

## Free Energy Estimation in Simulations

This is a brief overview of free energy estimation, much of which can be found in the article by D. Frenkel *Free-Energy Computation and First-Order Phase Transitions*, in **MD simulation of Statistical Mechanical Systems**, edited by Ciccotti and Hoover. Generic remarks are given in in *A&T* on pg. 40, otherwise for specific methods we give the various relevant pages.

The free energy is related to the partition function by:

$$e^{-\beta F} = \frac{1}{N! \Lambda^{dN}} \int d\mathbf{R} e^{-\beta V(\mathbf{R})} = Z \quad , \quad (1)$$

where  $N$  is number of particles,  $d$  is the dimension,  $\beta = (k_B T)^{-1}$ , and  $\Lambda = (4\pi\lambda\beta)^{-3/2}$ , the classical thermal de Broglie wavelength. The Gibbs' free energy is

$$G = \mu N = F + PV \quad , \quad (2)$$

where  $\mu$  is the chemical potential per particle. The entropy can be obtained from the free energy.

Free energy differences are needed to determine phase stability. Direct simulation of stability of a given phase is too expensive and unreliable. For example the different phases may have different size effects, there can be critical slowing down and hysteresis effects. In experiment, the free energy is determined experimentally by integrating over a reversible path; one cannot measure it directly.

### Seven Different Methods for Free-Energy Estimation

We will briefly describe seven methods that may be used in order to calculate the free energies. Each method has its own difficulties and, therefore, the choice of method for a particular problem should be chosen with the specific application in mind. Careful checks and analysis are required to have trustworthy results.

For lattice Hamiltonians some of the methods described are not applicable because many ground states may compete at finite temperature, depending on the range and types (i.e., pair, triplet, etc.) of interactions. Therefore you usually need to know which and how many Free Energies are needed *a priori* via simulation to construct the temperature versus composition phase diagram, beyond simply performing the simulation.

#### 1. Thermodynamic Integration

With this method one establishes the free energy in a reference state and then integrates the derivative of the free energy (which we can calculate) along a path from the reference state to the desired state with respect to the density or the temperature or even with respect to an artificial parameter. To integrate with respect to the density we have:

$$\left( \frac{dF}{dV} \right)_{NT} = -P \quad \Rightarrow F(V) = F_0(V) + \int_0^{\rho} d\rho' \frac{N\beta}{\rho'} \left( \frac{\beta P}{\rho'} - 1 \right) \quad . \quad (3)$$

Here,  $F_0$  is the free particle free energy, which is known exactly. For a fluid one is performing a density integration from  $\rho = 0$  to the density of interest.

$$\left( \frac{d(F\beta)}{d\beta} \right)_{NV} = E \quad \Rightarrow F(T) = F_0(T) + \beta^{-1} \int_0^{\beta} d\beta' (E(\beta') - E_0(\beta')) \quad . \quad (4)$$

Here,  $F_0$  and  $E_0$  are the free energy and energy of a reference system, which must be known exactly. For a solid, a temperature integration is performed from  $\beta = \infty$  to the temperature of interest.

When performing these thermal integrations, you must not cross a phase boundary, since the derivatives needed will depend very sensitively on the size of system and may be subject to critical slowing down. Notice that many careful simulations are required but this method always works! Clever choices of the path can save much computer time.

## 2. Direct Methods

**\*Hit or Miss Method** - Sample some probability distribution  $p(R)$  with  $\int dR p(R) = 1$

$$e^{-\beta F} = \frac{1}{N! \Lambda^{3N}} \int d\mathbf{R} p(R) \frac{e^{-\beta V(\mathbf{R})}}{p(R)} = e^{-\beta F_0} \left\langle \frac{e^{-\beta V(\mathbf{R})}}{p(R)} \right\rangle_p . \quad (5)$$

For non-trivial problems, this method gives a very large variance since it is very difficult to make  $p(R)$  large where  $V(R)$  is small.

**\*Sample  $\frac{e^{-\beta V}}{Z}$  Method** - Calculate the following estimator.

$$Z^{-1} = \langle e^{-\beta H} \rangle . \quad (6)$$

However, this has a very large variance unless  $V(R)$  is very weak.

