Computational Science for many-body problems

- This class is from 15:10 to 16:40 (90 min.)

- The recordings of the previous lectures have been uploaded on ITC-LMS.

- Lecture slides are also available on ITC-LMS and github: https://github.com/compsci-alliance/many-body-problems

拡張アンサンブル法によるモンテカルロ計算

Extended Ensemble method for Monte Carlo Methods

理学系研究科　大久保毅

Graduate school of science, **Tsuyoshi Okubo**

Today

Classical

1st: Many-body problems in physics and why they are hard to solve

2nd: Classical statistical models and numerical simulation

3rd: Classical Monte Carlo method

4th: Applications of classical Monte Carlo method

5th: Molecular dynamics simulation and its applications

Quantum

6th: Extended ensemble method for Monte Carlo methods

7th: Quantum lattice models and numerical simulation

8th: Quantum Monte Carlo methods

9th: Applications of quantum Monte Carlo methods

10th: Linear algebra of large and sparse matrices for quantum many-body problems

11th: Krylov subspace methods and their applications to quantum many-body problems

12th: Large sparse matrices, and quantum statistical mechanics

13th: Parallelization for many-body problems

Extended Ensemble method for Monte Carlo Methods

Contents

- Back ground
  - Density of states and the histogram method
- Extended ensemble methods using information of the density of states
  - Multi Canonical Method
  - Wang-Landau method
- Another extended ensemble: Replica exchange method
- Closing

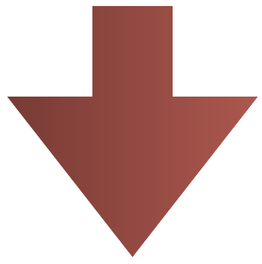
Back ground

Extended ensemble = general ensemble

In conventional MC or MD simulation:

We try to estimate expectation values under
"physically meaningful" ensembles.

NVE, NVT, NPT, …



Even if an ensemble is not directly connected to any physical systems, we can use is to enhance the efficiency of numerical calculations (MC, MD) for interested physical systems.

Large relaxation time in standard MC and MD

- Critical phenomena
  - $\tau \sim L^z$ with standard algorithm (critical slowing down)
  - z can be significantly reduced by using "global update"
- First order phase transition, Glass transition (structural glass, spin glass), protein foldings,
  - $\tau \sim \exp(\Delta E/T)$ or $\exp(\Delta E/|T\text{-}T_c|)$; Note $\Delta E \propto L^d$ $(or\ L^{d\text{-}1})$!
  - Exponential can be reduce to polynomial by using extended ensemble methods.

Density of state and histogram method

Origin of exponentially long relaxation time

Partition function of the canonical ensemble

$$Z = \int d\Gamma e^{-\beta \mathcal{H}(\Gamma)} = \int dE \underline{\rho(E)} e^{-\beta E} = \int dE e^{-\underline{\beta F(E)}}$$

Density of states

"Free energy"

$$F(E) \equiv E - k_B T \log \rho(E)$$

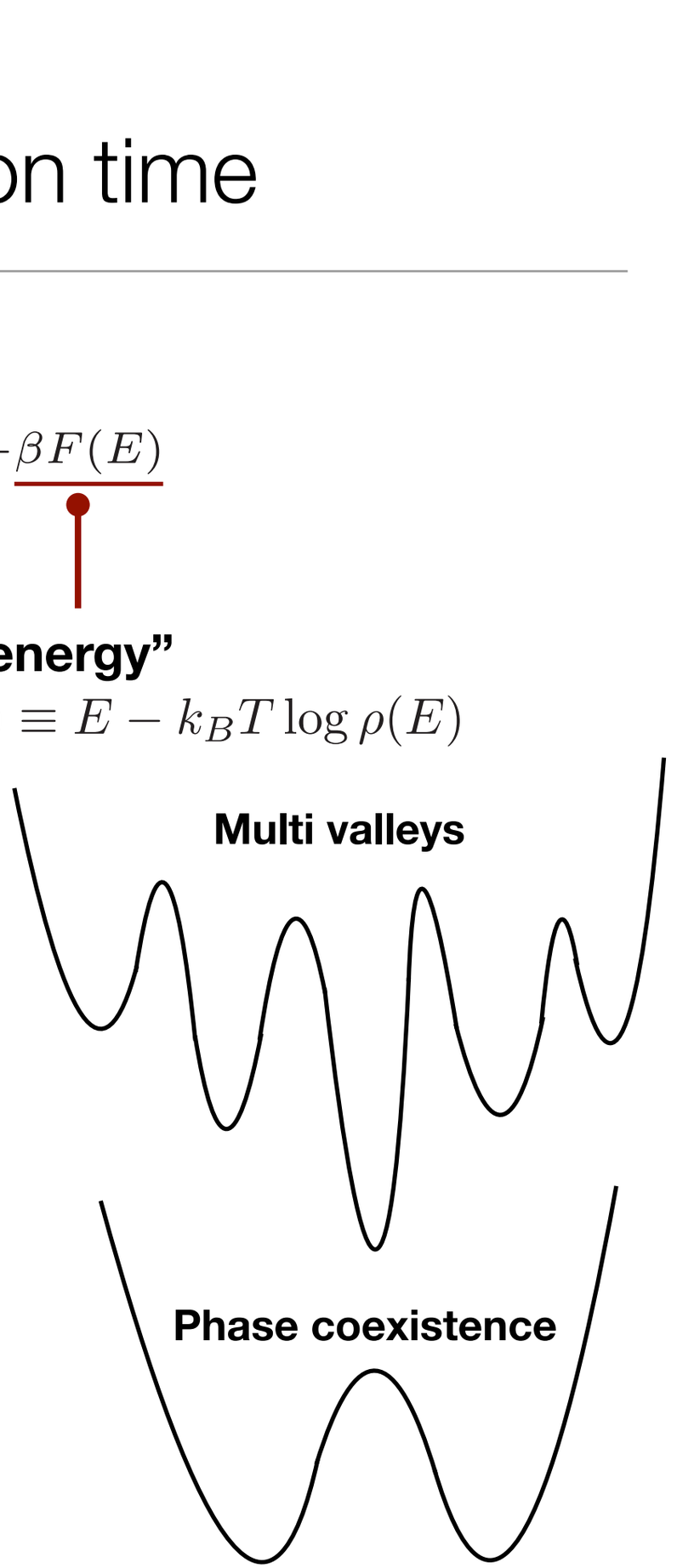
Probability distribution for energy

$$P(E) = \frac{1}{Z} e^{-\beta F(E)}$$

Note! Free energy is extensive: $F(E) \propto N$

- "Transition probability" is proportional to the exponential of Free-energy difference: exp(-ΔF/T)
- Usual algorithm of MC (and MD) changes the state (or the energy) gradually.

If there are local minima, the relaxation time could be exponentially large as the size is increased.



