

2024年3月15日 計算物理春の学校2024

モンテカルロ法入門から 効率的MCMCへ

諏訪秀磨

東京大学 理学系研究科 物理学専攻

研究紹介

- 数値計算手法開発
 - モンテカルロ法
 - 不可逆モンテカルロ (講義後半)
 - 量子モンテカルロ (量子スピン系、スペクトル計算)
 - 長距離相互作用系
 - 確率近似法
- 相関電子系に対する準古典ダイナミクス計算法
- 機械学習 (ニューラルネットワークポテンシャル)
- 量子古典ハイブリッドアルゴリズム
- フラストレート磁性体
 - 反強磁性スピネル (パイロクロア系、カゴメ系、etc)
- スピン軌道絶縁体 (スピン軌道結合と電子相関の協調現象)
 - 励起子絶縁体
 - トポロジカル磁性体
- スピンダイナミクス (準安定状態、非平衡現象)



諏訪 秀麿

講義内容

- ・ 講義資料：<https://github.com/suwamaro/Introduction-to-Monte-Carlo>
- ・ イントロダクション
 - モンテカルロ積分からマルコフ連鎖へ
- ・ 統計誤差の解析
 - 自己相関時間、漸近分散、Binning
- ・ クラスターアルゴリズム
 - Swendsen-Wang algorithm、Wolff algorithm
- ・ 相転移解析
 - 有限サイズスケーリング、ブートストラップ、確率近似法
- ・ ハイブリッド（ハミルトニアン）モンテカルロ
- ・ アニーリング（最適化手法）
- ・ 拡張アンサンブル法
 - シュミレーテッドテンパリング、レプリカ交換法、マルチカノニカル法
- ・ 不可逆モンテカルロ（確率の流れを操作するアプローチ）
 - リフティング、確率最適化

モンテカルロ法

- 「乱数を使って確率的に問題を解く」

乱数：でたらめな数、コンピュータで生成可能（擬似乱数）

- 最初の論文：**Nicholas Metropolis and S. Ulam, The Monte Carlo Method, J. Amer. Stat. Assoc. 44, 335-341 (1949).** 核物質中の中性子の拡散問題など
[S. Ulam, R. D. Richtmyer, and J. von Neumann, Los Alamos Scientific Laboratory. (Report LAMS-551) (1947).]
- アメリカ応用数学会[Society for Industrial and Applied Mathematics (SIAM)]による
20世紀の10大アルゴリズムのひとつに選出
(他は、線形計画法、クイックソート、高速フーリエ変換等)
- 愚直に計算すると絶対終わらないような問題に対して、サンプリングを用いることで
近似解をすばやく求める。
- 名前の由来は **Monte Carlo Casino in Monaco**

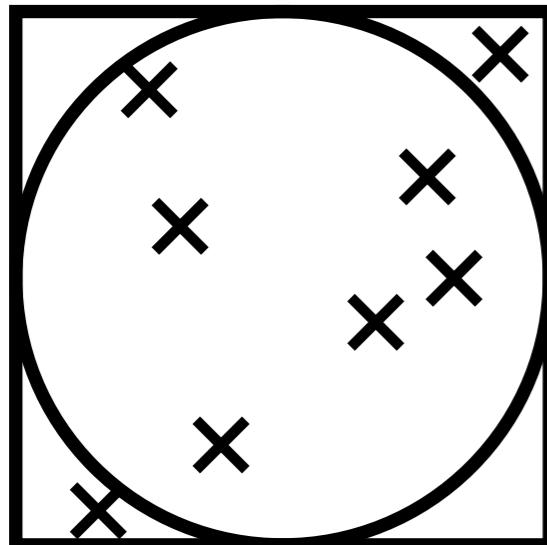
モンテカルロいろいろ

- 亂托アルゴリズム：トレースの計算 ($\text{tr}[H^M]$) など。賢くさぼる。
- モンテカルロ積分 (Monte Carlo integration)
低次元積分はO.K.だが、高次元積分は「次元の呪い」にかかる。
点列を工夫する準モンテカルロなど
- マルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo)
任意の確率分布からのサンプリング：物理現象のシミュレーション、多体問題計算、ベイズ推定（事後分布）、変分モンテカルロ
最適化（アニーリング）
- ポピュレーション型モンテカルロ法 (Population Monte Carlo)
サンプル分布からの再重み付けを行う手法：逐次モンテカルロ法 (Sequential Monte Carlo)、グリーン関数モンテカルロ、拡散モンテカルロ、射影モンテカルロ

伊庭幸人他、マルコフ連鎖モンテカルロ法とその周辺、岩波書店

モンテカルロ積分

- 円の面積の計算：一边の長さ2の正方形からランダムに点をサンプリング



円の面積 \approx 円の内部に入った点の割合 $\times 4$

サンプル数	平均値	誤差
10^2	3.08	0.062
10^4	3.133	0.0086
10^6	3.1424	0.00081
10^8	3.1419	0.00031

中心極限定理：
平均値の誤差はサンプル数 (M) に対して

$$\propto \frac{1}{\sqrt{M}}$$

1日やって全然ダメだったら3ヶ月は
計算しないと結果は良くならない！

次元の呪い

- 超立方体中の超球の占める割合は次元に対して指数関数的に小さくなる。

サンプル数 $M = 10^6$

次元	平均値	誤差
4	3.1404	0.001
8	3.147	0.006
12	3.125	0.016
16	3.184	0.043
20	0	3.14

半径 r の n 次元超球の体積

$$V_n(r) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} r^n$$

超球が占める割合

$$R_n = \frac{V_n(1)}{2^n}$$

次元	R _n
2	0.785
3	0.524
4	0.308

$\langle R_n \rangle$: 超球内のサンプル点の割合

$$\pi \simeq \left(2^n \Gamma \left(\frac{n}{2} + 1 \right) \langle R_n \rangle \right)^{2/n}$$

サンプルの確率分布と欲しい分布のズレが次第に大きくなり、
必要なサンプル数が次元に対して指数関数的に大きくなる。

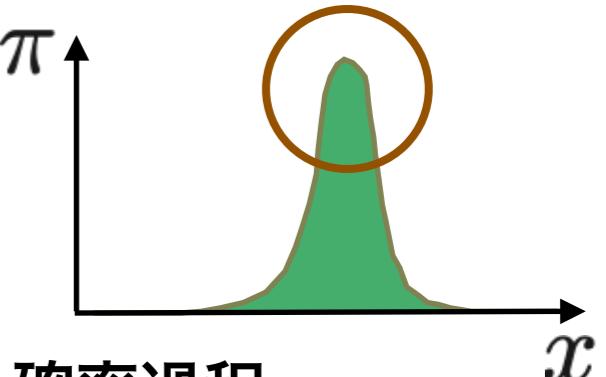
分布のズレを小さくする方法 = 重点サンプリング (importance sampling)

マルコフ連鎖モンテカルロ

- 任意の分布からサンプルを発生させる方法：Markov chain Monte Carlo (MCMC)

マルコフ連鎖

$$x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_k \rightarrow x_{k+1} \rightarrow \dots$$



マルコフ過程：前の状態（有限の過去）から次の状態が決まる確率過程

x_k ：状態（例えば粒子の位置）

π ：欲しい分布

簡単のため離散状態空間を考える

π_i ：状態 i の重み（分布の測度）

P_{ji} ：遷移行列（状態 j から i への遷移確率）

欲しい分布が不变分布となる条件（つりあい条件） = 固有値1の左固有ベクトル

$$\pi_i = \sum_j \pi_j P_{ji}$$

$$\pi = \pi P$$

$$\pi = (\pi_1 \ \pi_2 \ \dots \ \pi_n)$$

欲しい分布

- ボルツマン分布 (平衡状態の物理)

$$\pi = \frac{e^{-\beta H}}{Z}$$

逆温度 : $\beta = \frac{1}{k_B T}$

規格化定数 (分配関数) : $Z = \text{tr} e^{-\beta H}$

- ベイズ推定における事後確率分布

$$p(x, y) = p(y|x)p(x) = p(x|y)p(y)$$

$$\begin{aligned} p(x|y) &= \frac{p(y|x)p(x)}{p(y)} \\ &= \frac{p(y|x)p(x)}{\int dx p(y|x)p(x)} \end{aligned}$$

θ : 未知パラメータ、潜在変数

y : 観測値、データ

$$p(\theta|y) \sim p(y|\theta)p(\theta)$$

$p(\theta)$: 事前確率 (prior probability)、 θ に関する事前知識

$p(y|\theta)$: 尤度 (likelihood)、観測値の尤もらしさ

$p(\theta|y)$: 事後確率 (posterior probability)、観測結果を踏まえた θ の確率分布

MCMCの誕生とメトロポリス法

Metropolis, Rosenbluth, Rosenbluth, Teller, Teller, JCP 21, 1087 (1953)

THE JOURNAL OF CHEMICAL PHYSICS

VOLUME 21, NUMBER 6

JUNE, 1953

Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

AND

EDWARD TELLER,* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.

メトロポリス法
(メトロポリス・ヘイスティングス法)

Metropolis et al. (1953), Hastings (1970)

熱浴法 (ギブスサンプラー)

Baker (1965), Geman and Geman (1984)

$$p_{ij} = \min \left(1, \frac{\pi_j}{\pi_i} \right)$$

$$p_{ij} = \frac{\pi_j}{\sum_k \pi_k}$$

詳細つりあい条件 : $\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i, j$

- 確率分布の規格化定数は必要なく、重みの比から遷移確率を決められる。

MCMCの条件

- 次の2つの条件を課す

• S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer (1993)
• C. P. Robert and G. Casella, *Monte Carlo statistical methods*, Springer (2004)
• 鎌谷研吾他、モンテカルロ統計計算、講談社

全体つりあい (Global balance) : 欲しい分布が不变分布

$$\pi_i = \sum_j \pi_i p_{ij} = \sum_j \pi_j p_{ji} \quad \forall i$$

エルゴード性 (Ergodicity) : どんな初期分布からでも分布が収束 (離散変数では既約+非周期、ペロン・フロベニウスの定理 (Perron-Frobenius))

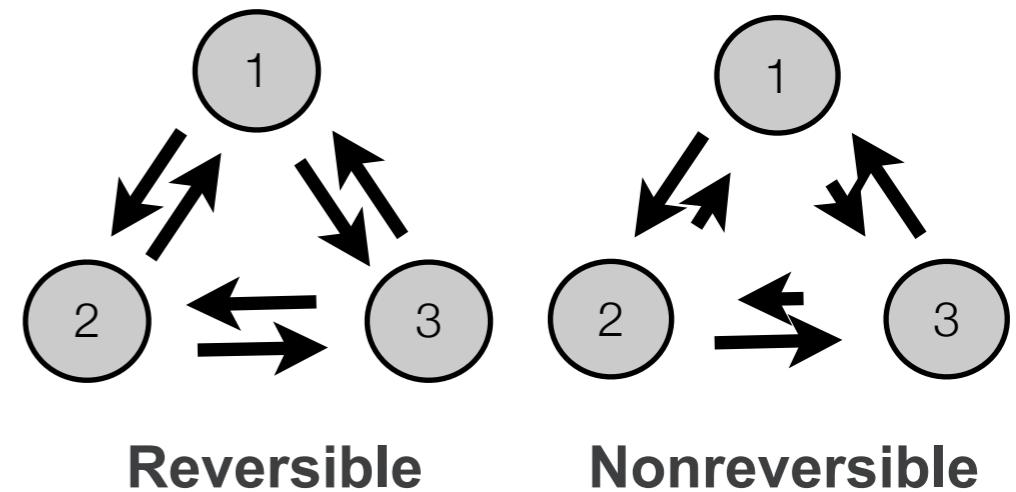
$$\exists k \in \mathbb{N} \text{ s. t. } P^k > 0$$

- 詳細つりあい (Detailed balance, reversibility)

全体つりあいの十分条件 (必要条件ではない)

シンプルかつ実用的

$$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i, j$$



- 詳細つりあいを破って計算効率を改善することが可能 (講義後半)

