

Statistical mechanics for molecular simulations (2015)

2/19/2015

Lecture 5

Free energy calculations

* free energy

* thermodynamic perturbation

* thermodynamic integration.

* histogram method; enhance sampling techniques.

Free energy plays an important role in

thermodynamics \Rightarrow all thermodynamic variables can

be calculated from free energy function.

here we consider

A exclusively,
but idea
 \Rightarrow general,
also applies
to G.

But calculating free energy via simulation is
not trivial!! Why? Let's check Helmholtz
free energy.

timeless

Recall

$$A_{total} = A_{ideal} + A_{excess}$$

with $\approx A$ in molecular simulation, we don't need to compute A_{ideal} .

$$A_{excess} = A_{total} - A_{ideal}$$

$$= -k_B T \cdot \ln Q_{excess}, \quad Q_{excess} = \int dr^N e^{-\beta U_{excess}}$$

Since the A_{ideal} part is trivial, hereafter

we define $A = A_{excess}$, i.e. only consider the excess free energy, configurational part.

At first glance it seems that the

configurational integral can be easily evaluated
← MC gives r^N only anyway.

via a MC or MD procedure. But this

is not true since:

averaged over canonical distribution!

$$\begin{aligned}
 A &= -k_B T \ln Q = k_B T \cdot \ln \left\{ \frac{1}{Q} \right\} = k_B T \ln \left[\int dr^N \frac{e^{-\beta U_{excess}}}{\int dr^N e^{-\beta U_{excess}}} \right] \\
 &= k_B T \cdot \ln \left[\int dr^N e^{-\beta U_{excess}} \cdot P(r^N) \right] = +k_B T \ln \left\langle e^{-\beta U_{excess}} \right\rangle_{\text{timeless}}
 \end{aligned}$$

To obtain A , we must evaluate the average

$$\langle e^{\beta U(\omega)} \rangle_{\text{eq.}} \quad \leftarrow \text{this form clearly}$$

This expectation value of $e^{\beta U(\omega)}$ is not easy to ~~average~~ calculate numerically because:

this is connected to the problem of sampling "rare" events

- ① high energy states ~~are~~ playing important part in the weighting, $e^{\beta U(\omega)}$, are not easy to sample in MD or MC !!
- ② the $e^{\beta U}$ series tends to converge slowly:

$$\langle e^{\beta U} \rangle = 1 + \beta \langle U \rangle + \frac{1}{2} \beta^2 \langle U^2 \rangle + \dots$$

unless $\beta U < 1$, otherwise the higher moments grow quickly and convergence is bad.

⇒ this is a hint that if $\beta U < 1$ is small then we can calculate this more accurately.

This is a vast field, there are many other methods not covered here!!

(4) (7)

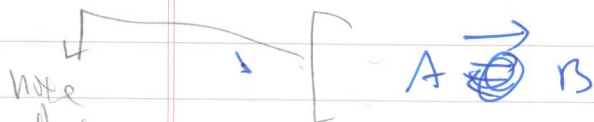
more seriously, one need to consider ~~these~~ for all practical use,

We need SA rigorous methods for calculations of free energy only anyway.

differences. There are two main methods and many variants!!

③ They all focus on aiming to calculate the free energy difference between

two states :



(now only) states (standard form can be difference).

$$\Delta A = A_{B,S} - A_{A,S} = -k_B T \cdot \ln \frac{Q_B}{Q_A}$$

$\leftarrow E \equiv U$

$$= -k_B T \ln \frac{\langle e^{+\beta E_B(\omega)} \rangle_B}{\langle e^{+\beta E_A(\omega)} \rangle_A}$$

notation only.

not necessary chemical,

A: solvated Cl^-
B: solvated Br^-

note that the difference is only in

the energy function that defines the two states,

diff. standard energy

mathematically, it is convenient to define:

timeless

λ = order parameter,

⑤ ⑥

$$E_x(\omega) = (1-\lambda) \cdot E_A(\omega) + \lambda \cdot E_B(\omega)$$

$$\therefore E_A(\omega) \rightarrow \lambda=0 \rightarrow E_0(\omega)$$

$$E_B(\omega) \rightarrow \lambda=1 \rightarrow E_1(\omega)$$

we can use λ to smoothly tune between A & B states,

okay, we are now ready to proceed for "correct ways" for free energy calculations.