## 3. Particle Insertion/Deletion Methods

Calculate the following average to obtain directly the chemical potential relative to the *ideal* gas (see *A & T* pgs. 49-50).

$$\mu = \mu_{ideal} - k_B T \ln \left\langle e^{-\beta V_N(r_N) - V_{N-1}(r)} \right\rangle_{N-1} . \quad (7)$$

Here we sample  $e^{-\beta V_{N-1}}$ , a  $N-1$  particle system, and add a particle uniformly at random and find the change in energy,  $V_N - V_{N-1}$ . Recall,  $\mu_{ideal} = k_B T \ln(\Lambda^{cl} \rho)$ . This method becomes very difficult as density increases and impossible for a solid. Also random deletion has infinite variance for hardcore systems.

## 4. Grand Canonical Ensemble Method

Given the Grand Canonical Ensemble of constant  $\mu$ ,  $T$ ,  $V$ , one may calculate from the Gibbs' Free Energy (see *A & T* pgs. 128) the Free Energy as

$$\frac{F}{N} = \mu - \frac{\langle P \rangle V}{\langle N \rangle} . \quad (8)$$

This technique had very slow convergence at high densities (i.e., solids) and roughly only good for  $\rho < \rho_{critical}$ . In addition, there is a large size dependence.

## 5. Artificial Thermodynamic Integration

Here one must sample from an artificial distribution  $(R, \lambda)$  such that

$$\Pi(R, 0) = \text{distribution with known } F_0 \quad (9)$$

$$\Pi(R, 1) = \frac{e^{-\beta V}}{Z} . \quad (10)$$

Here  $\lambda$  is an adjustable parameter to "turn on" the interactions:  $H(\lambda) = H_0 + \lambda H_1$ , where we know how to obtain the free energy, etc., for  $H_0$ . The partition function may be obtained as follows:

$$Z(\lambda) = \int d\mathbf{R} \Pi(R, \lambda) \quad (11)$$

$$Z(\lambda = 1) = Z_0 \exp \left[ \int_0^1 d\lambda \left\langle \frac{d \ln \Pi((R, \lambda))}{d\lambda} \right\rangle_\lambda \right] . \quad (12)$$

Care must be taken when using this method.

1. Carefully pick a reference state  $\Pi(R, 0)$  as close to  $\Pi(R, 1)$  as possible.
2. Make sure there are no *Phase Boundaries* between  $\Pi(R, 0)$  and  $\Pi(R, 1)$ .
3.  $\ln(\Pi(R, \lambda)) = \ln(\Pi(R, 0)) + \lambda \ln \left( \frac{\Pi(R, 1)}{\Pi(R, 0)} \right)$ .
4. simulate  $\Pi(R, \lambda)$  for several  $\lambda$ 's.

We now obtain the Free Energy of system by

$$F = F_0 + \beta^{-1} \int_0^1 d\lambda \left\langle \ln \left( \frac{\Pi(R, 1)}{\Pi(R, 0)} \right) \right\rangle_\lambda . \quad (13)$$

Beware of any non-analytic behavior as a function of  $\lambda$ . Also choose CPU time ( $\propto \lambda$ 's) so that uncertainties are independent of  $\lambda$ .

Let us consider the case where only the potential matters. Here  $V(\lambda) = V_0 + \lambda V_1$ , and, therefore,  $\lambda=0$  we have  $V(\lambda) = V_0$  with known  $F_0$ , whereas, for  $\lambda=1$  we have  $V(\lambda)$ . Generally, then

$$F(\lambda) = -k_B T \ln \left[ \frac{1}{N! \Lambda^{3N}} \int d^N q e^{-\beta V(\lambda)} \right] . \quad (14)$$

With  $F_0$  known,

$$\frac{dF(\lambda)}{d\lambda} = \frac{\int d^N q \frac{dV(\lambda)}{d\lambda} e^{-\beta V(\lambda)}}{\int d^N q e^{-\beta V(\lambda)}} = \left\langle \frac{dV(\lambda)}{d\lambda} \right\rangle_\lambda . \quad (15)$$

Notice that  $\left\langle \frac{dV(\lambda)}{d\lambda} \right\rangle_\lambda$ , a canonical average of a mechanical quantity for a system with  $H(\lambda)$ , can be obtained by a computer simulation, just as we have done for all other mechanical properties!

Examples of Artificial Integrations

1. Map a complicated liquid into a simple liquid (e.g., Car-Parinello Si  $\Rightarrow$  empirical force Si).
2. Realistic solid  $\Rightarrow$  Einstein solid, so as to use  $V_E = K \sum (r_i - r_{0,i})^2$ .
3. Perfect crystal  $\Rightarrow$  defective crystal, e.g., turn off one atom.
4. Bcc crystal  $\Rightarrow$  Fcc crystal, via distortion of simulation box.
5. Coiled protein  $\Rightarrow$  straight protein, by applying end-to-end force.