イジングモデル

- 統計力学における最も基本的な模型

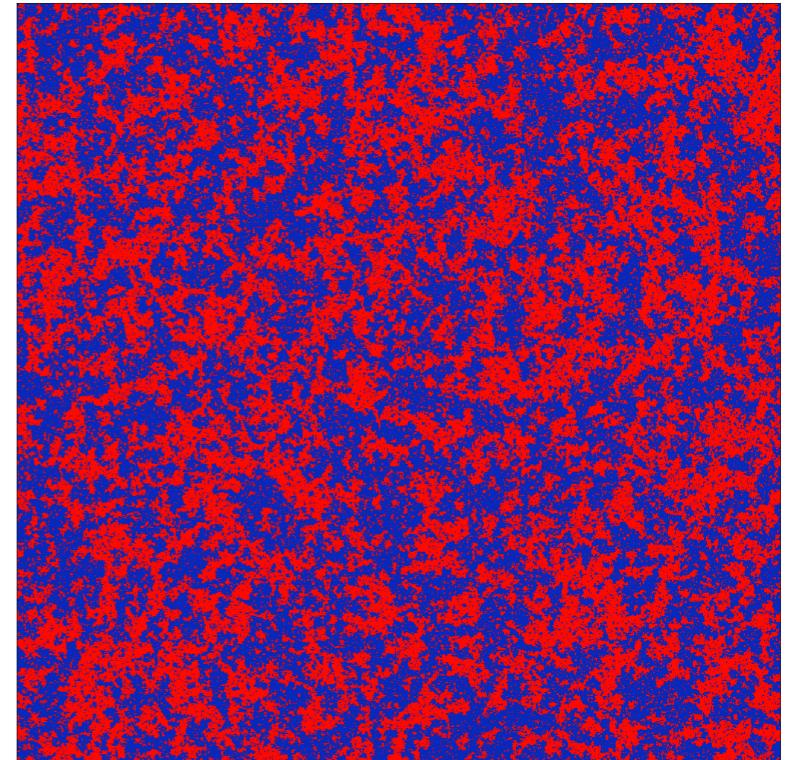
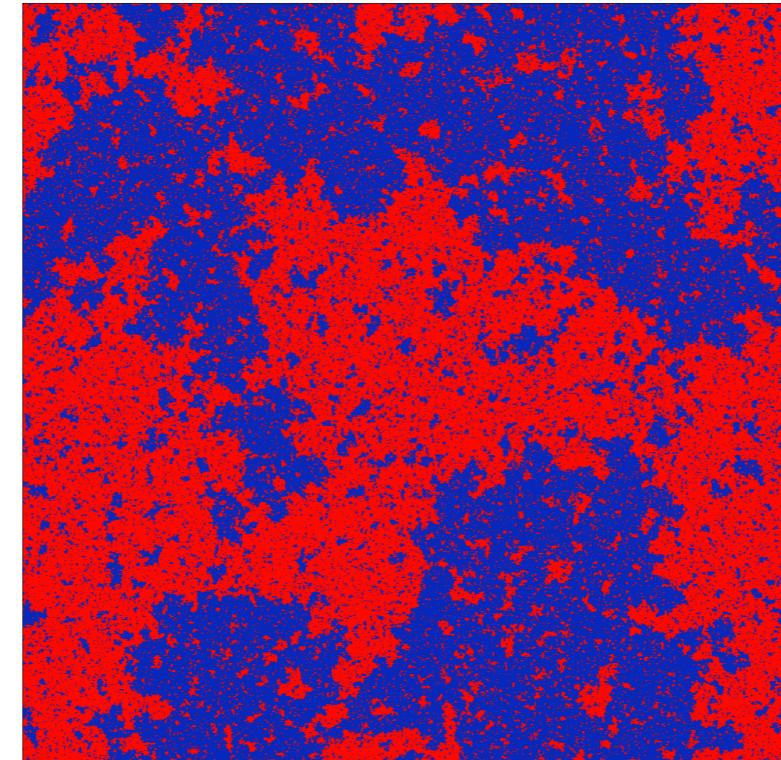
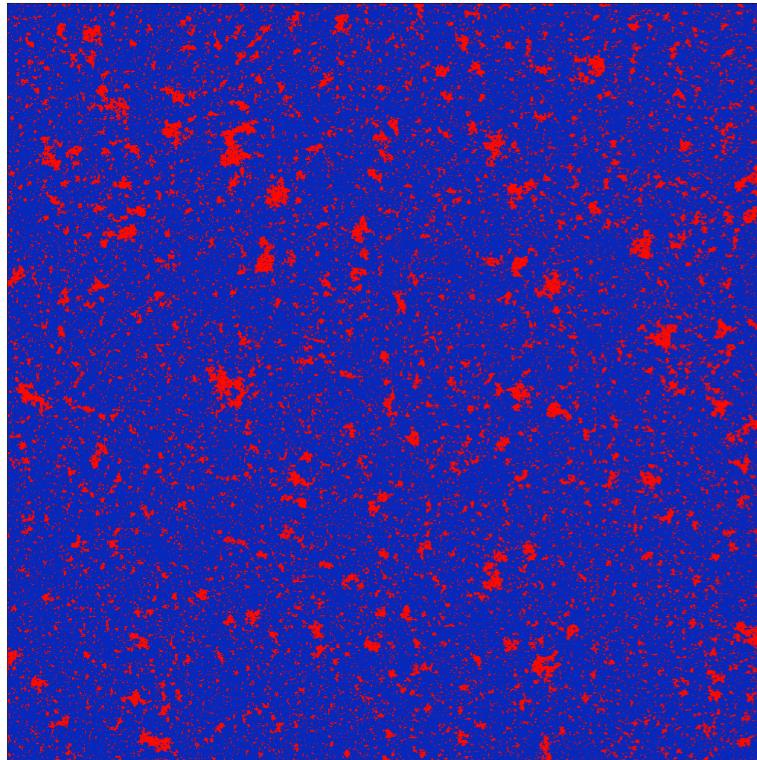
$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \quad \sigma_i = \pm 1$$

$$h = 0 \\ L = 1024$$

$$T = 0.99T_c$$

$$T = T_c$$

$$T = 1.05T_c$$



秩序相

エネルギー利得

臨界点

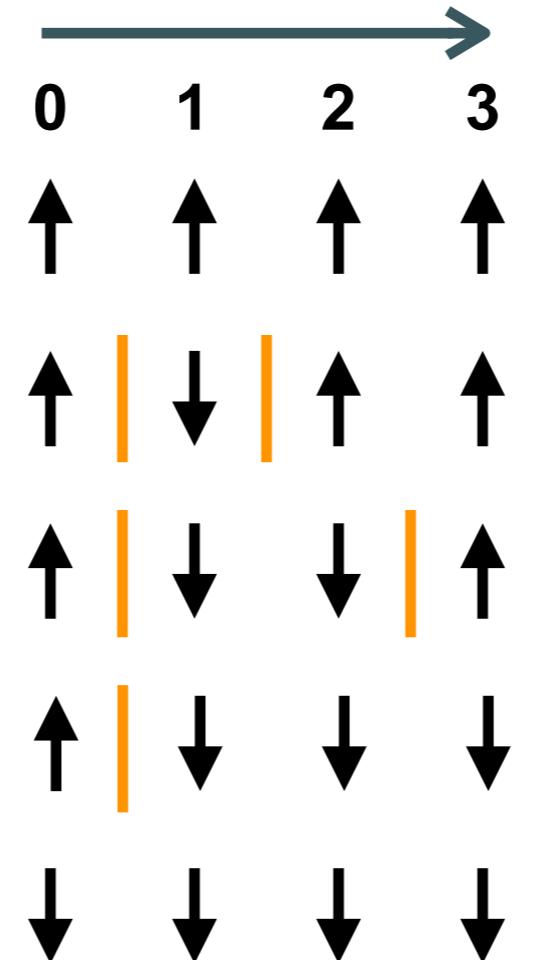
$$F = E - TS$$

無秩序相

エントロピー利得

エルゴード性が破れる例

- メトロポリス法を使ってもエルゴード的でない場合がある。
1次元イジング + 周期境界条件 + Sequential update + メトロポリス法



$$p_{ij} = \min \left(1, \frac{\pi_j}{\pi_i} \right)$$

他の例：
 2×2 イジング

- ドメインウォールが2つまでしか作られない。
→ 每回ランダムに更新する спинを選ぶ、熱浴法を使うなど、
いくつかの更新法を用いて double check する。

自己相関 (autocorrelation)

- MCMCではサンプル間の相関（自己相関）を減らすことが重要

自己相関関数 (autocorrelation function)

$$A_{\mathcal{O}}(t) = \frac{\langle \mathcal{O}_{i+t} \mathcal{O}_i \rangle - \langle \mathcal{O} \rangle^2}{\langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2} \sim e^{-t/\tau_{\text{exp}, \mathcal{O}}}$$

指数自己相関時間 (Exponential autocorrelation time)

$$\tau_{\text{exp}, \mathcal{O}} = -\frac{1}{\ln |\lambda_2|}$$

積分自己相関時間 (Integrated autocorrelation time)

$$\tau_{\text{int}, \mathcal{O}} = \frac{1}{2} + \sum_{t=1}^{\infty} A_{\mathcal{O}}(t)$$

モンテカルロ平均と真の期待値とのズレ

$$\sigma_{\mathcal{O}}^2 = \left\langle \left(\frac{1}{M} \sum_i \mathcal{O}_i - \langle \mathcal{O} \rangle \right)^2 \right\rangle = \left\langle \frac{1}{M^2} \sum_i (\mathcal{O}_i^2 - \langle \mathcal{O} \rangle^2) \right\rangle + \frac{2}{M^2} \sum_{i=1}^{M-1} \sum_{t=1}^{M-t} (\langle \mathcal{O}_{i+t} \mathcal{O}_i \rangle - \langle \mathcal{O} \rangle^2)$$

M : # Monte Carlo steps

$$\simeq \frac{2\tau_{\text{int}, \mathcal{O}}}{M} v_{\mathcal{O}}$$

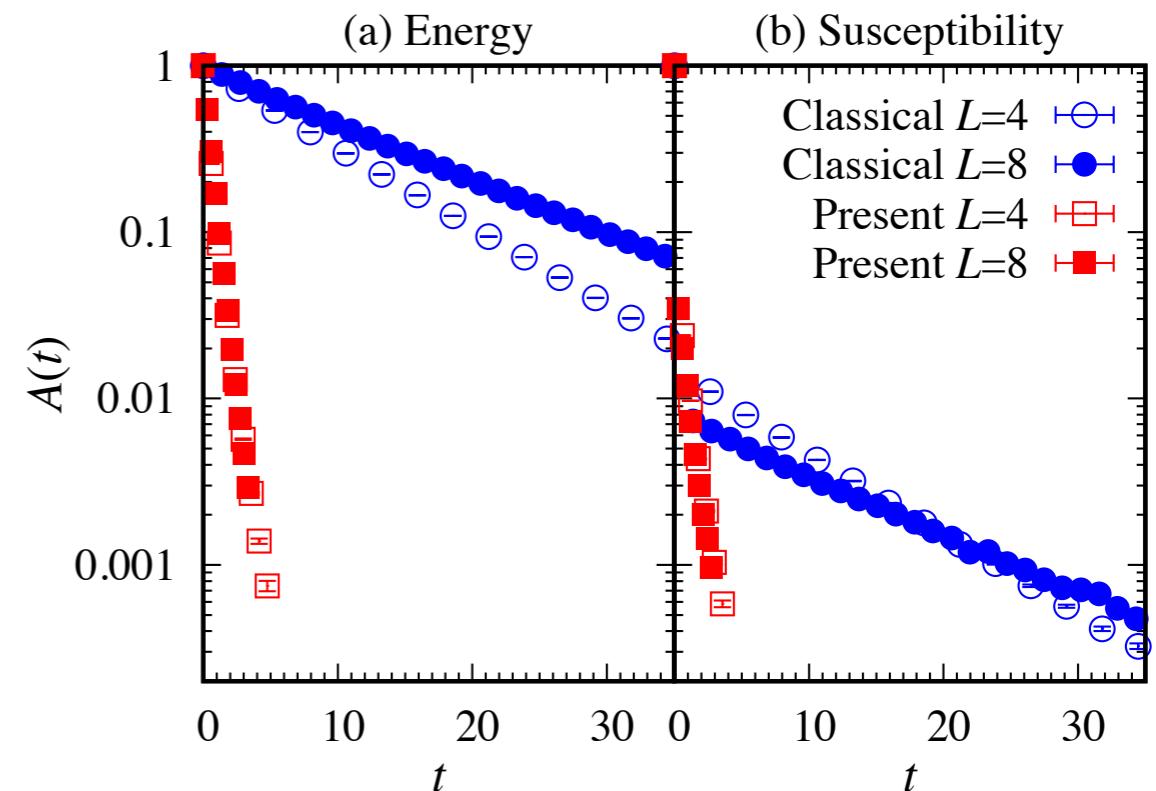
$$v_{\mathcal{O}} = \langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2$$

有効サンプル数 $M \rightarrow \frac{M}{2\tau_{\text{int}}}$

計算効率は自己相関時間と推定量の分散によって決まる

HS, Phys. Rev. E 103, 013308 (2021)

