Computational Methods for Quantum Many-Body Physics

Lecture 4

June 7, 2023

Principle of Monte Carlo algorithms

Importance sampling → Some dynamical process [Markov chain with extra conditions; lots of freedom to "choose" here]

→ Wait for "long enough"

→ Make "enough # of representative measurements"

→ Use these to get average values and their associated error bars

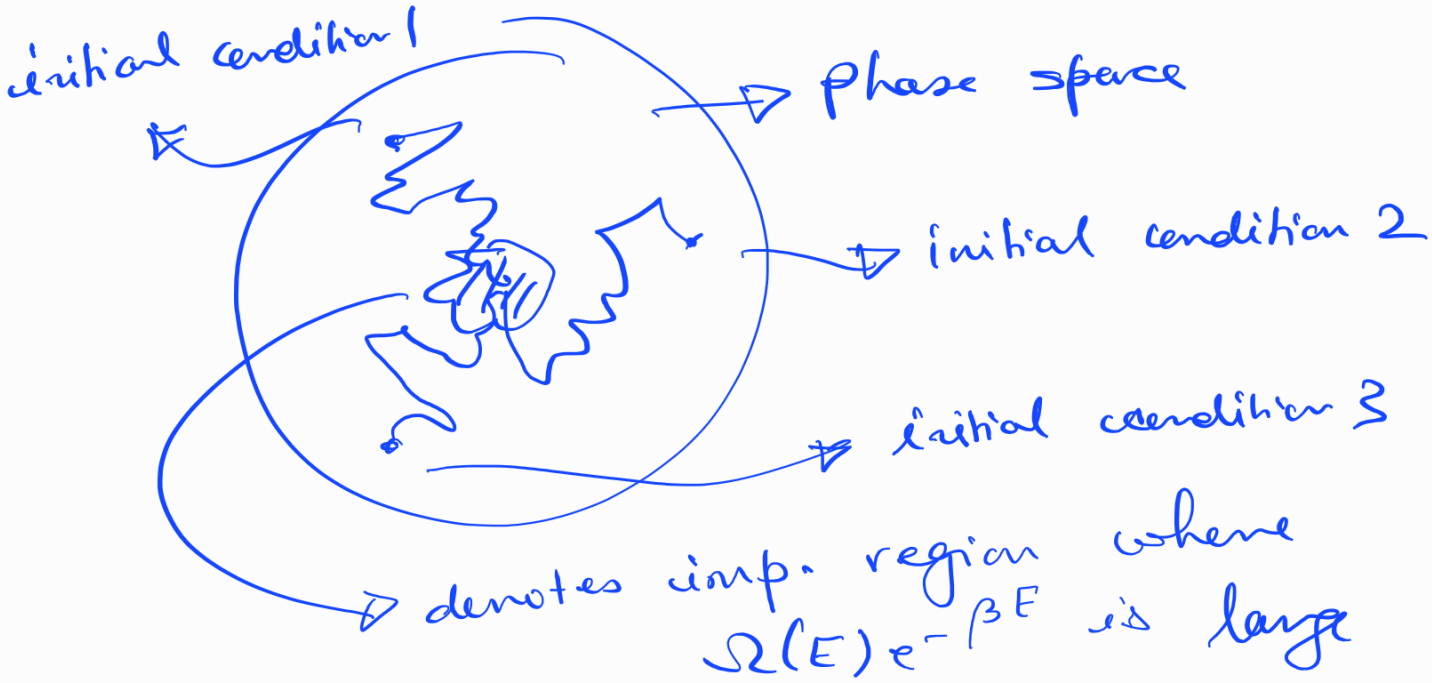
↳ Stochastic way to calculate

$$\langle Q \rangle = \frac{\sum_{\{\mu\}} Q_\mu e^{-\beta E_\mu}}{\sum_{\{\mu\}} e^{-\beta E_\mu}} \quad \text{(1)}$$

where μ denotes a microstate of the system

The RHS of (1) has an exponentially large # of terms in system size. However, not all terms equally important. → strongly peaked in energy

$$Z = \sum_{\{\mu\}} e^{-\beta E_\mu} = \int dE \, \boxed{e^{-\beta E} \Omega(E)} \quad \text{(2)}$$