① ② thermodynamic perturbation

now

$$\Delta A = -k_B T \ln \frac{Q(\lambda=1)}{Q(\lambda=0)}$$

$$= -k_B T \ln \left\{ \frac{\int d\omega \cdot e^{-\beta E_1(\omega)}}{Q(\lambda=0)} \right\}$$

determine of Q $e^{-\beta E_0(\omega)}$ $e^{-\beta E_1(\omega)}$

$$= -k_B T \ln \left\{ \int d\omega \cdot e^{\beta [E_1(\omega) - E_0(\omega)]} \cdot \frac{e^{-\beta E_0(\omega)}}{Q(\lambda=0)} \right\}$$

forward sampling

$$= -k_B T \ln \langle e^{\beta [E_1(\omega) - E_0(\omega)]} \rangle$$

OR average over $\lambda=0$, timeless

(MC or MD) $\lambda > 0$ (6) (9)

→ run a simulation use $E_0(r^N)$, then calculate the expectation value of $e^{-\beta[E_0(r^N) - E_1(r^N)]}$ to obtain ΔA .

E_1 using r^N generated in a $\lambda = 0$ simulation.

Note that the difference energy is smaller than the absolute ^{potential} energy, so the convergence is better.

Note that we can flip the $\frac{Q_1}{Q_0}$ to insert $e^{-\beta E_1(r^N)} e^{\beta E_0(r^N)}$ in the integrand for Q_0

to get a complementary expression.

$$\Delta A = k_B T \cdot \ln \left\langle e^{-\beta[E_0(r^N) - E_1(r^N)]} \right\rangle$$

calculate $E_0(r^N)$ using $\lambda = 1$ configuration.

this is called the

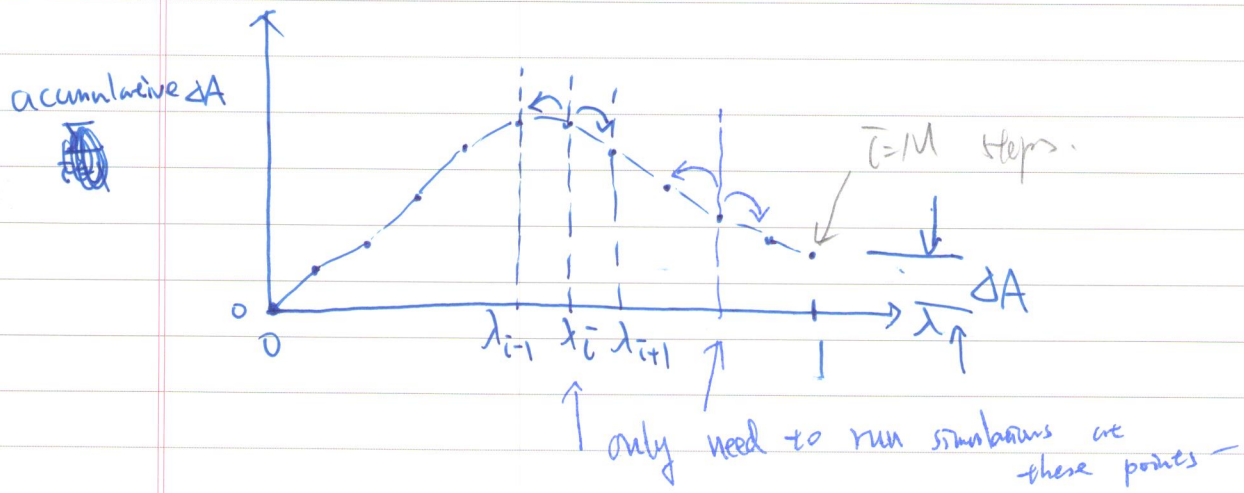
use $\lambda = 1$ simulation.

backward sampling because $\lambda = 1 \rightarrow \lambda = 0$.