Worm algorithm for 3D Ising model



漸近分散 (asymptotic variance)

$$v_{\text{asymp}, \mathcal{O}} = 2\tau_{\text{int}, \mathcal{O}} v_{\mathcal{O}}$$

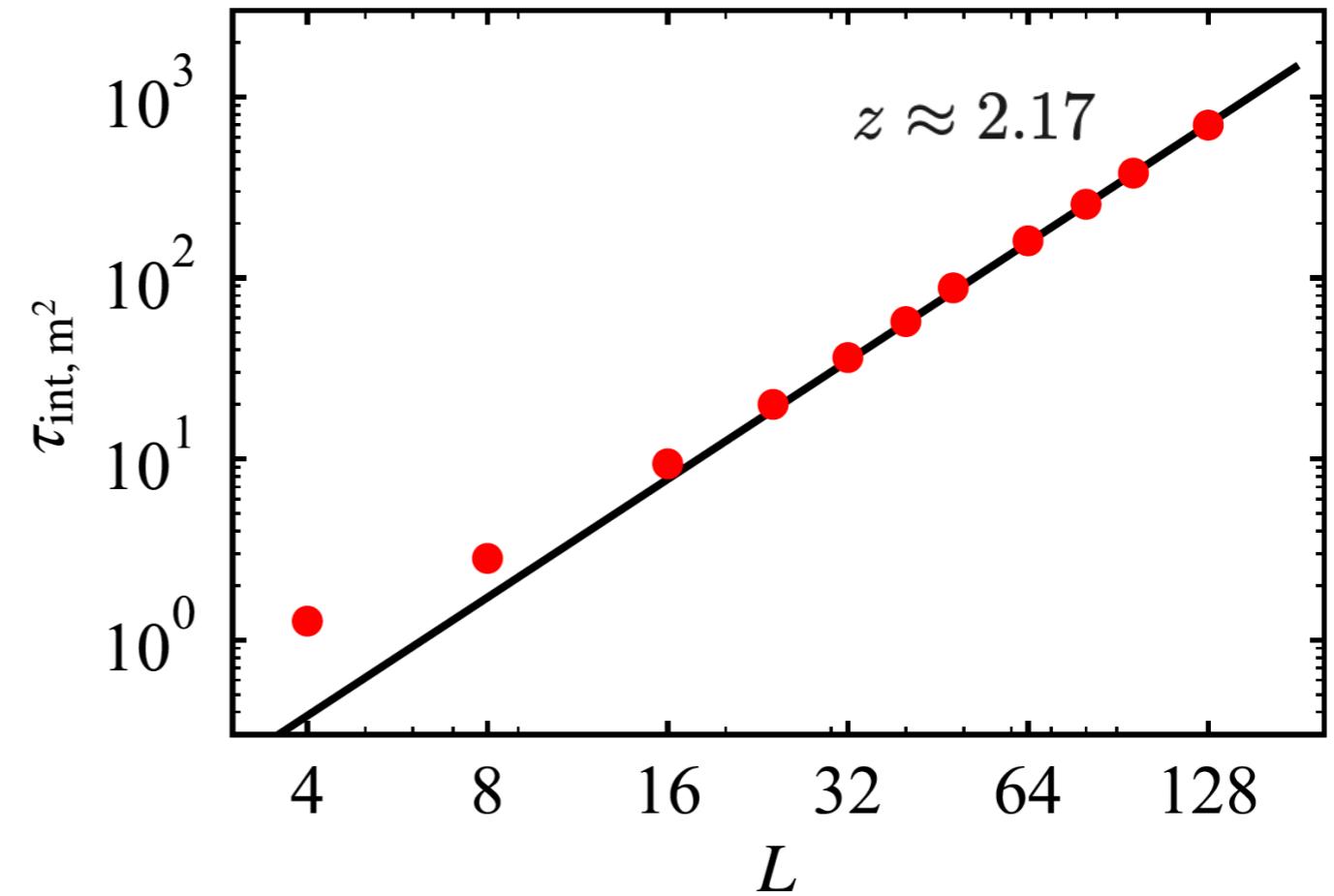
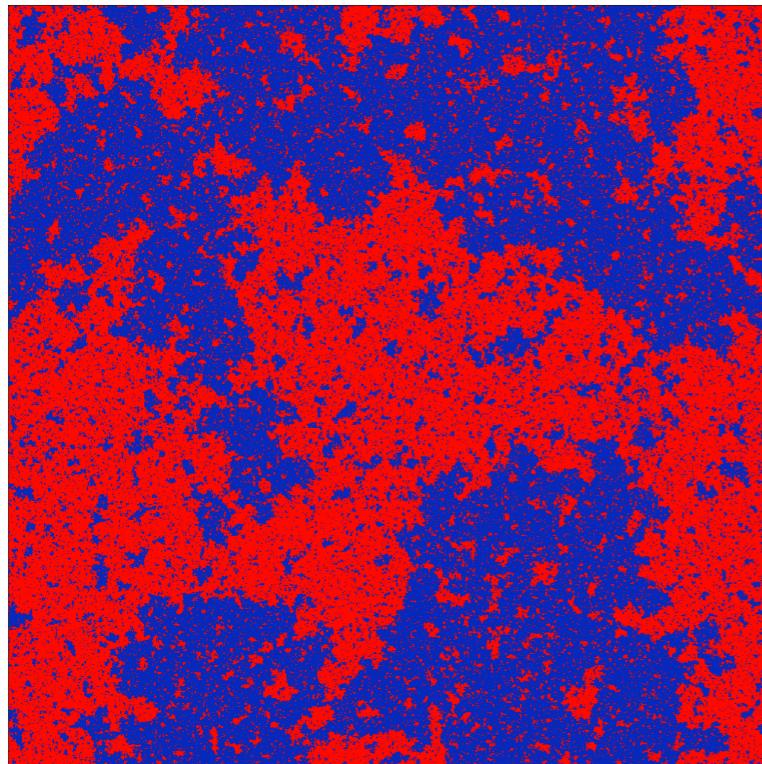
臨界緩和 (critical slowing down)

- 臨界点近傍では自己相関時間が系のサイズと共に増加

$$\tau \sim L^z$$

z : 動的臨界指数 (dynamic critical exponent)

2D イジング 臨界点 メトロポリス法



Binning 解析

- データを bin に分けて bin ごとの平均値を求める。

$$\underbrace{\mathcal{O}_1 \quad \mathcal{O}_2 \quad \mathcal{O}_3 \quad \cdots \quad \mathcal{O}_n}_{\bar{\mathcal{O}}_1^{(n)}} \quad \underbrace{\mathcal{O}_{n+1} \quad \mathcal{O}_{n+2} \quad \cdots \quad \mathcal{O}_{2n}}_{\bar{\mathcal{O}}_2^{(n)}} \quad \mathcal{O}_{2n+1} \quad \cdots$$

「bin ごとの平均値」の分散

$$v^{(1)} = v_{\mathcal{O}}$$

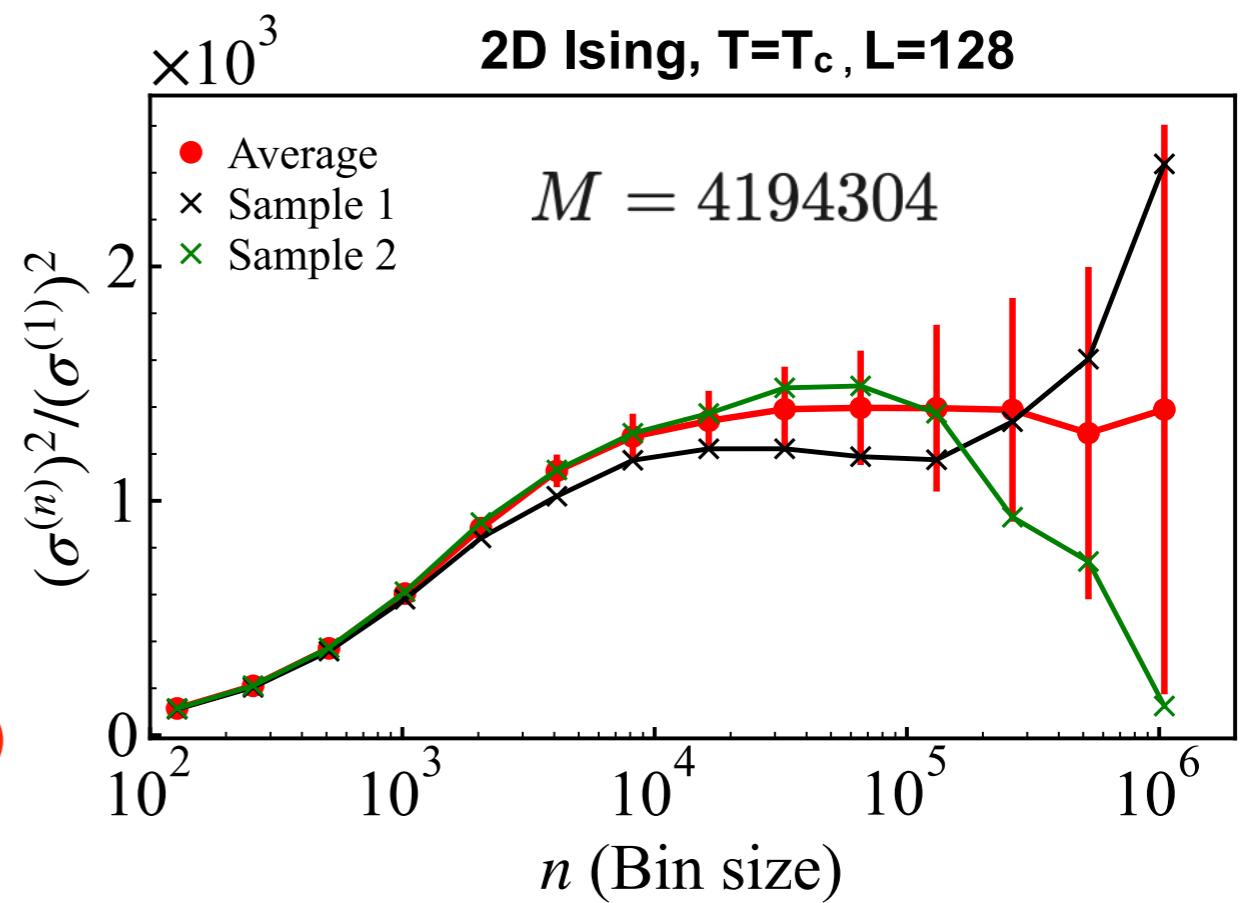
$$v^{(n)} = \frac{n}{M} \sum_{j=1}^{M/n} \left(\bar{\mathcal{O}}_j^{(n)} - \langle \mathcal{O} \rangle \right)^2 \approx \frac{2\tau_{\text{int}, \mathcal{O}}}{n} v_{\mathcal{O}} \quad (n \gg 1)$$

平均二乗誤差 [mean squared error (MSE)]

$$\sigma^2 = \lim_{n \rightarrow \infty} (\sigma^{(n)})^2 = \lim_{n \rightarrow \infty} \frac{n}{M} v^{(n)}$$