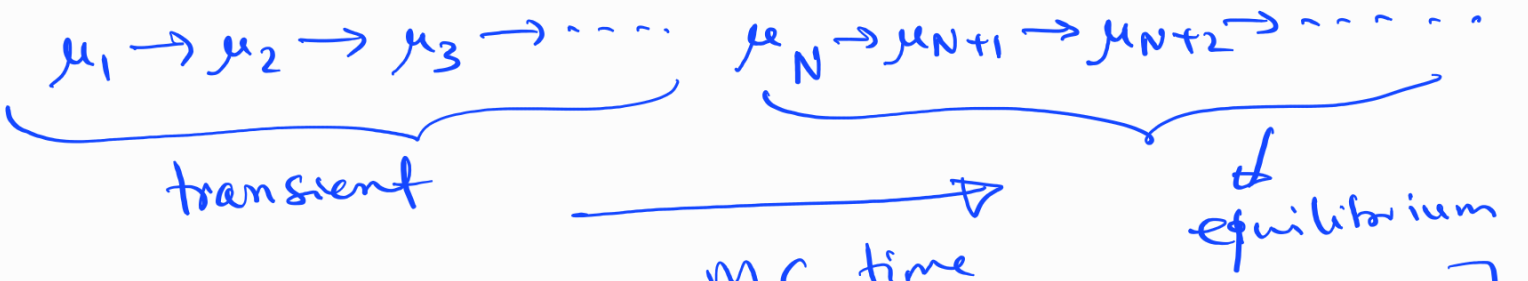
Monte Carlo algorithms use dynamical Markov processes.

- Define appropriate transition probabilities $P(\mu \to \nu)$ where μ and ν denote any two microstates of the system. Clearly, $P(\mu \to \nu)$ are non-negative numbers.

In a Markov process,
$\rightsquigarrow P(\mu \to \nu)$ do not vary with time
$\rightsquigarrow$ They only depend on μ and ν and not on any state that the system has encountered before.

Thus, $\sum_{\nu} P(\mu \to \nu) = 1$; note that $P(\mu \to \mu)$ may not be zero

Given this rules, a realization of this stochastic process (with two extra conditions that we state below) looks like



Two extra conditions ensure that the equilibrium distribution is the unique limiting distribution of this process.

- **Ergodicity**: It should be possible for the Markov process to reach any microstate ν starting from any microstate μ.

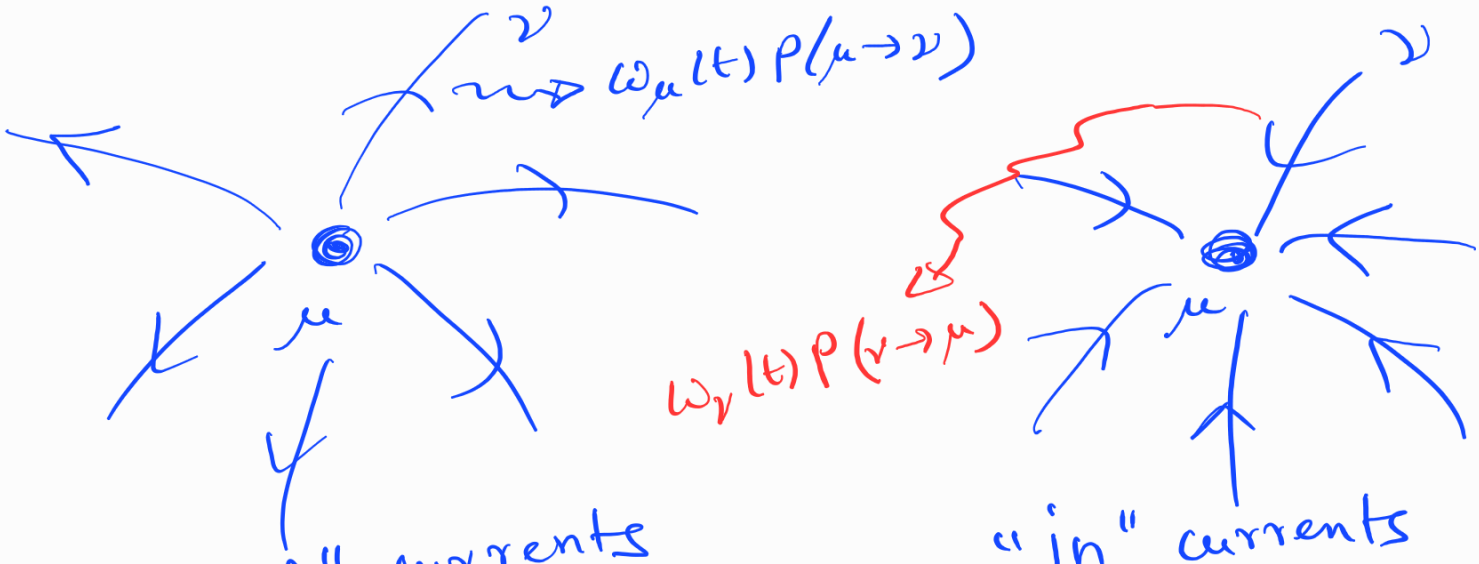
- **Detailed Balance**:

$$p_\mu^T P(\mu \to \nu) = p_\nu^T P(\nu \to \mu) \quad (3)$$

where p_μ^T specifies the target distribution which equals

$$p_\mu^T = \frac{e^{-\beta E_\mu}}{Z} \text{ in this case.}$$

Given this Markov process, let us now ask the following: Let $\omega_\mu(t)$ be the probability that the system is in microstate μ at time t. Then what is $\omega_\mu(t+1)$?



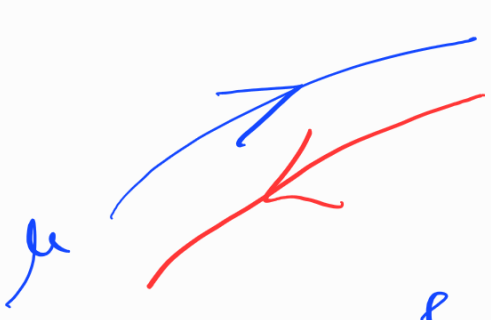
$$\Rightarrow \omega_\mu(t+1) - \omega_\mu(t) = \sum_\nu \Big[\underbrace{\omega_\nu(t)\,P(\nu\to\mu)}_{\text{in current}} - \underbrace{\omega_\mu(t)\,P(\mu\to\nu)}_{\text{out current}} \Big] \quad \text{— (4)}$$

Global balance condition for the limiting distribution $\omega_\mu(\infty)$. implies RHS of (4) equals zero for all μ given $\omega_\mu(\infty)$ and $P(\mu\to\nu)$

Since $\omega_\mu(\infty) = p_\mu^{eq}$ in our case, it gives (from RHS of Eq. (4))

$$\sum_\nu p_\nu^{eq}\, P(\nu\to\mu) = p_\mu^{eq} \quad \text{— (5)}$$

There are many ways to achieve this. Detailed balance is one way which involves a pairwise cancellation of terms in the RHS of Eq (4)



$$p_\nu^{eq} P(\nu \to \mu) = p_\mu^{eq} P(\mu \to \nu) \quad \text{— (6)}$$

where $\mu \neq \nu$

Eq. (4) can be simplified to

$$\omega_\mu(t+1) = \sum_\nu \omega_\nu(t) P(\nu \to \mu) \quad \text{— (7)}$$

Suppose, we interpret $P(\nu \to \mu)$ as a matrix where ν denotes the row and μ the column index j and denote $\omega_\nu(t)$ as a row vector. Then Eq (7) can be expressed as

$$\omega(t+1) = \omega(t) \cdot P \quad \text{— (8)}$$

in matrix notation.

Note that P is not a symmetric matrix in general $\therefore$ of Eq (6).