Note that accuracy can be improved by using multiple steps to reduce energy difference.

timeless

It is trivial to generalize to many steps:



here, a "double wide sampling" technique is often employed to generate two free energies from a single simulation.

$$\Delta A = \sum_i \Delta A_i$$

① Simulate on odd i λ_i step, potential function

$$\Delta A(\lambda_{i-1} \rightarrow \lambda_i) = k_B T \ln \left\langle e^{-\beta [E_{\lambda_{i-1}}(W) - E_{\lambda_i}(W)]} \right\rangle_{\lambda_{i-1}}$$

↖ using backward sampling

② $\Delta A(\lambda_i \rightarrow \lambda_{i+1})$ from forward sampling.
 ③ final accumulated results is ΔA timeless

2
 4 thermodynamic integration,

this other general approach is based on the following identity of an integral in elementary calculus:

$$\Delta A_{0 \rightarrow 1} = \int_0^1 \left(\frac{\partial A(\lambda)}{\partial \lambda} \right) \cdot d\lambda$$

recognize that $A(\lambda) = -k_B T \cdot \ln \left\{ \int dr^N \cdot e^{-\beta E_{\lambda}(r^N)} \right\}$

for our simple linear scaling

$$\begin{aligned} \therefore \left(\frac{\partial A(\lambda)}{\partial \lambda} \right) &= -k_B T \cdot \frac{1}{\int dr^N} \times \int dr^N \cdot \frac{1}{k_B} \left(\frac{\partial E_{\lambda}}{\partial \lambda} \right) \cdot e^{-\beta E_{\lambda}(r^N)} \\ &= \left\langle \left(\frac{\partial E_{\lambda}}{\partial \lambda} \right) \right\rangle_{\lambda} \quad \text{we potential at } E_{\lambda}(r^N) \\ &= \left\langle \left[E_1(r^N) - E_0(r^N) \right] \right\rangle_{\lambda} \end{aligned}$$

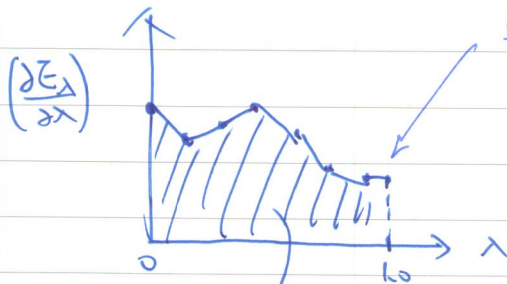
more complex dependence on λ is possible

the calculation of $\left(\frac{\partial A(\lambda)}{\partial \lambda} \right) \Big|_{\lambda}$ is the same as calculating average energy, double accuracy.

⇒ we change an average over $e^{-\beta E}$ to average $\langle E \rangle$ this of course converges more rapidly !!

timeless

naturally, the thermodynamical integration method is carry out using "summation" to approximate the integral :



still have to perform at several λ points, because we still need to have enough sample for this result to be accurate.

(i.e. overlap of distributions at different λ).

the area gives ΔA .

Final remarks on free energy calculations :

- ① Statistical mechanics are important !!
- ② accurate methods all require multiple simulations to get one ΔA \rightarrow very expensive !!
- ③ plan ahead & choose proper pathway, \Rightarrow

importance .

- ④ always evaluate the extent of "sampling"
 - \hookrightarrow see JPCB 2010, 114, 10235 & JCTC 2014, 10, 2632.

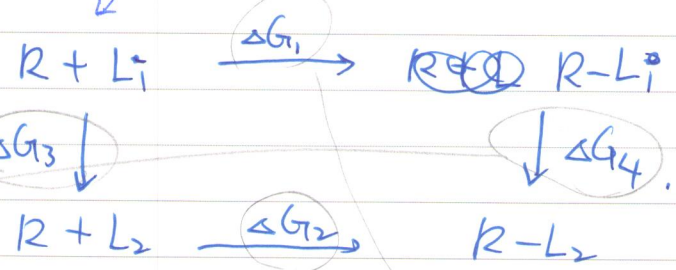
timeless

for example,
⊕ Considering thermodynamic cycles ~~that~~ is

very impressive, Assume we want to calculate free energy difference for

different binding instances of two Ligands

to a ~~acceptor~~ receptor R,
binding is "rare event"!!



unphysical but easier to compute

extremely difficult to calculate

We want to compare ~~ΔG~~ $\Delta \Delta G = \Delta G_2 - \Delta G_1$,

which is relevant in many many ~~scenarios~~ scenarios.

However, ΔG_1 & ΔG_2 is very difficult to calculate because "binding" is a rare event!!

