Statistical mechanics for molecular simulations (2015)

2/4/2015

Lecture 3

* Monte-Carlo Method.
* partition function & properties by integration
* Metropolis method.
* Monte Carlo simulation of molecules

So, we have covered basic statistical mechanics and molecular interactions, which provide the basic model to describe molecular systems. In the last lecture we have argued that in the perspective

PES is the key concept in Physical Chemistry. (2)

of classical mechanics, a system travels on the PES, which can exhibit many minima & transient structures. Molecular properties of course depends on the structures, and thus it is essential that an average over assessible structures must be performed in order to calculate macroscopic properties.

So, in the end it is a problem of samplings!!

In this lecture we will detail what exactly is to be sampled.

It should be clear to you already that two main sampling techniques exist in molecular simulations.

① Monte Carlo simulation : based on statistical principle, useful for equilibrium properties.
② Molecular dynamics simulation : based on Newtonian dynamics, not only useful for equilibrium properties.

We will first discuss the MC method.

Note that generally speaking, the MD approach should be the method of choice because it is based on "real" physics. However, in many cases MC could be more effective, i.e. no physical EOM (Ising model), unphysical moves useful for accerating sampling (glass ~~dyn~~ systems . .) . . . . .

* Properties by integration.

So, what exactly is this sampling problem?

Let's consider the calculation of the canonical partition function:

$$Q_{NVT} = \frac{1}{N!} \cdot \frac{1}{h^{3N}} \iint dp^N dr^N \cdot e^{-\frac{H(p^N, r^N)}{k_B T}}$$

↑ N particle system.

Now recognize that the Hamiltonian can be divide into the kinetic energy part and the potential:

$$H(p^N, r^N) = \underbrace{\sum_{i=1}^{N} \frac{p_i^2}{2m_i}}_{K: \text{dependent only on } p} + \underbrace{U(r^N)}_{\text{potential determined by } r^N}$$

Thus the $\{p^N\}$ part & the $\{r^N\}$ part can be factorized for normal molecular systems:

related to ideal gas!!

$$Q_{NVT} = \frac{1}{N!} \frac{1}{h^{3N}} \int dp^N \cdot e^{-\sum_i \frac{p_i^2}{2m_i k_B T}} \cdot \int dr^N \cdot e^{-\frac{U(r^N)}{k_B T}}$$

excess part due to molecular interactions!!

$$= Q_{NVT}^{ideal} \cdot Q_{NVT}^{excess}$$

a fixed $\{r^N\}$ → a configuration.

The excess partition function, the configurational contribution is then defined as

$$Q_{NVT}^{excess} = \frac{1}{V^N} \int dr^N \cdot e^{-\frac{U(r^N)}{k_B T}}$$

This is a configurational integral that sums up the

Boltzman weight of all configurations !!

Clearly $A^{excess} = -k_B T \cdot \ln Q_{NVT}^{excess}$

$$A = A^{ideal} + A^{excess}$$

Since the ideal gas part is trivial, we should focus on the configurational part.

If we define

$$Z = \int dr^N \, e^{-\beta U(r^N)}$$

then the probability of finding the configuration $\{r^N\}$ is $P(r^N) = \frac{e^{-\beta U(r^N)}}{Z}$

Any properties can then be calculated by averaging over configurations : ← Boltzman weight

$$\langle A \rangle = \frac{1}{Z} \int dr^N \cdot e^{-\beta U(r^N)} \cdot A(r^N) \quad \ldots\ldots \text{Eq. (1)}$$

↑ should integrate over all "configurational space volume" elements.

↑ A depends only on coordinates —

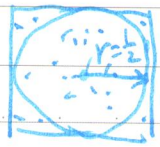
So it is clear now that we want to do "simulation" to evaluate this configurational integral. How to do numerical integration?

① use a numerical integration scheme such as a quadrature rule (trapezoid rule, Simpson rule, ... etc). This, of course, is not useful because of the curse of dimensionality. Think for a 100 atom system, each dimension with 5 points

$\Rightarrow 5^{300}$ points to evaluate !!

② use a random Monte Carlo scheme. ← Why this name?

Think finding the area of a unit circle? i.e. evaluation of π? randomly shot points !!

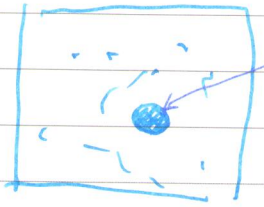


$A = \pi r^2$

$\pi \simeq 4 \times \frac{\#\text{ of points in the circle}}{\text{total } \#\text{ of points}}$

This "simple Monte Carlo" way of doing "numerical integration" works well when the region of interest is clear.

