古典モンテカルロ法とその応用

Classical Monte Carlo method and its Application

理学研究科　物理学専攻　大久保　毅

Outline

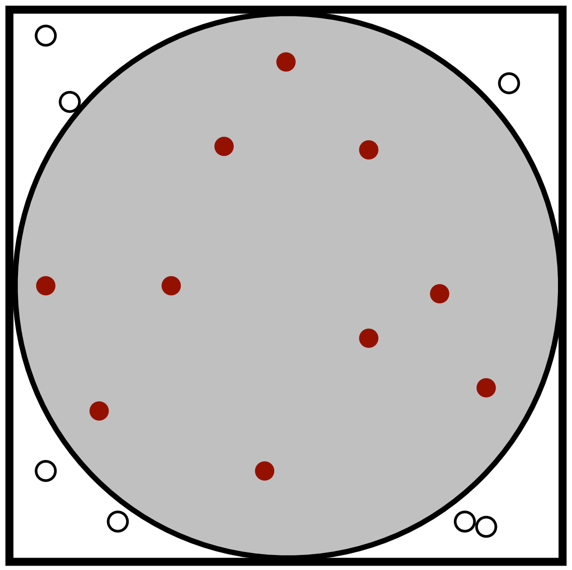
- Standard Monte Carlo method
  - Importance sampling, Markov Chain Monte Carlo
- Application to classical spin systems
  - Local update, Global update
- Computational Science using Monte Carlo method
  - Important tips to obtain reliable results
  - Application and analysis in the case of critical phenomena

Monte Carlo method: Randomized algorithm

Randomized algorithm: It changes its behavior depending on (psuedo) random numbers on execution

Example :

Area of a circle: $\int_{-1}^{1} dx \int_{-1}^{1} dy$



Algorithm

$N_a = 0 \quad N = 0$ initialize

loop i

$x_i \in [-1, 1]$
$y_i \in [-1, 1]$ take uniform random numbers

$N = N + 1$

if $x_i^2 + y_i^2 \leq 1$ then $N_a = N_a + 1$

end loop

$\frac{N_a}{N} \to \pi$

With statistical error proportional to $\frac{1}{\sqrt{N}}$

Monte Carlo Integration: General aspect

Monte Carlo Integration

$$\int d\Gamma f(\Gamma) = \int d\Gamma \frac{f(\Gamma)}{P(\Gamma)} P(\Gamma) = \left\langle \frac{f(\Gamma)}{P(\Gamma)} \right\rangle$$

$P(\Gamma)$: probability distribution

Estimate an integral as an expectation value under *P(Γ)*

Previous example: $P(\Gamma)$ = uniform distribution
obtained by a rejection sampling

Merit of Monte Carlo Integration

The error is basically independent on the dimension of *Γ*.

$\boxed{\epsilon \propto O(N^{-1/2})}$ N : sampling number

The error of usual numerical quadrature (eg. trapezoidal formula)
exponentially decreases as increase the dimension of *Γ*

eg. trapezoidal formula $\boxed{\epsilon \propto O(N^{-2/d})}$

Application to higher dimensions: The curse of dimensionality

Rejection sampling is inefficient for higher dimensions

Volume ratio between "*n*-dimensional hyper cubic" (with L=2) and "*n*-dimensional hyper sphere" (with r=1)

$$r = \frac{\pi^{n/2}/\Gamma(\frac{n}{2}+1)}{2^n} \sim \left(\frac{\pi}{en}\right)^{n/2}$$

Asymptotic form of *Γ*-function

$$\Gamma(x) \sim \left(\frac{x}{e}\right)^x$$

For larger *n*, the ratio exponentially decreases!

$\Rightarrow$ Error of the rejection sampling increases exponentially

Even if we can directly generate uniform distribution, uniform sampling is inefficient

$\int d\Gamma f(\Gamma)$: Several sampling points, $|f(\Gamma)| \ll 1$, don't contribute the integral so much.

If $|f(\Gamma)| \ll 1$

If we could pick up relevant points $|f(\Gamma)| \gg 1$ the efficiency largely increases! $\Rightarrow$ **Importance sampling**

Importance Sampling

Sampling the “important” points mainly

$$\int d\Gamma f(\Gamma) = \int d\Gamma \frac{f(\Gamma)}{P(\Gamma)} P(\Gamma) = \left\langle \frac{f(\Gamma)}{P(\Gamma)} \right\rangle$$

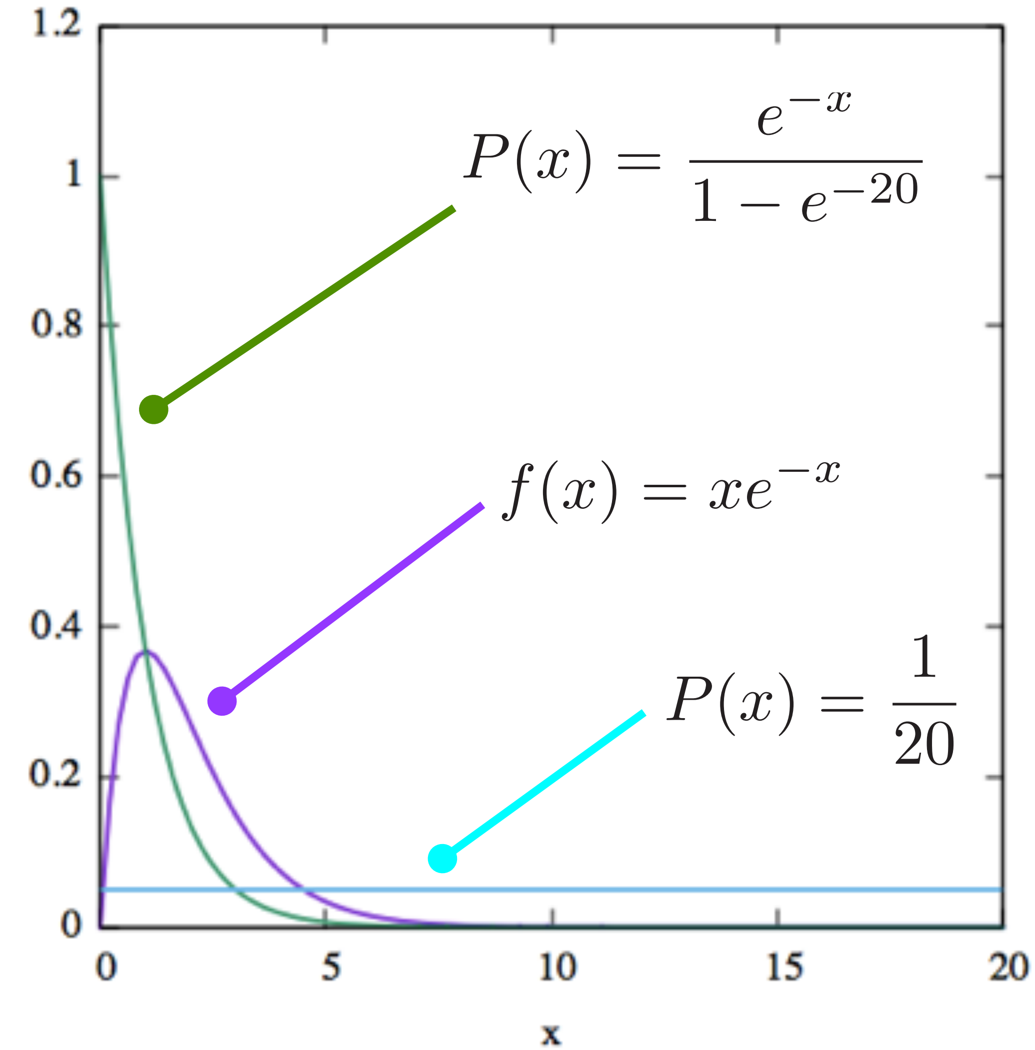
Chose $P(\Gamma)$ close to $f(\Gamma)$.

If we can choose $P(\Gamma) \propto f(\Gamma)$
it is the best.

However, it is unrealistic!

Because, in order to normalize $f(\Gamma)$, we have to know the value of integral, which is the answer we want to know.

Example of importance sampling



Markov Chain Monte Carlo

We can generate $P(\Gamma)$ as the steady state of a stochastic process

A sampling point move in Γ "randomly".

Master equation for general Markov process

$$\rho_{t+1}(\Gamma) = \rho_t(\Gamma) + \sum_{\Gamma'} W_{\Gamma'\to\Gamma}\rho_t(\Gamma') - \sum_{\Gamma'} W_{\Gamma\to\Gamma'}\rho_t(\Gamma)$$

$W_{\Gamma\to\Gamma'}$:transition probability from Γ to Γ'

$\rho_t(\Gamma)$: probability for appearance of Γ at time t

$$\sum_{\Gamma'} W_{\Gamma\to\Gamma'} = 1$$

$$\sum_{\Gamma} \rho_t(\Gamma) = 1$$

Markov process:

A future move depends only on the present state and independent of the past states.

If a Markov process becomes a steady state in the long time limit,

$$\lim_{t\to\infty} \rho_t(\Gamma) = P(\Gamma)$$

We can sample points with distribution $P(\Gamma)$ along this stochastic process

Markov Chain Monte Carlo: convergence condition

Conditions for transition probability for converging to $P(\Gamma)$

1. "**Ergodicity**"
   - Any two states Γ and Γ' are connected by W with finite steps
     - If we regard W as a matrix, this condition means

$$^{\exists}T > 0, {}^{\forall}(\Gamma, \Gamma'), \quad [(W)^T]_{\Gamma,\Gamma'} > 0$$

2. "**Balance Condition**"
   - The "flows" of probabilities are balanced for $P(\Gamma)$.