Density of states

$$Z = \int d\Gamma e^{-\beta \mathcal{H}(\Gamma)} = \int dE \rho(E) e^{-\beta E}$$
$$= \int dE \int dM \rho(E, M) e^{-\beta E}$$

$\int d\Gamma$ ~O(N)-dimensional integral

$\int dE$ ~1-dimensional

$\int dE \int dM$ ~2-dimensional

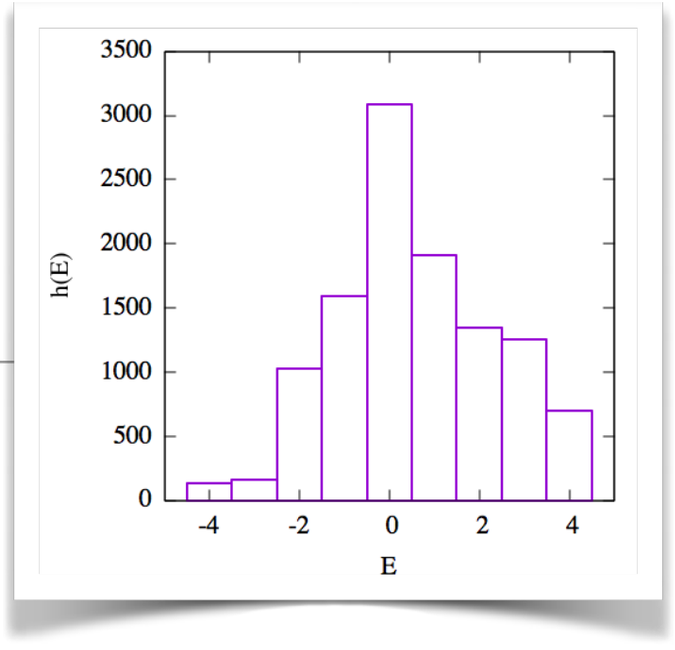
- If we know the exact $\rho(E)$(or $\rho(E, M)$), the calculation of partition function is reduced to 1 or (a few) -dimensional integral.
- Even if we only know an approximate density of states,

$$\tilde{\rho}(E) \simeq \rho(E)$$

we can improve the sampling efficiency by using its information.

- Histogram method
- Multi canonical method
- Wang-Landau method

Energy Histogram



Energy histogram:

In MC or MD calculations

$h(E_i)$:# of samples (snap shots) with energy in

$$E_i - \Delta E/2 \leq E < E_i + \Delta E/2$$

$$\Rightarrow \quad P(E)\Delta E \simeq \frac{1}{N_h} h(E)$$

Total # of samples

$$N_h \equiv \sum_i h(E_i)$$

e.*g.* NVT ensemble:

$$P(E) = \frac{1}{Z(\beta)} \rho(E) e^{-\beta E} \quad \Rightarrow \quad \rho(E) \simeq \frac{Z(\beta)}{N_h \Delta E} h(E) e^{\beta E}$$

We can calculate (approximate) density of states from usual MC or MD simulations!

Because we don't know the partition function,
DOS is determined up to the proportional coefficient.

Histogram method (reweighting method)

Energy expectation value of different temperatures

$$\langle E \rangle_{\beta'} = \frac{\int dE \rho(E) E e^{-\beta' E}}{\int dE \rho(E) e^{-\beta' E}} \simeq \frac{\sum_i E_i h(E_i) e^{-(\beta'-\beta)E_i}}{\sum_i h(E_i) e^{-(\beta'-\beta)E_i}}$$

Any expectation values can also be calculated by the histogram method.

$$\langle O \rangle_{\beta'} \simeq \frac{\sum_i O(E_i) h(E_i) e^{-(\beta'-\beta)E_i}}{\sum_i h(E_i) e^{-(\beta'-\beta)E_i}}$$

$$\rho(E) \simeq \frac{Z(\beta)}{N_h \Delta E} h(E) e^{\beta E}$$

Average at energy E_i

$$O(E_i) \equiv \sum_{E(\Gamma_j) \in E_i} O(\Gamma_j)$$

Limitation of histogram method

Reweighted histogram becomes
less accurate
when T' is far from the original T.

"Tail" of the original histogram has only small # of snapshots.→large noise

Central limit theorem

Width of energy distribution: $\propto \sqrt{N}$

Average of energy: $\propto N$

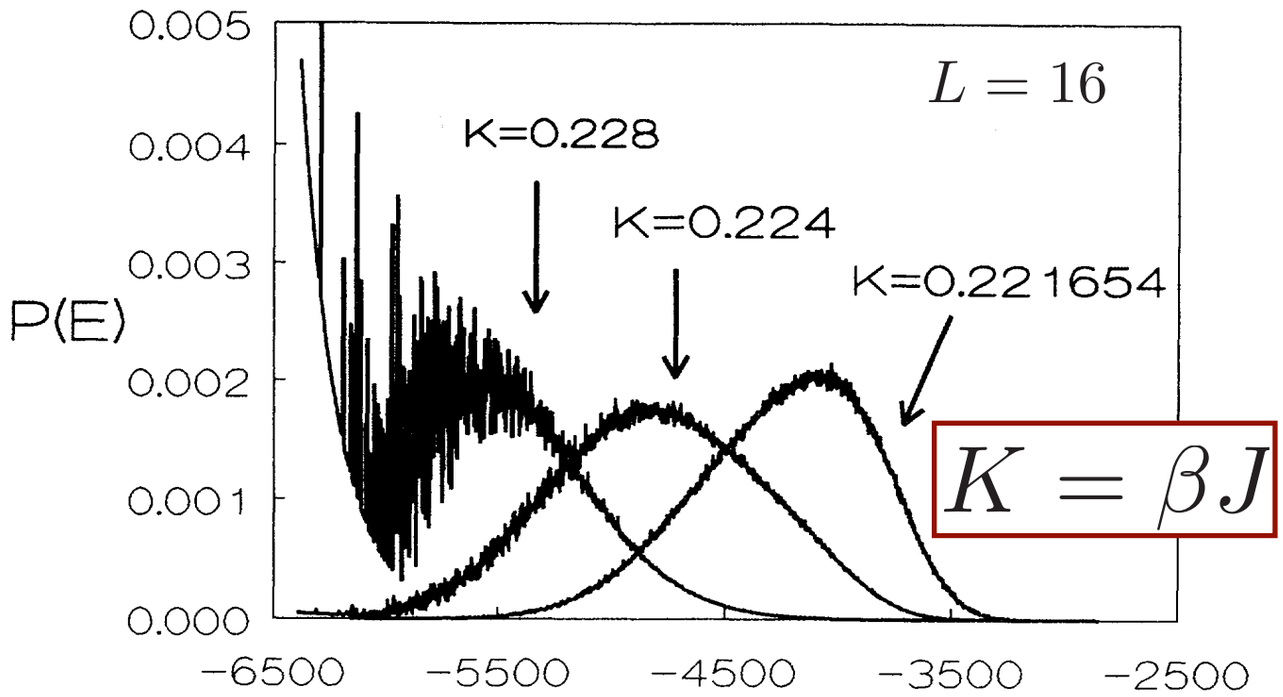
Distribution becomes narrower as *N* is increased!