~~However~~, Now, notice that ΔG_3 & ΔG_4 are easy to compute even though they are unphysical.

by recognizing

$$\Delta G_1 + \Delta G_4 = \Delta G_2 + \Delta G_3$$

$$\Rightarrow \Delta \Delta G = \Delta G_3 - \Delta G_1 = \Delta G_4 - \Delta G_2$$

So one should calculate ΔG_4 & ΔG_2

instead \Rightarrow morph a molecule into another
by changing the force fields

Thus we covered the fundamentals of
free energy calculations ~ .

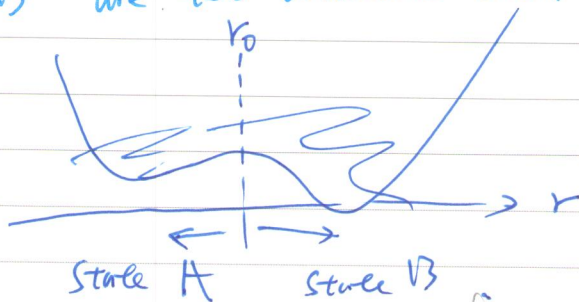
ref: Beveridge & DiCapua, Annu. Rev.
Biophys. Biophys. Chem. 1989, 18:431-92.

③ probability ratio method: (if Φ potential function the same).

Sometimes calculating difference free energy from probability ratios (population ratios) is

more accurate because of error canceling:

$A \rightleftharpoons B$ are two different local minima:



T. e. quasi-equilibrium problem

$$N_A = \sum_{i=1}^N [1 - \theta(r_i - r_0)] \quad , \quad N_B = \sum_{i=1}^N \theta(r_i - r_0)$$

now $A \rightleftharpoons B$

$$\therefore K = \frac{[A]}{[B]} = e^{-\Delta A / k_B T}$$

$$\therefore \Delta A = k_B T \cdot \ln \frac{N_A}{N_B}$$

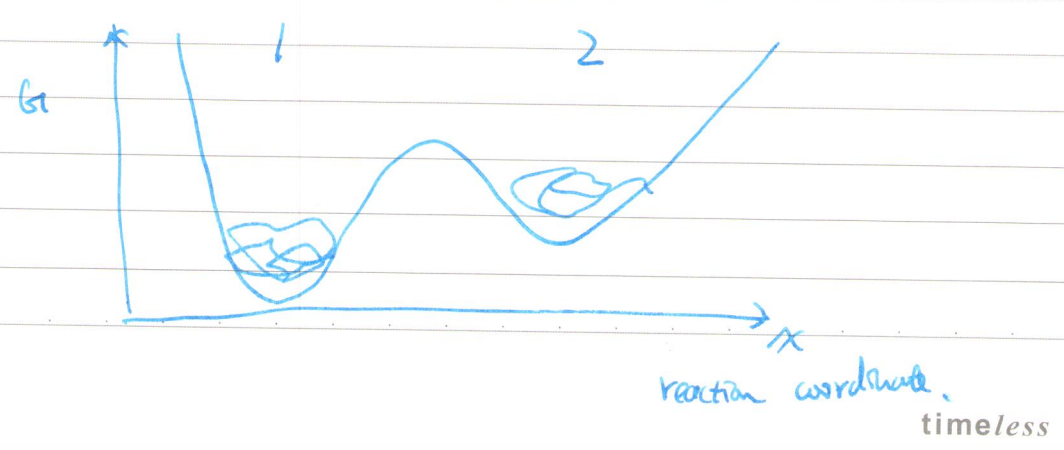
\Rightarrow if divides into multiple steps, a free energy curve can be drawn from the histogram of a long-long simulation timeless

requires many points, also suffer insufficient sampling problem

The problem is that sampling is never enough in reality. Especially:

- * barrier crossing → all reactions & large conformational changes.
 - * first encounter / binding → entropy controlled events.
 - * protein folding / global minima finding → ~~all~~ rare events.
- ⋮

This is a general problem of "enhance sampling". The general problem can be formulated into sampling in a double well problem:



We need to overcome the barrier !!

These methods are often used to locate/identify new states.