But how about this!



this area. Most points are wasted!!

This is not an effective scheme for evaluating Eq. (4) because most randomly generated configurations will have a high energy and very small Boltzman weight, meaning most points are wasted.

(3) Importance sampling.

Metropolis was the pioneer who recognized this problem and proposed a scheme for "Importance" sampling.

* Metropolis Monte Carlo

① pick a original configuration "0", calculate $U(0)$.

② make a random move to reach a new configuration "n", calculate $U(n)$.

③ If $U(n) < U(0)$, accept n, repeat from ①

④ calculate $f = e^{-\beta[U(n)-U(0)]}$

⑤ generate random number $r \in [0 \ldots 1]$,

⑥ if $r < f$ then n is accepted, otherwise n is rejected, 0 is accepted.

⑦ repeat from ①

each step generates a configuration.

repeat M times.

No momentum in MC!!

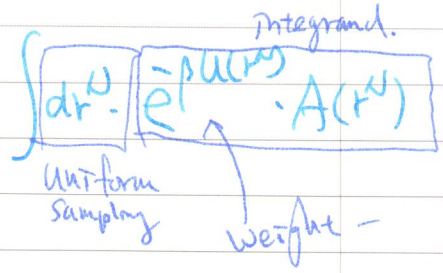
In the end we obtain a set of M configurations $\{C_j\}$. The integral is then evaluated:

$$\langle A \rangle = \frac{1}{M} \sum_{j=1}^{M} A(C_j)$$

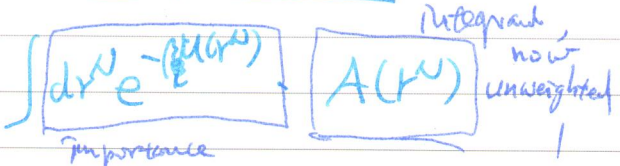
Note that this scheme generates more low energy configurations while allowing the sampling to reach higher energy states.

The basic idea:

Simple MC



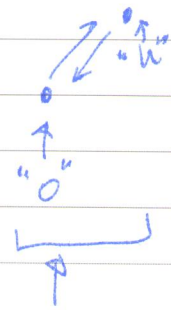
Metropolis MC



But how is the Metropolis scheme yields the Boltzmann weight? This can be proven mathematically by invoking a Markov chain model. Here I will skip the full mathematics and instead use the derivation in Frenkel & Smit, or many, many others.

Let's consider an ensemble of equilibrium MC trajectories. The Metropolis scheme deals with transition between two configurations

$$o \rightleftarrows n$$



In an ensemble, at a given "step", they all appear -

In the ensemble, there are $N(o)$ trajectories begins with "o" and $N(n)$ trajectories begins with "n".

The # of "o" & "n" should satisfy the Boltzman distribution: $\frac{N(o)}{N(n)} = \frac{e^{-\beta U(o)}}{e^{-\beta U(n)}} = e^{-\beta[U(o)-U(n)]}$.

After each step, the Boltzman distribution, i.e. the equilibrium ensemble, should be invariant.

This is guaranteed by the detailed balance condition

$$N(o) \cdot \pi(o \to n) = N(n) \cdot \pi(n \to o)$$

Where $\pi(o \to n)$ is the transition probability of $o \to n$.

Metropolis scheme uses two stages to control π.

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

(α: prob. of making the $o \to n$ move; P_{acc}: prob. of accepting the $o \to n$ move.)

Here α is the underlying matrix of the Markov chain. The Metropolis method adopted a symmetry Markov chain.

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

This "unbias" choice is usually effective.

Notice that biased α can be use to sample/explore rare events & subensembles (see Ch. 13).

With this α, we can determine the condition for the probability of accepting a trial move that leads to the detailed balance:

$$N(o) \cdot P_{acc}(o \to n) = N(n) \cdot P_{acc}(n \to o).$$

Therefore

$$\frac{P_{acc}(o \to n)}{P_{acc}(n \to o)} = \frac{N(n)}{N(o)} = e^{-\beta[U(n) - U(o)]}.$$

There are many ways to define P_{acc}, an obvious one is adopted by Metropolis!

$$P_{acc}(o \to n) = \begin{cases} N(n)/N(o) & \text{if } \underbrace{N(n) < N(o)}_{\text{i.e. } U(n) > U(o)} \\ 1 & \text{if } \underbrace{N(n) \geq N(o)}_{\text{i.e. } U(n) \leq U(o)} \end{cases}$$

