Hybrid Monte Carlo Methods

Hybrid Monte Carlo methods attempt to combine the best aspects of Molecular Dynamics and Monte Carlo sampling. The derivation of the acceptance test introduces some subtle requirements to the generation of the random velocities. This paper will derive the Hybrid Monte Carlo acceptance test, show how the random velocities should be generated, and will then explore alternative velocity generation methods and will derive their associated HMC acceptance tests.

Standard Hybrid Monte Carlo

The Hamiltonian, H, for a system is given by;

$$H(\mathbf{x},\mathbf{p}) = K(\mathbf{p}) + U(\mathbf{x})$$

$$H(\mathbf{x},\mathbf{p}) = \sum_{i=1}^{n} \frac{p_i^2}{2m_i} + \sum_{i<j} u_{ij}$$

where **x** is the vector of positions of the atoms, **p** is the vector of momentums of the atoms, K is the total kinetic energy of the system, U is the total potential energy, p_i is an individual atom's momentum, m_i is the mass of the atom, n is the number of atoms, and u_{ij} is the interaction energy between a pair of atoms.

The probability, **P**, of observing the configuration (**x**,**p**) is given by the Boltzmann equation;

$$\mathbf{P}(\mathbf{x},\mathbf{p}) = \frac{\exp(-\beta H(\mathbf{x},\mathbf{p}))}{Q}$$

where β is the inverse temperature ($1/k_B T$) and Q is the partition function. This can be split into separate terms for the kinetic and potential energy. The kinetic term can be integrated analytically, leaving just the potential term for numerical integration;

$$\mathbf{P}(\mathbf{x}) = \frac{\exp(-\beta U(\mathbf{x}))}{Z}$$

where Z is the configurational integral. A distribution with this probability density can be generated using Monte Carlo sampling. This uses a Markov chain defined by the conditional probability density of moving from one configuration in the chain (**x**) to the next configuration (**x′**) as $p_m(\mathbf{x} \rightarrow \mathbf{x}')$. To derive the form of this acceptance probability, we apply the over-strict detailed balance condition;

$$\mathbf{P}(\mathbf{x})p_m(\mathbf{x} \rightarrow \mathbf{x}') = \mathbf{P}(\mathbf{x}')p_m(\mathbf{x}' \rightarrow \mathbf{x})$$

This can be achieved by attempting a move from **x** to **x′** with the probability $p_s(\mathbf{x} \rightarrow \mathbf{x}')$, and accepting it with the probability;

$$P_A(\mathbf{x} \rightarrow \mathbf{x}') = \min\left\{1, \frac{\mathbf{P}(\mathbf{x}')p_s(\mathbf{x}' \rightarrow \mathbf{x})}{\mathbf{P}(\mathbf{x})p_s(\mathbf{x} \rightarrow \mathbf{x}')}\right\}$$

Hybrid Monte Carlo (HMC) uses NVE Molecular Dynamics to attempt the move from **x** to **x′**. Because of the deterministic nature of the integrator, for the given timestep, δt, the probability of moving from **x** to **x′** is given by the probability of choosing the momentum vector, **p**, that, when integrated from **x** using the deterministic integrator for a period of δt, results in configuration **x′**. Or in other words;

$$p_s(\mathbf{x} \rightarrow \mathbf{x}') = p_s(\mathbf{p})$$

Here is the subtle bit! We now choose the momentum vector from the following Gaussian distribution!

$$p_s(P) \propto \exp\left[-\beta \sum_{j=1}^{n} \frac{p_j^2}{2m}\right]$$

To do this, we must pick each momentum from the following Gaussian distribution;

$$p_s(\mathbf{P}) = \prod_{j=1}^{n} p_s(p_j) = \prod_{j=1}^{n} \exp\left[-\beta \frac{p_j^2}{2m}\right]$$

(e.g. generate the momentum for each component of the vector independently from the Gaussian distribution).