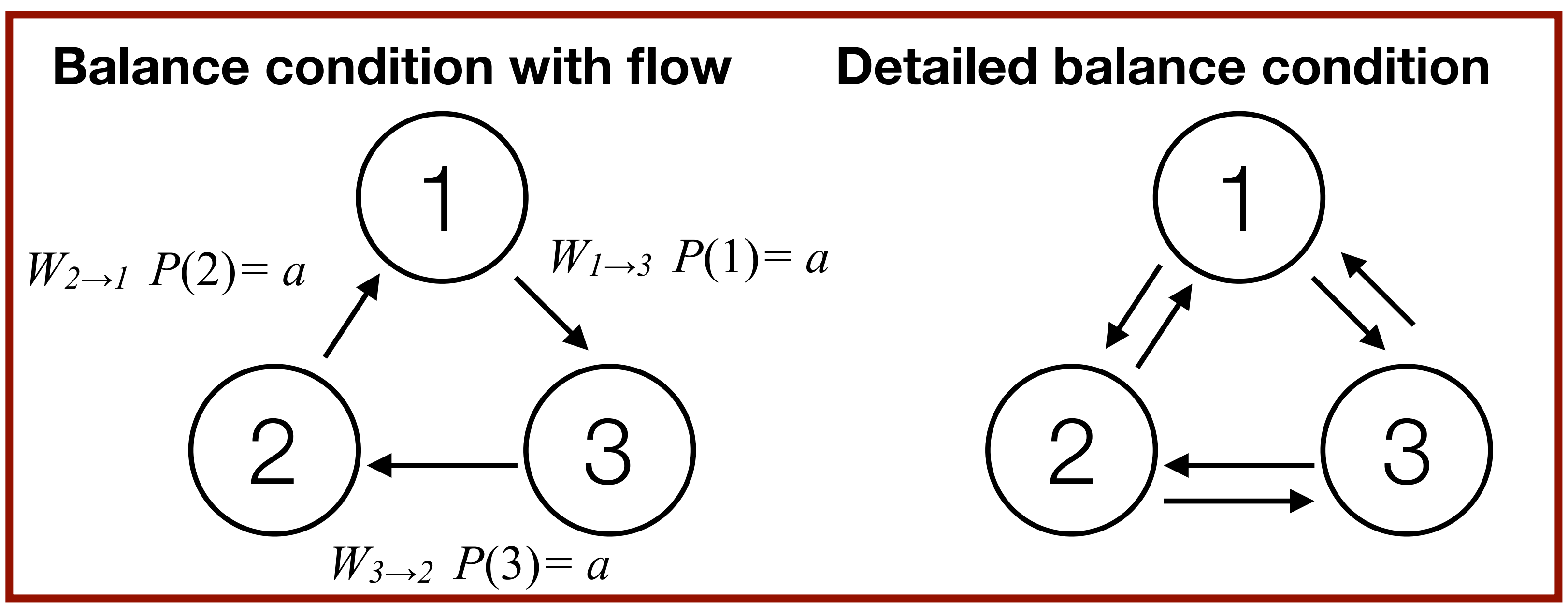
$$^{\forall}(\Gamma, \Gamma'), \quad \sum_{\Gamma'} W_{\Gamma\to\Gamma'} P(\Gamma) = \sum_{\Gamma'} W_{\Gamma'\to\Gamma} P(\Gamma')$$

Special case:

Detailed balance condition

$$W_{\Gamma\to\Gamma'} P(\Gamma) = W_{\Gamma'\to\Gamma} P(\Gamma')$$

No microscopic flow
in the steady state



Metropolis sampling

Example of transition probability satisfying detailed balance condition

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

* if $P(\Gamma') > P(\Gamma)$

$$W_{\Gamma\to\Gamma'} = 1$$
$$W_{\Gamma'\to\Gamma} = \frac{P(\Gamma)}{P(\Gamma')}$$

$\Rightarrow$ Satisfy the detailed balance condition

$$W_{\Gamma\to\Gamma'}P(\Gamma) = W_{\Gamma'\to\Gamma}P(\Gamma')$$

Algorithm based on Metropolis sampling

Step 0: Prepare an initial state $\Gamma_0 \in \{\Gamma\}$

loop *t*

1. Make next candidate state Γ' "randomly".
2. Calculate $P(\Gamma')/P(\Gamma_t)$
3. Make random number $r \in [0, 1]$
4. Select the next state Γ_{t+1} based on *r* as

$$\Gamma_{t+1} = \begin{cases} \Gamma', & r \leq P(\Gamma')/P(\Gamma) \\ \Gamma_t, & \text{otherwise} \end{cases}$$

$\Rightarrow$

$$I = \int d\Gamma f(\Gamma) P(\Gamma)$$
$$= \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} f(\Gamma_t)$$

Heat-bath sampling

Suppose we only change a part of variables in Γ

$$\Gamma = (\Gamma^1, \Gamma^2, \Gamma^3, \cdots, \Gamma^N) \Rightarrow \Gamma' = (\Gamma^{1\prime}, \Gamma^2, \Gamma^3, \cdots, \Gamma^N)$$

In this case, we may calculate "conditional" probability distribution of Γ^1

$$P(\Gamma^1 | \Gamma^2, \Gamma^3, \cdots \Gamma^N) = \frac{P(\Gamma)}{\int d\Gamma^1 P(\Gamma)}$$

Then we can chose a transition probability satisfying the detailed balance condition

$$W_{\Gamma \to \Gamma'} = P(\Gamma^{1\prime} | \Gamma^2, \Gamma^3, \cdots \Gamma^N)$$

We generate the next Γ' directly from the conditional probability!

- The transition probability is independent on the present Γ^1
- In general, it is not easy to produce the conditional probability distribution from uniform random numbers

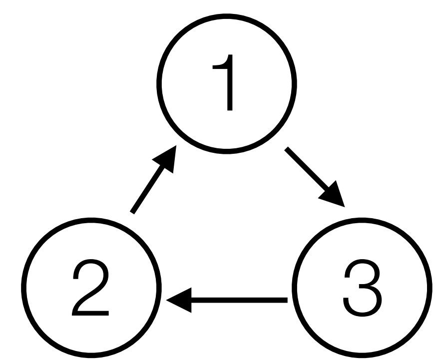
There is no general principle determining which of Metropolis and Heat-bath samplings are more efficient.

Sampling Based on balance condition

We do not necessarily use the detailed balance condition

By using more general balance condition,
we can make "rejection free" transition probabilities

$W_{\Gamma \to \Gamma} = 0$:The state necessarily changes to another state



eg. Suwa-Todo method

H. Suwa, and S. Todo, Phys. Rev. Lett. **105**, 120603 (2010).

"詳細釣り合いを満たさないモンテカルロ法"

諏訪秀麿, 藤堂眞治, 日本物理学会誌, **66**, 370 (2011).

Application to replica exchange Monte Carlo for molecular dynamics simulation

S. G. Itoh and H. Okumura, J. Chem. Theory Comput. **9**, 570 (2013).

Application to Classical spin system

Classical spin system

Model Hamiltonian

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i S_j - h \sum_i S_i^z$$

Nearest Neighbor interaction

e.g.

Ising spin: $S_i = \pm 1$

Heisenberg spin: $S_i = (S_i^x, S_i^y, S_i^z)$

Calculate expectation values under Canonical Ensemble

$$\langle \hat{O} \rangle = \frac{1}{Z} \int d\Gamma \hat{O}(\Gamma) e^{-\beta \mathcal{H}(\Gamma)}$$

e.g.

Energy: $\hat{E}(\Gamma) = \mathcal{H}$

Squared Energy: $\hat{E}^2(\Gamma) = (\mathcal{H})^2$

Heat capacity:

$$\frac{\langle \hat{E}^2 \rangle - \langle \hat{E} \rangle^2}{k_B T^2}$$

Squared Magnetization: $\hat{M}_z^2(\Gamma) = \left(\frac{1}{N} \sum_i S_i^z \right)^2$

MCMC method:

Target steady state is $P(\Gamma) = \frac{1}{Z} e^{-\beta \mathcal{H}(\Gamma)}$

$$\langle \hat{O} \rangle = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \hat{O}(\Gamma_t)$$

Γ_t: sampling points along Markov chain

Local update

Local update: We try to change a part of spins (typically single spin) at transitions along Markov chain

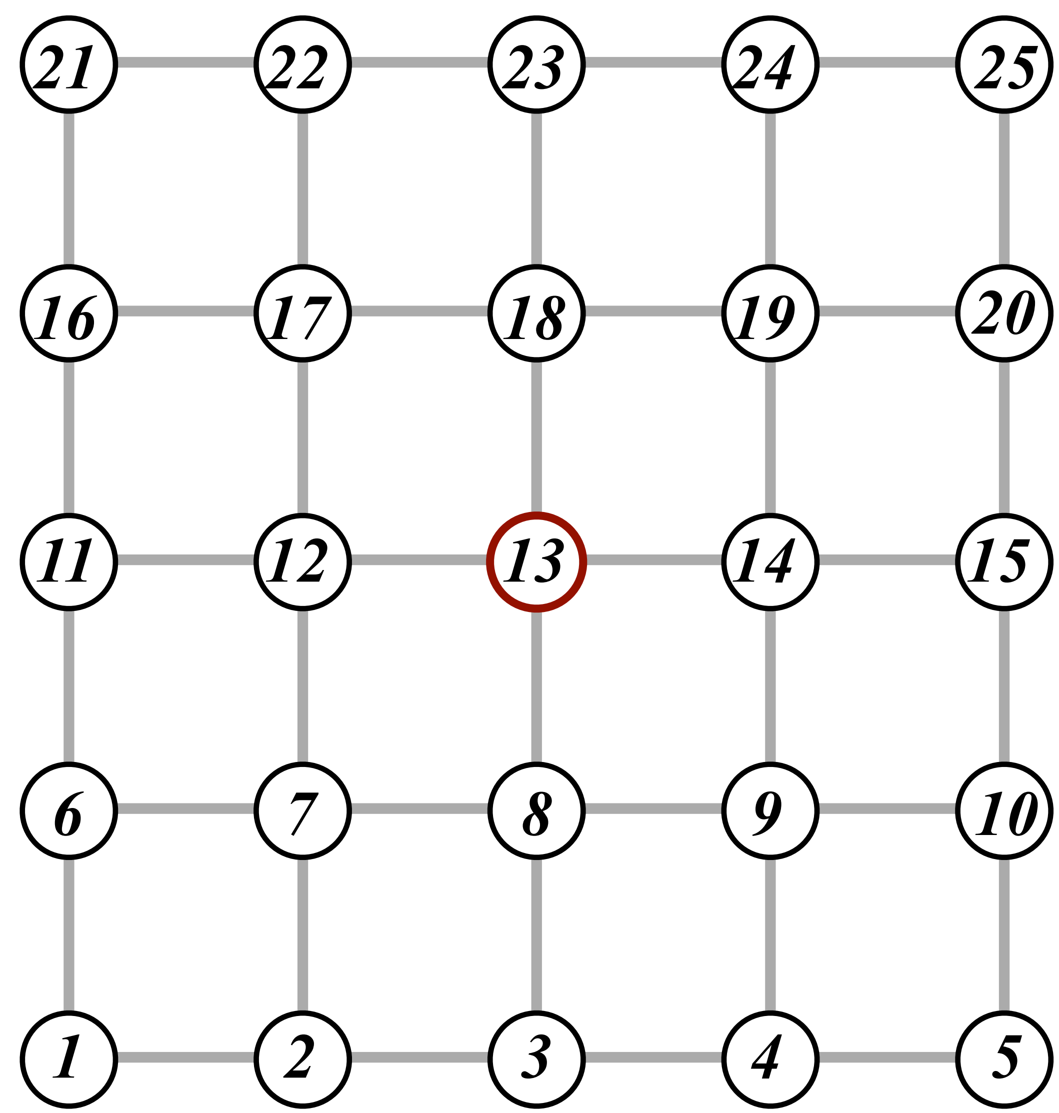
$$\Gamma = (S_1, S_2, \cdots, S_{13}, \cdots, S_{25})$$

$$\Rightarrow \Gamma' = (S_1, S_2, \cdots, S'_{13}, \cdots, S_{25})$$

From Γ to Γ', we fix $S_1, S_2, .. S_{12}, S_{14}, S_{15}, \ldots, S_{25}$, and try to change only S_{13}.