Reliable temperature region for reweighting:

$$\Delta T \propto \frac{1}{\sqrt{N}}$$

Energy distribution of 3d-Ising model

A. M. Ferrenberg and D. P. Landau, Phys. Rev. B **44**, 5081 (1991)



MC simulation at K=0.221654

Reweighting to
K=0.224 and K=0.228

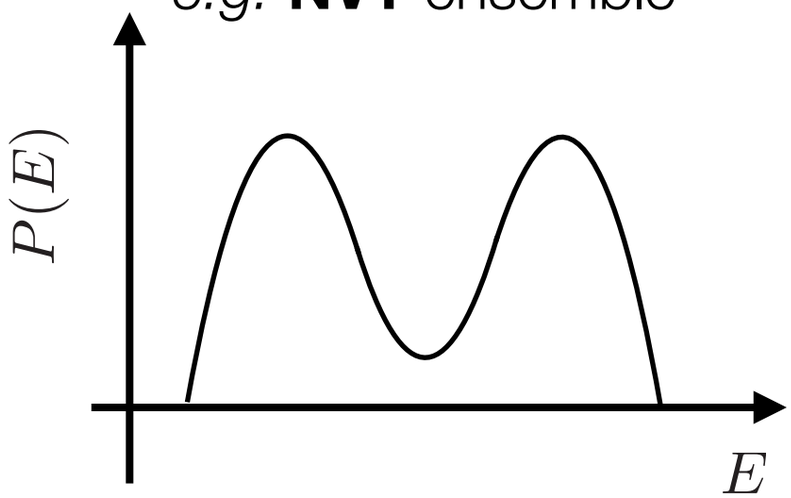
Multi Canonical methods

Idea of Multi-Canonical method

B.A. Berg and T. Neuhaus (1992)

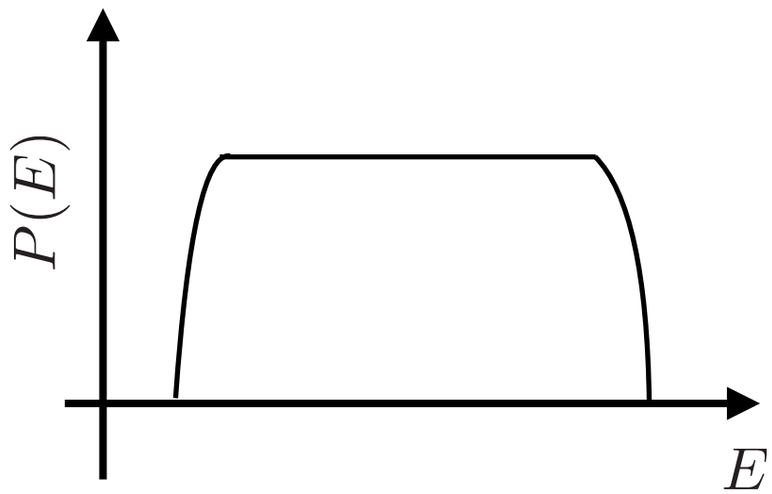
If we can prepare a special ensemble
where the energy distribution is “flat”,
we can efficiently sample all relevant states.

e.g. **NVT** ensemble



$$P(E) \propto \rho(E) e^{-\beta E}$$

“special” ensemble



$$\tilde{P}(E) \propto \rho(E) e^{-S(E)} = \text{const.}$$

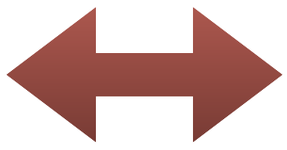
Special ensemble is related to log of DOS!

$$S(E) = \log \rho(E)$$

How to obtain the special ensemble?

Special ensemble is log of DOS!

$$S(E) = \log \rho(E)$$



DOS is unknown!

We can obtain $S(E)$ approximately by iterative calculations.

“Sketch” of an iterative algorithm:

1. Run MC simulation on a high temperature and calculate energy histogram.

$$h(E) \sim \rho(E) e^{-\beta E}$$

2. Based on the energy histogram, extract approximate $S(E)$.

$$S^0(E) = \beta E + \log h(E)$$

3. **Loop** n
    1. Run MC simulation under $S^{(n)}(E)$ and calculate histogram $h^{(n)}(E)$
    2. Calculate next $S^{(n+1)}(E)$ as

$$S^{(n+1)}(E) = S^{(n)}(E) + \log h^{(n)}(E)$$

How to obtain the special ensemble?

3. **Loop** n
   1. Run MC simulation under $S^{(n)}(E)$ and calculate histogram $h^{(n)}(E)$
   2. Calculate next $S^{(n+1)}(E)$ as

$$S^{(n+1)}(E) = S^{(n)}(E) + \log h^{(n)}(E)$$

- The histogram $h^{(n)}(E)$ is expected to be $h^{(n)}(E) \sim \rho(E)e^{-S^{(n)}(E)}$
- When $S^{(n)}(E)$ becomes close to $log\, \rho(E)$, the histogram becomes almost flat.

→ We can efficiently sample the histogram and DOS.

- By using accurate $S^{(n)}(E)$ we can calculate expectations values under the canonical ensemble by using reweighting technique.

$$\langle O \rangle_\beta = \frac{\int dE O(E)\rho(E)e^{-S(E)}e^{-\beta E+S(E)}}{\int dE \rho(E)e^{-S(E)}e^{-\beta E+S(E)}} = \frac{\langle Oe^{-\beta E+S(E)}\rangle_S}{\langle e^{-\beta E+S(E)}\rangle_S}$$

Detailed algorithm (will be skipped)

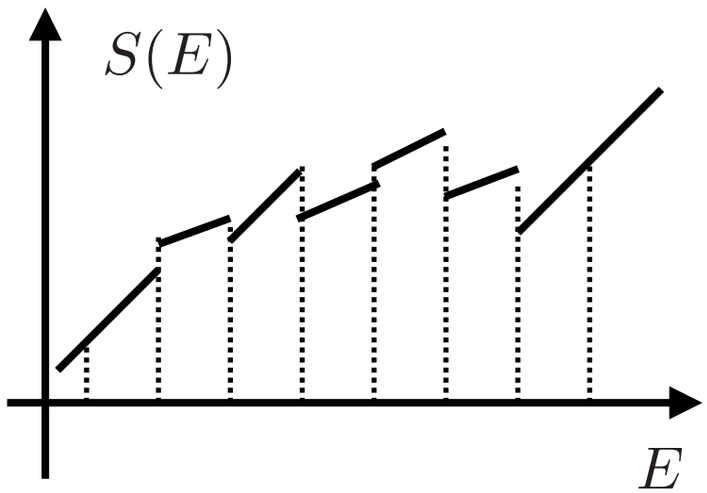
B.A. Berg, Nucl. Phys. B (Proc. Suple.) **63**A-C, 982 (1998)

Suppose $S(E)$ looks like: $S(E) = \beta(E)E - \alpha(E)$

(Energy dependent temperature)

$$S(E) \simeq \beta_i E - \alpha_i$$

for $E_i - \Delta E/2 \leq E \leq E_i + \Delta E/2$



In a specific interval, we want to optimize β and α,
i.e. $P(E)$ becomes flat.