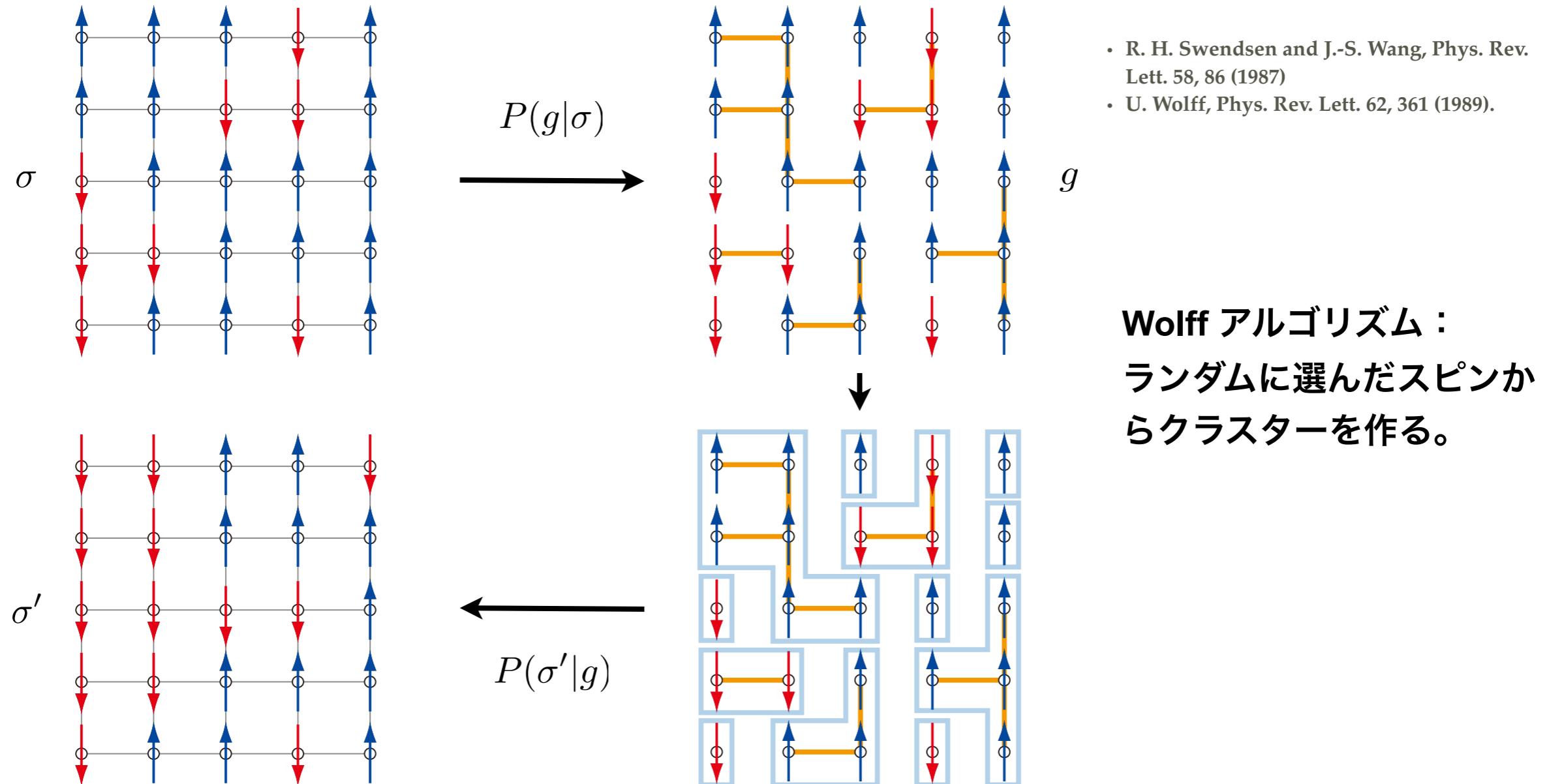
$$\tau_{\text{int}} = \frac{(\sigma^{(\infty)})^2}{2(\sigma^{(1)})^2}$$

- 相関時間の100倍程度は計算必要
- 最初の半分は捨てる (thermalization, burn-in)
- bin の数は10以上、100程度欲しい



Swendsen-Wang アルゴリズム

- 相関長に対応するサイズのクラスターを一度に更新することができる。
- 各ボンドについて、両端のスピンの向きが同じなら確率 $1 - e^{-2\beta J}$ でつなぐ。
- スピンの向きが反対ならつながない。
- つながっているスピンのクラスターをそれぞれ独立に確率 $1/2$ でフリップする。



カノニカルアンサンブル

- 熱浴とエネルギーをやりとりする系の平衡状態を記述する。

$$\text{逆温度} : \beta = \frac{1}{k_B T}$$

$$\text{期待値} : \langle \mathcal{O} \rangle = \frac{\sum_i \mathcal{O}_i e^{-\beta E_i}}{Z}$$

$$\text{規格化定数 (分配関数)} : Z = \text{tr} e^{-\beta H}$$

$$\text{比熱} : C = \frac{1}{N} \frac{\partial E}{\partial T} = - \frac{\beta^2}{N} \frac{\partial E}{\partial \beta} = \frac{\beta^2}{N} (\langle E^2 \rangle - \langle E \rangle^2)$$

$$\text{磁化} : m = \frac{1}{N} \left\langle \sum_i \sigma_i \right\rangle$$

$$\text{帶磁率} : \chi = \frac{\partial m}{\partial h} = N\beta (\langle m^2 \rangle - \langle |m| \rangle^2)$$

$$\text{相関関数} : G(r) = \langle \sigma_i \sigma_j \rangle \sim \begin{cases} e^{-r/\xi} & (T > T_c) \\ r^{-(d-2+\eta)} & (T = T_c) \\ m^2 + a e^{-r/\xi} & (T < T_c) \end{cases}$$

$$\text{構造因子} : S(q) = \frac{1}{N} \sum_{ij} \langle \sigma_i \sigma_j \rangle e^{-i(r_i - r_j) \cdot q} = \frac{S(0)}{1 + q^2 \xi^2} + O(q^4)$$

$$\text{“相関長”} : \xi = \frac{1}{|q|} \sqrt{\frac{S(0)}{S(q)} - 1}$$

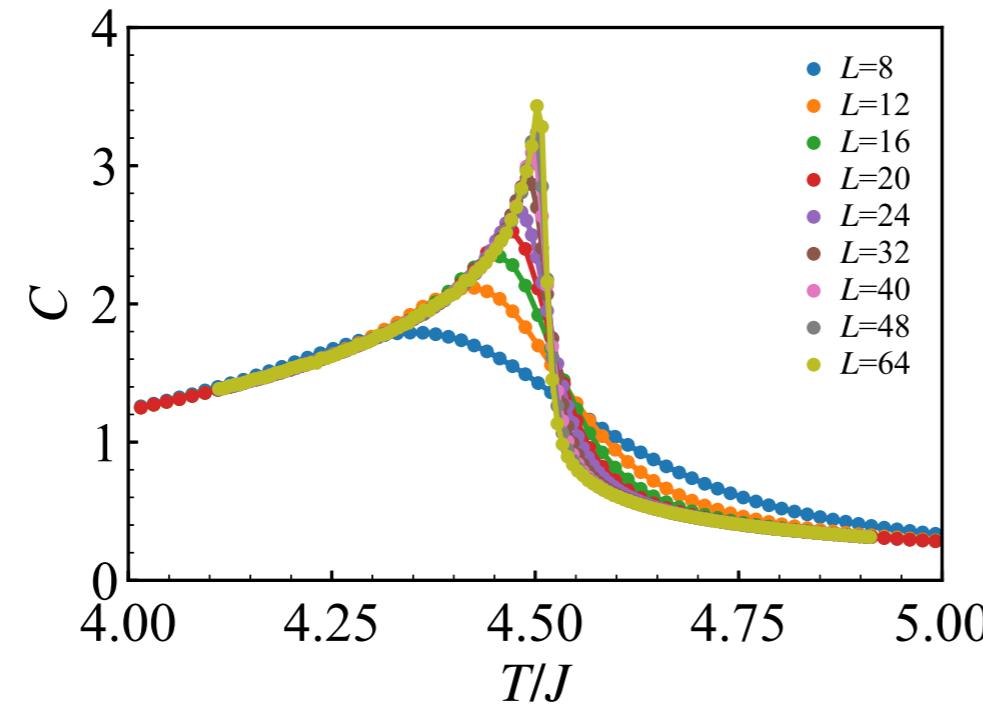
J. Cardy, *Scaling and Renormalization in Statistical Physics*, Cambridge University Press, (1996)

相転移

- 有限系は相転移しないが、熱力学的極限（無限系）で物理量が非解析な振る舞いを示す。

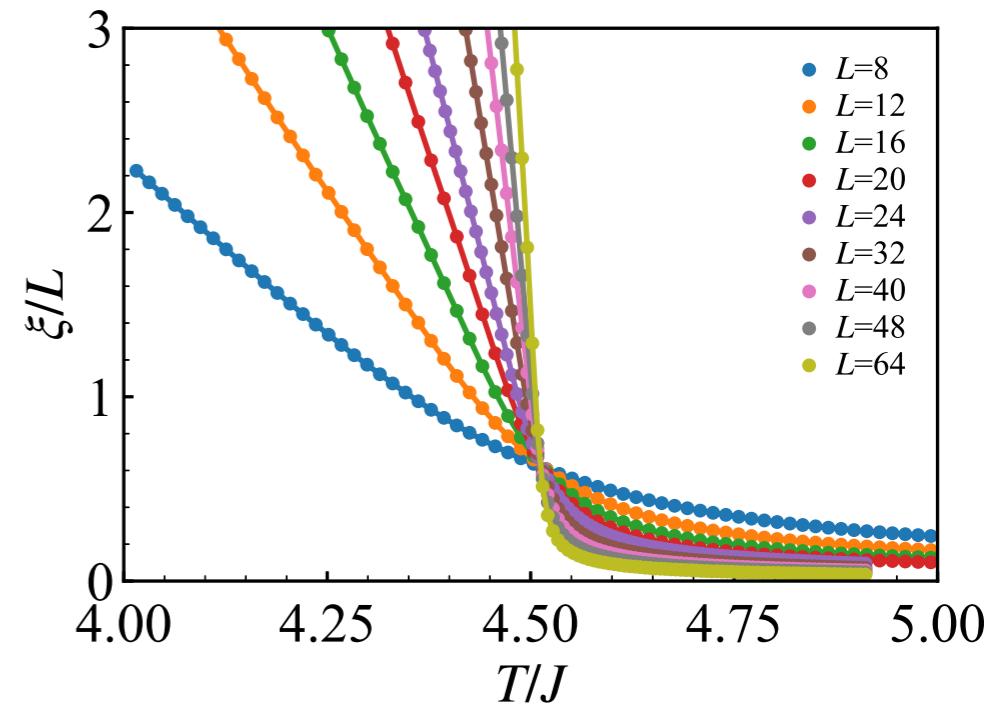
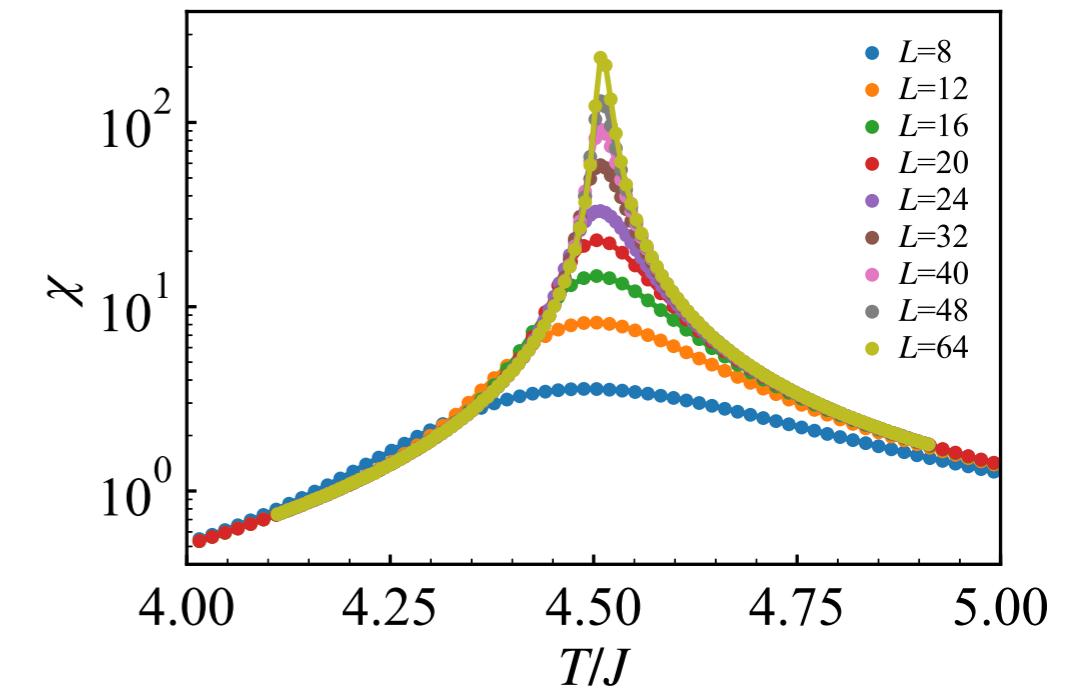
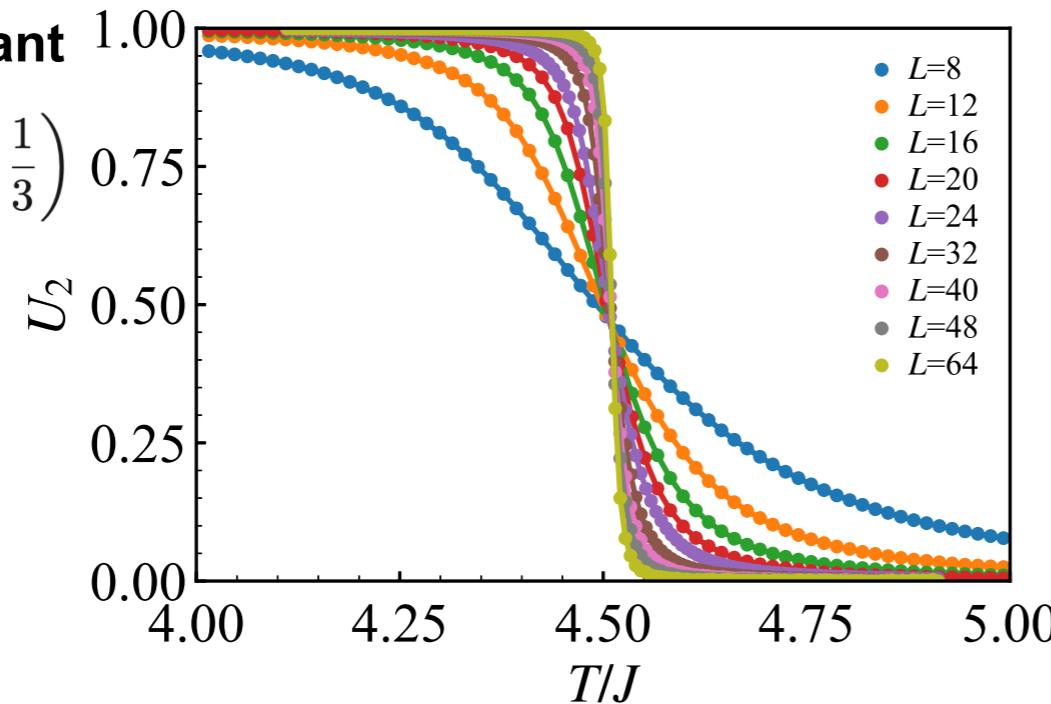
3D イジング

$$T_c/J \approx 4.511$$



Binder cumulant

$$U_2 = \frac{3}{2} \left(\frac{\langle m^2 \rangle^2}{\langle m^4 \rangle} - \frac{1}{3} \right)$$