They can be used to "explore" potential energy surfaces

There are many methods:

- * replica exchange
- * umbrella sampling
- * metadynamics
- * steered MD
- * dynamical flooding method
- * non-equilibrium methods

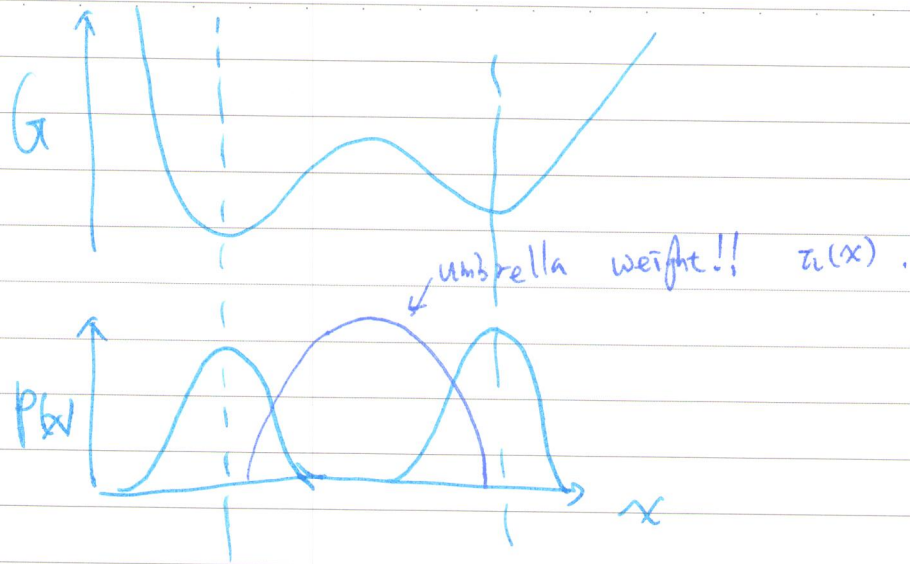
Here we will only briefly discuss umbrella sampling approach & the replica exchange method.

* umbrella sampling.

The umbrella sampling approach recognized that the insufficient sampling is due to the distribution function obtained within the available ^{simulation} time span at each well do not overlap with each other!

⇒ no ~~overlap~~ overlap, so ^{a traj.} from one well does not reach the other well !!

timeless



So the idea is to apply an additional weighting function to "bridge" the two distributions, thus the name "umbrella" ~.

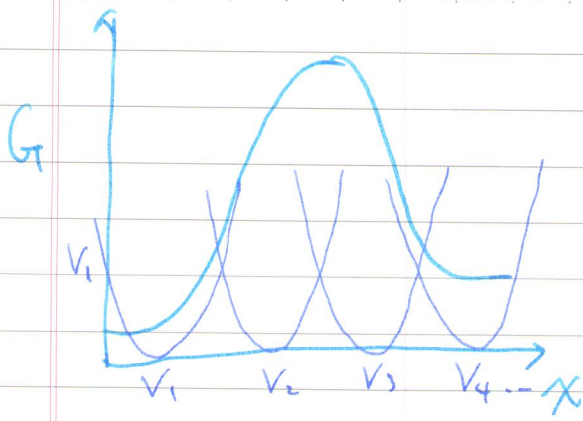
Now recognize that applying a new weight is ⁱⁿ equivalent to adding a biasing potential:

$$V(x) = -k_B T \cdot \ln \pi(x)$$

In principle, $V(x)$ can be anything,

but in reality, a barrier crossing event can be sampled with multiple US steps:

timeless

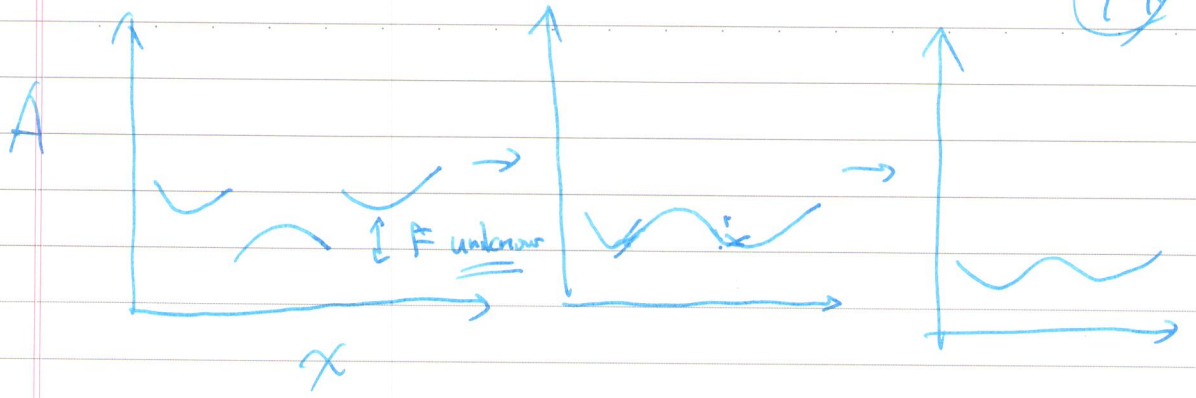


← This example use
4 simulations —
i.e., 4 windows.
harmonic potential
is often used!!