By defining $\beta_i \equiv \dfrac{S(E_{i+1}) - S(E_i)}{\Delta E}$

$$\alpha_{i-1} = \alpha_i + (\beta_{i-1} - \beta_i)E_i$$

We fix $\alpha_{i_{max}} = 0$

Detailed algorithm (will be skipped)

B.A. Berg, Nucl. Phys. B (Proc. Suple.) **63**A-C, 982 (1998)

Iteration :how to determine next β and α

In order to make the histogram flat, $S^{(n+1)}(E) = S^{(n)}(E) + \log h^{(n)}(E)$

$$\tilde{\beta}_i^{(n+1)} = \beta_i^{(n)} + \log \frac{h_{i+1}^{(n)}}{h_i^{(n)}}$$

This estimator could be suffered from large statistical error

Gradual change from the previous β

$$\beta_i^{(n+1)} = (1 - c_i)\beta_i^{(n)} + c_i \tilde{\beta}_i^{(n+1)}$$

*For optimal c_i , see the reference

α is calculated from β

$$\alpha_{i-1}^{(n+1)} = \alpha_i^{(n+1)} + (\beta_{i-1}^{(n+1)} - \beta_i^{(n+1)})E_i$$

Example of application

q-state Potts model on the square lattice

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \delta_{S_i, S_j} \qquad S_i = 0, 1, 2, \cdots, q-1$$

Phase transition at

$$T_c/J = \frac{1}{\log(1+\sqrt{q})}$$

$q = 2$: Equivalent to Ising model

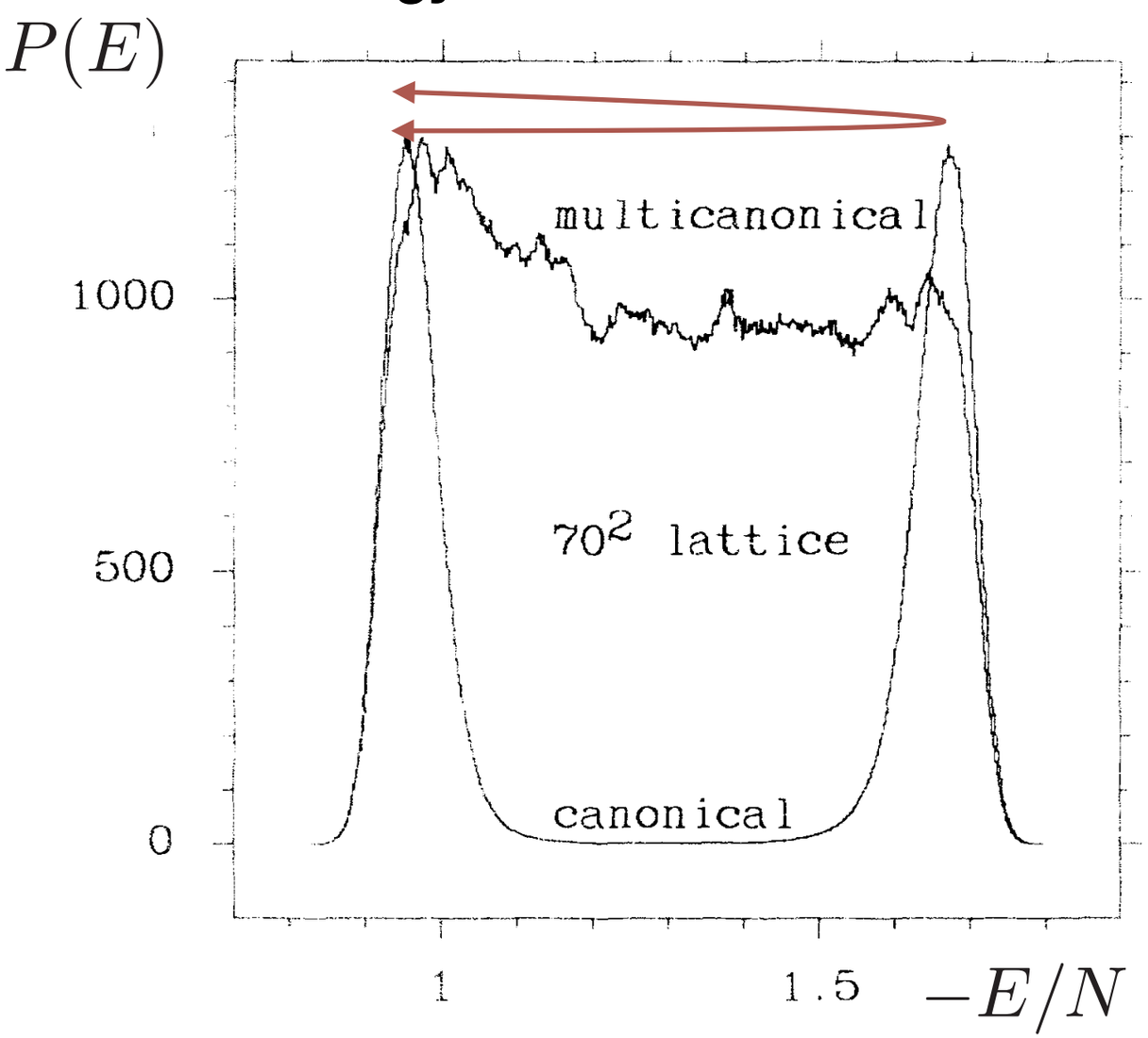
$q \leq 4$: Continuous phase transition

$q > 5$: 1st order phase transition

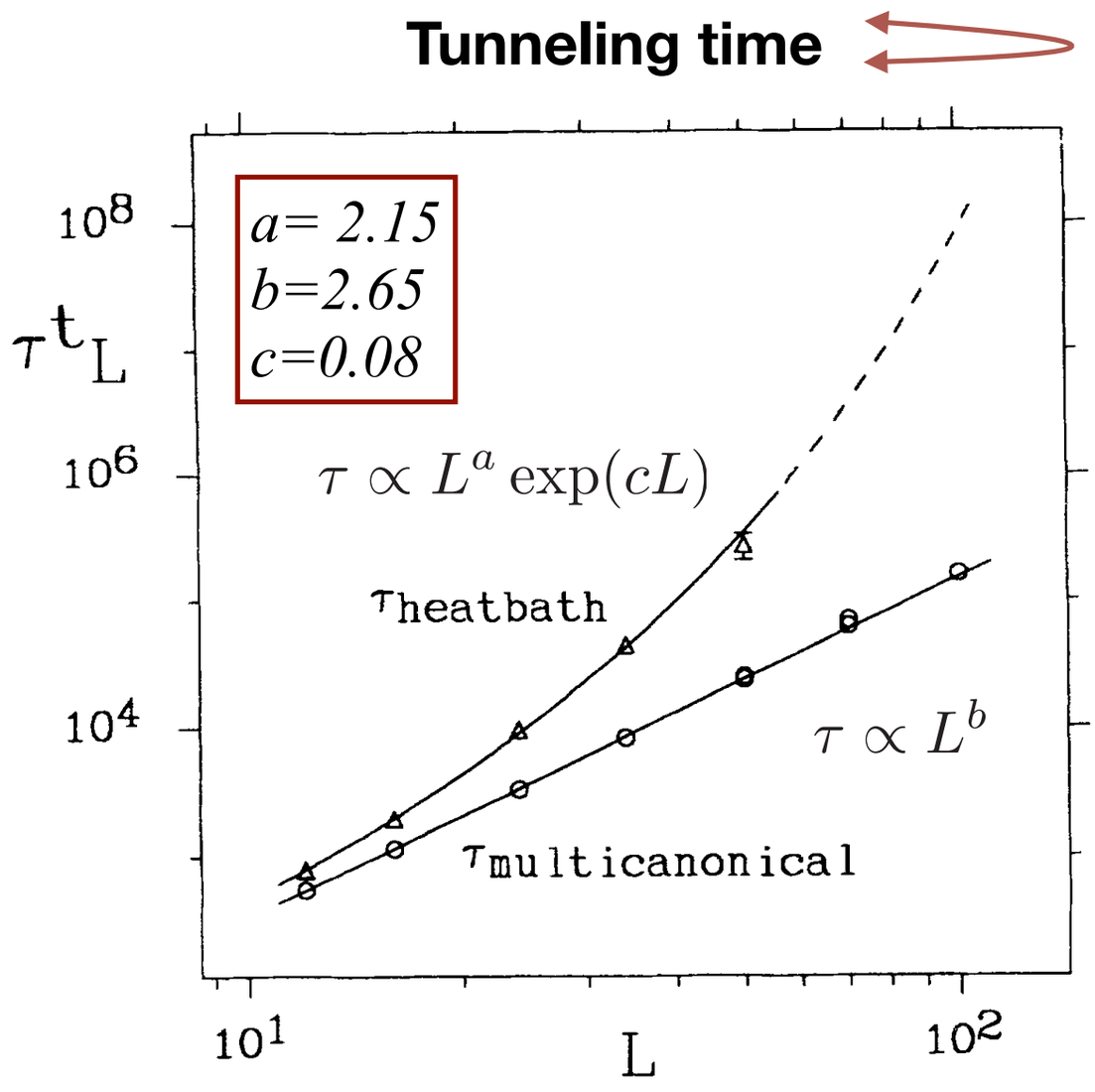
Multi Canonical method for q=10 Potts model