In this local update, we can easily estimate the transition probability W because the change of Hamiltonian (Energy) is determined only locally.

* If the Hamiltonian contains long range interactions, the energy estimation becomes more cost full.



Metropolis method:

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

$$\frac{P(\Gamma')}{P(\Gamma)} = e^{-\beta[\mathcal{H}(\Gamma')-\mathcal{H}(\Gamma)]} = e^{-\beta\Delta E}$$

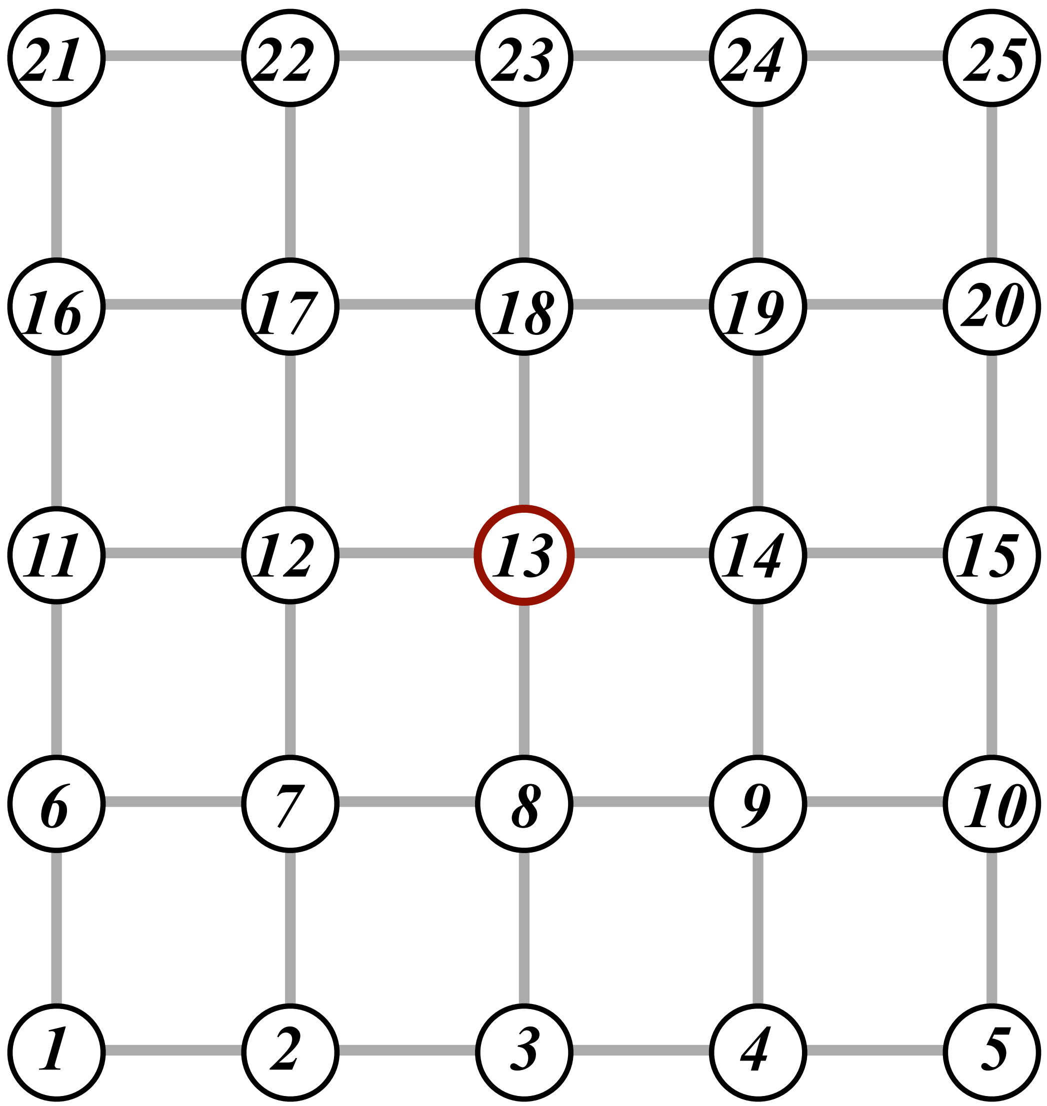
* We don't need partition function Z!

For local update on the square lattice,

$$\Gamma = (S_1, S_2, \cdots, S_{13}, \cdots, S_{25})$$

$$\Rightarrow \quad \Gamma' = (S_1, S_2, \cdots, S'_{13}, \cdots, S_{25})$$

$$\Delta E = -J(S_8 + S_{12} + S_{14} + S_{18})(S'_{13} - S_{13}) \\ - h[(S^z_{13})' - S^z_{13}]$$



Metropolis method with local update: summary

Step 0: Prepare an initial state $\Gamma_0 = (S_1, S_2, \cdots, S_N)$

loop t

select i-th site

1. Make next candidate state Γ' by changing S_i
   - Ising : $S_i' = -S_i$
   - XY, Heisenberg: $S_i' = S_i + \delta S$ or random unit vector
2. Calculate $\Delta E = \mathcal{H}(\Gamma') - \mathcal{H}(\Gamma)$
3. Make random number $r \in [0, 1]$
4. Select the next state Γ_{t+1} based on r as

$$\Gamma_{t+1} = \begin{cases} \Gamma' & r \leq e^{-\beta \Delta E} \\ \Gamma_t & \text{otherwise} \end{cases}$$

Calculate $O(\Gamma_t)$

Typically we choose
- random state ($T \to \infty$)
- ordered state ($T \to 0$)

If energy decreases ($\Delta E < 0$), we "accept" new state with probability 1.

It tends to sample low energy states.

Importance sampling in the canonical ensemble!

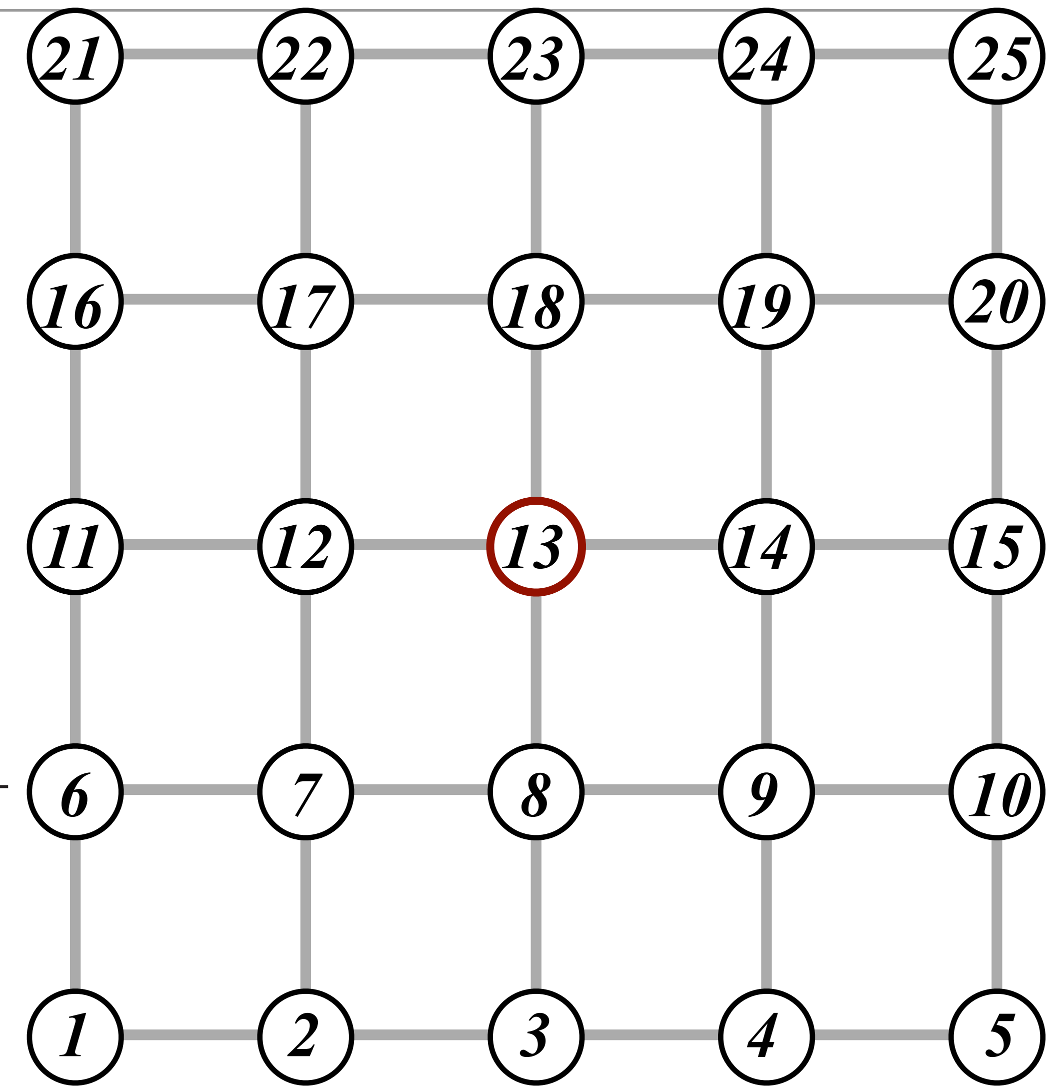
Usually, we observe quantities at least after N-spins are tried to change