The results from different simulations are then piece together by using the weighted histogram analysis method (Kuma et al, JCC, 13, 1011 (1992)).
↑ (WHAM)

So a general procedure is:

- ① Sample with umbrella potential $U_i' = U(r^N) + V_i'(r^N)$
(need to know the reaction coordinate first).
 $r^N \rightarrow x$.
- ② Compute biased probability $P_i'(x)$, note that $P_i'(x)$ are not accurate on the edges.
- ③ estimate unbiased free energy. \leftarrow a undetermined constant —
$$A(x) = -k_B T \ln P_i'(x) - V_i'(x) + F_c$$
- ④ Use WHAM to combine different windows.



There are generalizations to multidimensional reaction coordinates and multiple temperatures, but that is out of the scope here.

* replica exchange

The replica exchange, ~~with~~ which belongs to a large class of parallel tempering method, is another popular approach that can be used to overcome free energy barrier. Note that

but then you don't need to know R.C. to begin with.

this method use temperature to enhance sampling, and it is not directed, i.e. less efficient in ~~finding~~ exploring reaction coordinates.

Here, one launches simulations at different

replica at different temperatures!!

temperatures, and adding Monte Carlo steps to swap configurations based on a ^{exactly the} ~~similar to~~ Metropolis criterion.

Boltzmann Criterion:

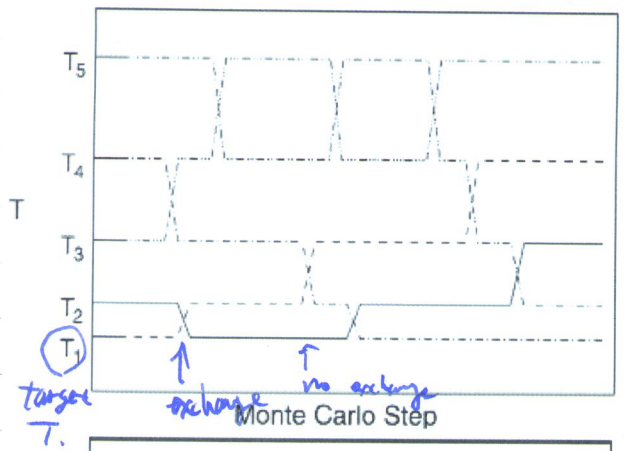
$$p = \min \left(1, e^{(\epsilon_i - \epsilon_j) \left(\frac{1}{kT_i} - \frac{1}{kT_j} \right)} \right)$$

↪ prob. to swap $(\nu)_i$ at T_i and $(\nu)_j$ at T_j .

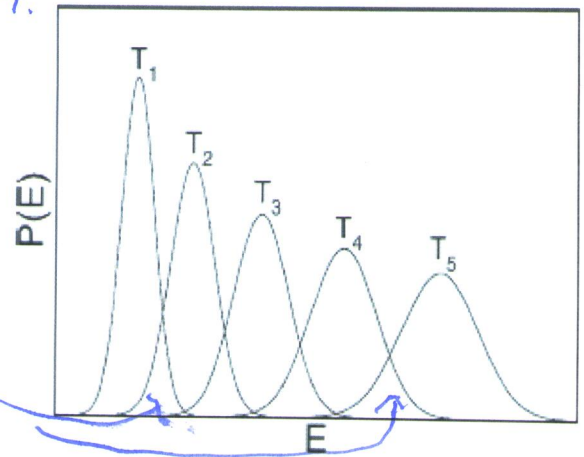
timeless

The process is explained by the following figures:

Example using 5 replicas



overlap is essential for efficiency



Note that:

- * overlap of distributions is required for this to work.
- * the lowest temperature (target) "trajectory" samples the Boltzmann dist. at that T .
- * the MC steps are unphysical!!