B.A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992)

Energy distribution around Tc



Tunneling time



By Multi canonical method, the tunneling time is reduced to the power of L!

Wang-Landau method

Need to update

Wang-Landau method

F. Wang and D. P. Landau (2001)

Another method to obtain the density of state:

Random walk on the energy space

Markov Chain Monte Carlo with the transition probability

$$W_{\Gamma\to\Gamma'} = \min\left(\frac{g(E(\Gamma))}{g(E(\Gamma'))}, 1\right)$$

$g(E)$:estimate of DOS

→ Its steady state is

$$P(\Gamma) \propto \frac{1}{g(E(\Gamma))}$$

The energy distribution (histogram) :

$$P(E)dE = P(\Gamma)d\Gamma = P(\Gamma)\rho(E)dE \propto \frac{\rho(E)}{g(E)}dE$$

if $g(E) = \rho(E)$ → This MCMC gives us a completely flat histogram!

Wang-Landau method:update of g(E)

F. Wang and D. P. Landau (2001)

Initially, we don't know DOS. → Set an initial guess, e.g. $g(E) = 1$

Along MCMC, we update $g(E)$ of the $E(\Gamma)$ as

$$g_{new}(E) = g(E) \times f \qquad (\log g_{new}(E) = \log g(E) + \log f)$$

*Note after N step, g(E) changes like

$$g_{new}(E) \sim g(E) f^{N \frac{\rho(E)}{g(E)}}$$

g(E) increases when $g(E) < \rho(E)$ for $f > 1$.

If the multiplication factor is "gradually" reduced to $f = 1$,

$g(E)$ eventually converges to $\rho(E)$.

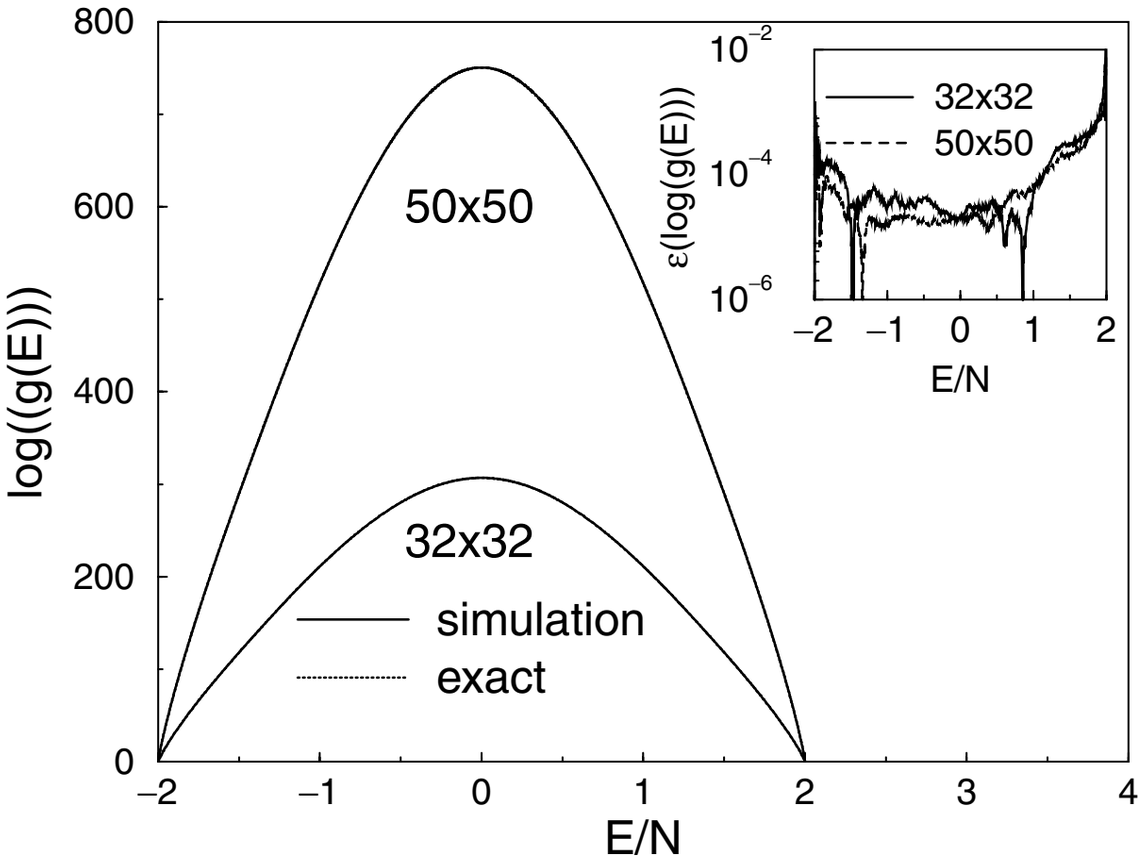
"gradual" change of f :

1. Initially $f = f_0$ (e.g. $f_0 = e^1$)
2. Loop i
   - If (the histogram $h(E)$ becomes "flat"?)
     - Then, we decrease f_i as
       $f_{i+1} = (f_i)^x$ (e.g. $x = 1/2$),
       and reset the histogram.
3. Repeat until f_i becomes enough small (e.g. $f \sim \exp(10^{-8})$)