Heat-bath method:

$$\Gamma = (S_1, S_2, \cdots, S_{13}, \cdots, S_{25})$$

$$\Gamma' = (S_1, S_2, \cdots, S'_{13}, \cdots, S_{25})$$

$$W_{\Gamma\to\Gamma'} = P(S'_{13}|S_1, S_2, \cdots S_{12}, S_{14}, \cdots, S_{25})$$
$$= \frac{P(\Gamma')}{\int dS_{13} P(\Gamma)}$$
$$= \frac{e^{\beta[J(S_8+S_{12}+S_{14}+S_{18})S'_{13}+h(S^z_{13})']}}{\int dS_{13} e^{\beta[J(S_8+S_{12}+S_{14}+S_{18})S'_{13}+h(S^z_{13})']}}$$



In the case of Ising or Heisenberg spins,
we can easily generate this probability distribution

Ising: $$W_{\Gamma\to\Gamma'} = \frac{e^{\beta h_{\rm eff} S'_{13}}}{e^{\beta h_{\rm eff} S'_{13}} + e^{-\beta h_{\rm eff} S'_{13}}}$$

Scaler value

$$h_{\rm eff} \equiv J(S_8 + S_{12} + S_{14} + S_{18}) + h$$

Heisenberg: $$W_{\Gamma\to\Gamma'} = \frac{(\beta|h_{\rm eff}|)e^{\beta h_{\rm eff}\cdot S'_{13}}}{[e^{\beta|h_{\rm eff}|} - e^{-\beta|h_{\rm eff}|}]}$$

Three component vector

$$h_{\rm eff} \equiv J(S_8 + S_{12} + S_{14} + S_{18}) + h\hat{e}_z$$

Heat-bath method with local update: summary

Step 0: Prepare an initial state $\Gamma_0 = (S_1, S_2, \cdots, S_N)$

loop t

select i-th site

1. Calculate effective field h_{eff}
2. Generate S'_i based on the probability

$$P(S'_i) \propto e^{\beta h_{\mathrm{eff}} S'_i}$$

(for Ising and Heisenberg spins, it can be generated from uniform random number)

3. The next state Γ_{t+1} is Γ'

Calculate $O(\Gamma_t)$

Typically we choose
- random state ($T \to \infty$)
- ordered state ($T \to 0$)

$r \in [0,1]$:uniform random number

Ising : $S'_i = \begin{cases} 1 & r \leq P(1) \\ -1 & \text{otherwise} \end{cases}$

Heisenberg:

(in polar co-ordinate with z // h_{eff})

$$S'_x = \sin\theta\cos\phi$$
$$S'_y = \sin\theta\sin\phi$$
$$S'_z = \cos\theta$$

$r_1, r_2 \in [0,1]$:uniform random number

$$\phi = 2\pi r_1$$
$$\cos\theta = -1 + \frac{1}{\beta|h_{\mathrm{eff}}|} \times \ln[r_2 + (1-r_2)e^{2\beta|h_{\mathrm{eff}}|}]$$

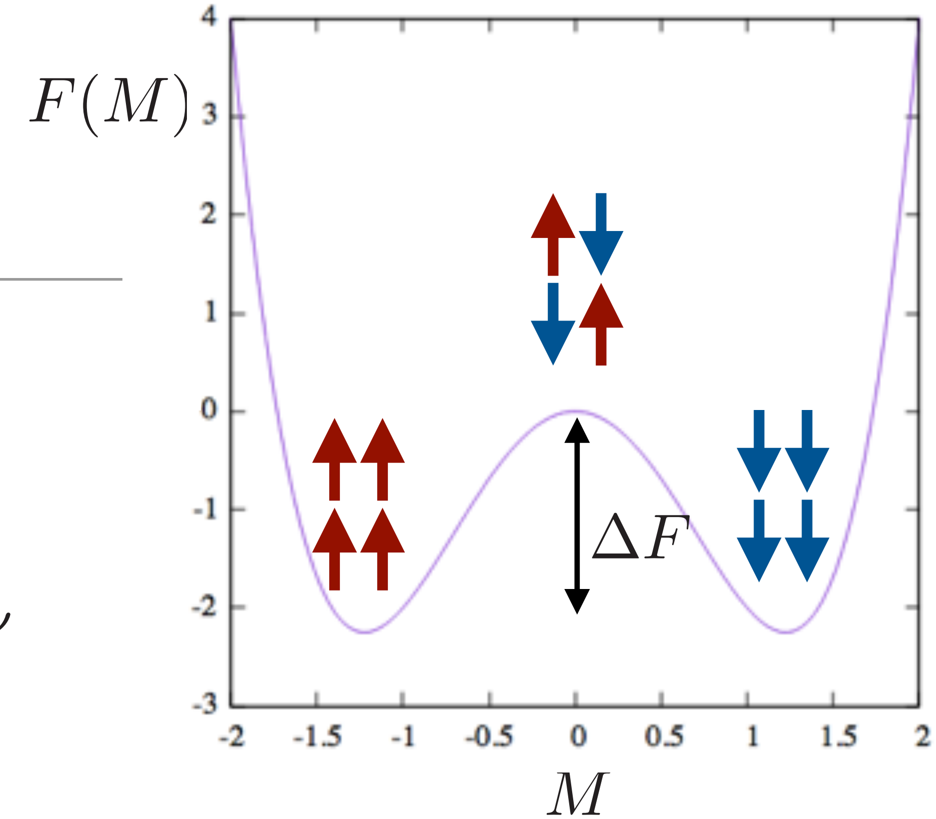
Usually, we observe quantities at least after N-spins are tried to change

Problems in local update

Sampling efficiency largely decreases for

1. Critical phenomena
   - Divergence of relaxation time: $\tau \propto |T - T_c|^{-z\nu}$
2. 1st order phase transition (phase coexistence)
3. Low temperature phase with discrete symmetry (e.g. Ising model)
   - Exponentially small probability to move other local minima: $\tau \propto \exp\left[\frac{\Delta F}{T}\right]$

Free energy landscape



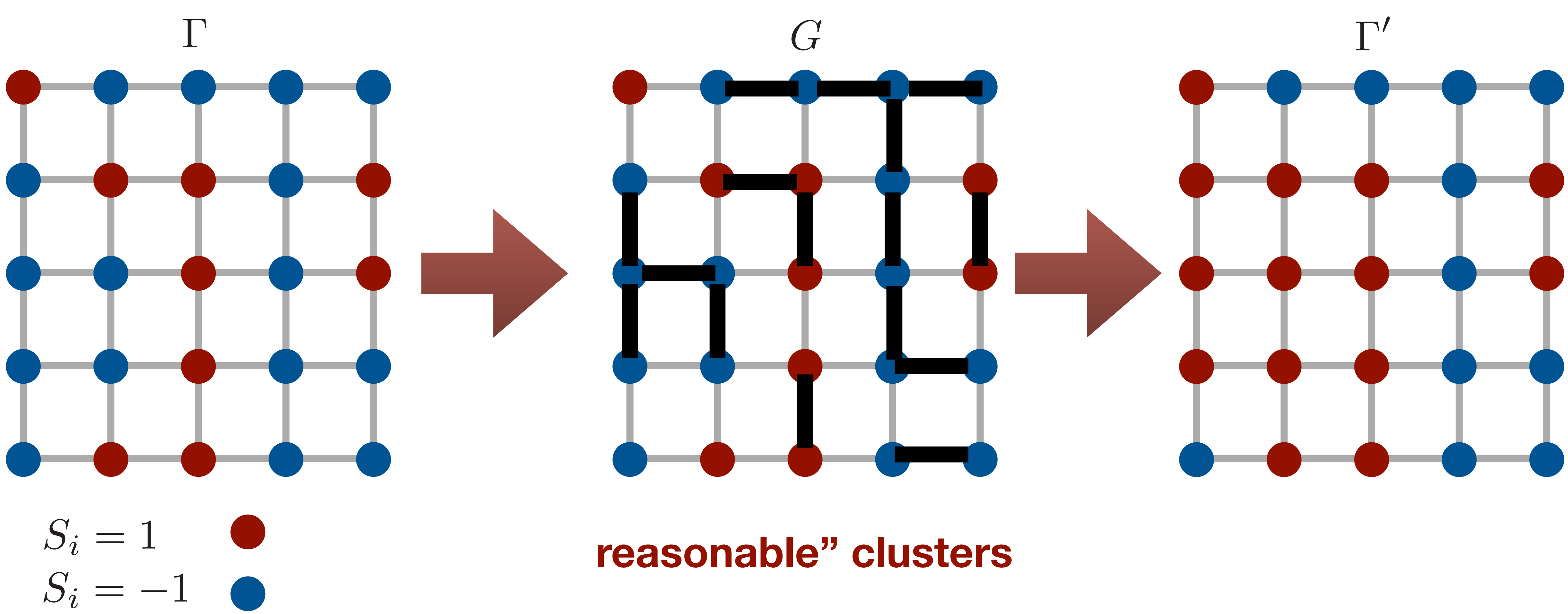
Part of these difficulties can be reduced by using "global update"

Simultaneous change of spins in "large cluster"

Cluster update method

Idea of cluster updates

- From a spin configuration Γ, we can define "reasonable" clusters G
- When we "flip" all spins on a cluster G and make new configuration Γ', the free energy difference between Γ and Γ' is not so large
- We can change the configuration drastically with higher probability



How to make a cluster configuration?