Why do we choose this momentum vector? Note how the probability is very close in form to that of the kinetic energy. If we substitute back into the acceptance probability we get;

$$P_A(\mathbf{x} \rightarrow \mathbf{x}') = \min\left\{1, \frac{\mathbf{P}(\mathbf{x}')\exp\left[-\beta \sum_{j=1}^{n} \frac{p_j'^2}{2m}\right]}{\mathbf{P}(\mathbf{x})\exp\left[-\beta \sum_{j=1}^{n} \frac{p_j^2}{2m}\right]}\right\}$$

Substituting in for **P(x)** and **P(x′)** gives;

$$P_A(\mathbf{x} \to \mathbf{x}') = \min\left\{1, \frac{\exp(-\beta U(\mathbf{x}'))\exp\left[-\beta \sum_{j=1}^{n} \frac{p_j'^2}{2m}\right]}{\exp(-\beta U(\mathbf{x}))\exp\left[-\beta \sum_{j=1}^{n} \frac{p_j^2}{2m}\right]}\right\}$$

Recognising that $\sum_{j=1}^{n} \frac{p_j'^2}{2m}$ is equal to the kinetic energy, K(**p**), and collapsing the exponentials together gives us;

$$P_A(\mathbf{x} \to \mathbf{x}') = \min\left\{1, \frac{\exp(-\beta U(\mathbf{x}') - \beta K(\mathbf{p}'))}{\exp(-\beta U(\mathbf{x}) - \beta K(\mathbf{p}))}\right\}$$

$$P_A(\mathbf{x} \to \mathbf{x}') = \min\left\{1, \frac{\exp(-\beta H(\mathbf{x}', \mathbf{p}'))}{\exp(-\beta H(\mathbf{x}, \mathbf{p}))}\right\} = \min\{1, \exp(-\beta \Delta H(\mathbf{x}, \mathbf{p}))\}$$

Notice how the acceptance test is now on the **total energy** of the system, not just on the potential energy. This is why we must choose the momentum vector as described, so that this convenient result may emerge!

Also note that **p**′ is the negative of the momentum vector at the end of the integration from **x** to **x**′ (as it is the momentum vector required to integrate back to **x** from **x**′).

Momentum Enhanced Hybrid Monte Carlo

Momentum Enhanced Hybrid Monte Carlo (MEHMC) biases the generation of the random momentum vector such that the chosen vector advances the sampling along the slow degrees of freedom. It achieves this by forming the biasing momentum vector, $\mathbf{p_0}$, as the average momentum from a large number of time steps. The random momentum vector is generated according to the probability distribution;

$$p_s(\mathbf{P}) \propto \exp\left(-\beta \sum_{j=1}^{n} \frac{p_j^2}{2m} - \beta \sum_{j=1}^{n} B_{0j} p_j\right) + \exp\left(-\beta \sum_{j=1}^{n} \frac{p_j^2}{2m} + \beta \sum_{j=1}^{n} B_{0j} p_j\right)$$

where $\mathbf{B} = \zeta \mathbf{p_0}$ and ζ is a multiplicative constant. The exponentials can be factored out to get;

$$p_s(\mathbf{P}) \propto \exp\left(-\beta \sum_{j=1}^{n} \frac{p_j^2}{2m}\right)\left[\exp\left(\beta \sum_{j=1}^{n} B_{0j} p_j\right) + \exp\left(-\beta \sum_{j=1}^{n} B_{0j} p_j\right)\right]$$

substituting this back into the equation for the acceptance probability we get;

$$P_A(\mathbf{x} \to \mathbf{x}') = \min\left\{1, \frac{\exp(-\beta U(\mathbf{x}'))\exp\left[-\beta\sum_{j=1}^{n}\frac{p_j'^2}{2m}\right]\left[\exp\left(\beta\sum_{j=1}^{n}B_{0j}p_j'\right)+\exp\left(-\beta\sum_{j=1}^{n}B_{0j}p_j'\right)\right]}{\exp(-\beta U(\mathbf{x}))\exp\left[-\beta\sum_{j=1}^{n}\frac{p_j^2}{2m}\right]\left[\exp\left(\beta\sum_{j=1}^{n}B_{0j}p_j\right)+\exp\left(-\beta\sum_{j=1}^{n}B_{0j}p_j\right)\right]}\right\}$$

recognising the kinetic energy in this equation, and replacing $\sum_{j=1}^{n}B_{0j}p_j$ with $\mathbf{Bp}$ gives;