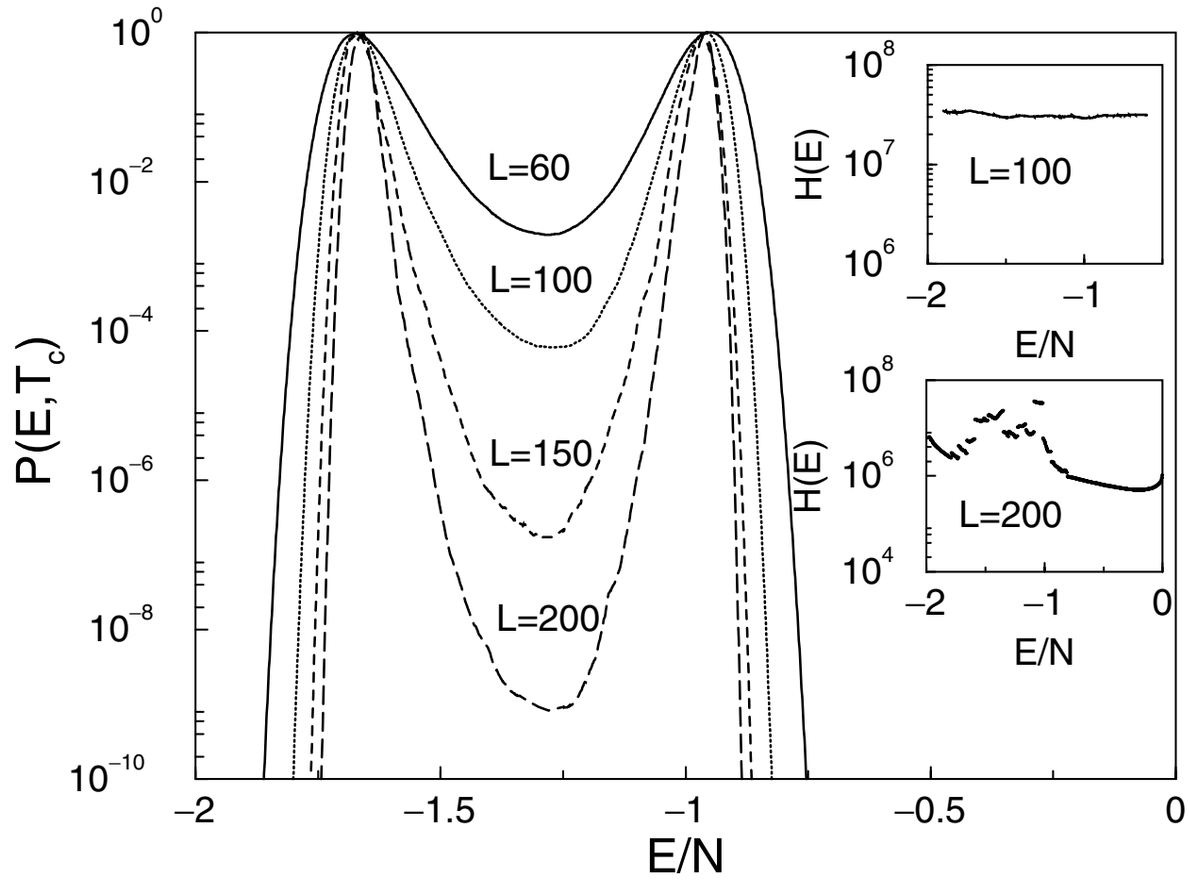
Power of Wang-Landau method

F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001)

Density of state of 2D-Ising model



Density of state of q=10 Potts model



We can obtain very accurate density of state by Wang-Landau method!

Replica Exchange method

K.Hukushim and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).

Replica exchange (parallel tempering)

A different type of extended ensemble:

Usual MC or MD considers one parameter and one realization:

$$T,\ \Gamma = \{S_i\}, \{\boldsymbol{q}_i, \boldsymbol{p}_i\}$$

Replica exchange method considers
multiple parameters together with multiple realizations:

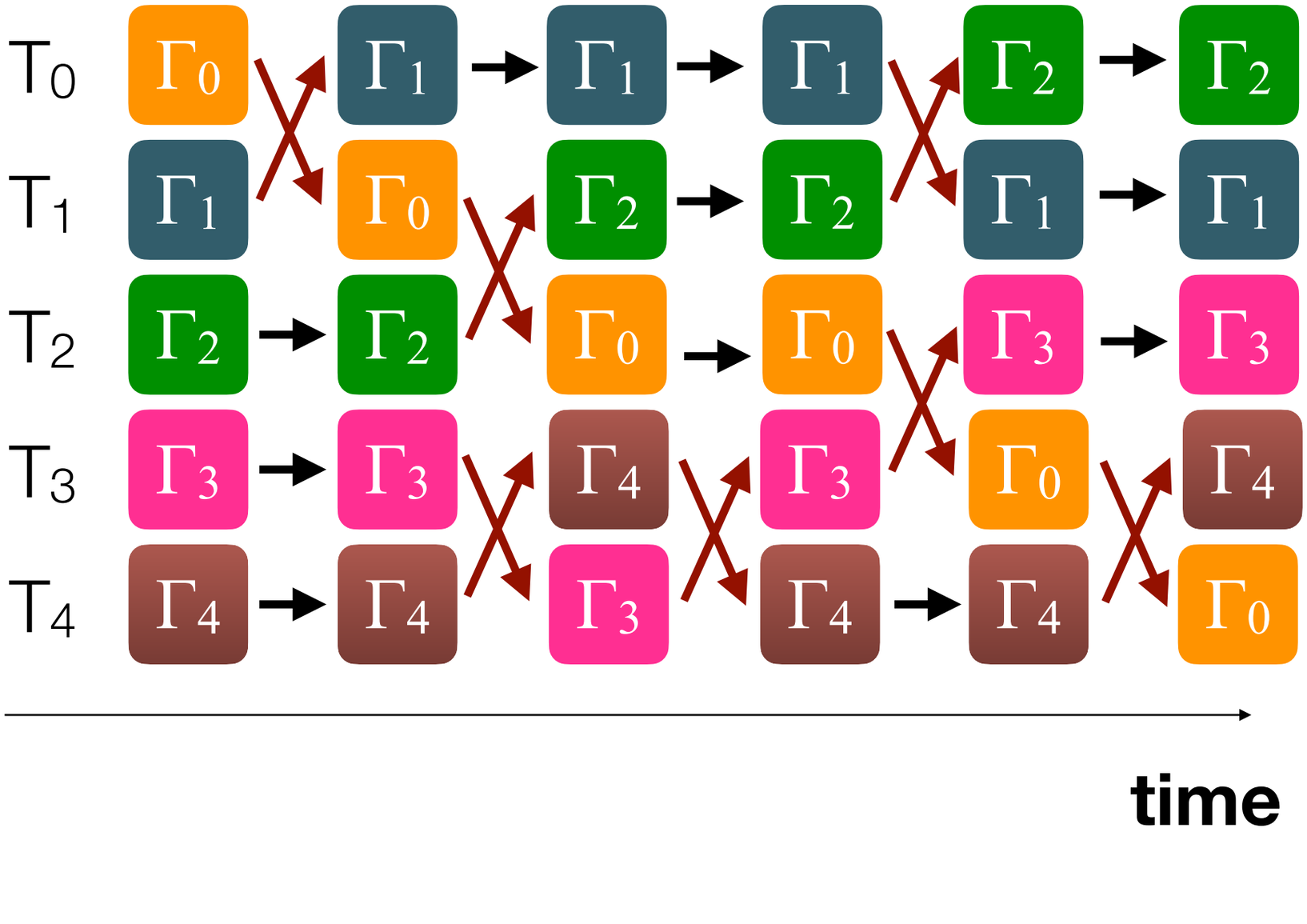
$$\{T_0, T_1, \cdots, T_M\},\ \{\Gamma_0, \Gamma_1, \cdots, \Gamma_M\},$$