Clearly, $\omega(\infty) = \omega(\infty) \cdot P$

The matrix P (which is $\Omega \times \Omega$ dim. where Ω = # of microstates) has a unique non-degenerate eigenvalue $\lambda_0 = 1$ and all other eigenvalues $|\lambda_i| < 1$.

(*) This unique $\lambda_0 = 1$ corresponds to the right eigenvector $(1\ 1\ 1 \cdots 1)^T$ and the left eigenvector $(p_1^{eq}\ p_2^{eq}\ p_3^{eq} \cdots p_N^{eq})$

(@) Convergence to the fixed point $\omega(\infty)$

$$\omega(t+1) = \omega(t) \cdot P = \omega(t-1) \cdot P \cdot P = \cdots$$

$$\Rightarrow \omega(t) = \omega(0) \cdot P^t$$

Express $\omega(0)$ as the linear combination of the left eigenvectors v_i of P

$$\omega(0) = \sum_i a_i v_i$$

$$\Rightarrow \omega(t) = \left(\sum_i a_i v_i\right) \cdot P^t = \sum_i a_i \lambda_i^t v_i \quad \text{---} (9)$$

As $t \to \infty$, in RHS of (9), only $\lambda_0 = 1$ contribution survives.

(@) The $|\lambda_i| < 1$ can be thought of as transients that eventually die out when the system approaches thermal equilibrium

(@) The slowest transient timescale is set by $\tau_1 = -\frac{1}{\log|\lambda_1|}$ --- (10)

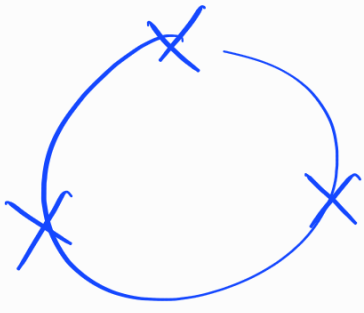
where the ordering implied is

$$|\lambda_0| > |\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \cdots \cdots$$

Let us give a simple example of this P matrix using the Metropolis algorithm for the Ising model.

We take a simple 3 Ising spins system on a ring (PBC) with the Hamiltonian

$$H = -J \sum_{\langle ij \rangle} s_i s_j$$



0 0 0	→ 0	→ $E = -3J$
0 0 1	→ 1	→ $E = J$
0 1 0	→ 2	
0 1 1	→ 3	
1 0 0	→ 4	
1 0 1	→ 5	
1 1 0	→ 6	
1 1 1	→ 7	→ $E = -3J$

Metropolis algorithm states that a spin on the lattice is selected at random and then a flip-attempt is made for that spin. If the flipped microstate has an equal or lower energy, then the flip move is accepted. If the energy of the microstate is higher than the previous microstate, then the flip is attempted only

with the probability $e^{-\beta \Delta E}$ where ΔE is the energy difference between the two microstates.

For what follows, it is useful to write $P(\mu \to \nu)$ as

$$P(\mu \to \nu) = g(\mu \to \nu)\, A(\mu \to \nu) \quad \text{— (11)}$$

where $g(\mu \to \nu)$ = probability of considering the move from microstate μ to ν

and $A(\mu \to \nu)$ = probability of accepting this move

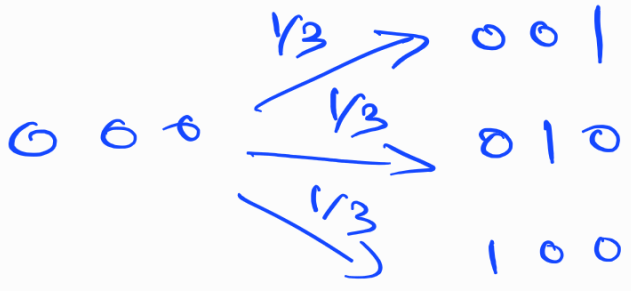
If $\mu \to \nu$ is rejected after having being considered, then the move $\mu \to \mu$ is chosen instead and the system remains in its original microstate.

For the Metropolis algorithm,

$$g(\mu \to \nu) = \frac{1}{N} \quad \text{(where } N = \#\text{ of sites)} \quad \text{— (12)}$$

for two microstates μ and ν connected by a single spin flip since the site where the spin flip attempt is made is chosen completely randomly. All other $g(\mu \to \nu) = 0$

E.g.



