Enhanced Sampling and Rare Events

Enhanced Sampling

We said before that the time avg

$$\langle A \rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} A(X_i) \quad \text{for } N \text{ mc samples}$$

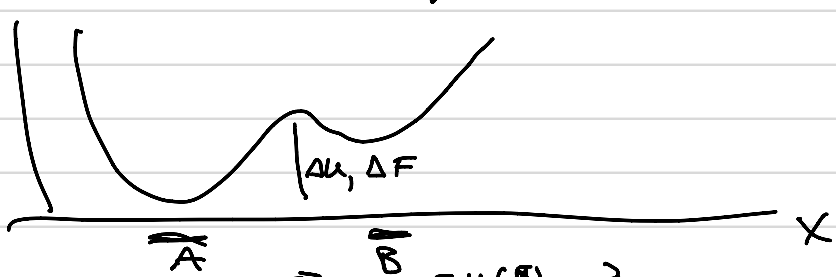
or $N = T/\Delta t$ md time steps is true if the system is ergodic, ie sees all the states

The problem in real simulation is $N \neq \infty$,

$N \sim (1 - 10^{10})$

This works for some problems, but there is a very common problem.

Suppose $U(x)$ or $F(x)$



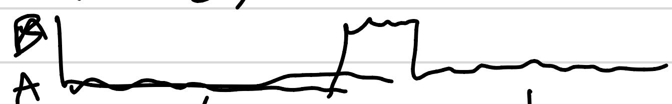
$$F(\vec{x}) = -kT \log \int \delta(M(\vec{q}) - \vec{X})\, e^{-\beta U(\vec{q})}\, d\vec{q} - F_0$$

(Potential of mean force)

$$\text{Rate}_{A \to B} \propto e^{-\beta \Delta u^{\neq}} \text{ or } e^{-\beta \Delta F^{\neq}}$$

So if $1/\text{rate} \gg N\Delta t$, then you will be trapped in A (or B)

(rare event problem)

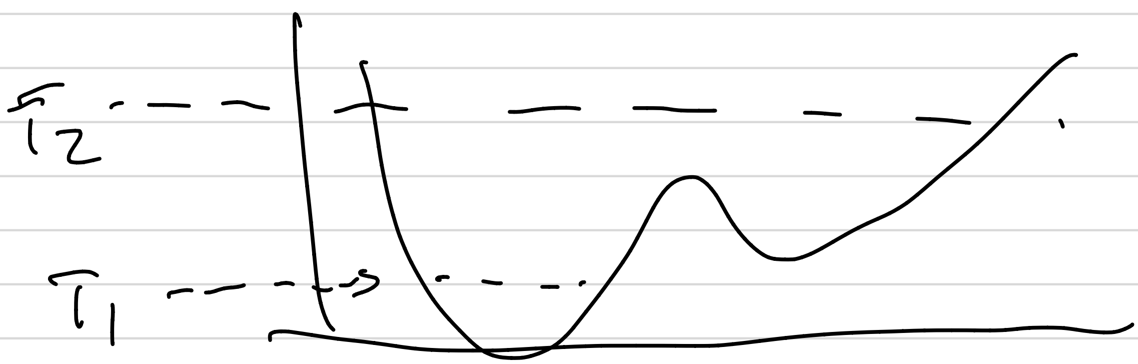


We need tricks to overcome this problem!

These are called enhanced or accelerated sampling methods, and generally estimate the (free)-energy diff between A & B too

$\left(\Delta A = -kT \log P_B/P_A \right.$ or like equilibrium const

Idea! increase temp → rate faster



$T \propto \langle KE \rangle$

$$\langle A \rangle_{T_1} = \int dx\, P_1(x)\, A(x)$$

$$P_1(x) = \omega_1(x) / Z_1 \quad , \quad \text{eg } \frac{e^{-U(x)/k_B T_1}}{\int dx\, e^{-U(x)/k_B T_1}}$$

$$\langle A \rangle_{T_1} = \int dx\, A(x) \frac{w_1(x)}{Z_1} \cdot \frac{\frac{w_2(x)}{Z_2}}{\frac{w_2(x)}{Z_2}}$$

$$= \int dx\, A(x) \frac{w_2(x)}{Z_2} \cdot \left(\frac{w_1(x)}{w_2(x)}\right) \cdot \frac{Z_2}{Z_1}$$

$$= \frac{Z_2}{Z_1} \cdot \int dx\, A(x) \frac{w_1(x)}{w_2(x)} \cdot \underbrace{\left[\frac{w_2(x)}{Z_2}\right]}_{\sim P_2(x)}$$

$$= Z_2/Z_1 \langle A\, w_1/w_2 \rangle_{T_2}$$

for N, V, T $$= Z_2/Z_1 \left\langle A\, e^{-\frac{1}{k_B T_1} U(x)} e^{+\frac{1}{k_B T_2} U(x)} \right\rangle_{T_2}$$

$$= Z_2/Z_1 \left\langle A e^{-\frac{U(x)}{k_B}\left(\frac{1}{T_1} - \frac{1}{T_2}\right)} \right\rangle_{T_2}$$

$$Z_1/Z_2 = \frac{\int dx\, w_1(x)}{Z_2} = \frac{\int dx\, w_1(x) \cdot \frac{w_2(x)}{w_2(x)}}{Z_2} = \left\langle \frac{w_1(x)}{w_2(x)} \right\rangle_{T_2}$$

$$= \langle \exp(-U(x)/k_B (1/T_1 - 1/T_2)) \rangle$$

if $T_2 \gg T_1$, weights very small

problem numerically

Solution, run many sims @ diff T so
$1/T_i - \frac{1}{T_{i+1}}$ not too big

Replica exchange MD/ Parallel tempering [didn't talk about const T yet?]