Try to sample "(M+1)-dimensional" joint-distribution

$$P(\Gamma_0, \Gamma_1, \cdots, \Gamma_M; T_0, T_1, \cdots, T_M)$$

“Replica exchange”

Along simulation, we “exchange” the relationship between parameter and realization

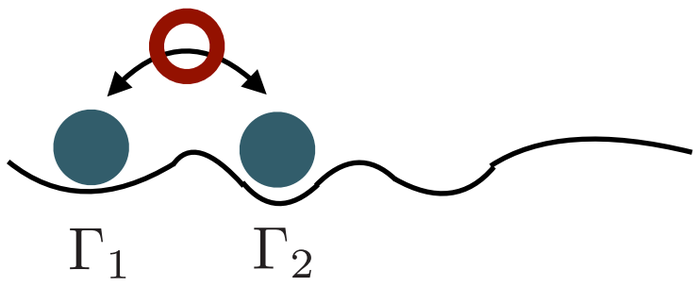


Purpose of replica exchange

Free energy landscape depends on the parameter

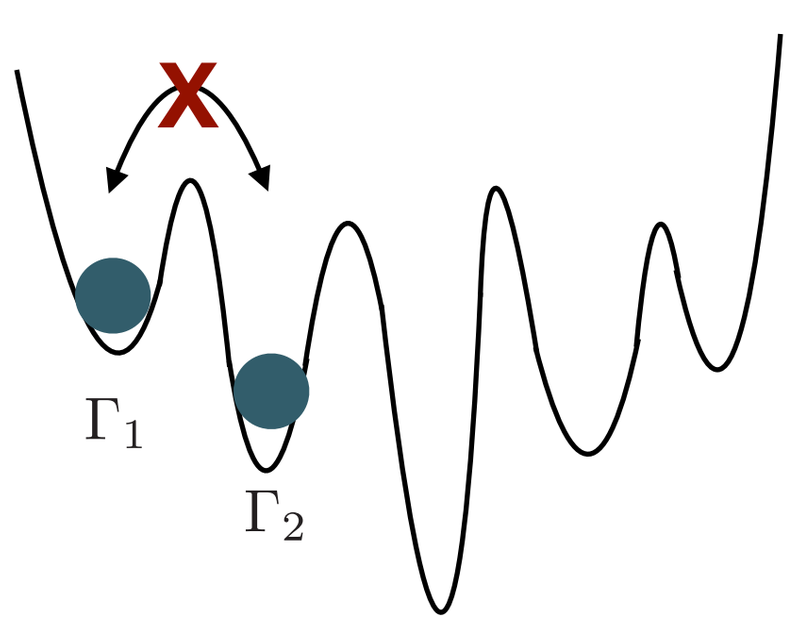
High temperature:T_h

Γ easily moves to other points!



Low temperature:T_l

Γ hardly moves to other minima!



Make a pass like:

$$\{\Gamma_1, T_l\} \rightarrow \{\Gamma_1, T_h\} \rightarrow \{\Gamma_2, T_h\} \rightarrow \{\Gamma_2, T_l\}$$

low **high** **high** **low**

* Parameter is not necessary a temperature.

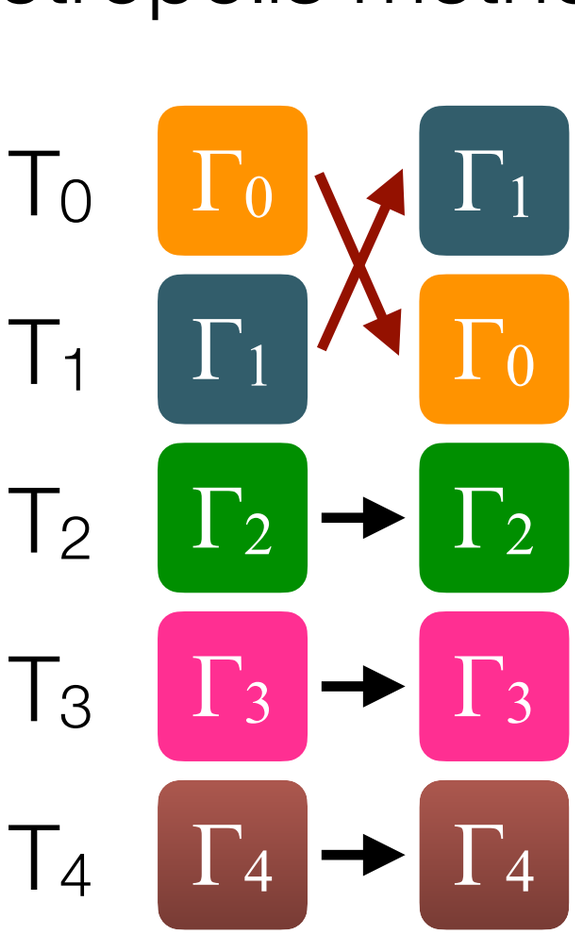
Markov Chain Monte Carlo for Replica Exchange

Target steady state distribution:

$$P(\Gamma_0, \Gamma_1, \cdots, \Gamma_M; T_0, T_1, \cdots, T_M) \propto e^{-\sum_i^M \beta_i E_i}$$

$$E_i \equiv \mathcal{H}(\Gamma_i)$$

Metropolis method:



$\mathcal{T}$ **:sequence of temperatures**

$$\mathcal{T} = \{T_1, T_0, T_2, \cdots\}$$

$$\{T_0, \Gamma_0\}, \{T_1, \Gamma_1\} \to \{T_1, \Gamma_0\}, \{T_0, \Gamma_1\}$$

$$\mathcal{T}_{01} \qquad\qquad \mathcal{T}_{10}$$

Transition probability

$$W_{\mathcal{T}_{01} \to \mathcal{T}_{10}} = \min\left(1, \frac{P(\{\Gamma_i\}; \mathcal{T}_{10})}{P(\{\Gamma_i\}; \mathcal{T}_{01})}\right)$$

$$\frac{P(\{\Gamma_i\}; \mathcal{T}_{10})}{P(\{\Gamma_i\}; \mathcal{T}_{01})} = \frac{e^{-\beta_1 E_0 - \beta_0 E_1}}{e^{-\beta_0 E_0 - \beta_1 E_1}}$$