Fortuin-Kasteleyn mapping (for Ising model)

P. W. Kasteleyn and C. M. Fortuin, J. Phys. Soc. Jpn, Suppl. **26**, 11 (1969)

C. M. Fortuin and P. W. Kasteleyn, Physica **57**, 536 (1972)

Ising model

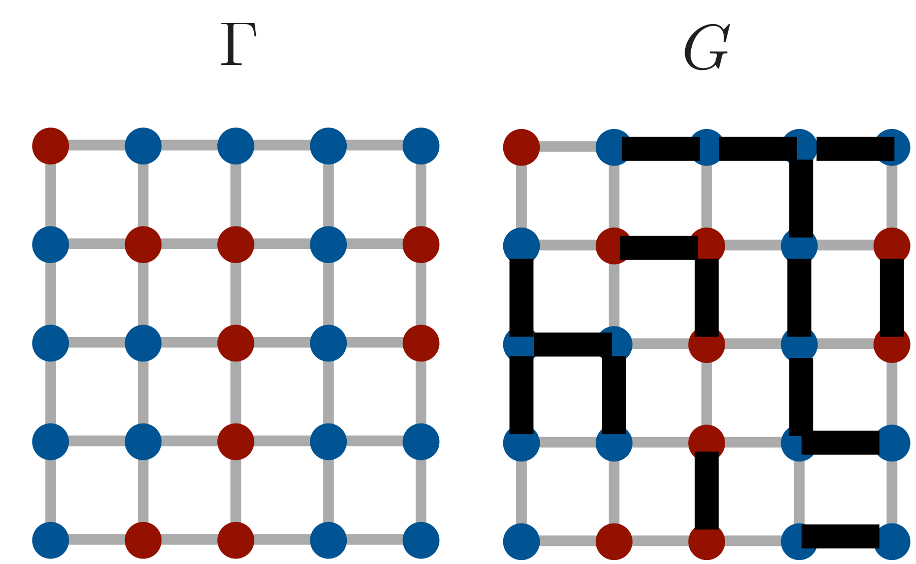
$$\boxed{\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i S_j \ , S_i = \pm 1}$$

$$e^{\beta J S_i S_j} = \underbrace{\boxed{e^{-\beta J}}}_{g=0} + \underbrace{\boxed{\delta_{S_i,S_j}(e^{\beta J} - e^{-\beta J})}}_{g=1} = \sum_{g=0,1} w(g, S_i, S_j)$$

$$Z = \sum_{\Gamma} e^{\beta J \sum_{\langle i,j \rangle} S_i S_j}$$

$$= \sum_{G} \sum_{\Gamma} \prod_{\langle i,j \rangle} w(g_{i,j}, S_i, S_j)$$

$G = \{g_{i,j}\}$ $\Gamma = \{S_i\}$



$S_i = 1$ (red) $g_{i,j} = 1$ (black)

$S_i = -1$ (blue) $g_{i,j} = 0$ (grey)

Markov chain in extended (G, Γ) space

$$Z = \sum_G \sum_\Gamma \prod_{\langle i,j \rangle} w(g_{i,j}, S_i, S_j) = \sum_G \sum_\Gamma W(G, \Gamma)$$

$$\cdots \to \Gamma_t \to G_t \to \Gamma_{t+1} \to G_{t+1} \to \cdots$$

$$W(\Gamma) = \sum_G W(G, \Gamma)$$

$$W(G) = \sum_\Gamma W(G, \Gamma)$$

Transition probabilities

$$W_{\Gamma \to G} = \frac{W(G, \Gamma)}{W(\Gamma)}, W_{G \to \Gamma} = \frac{W(G, \Gamma)}{W(G)}$$

$$= \prod_{\langle i,j \rangle} w_{(S_i, S_j) \to g_{ij}} \quad = \prod_{C_j} P(\{S_i \in C_j\})$$

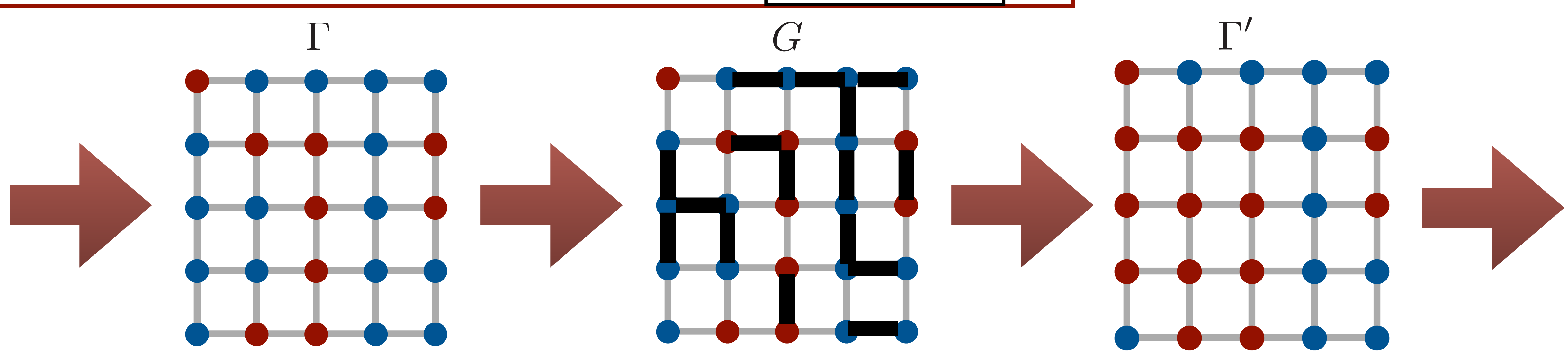
cluster formed from g=1 links

$$w_{(S_i, S_j) \to 0}$$

$$= \begin{cases} 1 & (S_i \neq S_j) \\ e^{-2\beta J} & (S_i = S_j) \end{cases}$$

$$P(\{S_i \in C_j\}) = 1$$

(If all spin in cluster is pointing same direction)



Swendsen-Wang algorithm

Swendsen-Wang algorithm

R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987)

Step 0: Prepare an initial state $\underline{\Gamma_0 = (S_1, S_2, \cdots, S_N)}$

loop *t*

loop <i,j>

- if $S_i = S_j$, generate a random number
  - if $r \leq 1 - e^{-2\beta J}$ connects i and j ($g_{ij}=1$)

end loop <i,j>

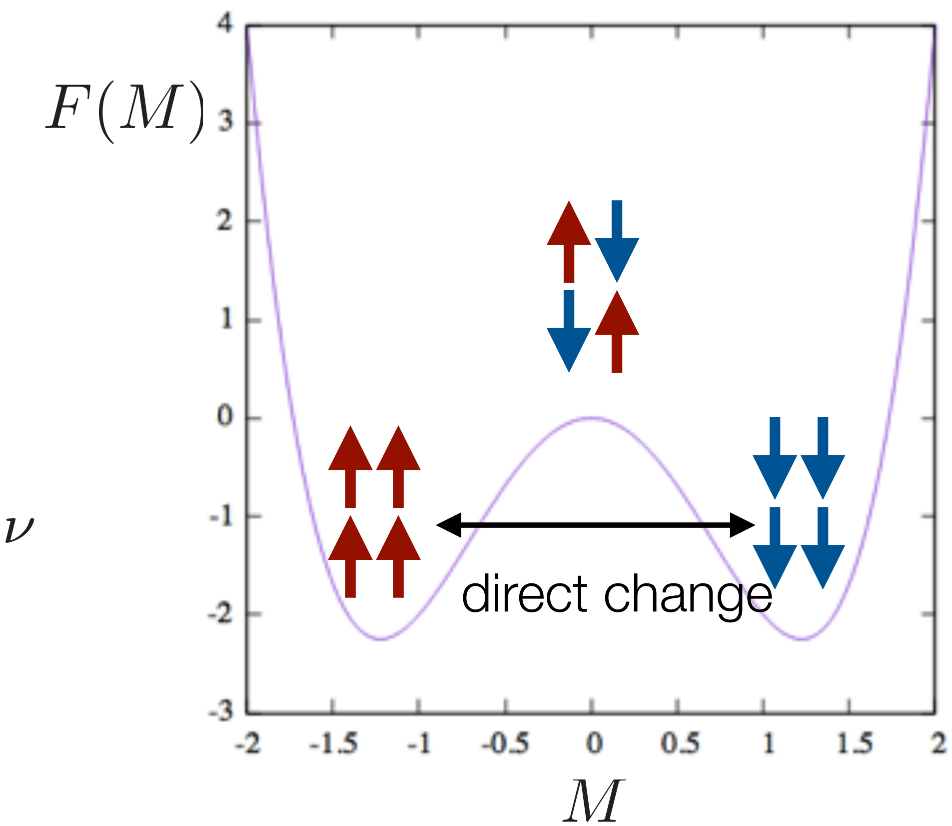
- Make clusters using algorithms (e.g. union find)
- Change spins on the same clusters simultaneously with probability 1/2 (using random number)

Calculate $O(\Gamma_t)$

Merit of cluster update

1. For low temperature phase, the system easily transit other minima
    - Minima are related to the symmetry of the Hamiltonian
2. For critical phenomena “the dynamical critical exponent become much smaller
    - Swendsen-Wang : $z \simeq 0 \quad \tau \propto |T - T_c|^{-z\nu}$
3. Graph representation closely related to physics
    - e.g. Magnetic susceptibility: $\chi = \beta \langle |C| \rangle$
    - By using observable based on graph, statistical error is largely reduced “Improved estimator”