## 6. Acceptance Ratio Method (Bennett)

Suppose you want to calculate the free energy distribution between  $\Pi_0(R)$  and  $\Pi_1(R)$ , where  $Z_k = \int \Pi_k$  for  $k=0,1$ . We then are considering an enlarged sample space, i.e.  $(R,k)$ . System is allowed to hop from 0 to 1. Thus, there are two types of moves: 1) usual  $R \rightarrow R'$ , and 2) transition between  $k \rightarrow k'$ , given by T. Here,  $P_k$  = probability that system is in state  $k$  (irrespective of  $R$ ).

$$\Pi(R, k) = \begin{cases} \Pi_0(R) & \text{for } k=0 \\ c\Pi_1(R) & \text{for } k=1 \end{cases} , \quad (16)$$

and “c” is the importance sampling parameter. Therefore, we find that

$$\frac{P_0}{P_1} = \frac{Z_0}{cZ_1} \quad (17)$$

which is how we can calculate  $Z_0/Z_1$ . The optimal value of the importance sampling parameter is usually  $c \approx Z_0/Z_1$ . Let  $T_{0 \rightarrow 1}$  and  $T_{1 \rightarrow 0}$  be the transition rates of hopping from  $0 \rightarrow 1$ , or vice-versa. We then have

$$\frac{P_0}{P_1} = \frac{\langle T_{1 \rightarrow 0} \rangle_1}{\langle T_{0 \rightarrow 1} \rangle_0} . \quad (18)$$

Use a heat bath for T:

$$T_{1 \rightarrow 0} = \frac{c\Pi_1}{\Pi_0 + c\Pi_1} = \frac{1}{1 + \frac{\Pi_0}{c\Pi_1}} . \quad (19)$$

Finally, we see that

$$\frac{Z_0}{Z_1} = c \left\langle \left( 1 + \frac{\Pi_0}{c\Pi_1} \right) \right\rangle_0^{-1} \left\langle \left( 1 + \frac{c\Pi_1}{\Pi_0} \right) \right\rangle_1 \quad (20)$$

whose variance will be large unless we can hop from  $0 \leftrightarrow 1$ . Notice that by introducing more states  $0 \leq \lambda_i \leq 1$  leads to a continuous transition and, therefore, to exact thermodynamic integration.

## 7. Umbrella Sampling/Histogram Method

Within this approach one applies importance sampling to

$$\frac{Z_1}{Z_0} = \frac{\int d\mathbf{R} \Pi_1(R)}{\int d\mathbf{R} \Pi_0(R)} . \quad (21)$$

Sample R from  $\Pi^*$  within Metropolis Algorithm, thereby giving the estimator

$$\frac{Z_1}{Z_0} = \frac{\langle \Pi_1/\Pi^* \rangle_{\Pi^*}}{\langle \Pi_0/\Pi^* \rangle_{\Pi^*}} . \quad (22)$$

Here is an example with  $\Pi^* = \sqrt{\Pi_0 \Pi_1}$ . In this case,

$$\frac{Z_1}{Z_0} = \frac{\langle \sqrt{\Pi_1/\Pi_0} \rangle_{\Pi^*}}{\langle \sqrt{\Pi_0/\Pi_1} \rangle_{\Pi^*}} . \quad (23)$$

Thus, the optimal  $\Pi^*$  is found by usual method:

$$\Pi^* \propto \sqrt{\Pi_0 \Pi_1} |f(R) - f(R)^{-1}| \quad (24)$$

$$= |\Pi_0/Z_0 - \Pi_1/Z_1| , \quad (25)$$

where  $f = \sqrt{(\Pi_1 Z_0)/(\Pi_0 Z_1)}$ . We can introduce an uncertainty in  $Z_1/Z_0$  as

$$\Pi^* \approx \sqrt{(\Pi_0 - c\Pi_1)^2 + \epsilon c\Pi_0\Pi_1} \quad , \quad (26)$$

where  $c$  is an estimator of  $Z_0/Z_1$  and  $\epsilon$  is chosen to minimize variance (arising from under-sampling and convergence).

In lattice models this method is known as the histogram method.

©*Duane Johnson and David Ceperley, 2000*