Also, $A(\mu \to \nu) = \begin{cases} e^{-\beta(E_\nu - E_\mu)} & \text{if } E_\nu - E_\mu > 0 \\ 1 & \text{otherwise} \end{cases}$

(13)

- Ergodicity is straightforward to see since any microstate can be connected to any other microstate through a series of spin flips.
- For detailed balance, let us consider any two distinct microstates μ and ν. Further, consider that $E_\mu \geq E_\nu$

Then, we need to show that

$$e^{-\beta E_\mu} g(\mu \to \nu) = e^{-\beta E_\nu} g(\nu \to \mu)$$

$$\Rightarrow e^{-\beta E_\mu} g(\mu \to \nu) A(\mu \to \nu) = e^{-\beta E_\nu} g(\nu \to \mu) A(\nu \to \mu)$$

$$\Rightarrow e^{-\beta E_\mu} \frac{1}{N} A(\mu \to \nu) = e^{-\beta E_\nu} \frac{1}{N} A(\nu \to \mu)$$

Since $E_\mu \geq E_\nu$, $A(\mu \to \nu) = 1$

and $A(\nu \to \mu) = e^{-\beta(E_\mu - E_\nu)}$

$$\Rightarrow e^{-\beta E_\mu} \cdot 1 = e^{-\beta E_\nu} e^{-\beta E_\mu} e^{+\beta E_\nu}$$

(Hence proved)

Then, the Markov process defined by the Metropolis algorithm is guaranteed to lead to the limiting distribution $\omega_\mu(\infty) = p_\mu^{eq}$. Let us write down the P matrix for this Markov process explicitly for the 3-Ising spins example.

$$
\begin{bmatrix}
1-e^{-4\beta J} & \frac{e^{-4\beta J}}{3} & \frac{e^{-4\beta J}}{3} & 0 & \frac{e^{-4\beta J}}{3} & 0 & 0 & 0 \\
\frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\
\frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\
0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} \\
\frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\
0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\
0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\
0 & 0 & 0 & \frac{e^{-4\beta J}}{3} & 0 & \frac{e^{-4\beta J}}{3} & \frac{e^{-4\beta J}}{3} & 1-e^{-4\beta J}
\end{bmatrix}
$$

One can plug in various values of βJ and explicitly check that $\lambda_0 = 1$ and all other $|\lambda_i| < 1$ for $\beta J \in (0, \infty)$

Also, one can verify that

$(1\ 1\ 1\ \ldots\ \ldots\ 1)^T$ has $\lambda_0 = 1$

and $(1\ e^{-4\beta J}\ e^{-4\beta J}\ e^{-4\beta J}\ e^{-4\beta J}\ e^{-4\beta J}\ e^{-4\beta J}\ 1)$

has $\lambda_0 = 1$

h/t (Anupam Kundu, ICTS)

Let us prove that all eigenvalues of matrix $P(\mu \to \nu)$ are real even though it is not symmetric.

Define the following diagonal matrix D

$$D_{\mu\nu} = \sqrt{p^{eq}_\mu}\, \delta_{\mu\nu} \qquad (14)$$

Then, define the following similarity transformation

$$M = D \cdot P \cdot D^{-1} \qquad (15)$$

$$M_{\mu\nu} = \sqrt{p^{eq}_\mu}\, P(\mu \to \nu) \frac{1}{\sqrt{p^{eq}_\nu}} = \sqrt{p^{eq}_\mu}\, \frac{P(\nu \to \mu) p^{eq}_\nu}{p^{eq}_\mu} \frac{1}{\sqrt{p^{eq}_\nu}}$$

$$= \sqrt{p^{eq}_\nu}\, P(\nu \to \mu) \frac{1}{\sqrt{p^{eq}_\mu}} = M_{\nu\mu}$$

Since M is a real symm. matrix, its eigenvalues are real. The eigenvalues of M and P are identical since they are related by Eq (15).