Free energy landscape



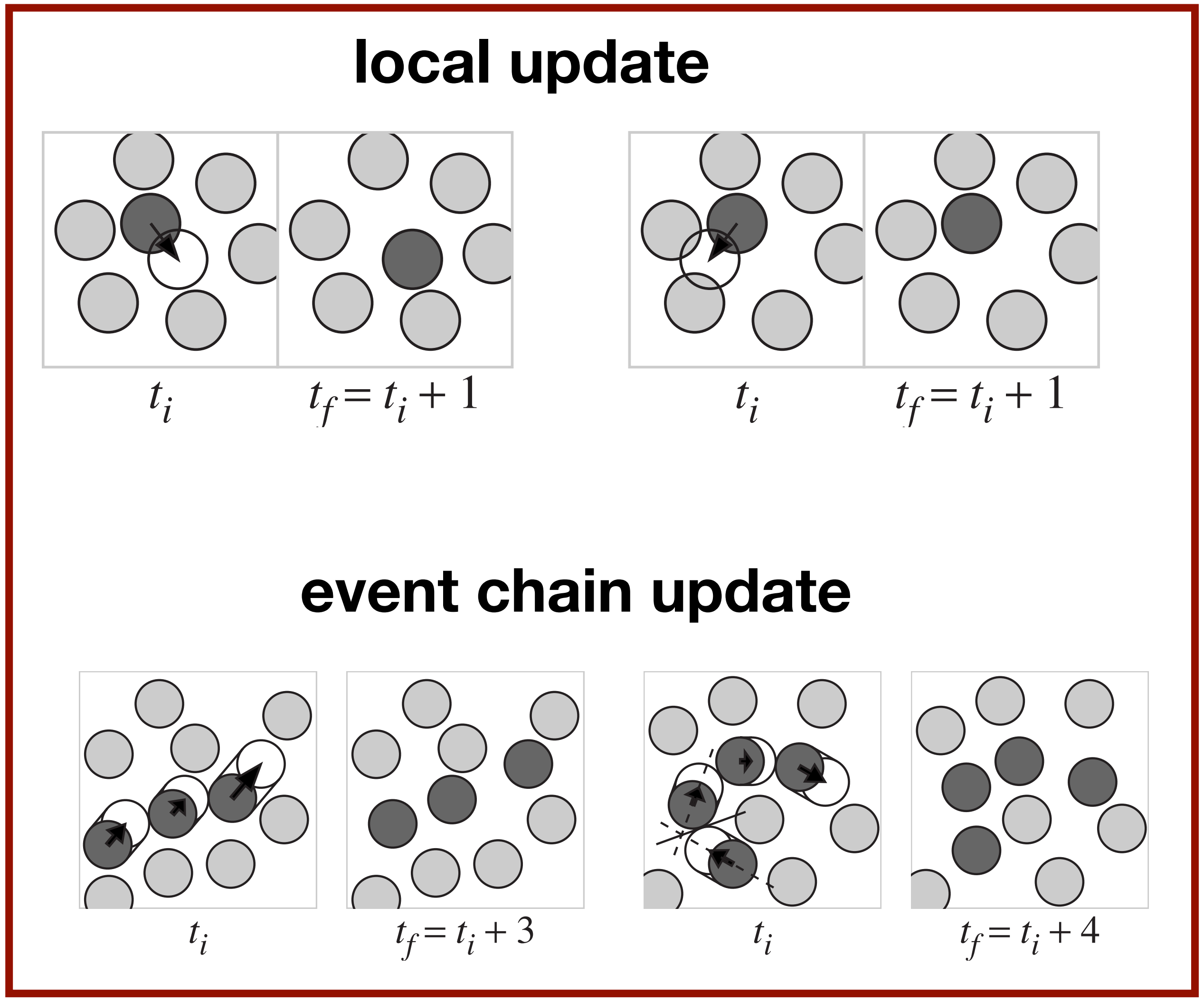
$|C|$:Cluster size

*Linear size of cluster ~ ξ

Event-chain Monte Carlo

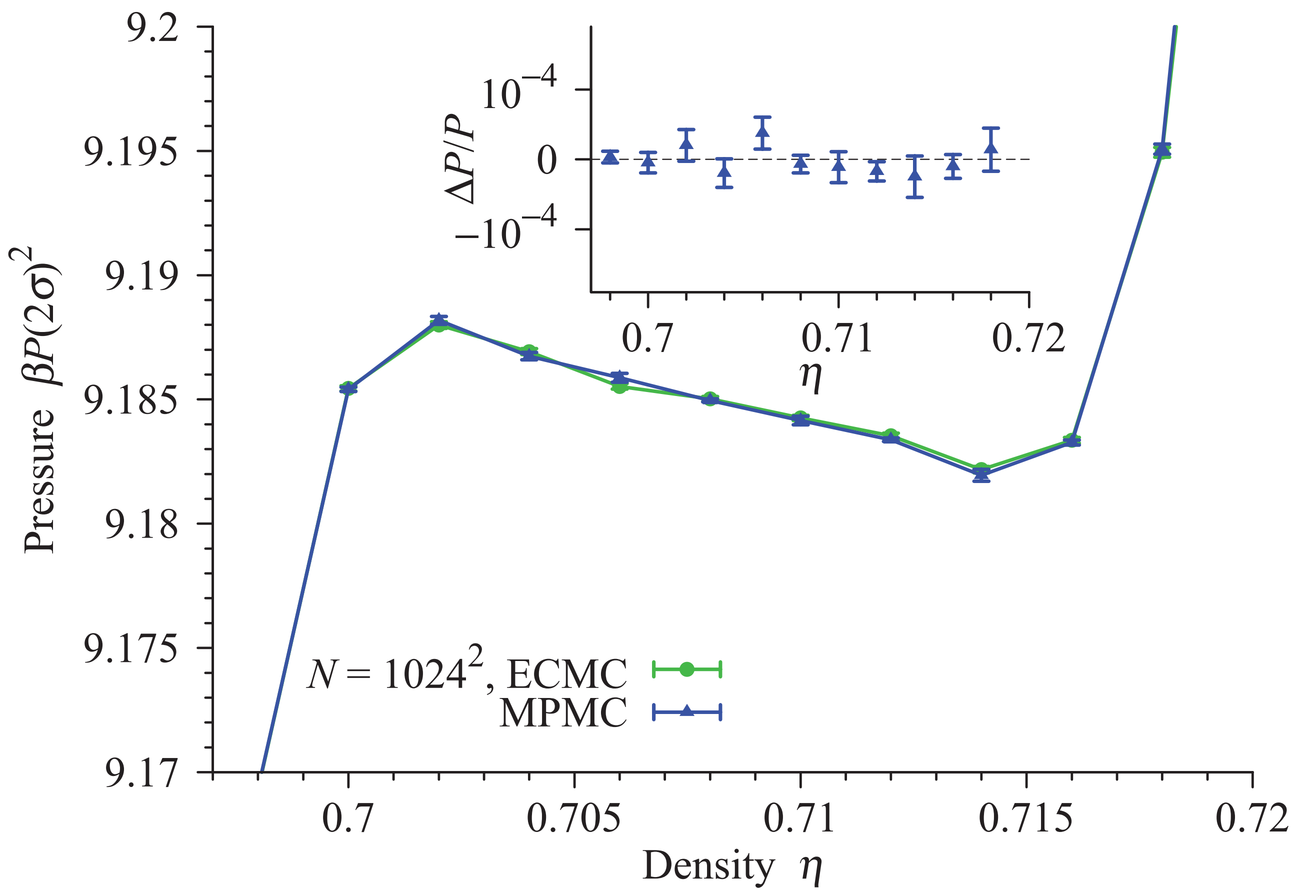
A global update for particle system (hard spheres)

E. P. Bernard, W. Krauth, and D. B. Wilson, Phys. Rev. E **80**, 056704 (2009)



Application to 2d melting

M. Engel *et al,* Phys. Rev. E **87**, 042134 (2013)



Computational science using MC methods

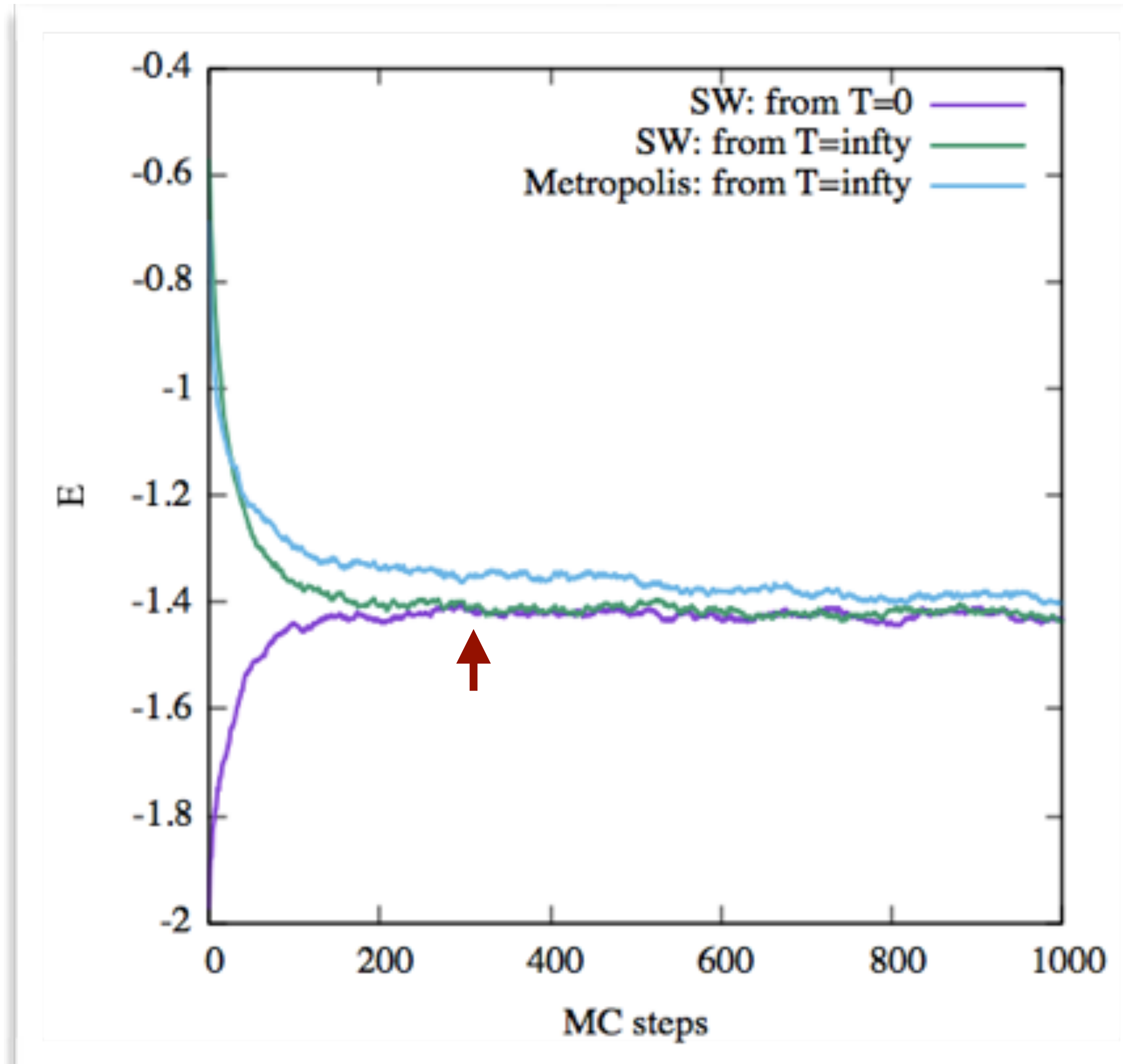
Important tips for real calculations 1

In each calculations, we have to check the convergence

If the correlation time is very long, obtained data (expectation values) might be biased from the initial state Γ_0

Usual procedure:

- Discard initial several MC steps
- Change MC steps and compare results
- Change Initial state
- ….