$$P_A(\mathbf{x} \to \mathbf{x}') = \min\left\{1, \frac{\exp(-\beta U(\mathbf{x}'))\exp(-\beta K(\mathbf{p}'))[\exp(\beta\mathbf{Bp}')+\exp(-\beta\mathbf{Bp}')]}{\exp(-\beta U(\mathbf{x}))\exp(-\beta K(\mathbf{p}))[\exp(\beta\mathbf{Bp})+\exp(-\beta\mathbf{Bp})]}\right\}$$

$$P_A(\mathbf{x} \to \mathbf{x}') = \min\left\{1, \frac{\exp(-\beta H(\mathbf{x}',\mathbf{p}'))}{\exp(-\beta H(\mathbf{x},\mathbf{p}))}\frac{\exp(\beta\mathbf{Bp}')+\exp(-\beta\mathbf{Bp}')}{\exp(\beta\mathbf{Bp})+\exp(-\beta\mathbf{Bp})}\right\}$$

$$P_A(\mathbf{x} \to \mathbf{x}') = \min\left\{1, \exp(-\beta\Delta H(\mathbf{x},\mathbf{p}))\frac{\exp(\beta\mathbf{Bp}')+\exp(-\beta\mathbf{Bp}')}{\exp(\beta\mathbf{Bp})+\exp(-\beta\mathbf{Bp})}\right\}$$

Note that the aim of this distribution is to enhance the probability of choosing a momentum vector that is close to the average momentum vector, or close to the negative of the average momentum vector. This distribution is not really the true binormal distribution, as the paper takes a few shortcuts…

The true binormal distribution would be (see random.pdf);

$$p_s(\mathbf{p}) \propto \exp\left(-\beta\sum_{j=1}^{n}\frac{(p_j - p_0)^2}{2m_j}\right) + \exp\left(-\beta\sum_{j=1}^{n}\frac{(p_j + p_0)^2}{2m_j}\right)$$

if we expand this, and collect together the common factors, we get;

$$p_s(\mathbf{p}) \propto \exp\left(-\beta\sum_{j=1}^{n}\frac{p_j^2}{2m_j} - \frac{p_0 p_j}{2m_j} + \frac{p_0^2}{2m_j}\right) + \exp\left(-\beta\sum_{j=1}^{n}\frac{p_j^2}{2m_j} + \frac{p_0 p_j}{2m_j} + \frac{p_0^2}{2m_j}\right)$$

$$p_s(\mathbf{p}) \propto \exp\left(-\beta\sum_{j=1}^{n}\frac{p_j^2}{2m_j}\right)\exp\left(-\beta\sum_{j=1}^{n}\frac{p_0^2}{2m_j}\right)\left[\exp\left(-\beta\sum_{j=1}^{n}-\frac{p_0 p_j}{2m_j}\right)+\exp\left(-\beta\sum_{j=1}^{n}+\frac{p_0 p_j}{2m_j}\right)\right]$$

$\exp\left(-\beta\sum_{j=1}^{n}\frac{p_0^2}{2m_j}\right)$ is a constant, so it can be dropped. We can also recognise the kinetic energy in the first term. Making these changes leaves us with;

$$p_s(\mathbf{p}) \propto \exp(-\beta K(\mathbf{p}))\left[\exp\left(-\beta\sum_{j=1}^{n}-\frac{p_0 p_j}{2m_j}\right)+\exp\left(-\beta\sum_{j=1}^{n}+\frac{p_0 p_j}{2m_j}\right)\right]$$

If we now say that $\mathbf{B} = \frac{\mathbf{p_0}}{\mathbf{2m}}$ then we get;

$$p_s(\mathbf{p}) \propto \exp(-\beta K(\mathbf{p}))[\exp(-\beta\mathbf{Bp}) + \exp(-\beta\mathbf{Bp})]$$

which was exactly what was presented in the original MEHMC paper. We also now see that the multiplicative constant is 1/2**m**.

This means that we can generate random momentum vectors using the binormal distribution (as described in random.pdf) and then accept or reject the moves using the same acceptance test as presented above and in the original paper.