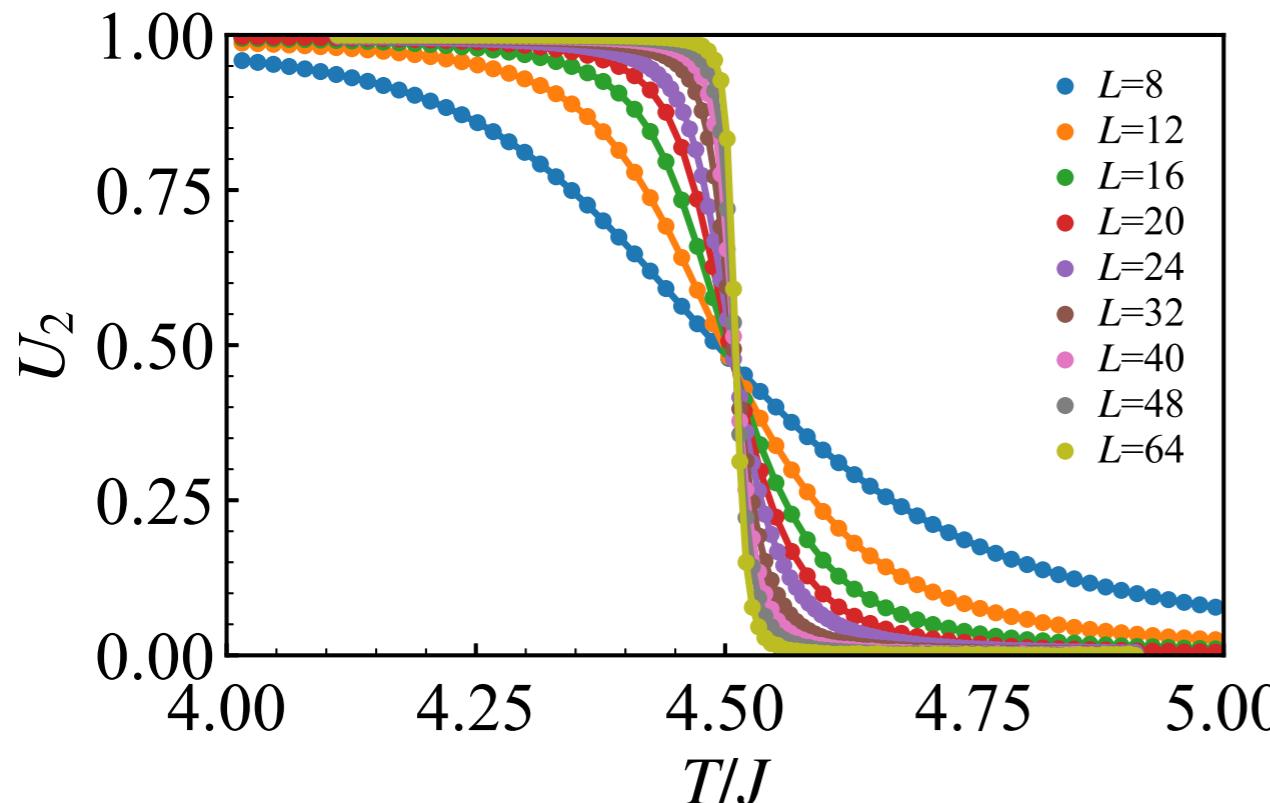
有限サイズスケーリング：ビンダー比（尖度）

- スケーリング仮設に基づき、転移点と臨界指数を求める手法。
- 異なるサイズのデータが一つの曲線に乗るようにパラメータを決める。
- サイズと温度範囲の選択は非自明

$$U_2(t, L) = \tilde{U}_2(tL^{1/\nu})$$

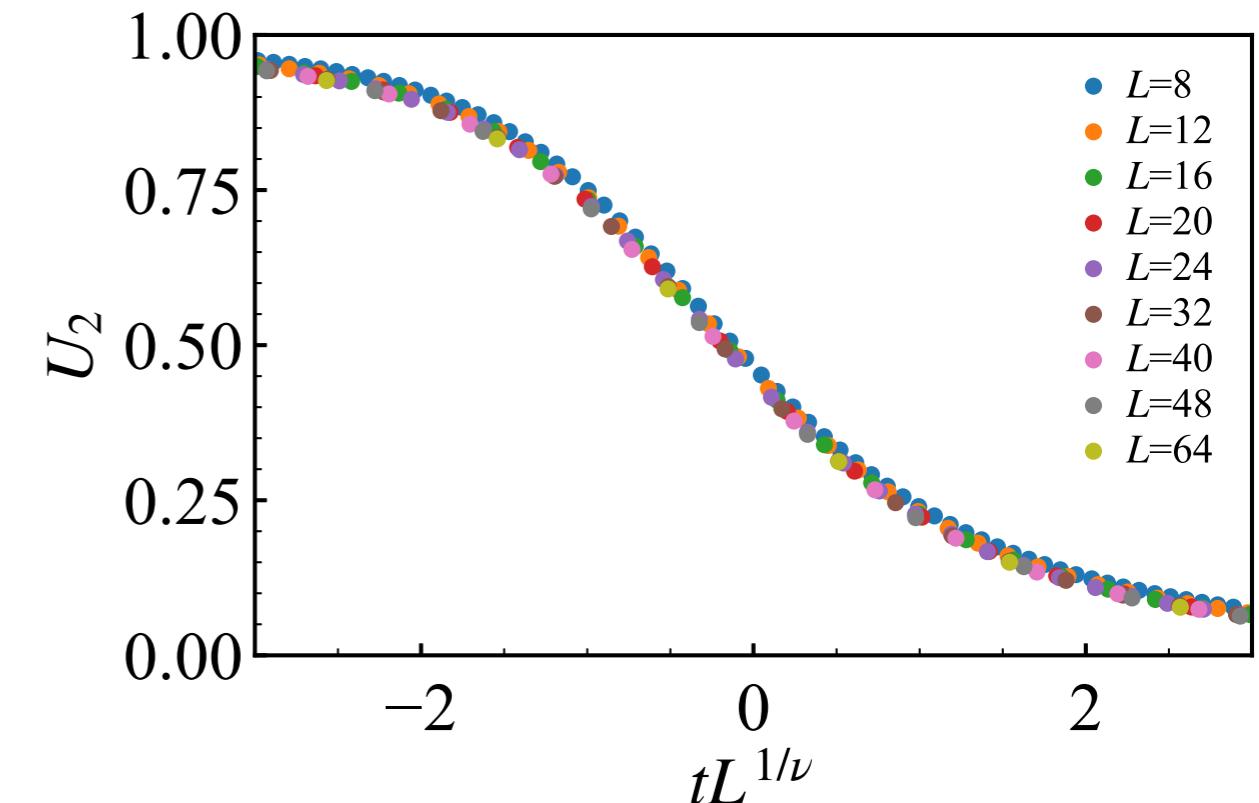
ガウス過程回帰

K. Harada, Phys. Rev. E 84, 056704 (2011)



Binder cumulant

$$U_2 = \frac{3}{2} \left(\frac{\langle m^2 \rangle^2}{\langle m^4 \rangle} - \frac{1}{3} \right)$$

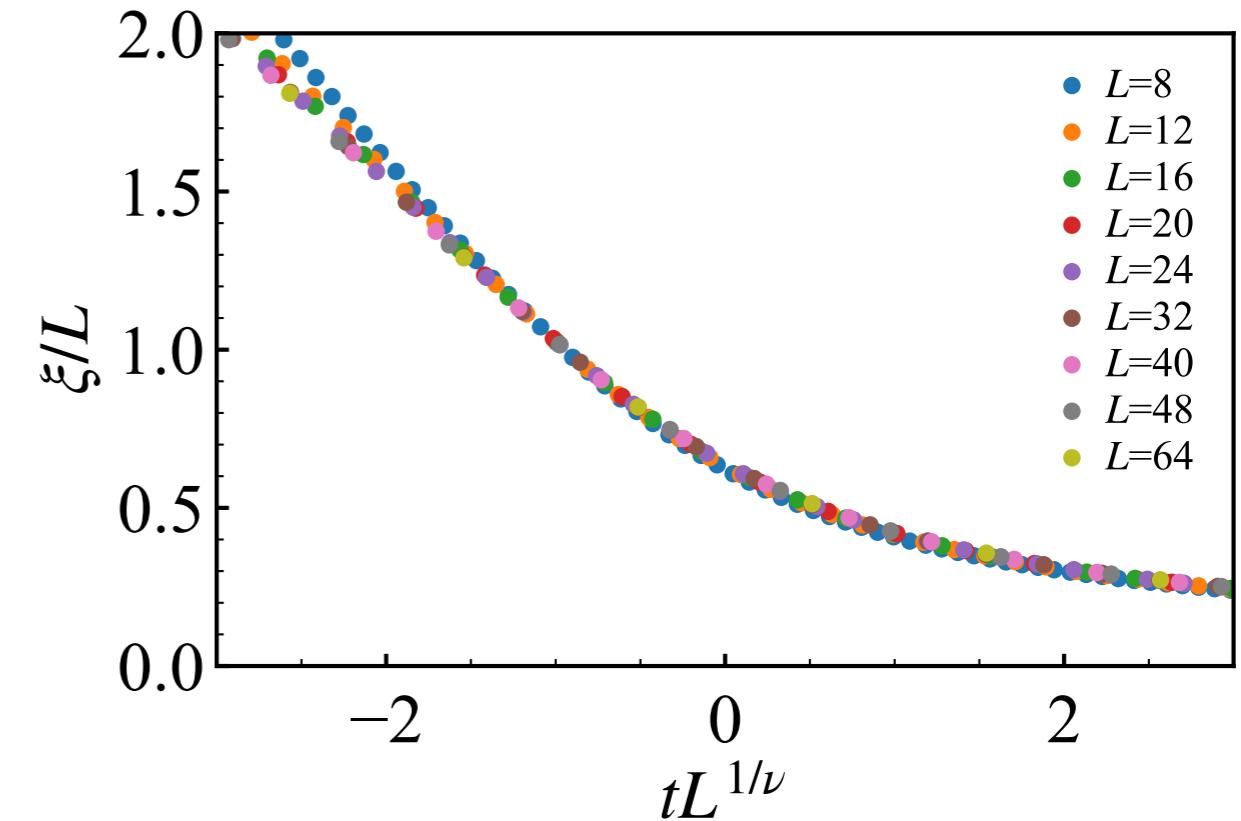
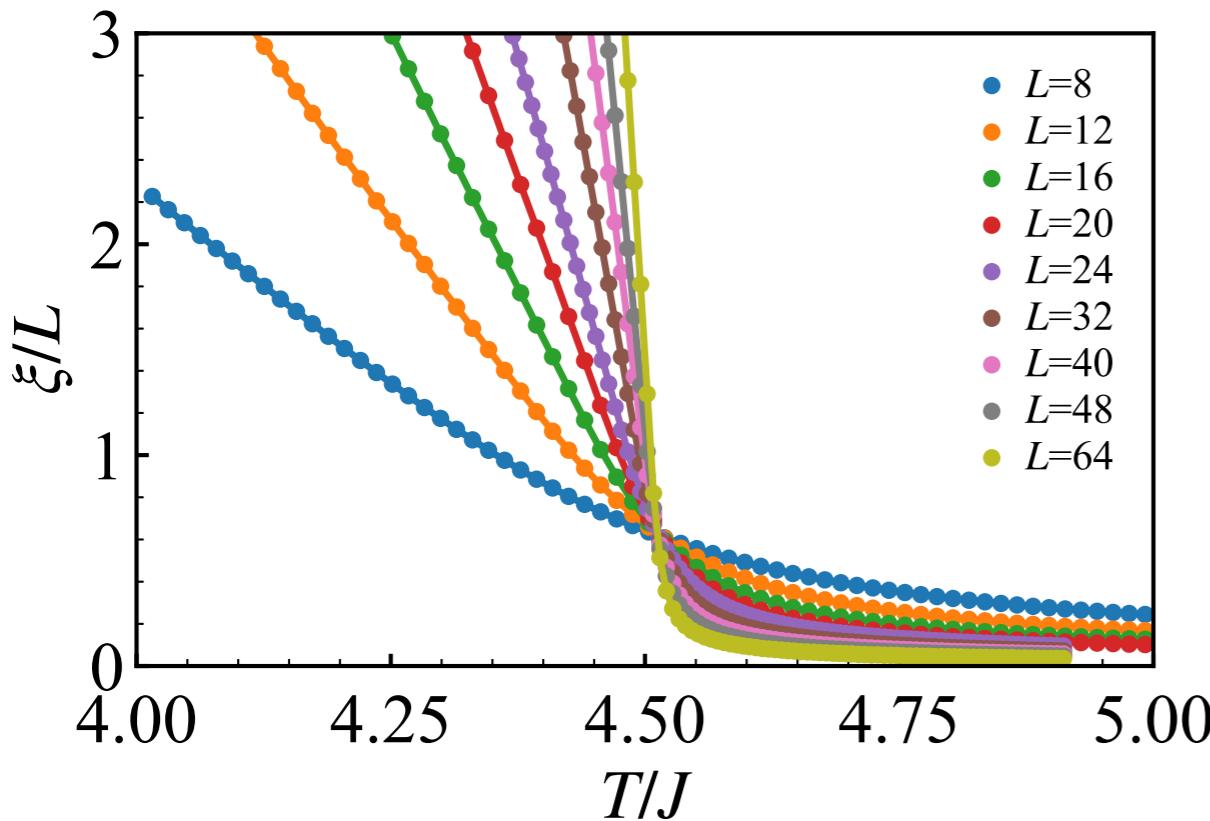


