute the ensemble average of a quantity,

$$< A >= \int d\mathbf{X} A(\mathbf{X}) P(\mathbf{X})$$

all we have to do is constructing an algorithm for evaluating this multi-dimensional integral very efficiently - e.g., focus on regions that have high $P(\mathbf{X})$ instead of uniformly. This is the basic idea of "Importance Sampling".

If we know $P(\mathbf{X})$, the integral is straightforward to evaluate. Consider a one-dimensional example,

$$I = \int_0^1 f(x)dx = \int_0^1 dx w(x) \frac{f(x)}{w(x)}$$

Let $du(x)/dx = w(x)$, then we have

$$I = \int_{u(0)}^{u(1)} du \frac{f[x(u)]}{w[x(u)]}$$

So we can generate a uniform distribution for $u(x)$, then evaluate the average of $\frac{f[x(u)]}{w[x(u)]}$ over those points. One can show that the statistical uncertainty associated with this average over L points is

$$\sigma_I^2 = \frac{1}{L}\left[< (f/w)^2 > - < f/w >^2\right]$$

Therefore, if $w(x)$ is chosen such that f/w is a smooth function, the variance would be much smaller than that for $w(x) = 1$ (uniform distribution in x).

1.2 Metropolis scheme

A slight problem is that we do not know the probability density, $P(\mathbf{X})$, itself. Rather, we only know $exp[-\beta U(\mathbf{X})]$ and not the configuration integral, $Z = \int d\mathbf{X} exp[-\beta U(\mathbf{X})]$. So we have to develop an algorithm such that configurations are generated with the **relative** probability proportional to the Boltzmann factor.

Starting from a configuration o ("old") with Boltzmann factor $exp[-\beta U(o)]$, add a displacement to reach configuration n ("new") with Boltzmann factor $exp[-\beta U(n)]$. We have to develop a rule to reject/accept n such that on average the relative probability of finding the system in n and o is equal to the ratio of the Boltzmann factors.

The transition probability from o to n is written as $\pi(o \to n)$, which involves two steps - pick a move, then reject/accept the move,

$$\pi(o \to n) = \alpha(o \to n) \times acc(o \to n)$$

$\alpha(o \to n)$ is called the underlying matrix for the Markov chain. In general, it can be non-symmetric (i.e.,$\alpha(o \to n) \neq \alpha(n \to o)$), although often it is symmetric (e.g., uniformly random in x).

The (strong) condition that the transition probability has to satisfy under equilibrium is *detailed balance*:

$$P(o)\pi(o \to n) = P(n)\pi(n \to o)$$

Note: detailed balance comes from the discussion of Master Equation, which can be written generally as:

$$\frac{dP_k}{dt} = \sum_{l \neq k} [\pi_{kl} P_l - \pi_{lk} P_k] \tag{1}$$

where P_k is the probability (population) for state k and π_{kl} is the transition probability from state l to k. Clearly, if detailed balance is satisfied for all state pairs, we have a stationary distribution, i.e., $dP_k/dt = 0$ for all k. But you can also see, this is a strong condition; i.e., a stationary distribution does NOT require detailed balance to be satisfied!

The *Metropolis scheme*:

$$\pi(o \to n) = \alpha(o \to n), P(n) \geq P(o)$$

$$\pi(o \to n) = \alpha(o \to n)[P(n)/P(o)], P(n) \leq P(o)$$

$$\pi(o \to o) = 1 - \sum_{n \neq o} \pi(o \to n)$$

or equivalently,

$$acc(o \to n) = min[1, P(n)/P(o)]$$

Does the *Metropolis scheme* satisfy detailed balance? Imagine that $U(n) < U(o)$, then

$$\pi(o \to n) = \alpha(o \to n)$$

$$\pi(n \to o) = \alpha(n \to o)[P(o)/P(n)]$$

which means that, if α is symmetric,

$$\frac{\pi(o \to n)}{\pi(n \to o)} = \frac{P(n)}{P(o)}$$

Accepting move with a probability $P(n)/P(o) < 1$: generate a random number uniformly in $[0, 1]$. If it is smaller than $P(n)/P(o)$, accept the move. Otherwise, reject. For a long MC run, a good random generator routine is useful.

1.3 A few practical points

• Displacements: single or all?

Since the probability of accepting a move is small if the energy rise is much larger than $k_B T$, it is much better to move one particle at a time than moving all particles at the same time in a random fashion. However, one can generate collective moves using MD (with large step size) which does not change energy much. This is the so-called hybrid Monte Carlo approach, which was shown to sample better than the standard MC.

• Size of displacement - optimal acceptance ratio?

Clearly the size of displacement should be adjusted such that the acceptance ratio is reasonable. One considers both the efficiency of sampling (how far we move around) and ergodicity of sampling (sample very different configurations). Previous work, for example that of Thirumalai et al. (Physica, A, 210, 453, 1994). showed that the optimal acceptance ratio is 20% instead of the naive 50%.

• Importance of detailed balance

Should avoid: changing maximum stepsize after equilibration.

Even if a move is rejected, the old configuration should still be counted in the average! because $\pi(o \rightarrow o)$ is in general not zero.

• Unphysical moves can significantly speed things up! Especially for highly packed systems such as solid mixtures, dense polymers and proteins.