Important tips for real calculations 2

We need to estimate the statistical errors

$$\bar{A} \equiv \frac{1}{T}\sum_{t=1}^{T}\hat{A}(\Gamma(t))$$

Standard error: $\epsilon^2 = \langle \bar{A}^2 \rangle - \langle \bar{A} \rangle^2 \qquad \epsilon \propto \sqrt{\frac{\tau}{T}}$

Maximum likelihood estimation for standard error

Prepare "independent" M samples for $\bar{A}$: $\{\bar{A}_1, \bar{A}_2, \cdots, \bar{A}_M\}$

$$\sigma^2(M) = \frac{\frac{1}{M}\sum_i \bar{A}_i^2 - \left(\frac{1}{M}\sum_i \bar{A}_i\right)^2}{M-1}$$

$$\lim_{M\to\infty} \sigma^2(M) = \epsilon^2$$

Make "error bar" based on σ, and use it for data analysis

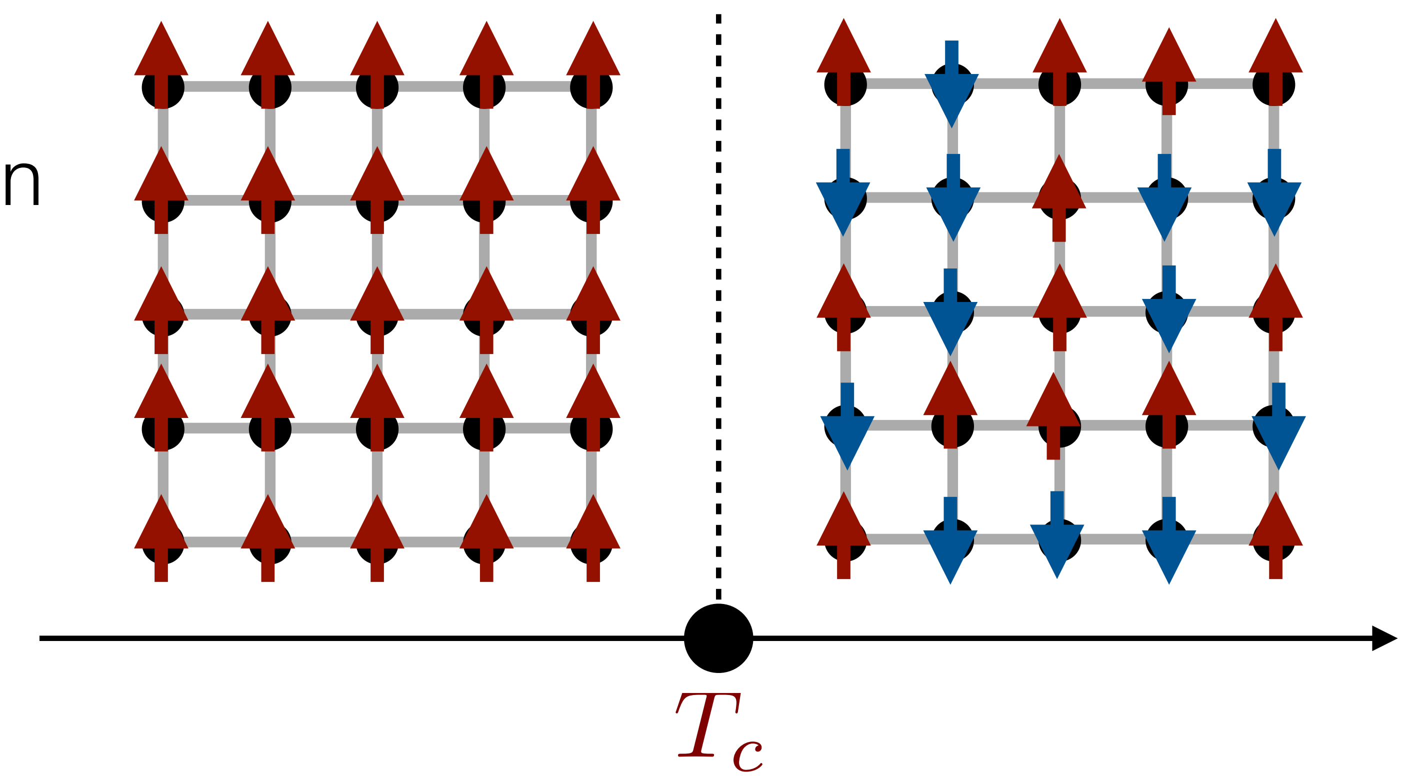
Example: Application for critical phenomena

- Square lattice Ising model

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i S_j$$

  - Continuous phase transition at $T=T_c$

$$T_c/J = \frac{2}{\ln(1+\sqrt{2})} = 2.26918531\ldots$$

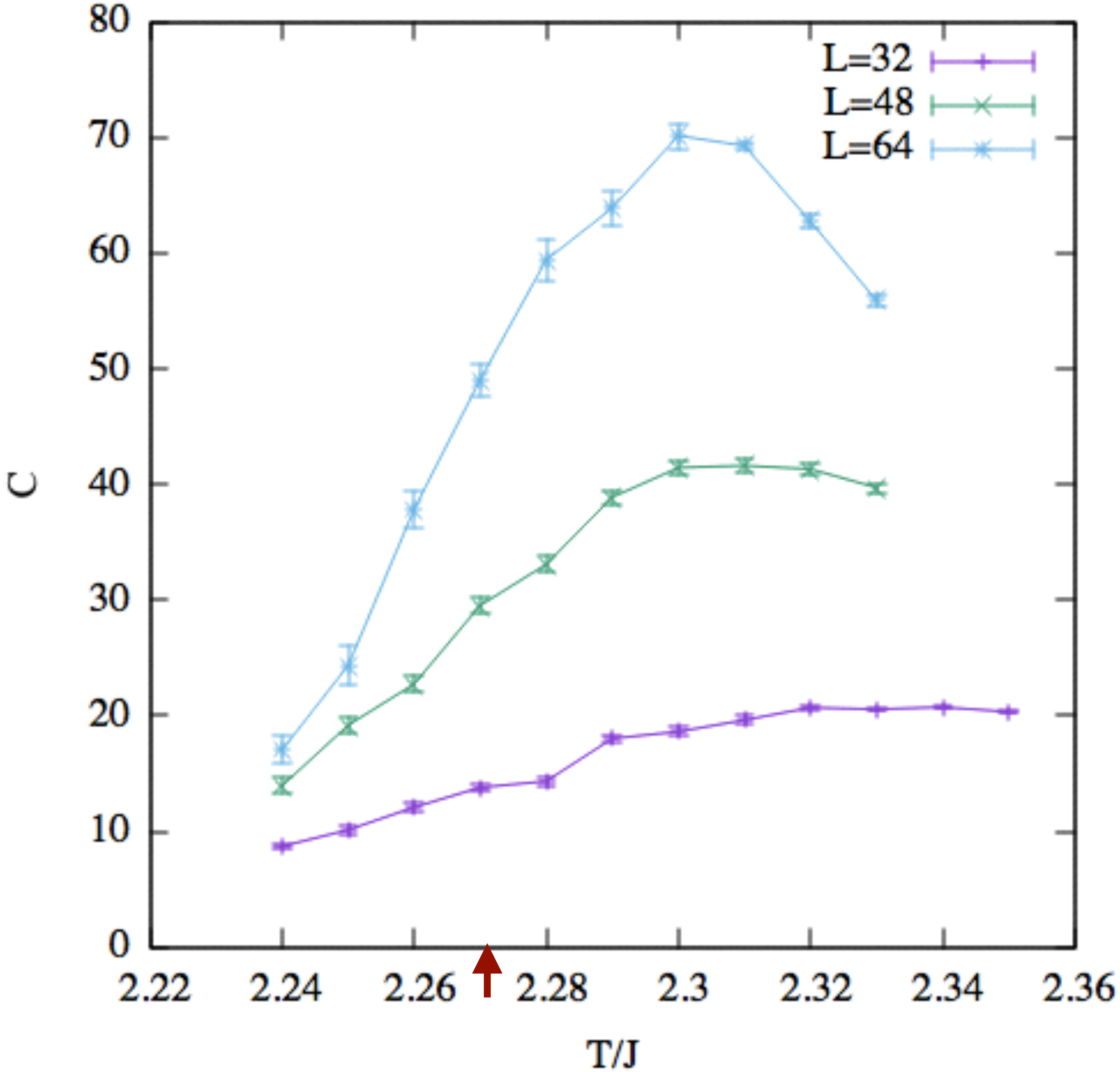


    - $T>T_c$: Paramagnetic
    - $T<T_c$: Ferromagnetic

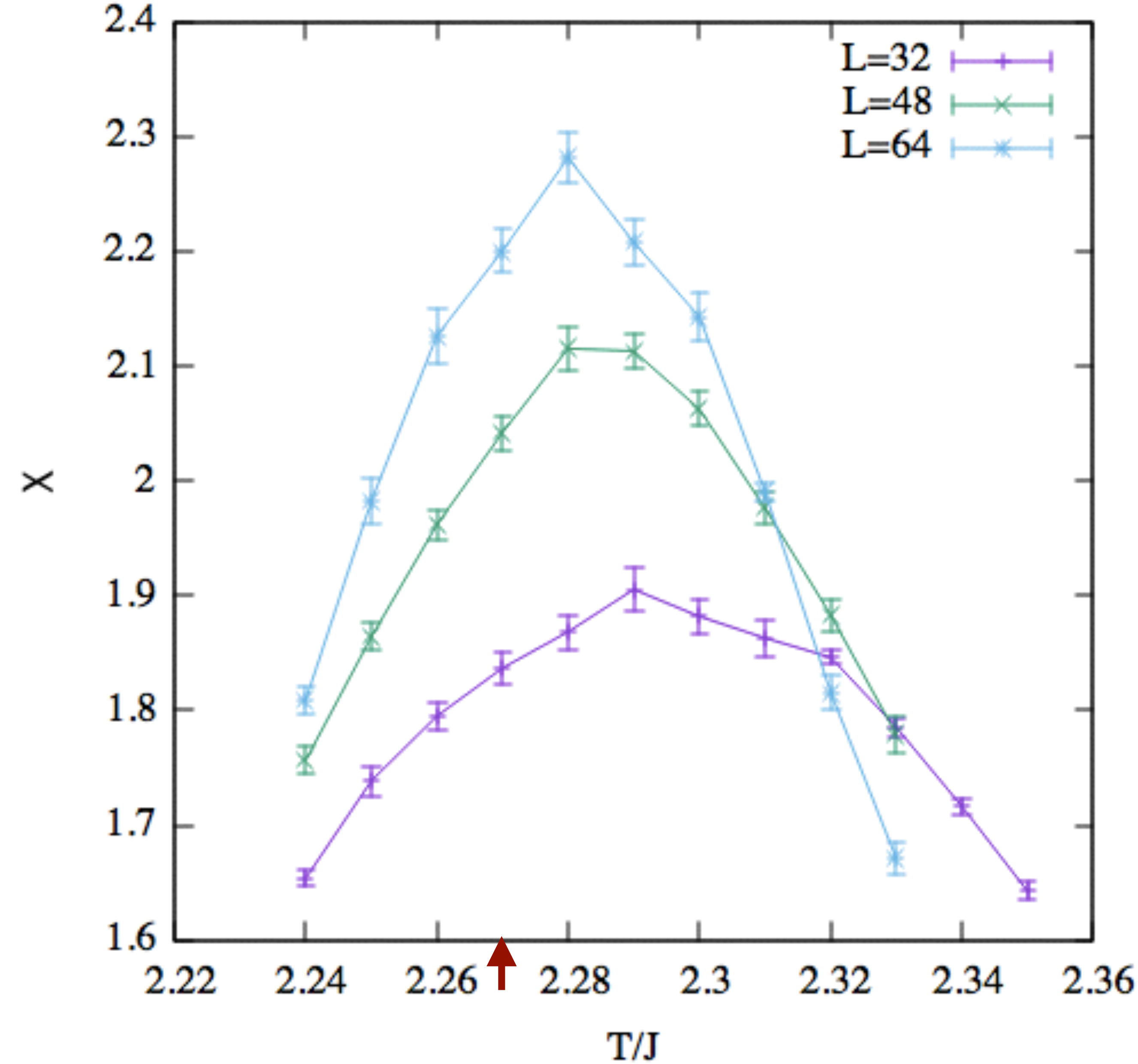
  - Monte Carlo Simulation using ALPS (spinmc)
    - spinmc: Simulator for classical spin system by MCMC