$$t = \frac{T - T_c}{T_c} \quad T_c/J \approx 4.511$$
$$\nu \simeq 0.63002$$

有限サイズスケーリング：相関長

- 多くの場合にビンダー比より精度良く求まる

$$\xi(t, L)/L = \tilde{\xi}(tL^{1/\nu})$$



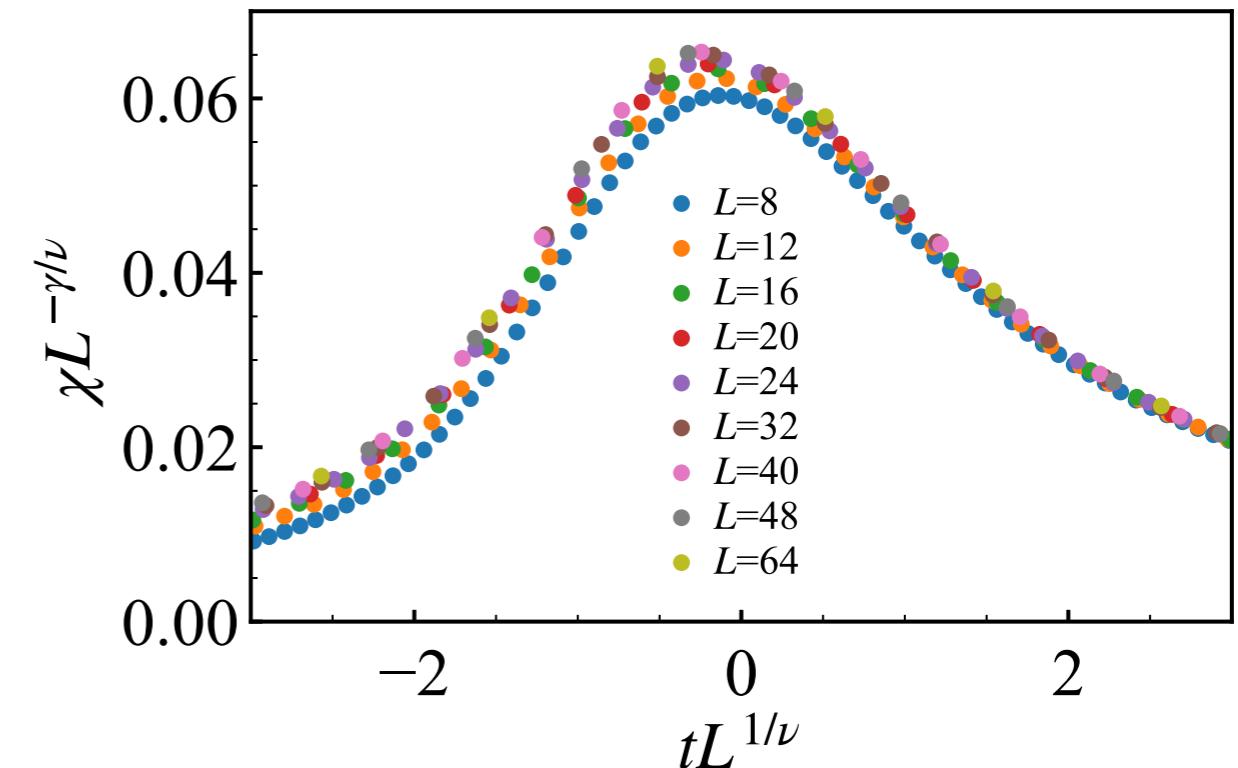
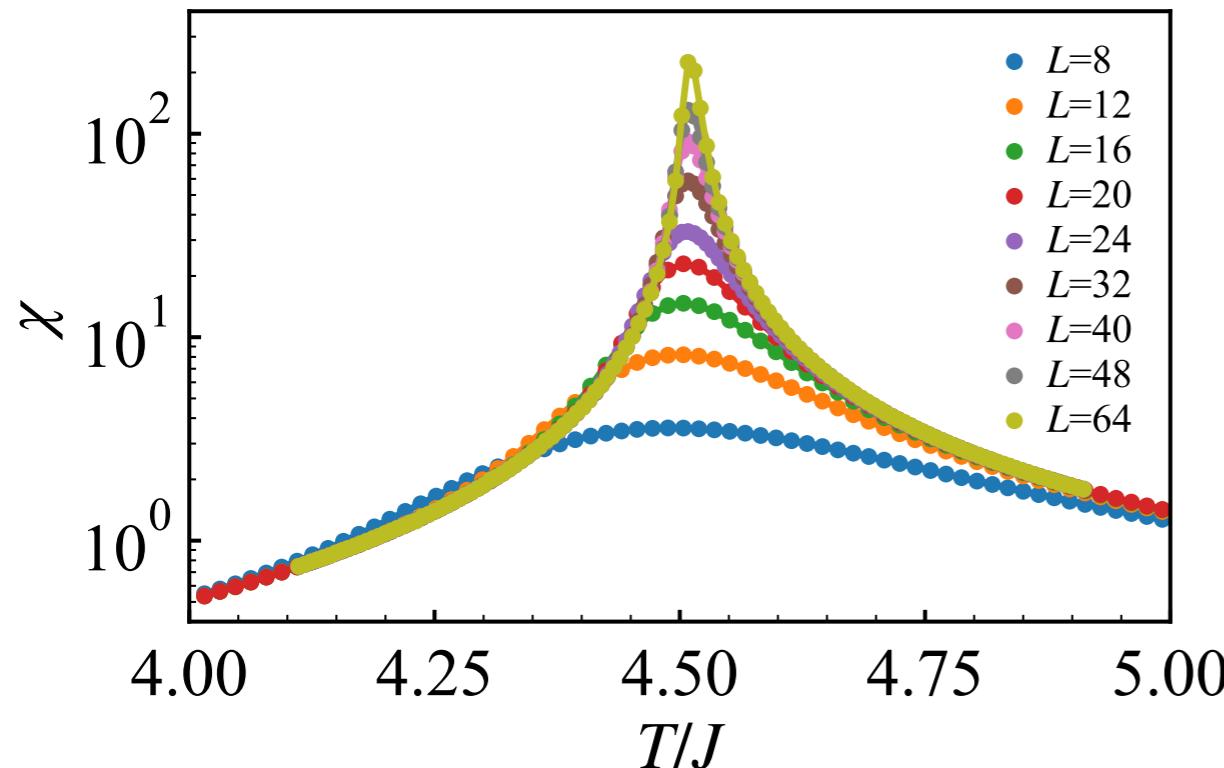
$$t = \frac{T - T_c}{T_c} \quad T_c/J \approx 4.511$$

$$\nu \simeq 0.63002$$

有限サイズスケーリング：帯磁率

- 2つの臨界指数を求めることができる。
- ただしパラメータが増えるとフィッティングは難しくなる。

$$\chi(t, L) = L^{\gamma/\nu} \tilde{\chi}(tL^{1/\nu})$$



$$t = \frac{T - T_c}{T_c} \quad T_c/J \approx 4.511$$

ブートストラップ法

- サンプル平均の非線形関数の誤差を信頼度高く計算する手法：
ビンダー比、相関長、フィッティングなど

(ビンごとの平均) データ : x_1, x_2, \dots, x_N
 y_1, y_2, \dots, y_N

ビンダー比の場合

$$f(x, y) = \frac{x^2}{y}$$
$$x = m^2 \quad y = m^4$$

- N個のもの（ビン）から重複を許してN個選ぶ（リサンプリング）

$$x_k^B = \frac{1}{N} \sum_{i=1}^N n_{i,k} x_i$$

P. Young, arXiv:1210.3781

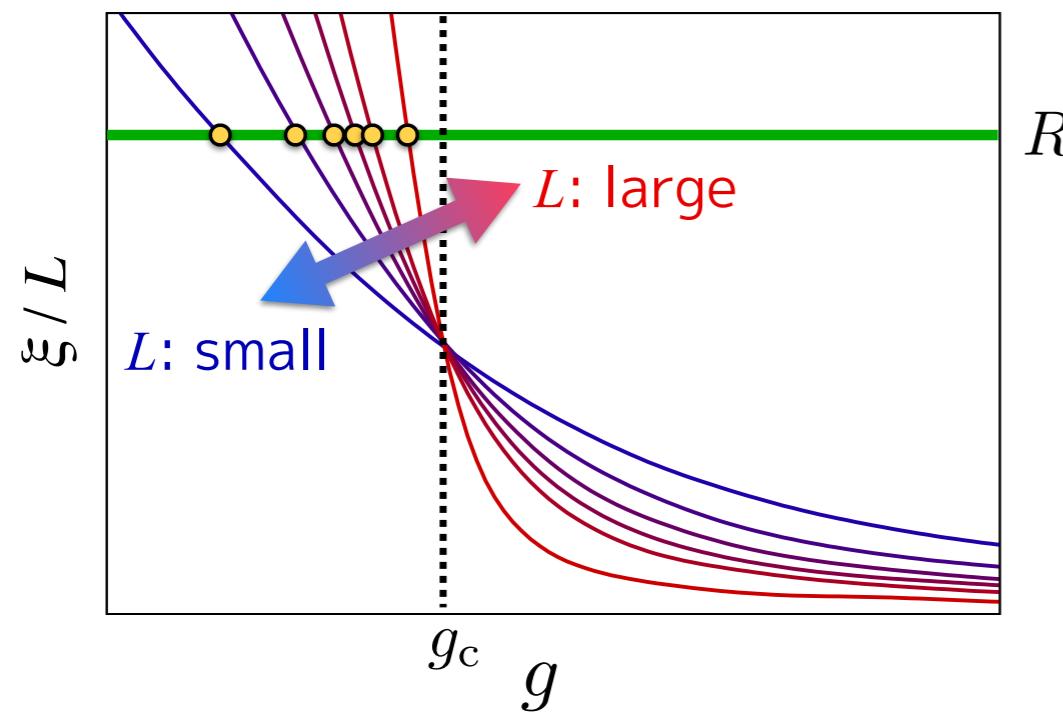
平均値 $\overline{f^B} = \frac{1}{N_{\text{boot}}} \sum_{k=1}^{N_{\text{boot}}} f_k^B$ $f_k^B = f(x_k^B, y_k^B)$ バイアス $\sim \frac{1}{N}$