In summary, the Metropolis scheme adopts the following transition probability:

$$\pi(o \to n) = \alpha(o \to n) \quad \text{if } U(n) \leq U(o) \text{ (down hill)}$$
$$= \alpha(o \to n) \times \frac{N(n)}{N(o)} \quad \text{if } U(n) > U(o) \text{ (up hill)}$$

$$\pi(o \to o) = 1 - \sum_{n \neq o} \pi(o \to n) \quad \text{(the "o" must be retained if } o \to n \text{ is rejected)}$$

This scheme satisfies the detailed balance, hence will lead to the Boltzman distribution.

⇒ as long as M is a large number !!

⇒ a bonus is the "trajectories"; They are not real realtime dynamics, but can be interpreted in a stochastic/random walk manner.

* Technical issues in MC.

① random number generator

⇒ must be good!!

③ effective moves.

- sequential / non-random moves: can have non-ergodic problem, always using a "random move" is a good practice.

- step size & rejection rate

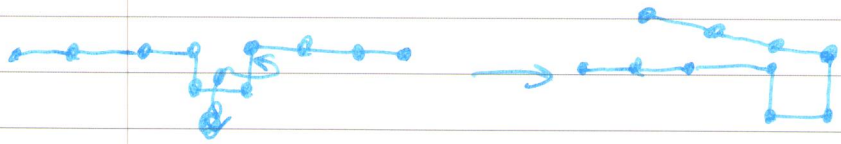
small size: high Pacc, but explores the pts slowly, computationally expensive.

large size: low Pacc, hard to move.

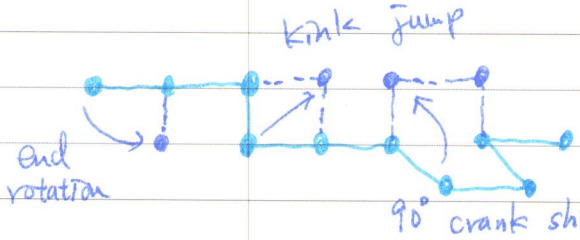
empirically, a accepting rate of 20% - 50% could be a good choice, but this is model dependent.

- large molecules :

a torsional motion can lead to very large change, hence is not good.



local moves are better :

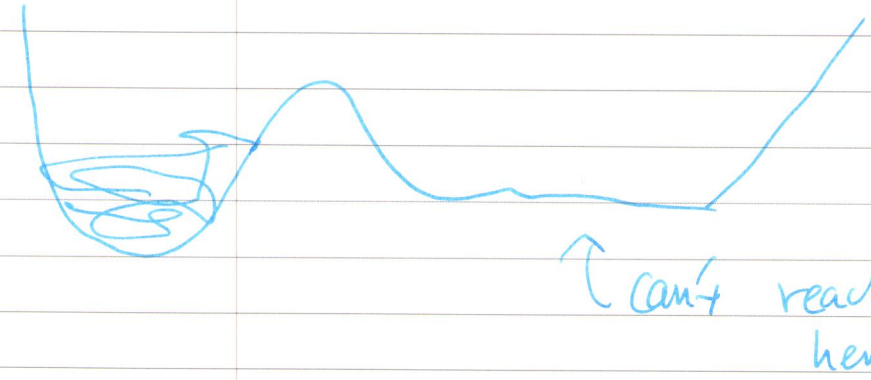


~~Excluded~~

excluded volume problem can be a trouble!!

③ barrier crossing / quasi-ergodicity.

a problem about the Metropolis scheme is that crossing a high energy barrier is hard → quasi-ergodicity.



There are many ways to overcome this,

and two simplest ways are

- J-walking (jumping; high-temperature J-walker)

- multicanonical Monte Carlo.
  ↑ use a multicanonical weight factor that is a non-Boltzmann

We won't have time to cover the details here. A good intro is in Ch. 8.8 of Leach's book.

④ Initial guess & equilibrium.

formally, given a initial population distribution in the configurational space, the MC iteration is represented by a Markov chain:

$$\rho_n = T^n \cdot \rho(0)$$

In principle $\rho(0)$ should be the equilibrium distribution, i.e. sample many trajectories from Boltzman distribution. But this itself can be problematic.

Fortunately, the Markov chain will reach equilibrium by itself independent of $\rho(0)$.

given the system is ergodic!!

So, one can choose $\rho(0)$, run MC, get rid of the initial transient period, and use time-average instead of ensemble average.