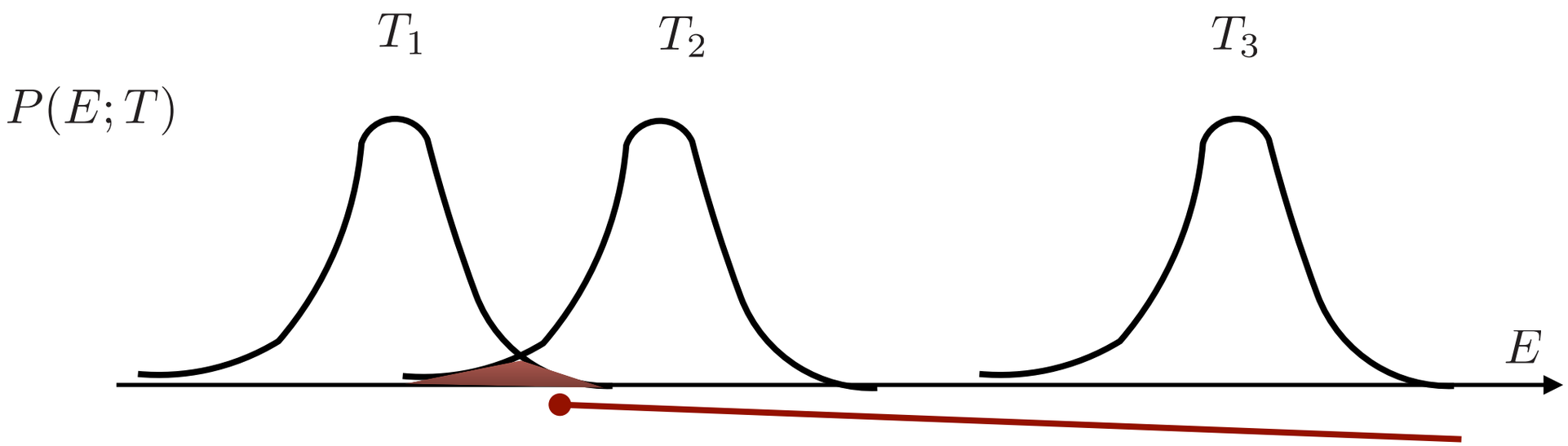
$$= e^{(\beta_0 - \beta_1)(E_0 - E_1)}$$

Select of temperature sequence

$$\frac{P(\{\Gamma_i\};\mathcal{T}_{10})}{P(\{\Gamma_i\};\mathcal{T}_{01})} = e^{(\beta_0-\beta_1)(E_0-E_1)} = \frac{P(E_1;T_0)P(E_0;T_1)}{P(E_0;T_0)P(E_1;T_1)}$$

Energy distribution at T

$P(E;T)$



Almost all exchange occurs the energy region of “overlap”.

$\{\Gamma_1, T_1\}, \{\Gamma_2, T_2\} \rightarrow \{\Gamma_1, T_2\}, \{\Gamma_2, T_1\}$:acceptable!

$\{\Gamma_2, T_2\}, \{\Gamma_3, T_3\} \rightarrow \{\Gamma_2, T_3\}, \{\Gamma_3, T_2\}$:almost rejected!

For efficient exchange, we have to choose a sequence of temperatures so that the energy distributions have finite overlap!

Usually we only exchange the nearest neighbor pairs of temperatures

Select of temperature sequence: Example

Suppose $C = \dfrac{dE}{dT} = \text{const.}$

$$\frac{P(\{\Gamma_i\}; \mathcal{T}_{10})}{P(\{\Gamma_i\}; \mathcal{T}_{01})} = e^{(\beta_0 - \beta_1)(E_0 - E_1)}$$

➡ Temperature sequence satisfying almost "flat" transition probability

$$(\beta_i - \beta_{i+1})(E_i - E_{i+1}) = \text{const.}$$

$$\Leftrightarrow \quad C\frac{(T_{i+1} - T_i)^2}{T_{i+1}T_i} = \text{const.}$$

$$\Rightarrow \quad T_{i+1} = \alpha T_i \quad \textbf{:Temperatures are geometric sequence!}$$

$$\alpha \sim 1 + O(1/\sqrt{C})$$

Important notice:

Heat capacity C is an extensive quantity: $C \sim O(N)$

➡ In order to keep finite overwrap, we need to increase temperature point M as

$$M \propto \sqrt{N} \qquad (T_{max} = T_M = \alpha^M T_{min})$$

Relaxation time of the replica exchange

In order to confirm the equilibration of the whole system,
usually we need two criterions.

1. The correlation time at the highest temperature is sufficiently short,
   e.g. $\tau=O(1)$

   → If a replica visits the highest temperature, it can easily change its state Γ.

2. All replicas make several (~O(10)) round trips between
   the lowest and the highest temperatures

   → The ensemble at the lower temperature is in the equilibrium.

The second part determines the relaxation time of the method.

$$\tau_{\mathrm{RE}} \sim \text{round trip time}$$

* If the replica exchange is an random walk:

$$\text{round trip time} \propto M^2$$

Summary of replica exchange

Algorithm:

1. Make a temperature set $\{T_1, T_2, \ldots, T_M\}$
2. Loop n
   (1) Do MC or MD for M replicas: $\{\Gamma_1, \Gamma_2, \ldots \Gamma_M; T_1, T_2, \ldots, T_M\}$
   (2) Calculate the energies of replicas
   (3) Try replica exchange based on, e.g. Metropolis method
   - Usually we alternatively try replica exchange such as
     even n; $\{1\leftrightarrow2\},\{3\leftrightarrow4\},\{5\leftrightarrow6\},\ldots$
     odd n; $\{2\leftrightarrow3\},\{4\leftrightarrow5\},\{6\leftrightarrow7\},\ldots$
     Note: each exchange trial is independent
   (4) Observe the quantities for $\{\Gamma_1, \Gamma_2, \ldots \Gamma_M; T_1, T_2, \ldots, T_M\}$

If we already have a MC or MD programs,
it is very easy to introduce the replica exchange method!

Closing

Summary

- Numerical approach to calculate the ensemble average as time average
  - Markov Chain Monte Carlo
    - By the importance sampling, we can realize stochastic process those steady state is the target distribution.
    - By the global update technique, we can largely reduce the correlation time
  - Molecular dynamics simulation
    - By the symplectic integral, we can perform long time simulate stably.
    - We can control temperature and pressure
  - We can use generalized ensemble to accelerate the sampling efficiency.
    - Multi canonical and Wang-Landau methods: Efficient sampling for the density of states
    - Replica exchange method: Simultaneous simulation with different parameter
  - We can also combine MCMC and MD to investigate your system
    - Hybrid Monte Carlo (or Hamiltonian Monte Carlo)
      - By introducing momentum degrees, we perform dynamical simulation for, e.g., Ising model.
    - Spin dynamics simulation
      - Generate initial state from MCMC, and perform dynamical simulation as like MD.
- Data analysis
  - Estimate of the error (both of statical and systematic errors) are important
  - Finite size scaling is powerful to investigate, e.g., critical phenomena.

Next (5/31) (will be given by Yamaji-sensei)

Classical

1st: Many-body problems in physics and why they are hard to solve

2nd: Classical statistical models and numerical simulation

3rd: Classical Monte Carlo method

4th: Applications of classical Monte Carlo method

5th: Molecular dynamics simulation and its applications

6th: Extended ensemble method for Monte Carlo methods

Quantum

7th: Quantum lattice models and numerical simulation

8th: Quantum Monte Carlo methods

9th: Applications of quantum Monte Carlo methods

10th: Linear algebra of large and sparse matrices for quantum many-body problems

11th: Krylov subspace methods and their applications to quantum many-body problems

12th: Large sparse matrices, and quantum statistical mechanics

13th: Parallelization for many-body problems