誤差 $\sigma_f = \sqrt{\frac{N}{N-1}} s_{fB}$ $s_{fB}^2 = \overline{(f^B)^2} - \left(\overline{f^B}\right)^2$

$$\sim \frac{1}{\sqrt{N}}$$

確率近似法 (Stochastic approximation)

- 有限サイズスケーリングは転移点近くにたくさんのデータが必要
- 相転移点をほぼ全自動で見つけたい
- 確率近似法はノイズがある場合に非線形方程式の解を求める方法

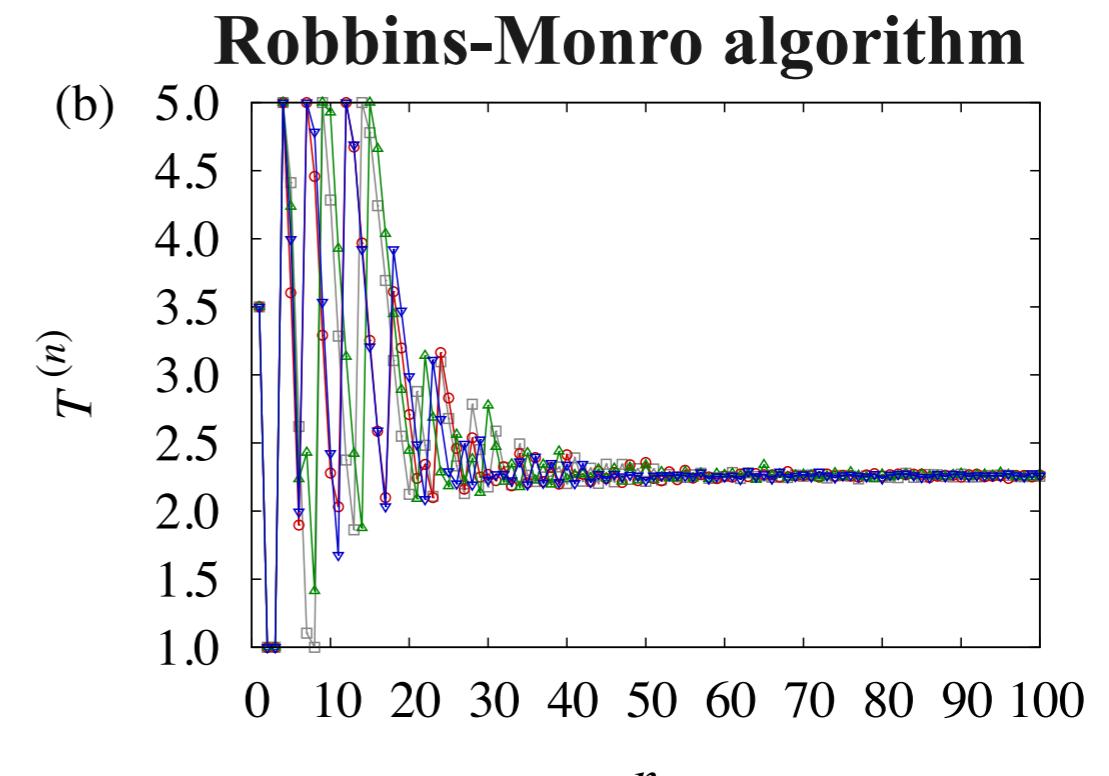


$$g_c(L) - g_c \propto L^{-1/\nu}$$

- べき関数のフィティングだけで転移点と臨界指数が求まる。

$$f(T) = \frac{\xi(T)}{L} - R$$

$$f(T_c) = 0$$



$$T^{(n+1)} = T^{(n)} - \frac{p}{n} f(T^{(n)})$$

メトロポリス・ヘイスティングス法

- 提案確率も考慮して遷移確率を決める。
- 提案分布を状態によって「うまく」変えようとすると、つりあいが破れることがある。

提案分布 $q(x'|x)$

採択率 $a(x', x) = \min \left(1, \frac{\pi(x')q(x|x')}{\pi(x)q(x'|x)} \right)$

詳細つりあい $\pi(x)q(x'|x)a(x', x)dx dx' = \pi(x')q(x|x')a(x, x')dx dx'$
 $= \min(\pi(x)q(x'|x), \pi(x')q(x|x'))dx dx'$

ランジュバンダイナミクス

- ・ ポテンシャルの微分（力）が計算できるときは、多くの場合利用した方が良い。

$$m \frac{d^2\mathbf{x}}{dt^2} = -\nabla U(\mathbf{x}) - \gamma \frac{d\mathbf{x}}{dt} + \sqrt{2\gamma k_B T} \mathbf{R}(t)$$

m : 質量

$U(\mathbf{x})$: ポテンシャルエネルギー

γ : 摩擦係数

White noise

$$\langle R_i(t) \rangle = 0$$

$$\langle R_i(t) R_j(t') \rangle = \delta_{ij} \delta(t - t')$$

過減衰 (overdamped) ランジュバンダイナミクス：スピン系など

$$\gamma \frac{d\mathbf{x}}{dt} = -\nabla U(\mathbf{x}) + \sqrt{2\gamma k_B T} \mathbf{R}(t)$$

慣性項

$$\left| m \frac{d^2\mathbf{x}}{dt^2} \right| \ll \left| \gamma \frac{d\mathbf{x}}{dt} \right|$$

摩擦項

c.f. Velocity Verlet algorithm

ハイブリッド（ハミルトニアン）モンテカルロ

- 仮想的な運動量を導入し、エネルギー一定の位相空間で遷移させる (MD)
- エネルギーに離散化誤差が生じるが、メトロポリス法で採択 or 棄却する (MC)

もともとの状態変数： x_1, x_2, \dots, x_N

仮想的な運動量： p_1, p_2, \dots, p_N

$$\pi(x) = \frac{e^{-E(x)}}{Z} \quad \beta \text{ を } E(x) \text{ に含める}$$

$$\mathcal{H}(x, p) = \sum_i \frac{p_i^2}{2m_i} + E(x)$$

ハミルトニアンダイナミクス（古典力学）

A. Duane et al., Phys. Lett. B 195, 216 (1987)

$$\begin{aligned}\frac{dx_i}{dt} &= \frac{\partial \mathcal{H}}{\partial p_i} \\ \frac{dp_i}{dt} &= -\frac{\partial \mathcal{H}}{\partial x_i}\end{aligned}$$

ハイブリッド（ハミルトニアン）モンテカルロ

ハミルトニアンダイナミクスの特徴

- 全エネルギーが保存される

$$\frac{d\mathcal{H}}{dt} = \sum_i \left\{ \frac{\partial \mathcal{H}}{\partial x_i} \frac{dx_i}{dt} + \frac{\partial \mathcal{H}}{\partial p_i} \frac{dp_i}{dt} \right\} = \sum_i \left\{ \frac{\partial \mathcal{H}}{\partial x_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial \mathcal{H}}{\partial x_i} \right\} = 0$$

- 位相空間の体積が保存される (Liouvilleの定理)

位相空間上の流れ場 $\vec{v} = \left(\frac{d\vec{x}}{dt}, \frac{d\vec{p}}{dt} \right)$

$$\operatorname{div} \mathbf{v} = \sum_i \left\{ \frac{\partial}{\partial x_i} \frac{dx_i}{dt} + \frac{\partial}{\partial p_i} \frac{dp_i}{dt} \right\} = \sum_i \left\{ \frac{\partial}{\partial x_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial}{\partial p_i} \frac{\partial \mathcal{H}}{\partial x_i} \right\} = 0$$

- ⇒ 位相空間上の同時分布 $\pi(x, p) = \frac{e^{-\mathcal{H}(x, p)}}{Z_H}$ が時間発展の下で不变
- もともと欲しいボルツマン分布は $\pi(x, p)$ の周辺分布

ハイブリッド（ハミルトニアン）モンテカルロ

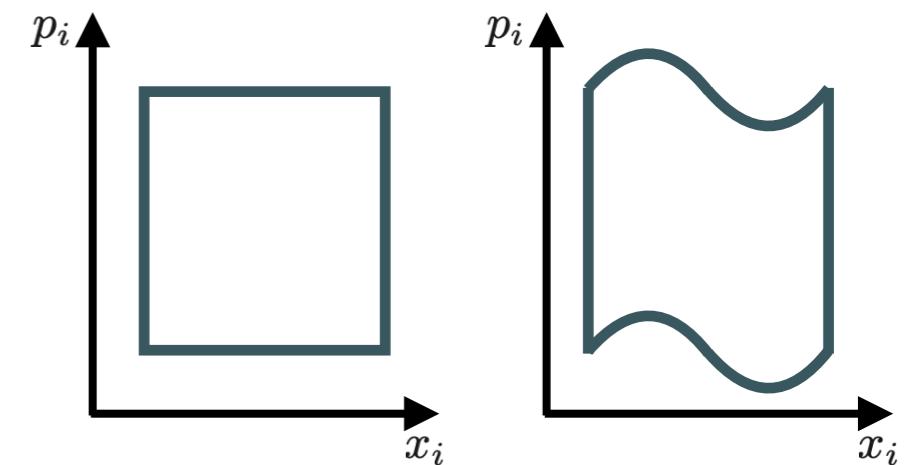
- ・ハミルトニアンドイナミクスに従い、適当な時間だけ時間発展する (MD)
- ・エネルギーに離散化誤差が生じるが、メトロポリス法で採択 or 棄却する (MC)
- ・位置を固定して運動量を確率的に更新する (全エネルギーも更新)
- ・詳細つりあいを満たす

リープfrog法 (Leapfrog) : 一度に p か x のどちらかだけを更新 \Rightarrow 位相空間の体積保存

$$\underline{p_i(t + dt/2)} = p_i - \frac{dt}{2} \frac{\partial E(\mathbf{x}(t))}{\partial x_i}$$

$$\underline{x_i(t + dt)} = x_i(t) + dt \underline{p_i(t + dt/2)}$$

$$p_i(t + dt) = \underline{p_i(t + dt/2)} - \frac{dt}{2} \frac{\partial E(\mathbf{x}(t + dt))}{\partial x_i}$$

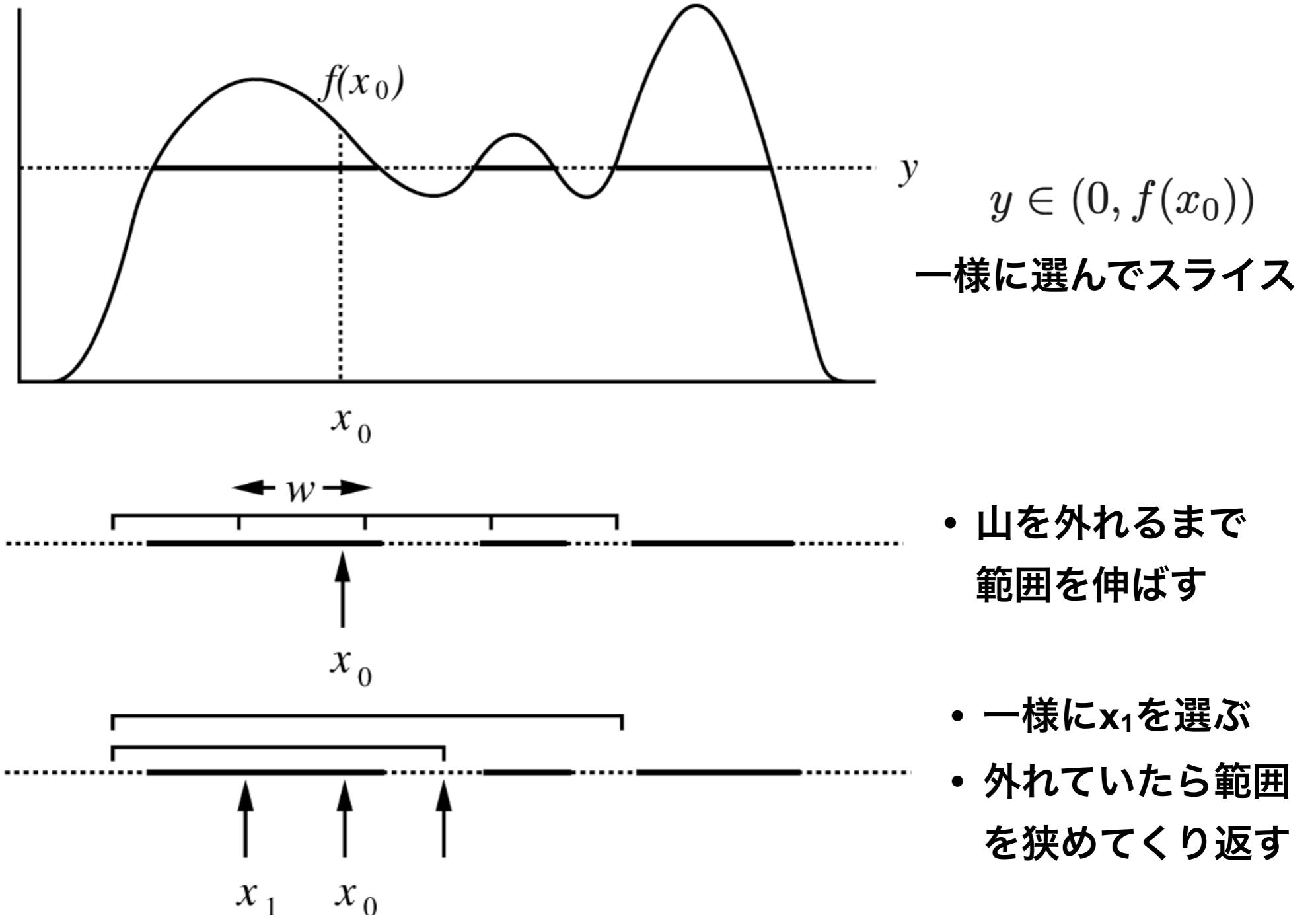


- ・パラメータが多いのが難点 (質量m、ステップ時間dt、ステップ数)