Calculated data (ALPS tutorial 7b)

Specific heat



Susceptibility



$$T_c/J \simeq 2.269$$

Binder ratio

Binder ratio

$$b = \frac{\langle M^4 \rangle}{\langle M^2 \rangle^2}$$

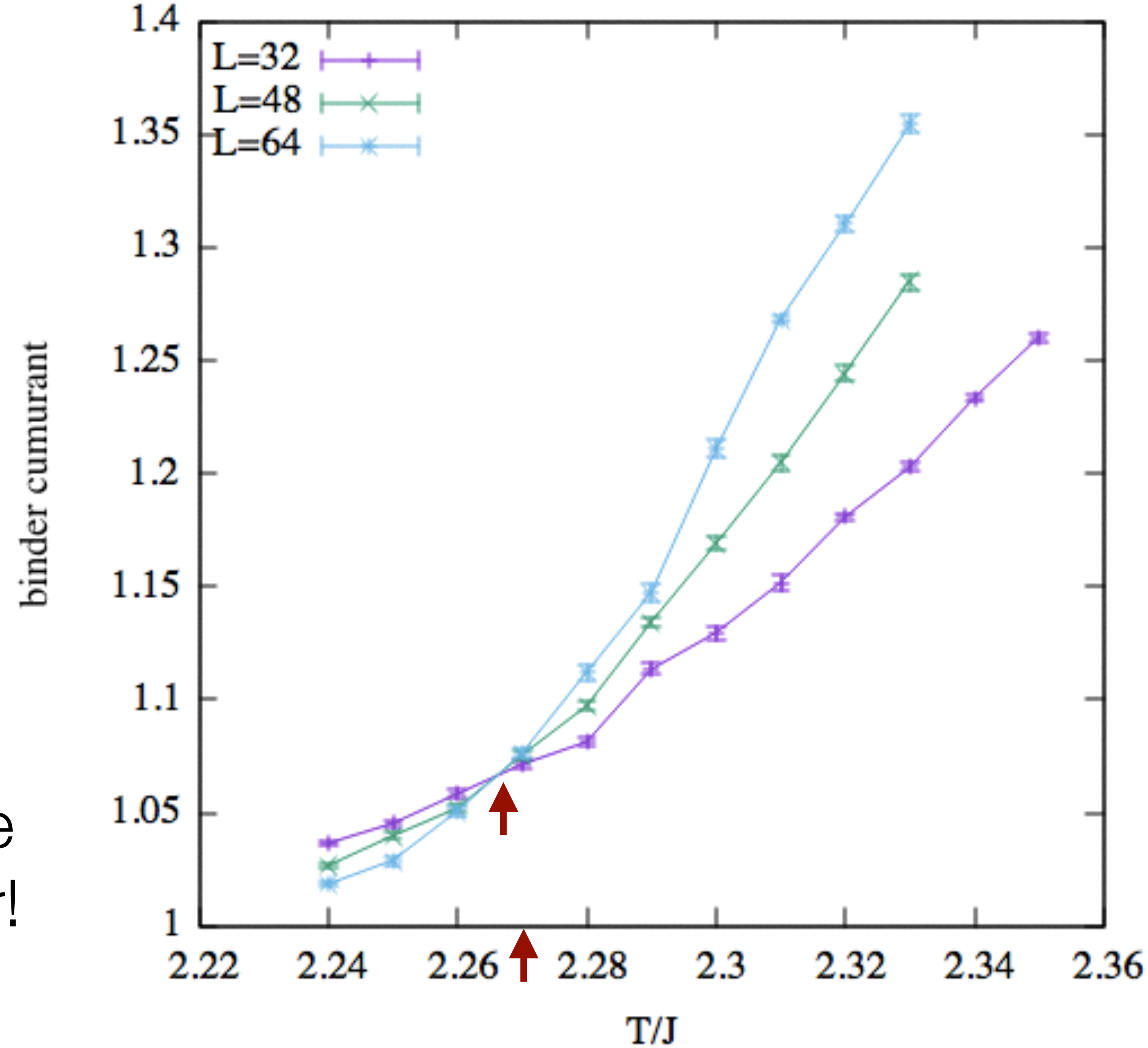
$$b = 3 \quad (T \to \infty)$$

$$b = 1 \quad (T \to 0)$$

The scaling dimension of b is exactly zero

At Tc, the size dependence disappears in leading order!

Binder ratio



$$T_c/J \simeq 2.269$$

Finite size scaling

Binder ratio

$$b = \frac{\langle M^4 \rangle}{\langle M^2 \rangle^2}$$

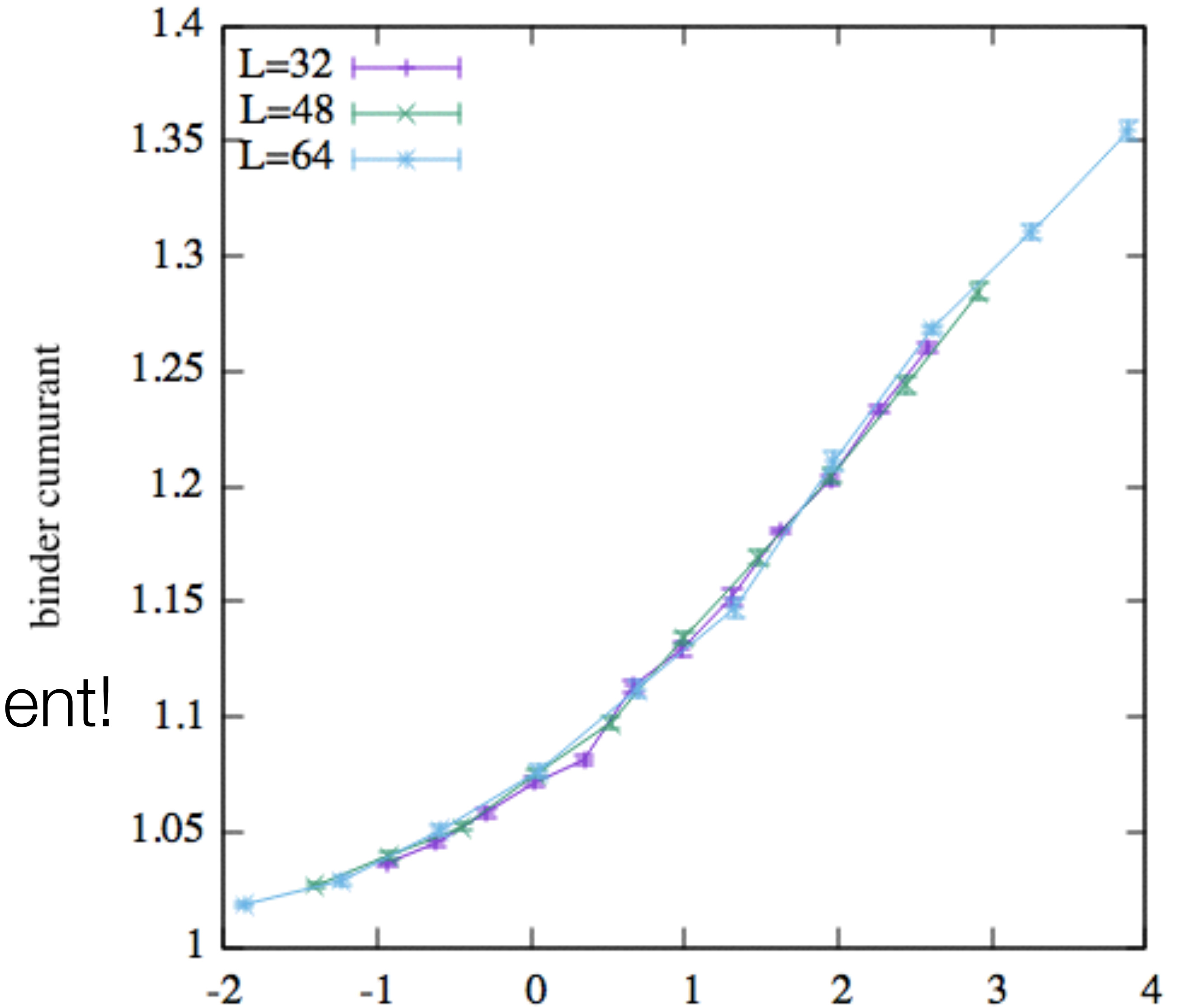
finite size scaling around Tc

$$b = f((T - T_c)L^{1/\nu})$$

We can determine critical exponent!

$$\nu = 1$$

Binder ratio



$$(T - T_c)L^{1/\nu}$$