Every τ steps, try to swap configs

Then sample can be heated up & cooled down to T_1, overcoming barrier but still sampling @ T_1

Exchange prob? $P(A\to B)P(A) = P(B\to A)P(B)$

$A = \{\vec{x}@T_l, \vec{y}@T_h\}$

$B = \{\vec{x}@T_h, \vec{y}@T_l\}$

$$P(A\to B) = \min\left(1, P(B)/P(A)\right) = \min\left(1, \frac{e^{-U(x)/kT_h}\, e^{-U(y)/kT_l}}{e^{-U(x)/kT_l}\, e^{-U(y)/kT_h}}\right)$$

$$\Rightarrow P(A \to B) = \min\left(1, e^{-U(x)/k\left(\frac{1}{T_h} - \frac{1}{T_\ell}\right)} \cdot e^{-\frac{U(y)}{k}\left(\frac{1}{T_\ell} - \frac{1}{T_h}\right)}\right)$$

$$= \min\left(1, e^{-\left[\frac{U(x)-U(y)}{k} \cdot \left(\frac{1}{T_h} - \frac{1}{T_\ell}\right)\right]}\right)$$

if $T_\ell < T_h$, $\frac{1}{T_h} - \frac{1}{T_\ell} < 0$

and $U(x) - U(y)$ prob < 0

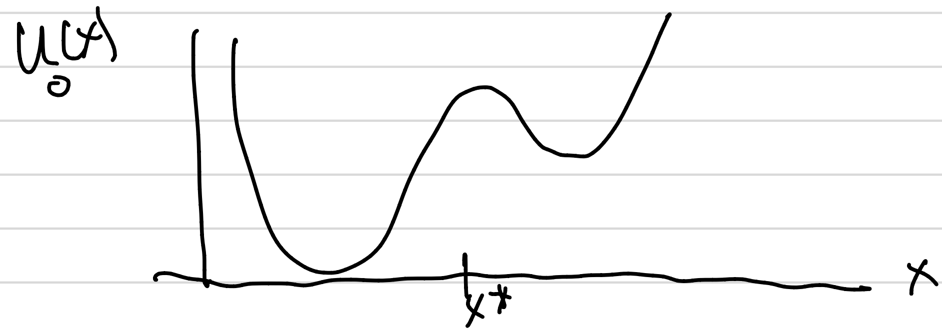
so swaps usually have prob < 1

Now since swaps satisfy detailed balance and MD or MC @ each temp satisfies detailed balance, have a chain of

$X_i @ T_1$ st $P(X_i) \to e^{-U(x_i)/k_B T_1}$ &

$$\langle A \rangle \sim \lim_{N \to \infty} \sum_{i=1}^{N} A(x_i^{T_1})$$

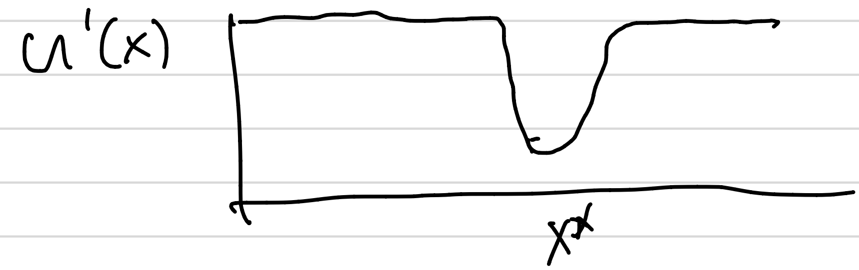
Idea #2 Go back to this picture



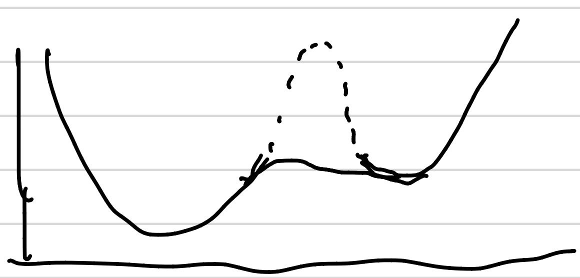
Idea: Torrie, Valleau 1977

what if we add a potential to this to reduce the barrier and correct for the effect

Example



Then $U_1(x) = U_0(x) + U'(x)$



This will give fast transitions from $L \to R$

But how do we get

$\langle A \rangle = \int A(x)\, P_0(x)$ when we are simulating with u_1 and

hence $P_1(x) = e^{-\beta u_1(x)}/Z_1 = \omega_1(x)/Z_1$

We actually did this before, perturbation theory

$$\langle A \rangle_0 = \frac{1}{Z_0} \int dx\, A(x)\, e^{-\beta u_0(x)}$$

like before, mult by $e^{-\beta u_1(x)}/Z_1$ in top and

$$= \frac{1}{Z_0} \int dx\, A(x)\, e^{-\beta u_0(x)} \cdot \frac{e^{-\beta u_1(x)}}{e^{-\beta u_1(x)}} \cdot \frac{Z_1}{Z_1}$$

$$\exp(-\beta u_0(x))/\exp(-\beta u_1(x)) = \exp(-\beta u_0(x) + \beta u_1(x))$$
$$= \exp(-\beta u_0(x) + \beta(u_0(x) + u'(x)))$$

$$= Z_1/Z_0 \int dx\, A(x) e^{+\beta u'(x)} \cdot \frac{e^{-\beta u_1(x)}}{Z_1}$$

$$= \langle A e^{+\beta u'(x)} \rangle_1 / \langle e^{+\beta u'(x)} \rangle_1 = \frac{\langle A/\omega' \rangle_1}{\langle 1/\omega' \rangle_1}$$

what is this weight doing? correcting for every time there is something near ** it should have less weight