スライスサンプリング

- 「分布関数の値の次元」をひとつ足して一様サンプリングする
- 山から出発して離れた山にもいけるようにする

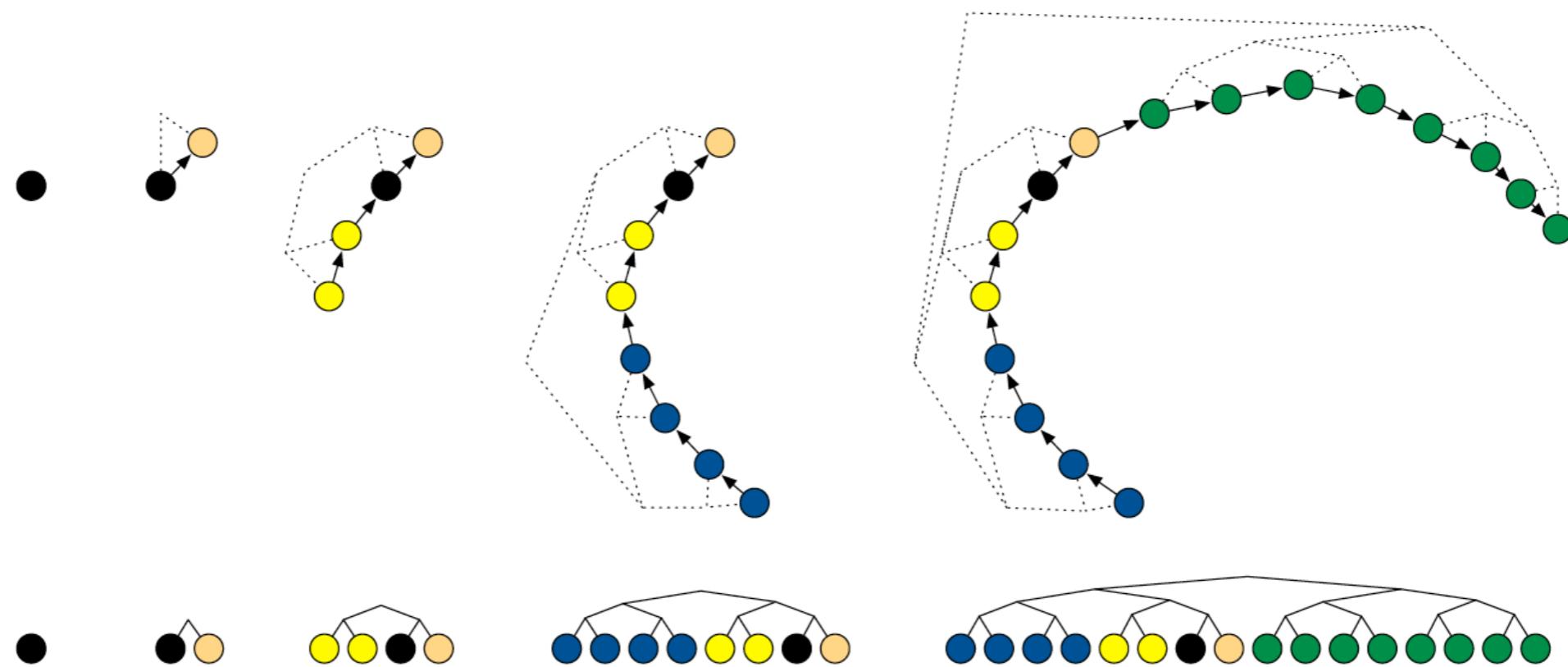
R. M. Neal, Ann. Statist. 31, 705 (2003)



No-U-Turn sampler (NUTS)

- ・ ハミルトニアンモンテカルロでステップ数を最適化する手法
- ・ 「Uターン」し始めたらストップ
- ・ 詳細つりあいを満たすように時間対称にパスを生成
- ・ 「スライス」の上にある点から一様に次の状態を選択

M. D. Homan and A. Gelman,
J. Mach. Learn. Res. 15, 1593 (2014)

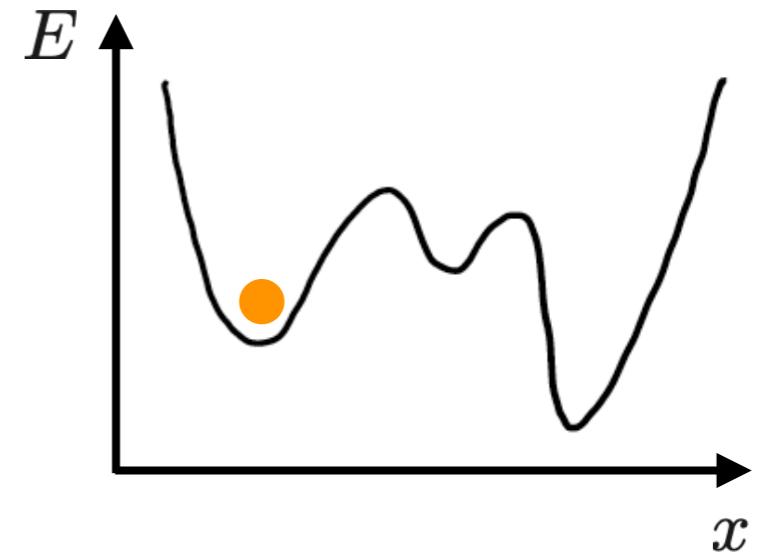


Uターン条件

$$(x^+ - x^-) \cdot p^+ < 0 \quad \text{or} \quad (x^+ - x^-) \cdot p^- < 0$$

MCMCの苦手なこと

- 自己相関時間が長い系（緩和が遅い系）
 - 臨界緩和（連続転移点近傍）
 - 一次転移点近傍
 - マルチモーダル（多峰的）な分布
- 規格化定数（分配関数）、エントロピーの計算

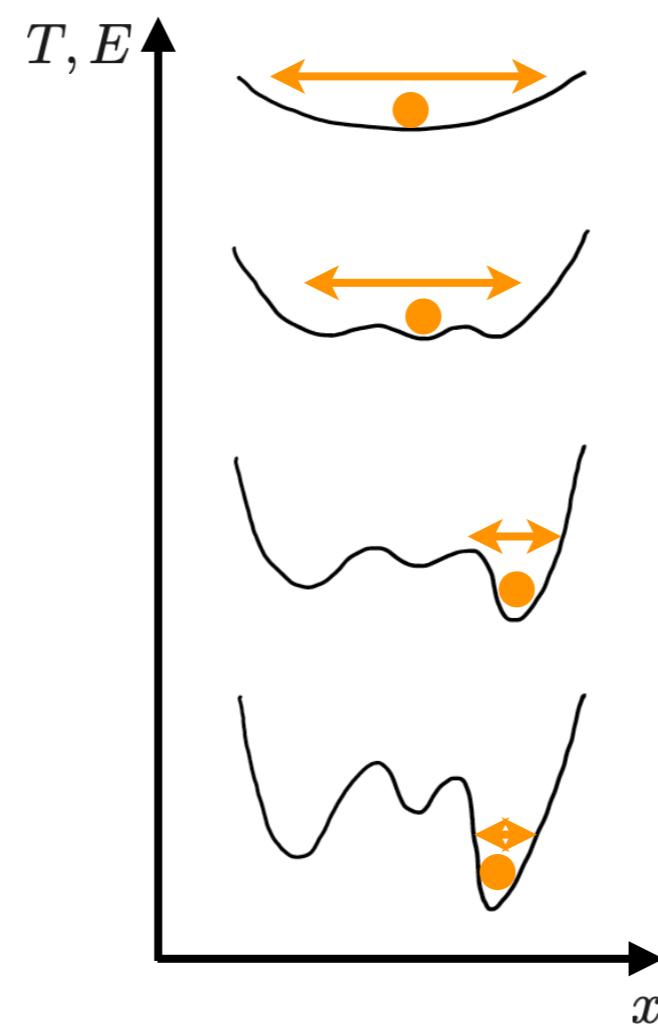


→ 拡張アンサンブル法

- シミュレートドテンパリング
- レプリカ交換法（パラレルテンパリング）
- マルチカノニカル法

シミュレーテッドアニーリング（焼きなまし法）

- 高温からゆっくり温度を下げていき、基底状態を見つける手法。
- 各温度でMCMCを行い、熱ゆらぎで局所安定状態から脱することができる。
- ポテンシャルの勾配が求まらなくても使える汎用的な最適化手法。
- 温度を変えるときにつりあい条件を破っているから平衡状態ではない。
- ボルツマン分布を用いずに速く収束させることも可能。



$$T(t) \geq \frac{c}{\log(t+2)}$$

$t \rightarrow \infty$

最適解に収束

- S. Kirkpatrick et al., Science 220, 671 (1983)
- S. Geman and D. Geman, IEEE Trans. PAMI 6, 721 (1984)
- C. Tsallis and D.A. Stariolo, Physica A 233, 395 (1996)

量子アニーリング

- 熱ゆらぎではなく量子ゆらぎを用いる最適化手法
- 多くの組み合わせ最適化問題はイジングモデルの基底状態を求める問題に変換可能
- 量子ゆらぎ（横磁場等）をゆっくり小さくしていき解きたい問題の基底状態を得る
- 断熱条件が満たされれば基底状態を連続的にたどることができる
- D-Wave など実機も存在する
- 量子モンテカルロでシミュレート可能

• T. Kadowaki and H. Nishimori, Phys. Rev. E 58, 5355 (1998)

横磁場イジングモデル $\mathcal{H} = \sum_{\langle ij \rangle} J_{ij} \sigma_i^z \sigma_j^z + \sum_i h_i \sigma_i^z - \Gamma \sum_i \sigma_i^x$

初期状態 (large Γ) $|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$ 時間と共にゆっくりと $\Gamma \rightarrow 0$

断熱条件

$$\frac{\max \left| \langle \Psi_1(t) | \frac{d\mathcal{H}(t)}{dt} | \Psi_0(t) \rangle \right|}{\min \Delta E(t)^2} \ll 1$$

$$\Delta E(t) = E_1(t) - E_0(t)$$

シミュレーテッドテンパリング（焼き戻し法）

- 温度を上げ下げすることによって局所安定にトラップされないようにする。
- 温度を変えても常に「平衡状態」となるようにする。
- 分配関数（規格化定数）は温度に依存する。
- 規格化定数の良い近似が必要となるが、計算は簡単でない。

温度セット

$$\beta_1, \beta_2, \dots, \beta_K$$

- A. P. Lyubartsev et al., J. Chem. Phys. 96, 1776 (1992)
- E. Marinari and G. Parisi, Europhys. Lett. 19, 451 (1992).

温度に依存する重み g を導入する

$$\pi(x, \beta_k) = \frac{e^{-\beta_k E(x) + g_k}}{Z_{\text{ST}}}$$

全体の規格化定数

$$Z_{\text{ST}} = \sum_{k=1}^K \sum_x e^{-\beta_k E(x) + g_k} = \sum_{k=1}^K Z(\beta_k) e^{g_k}$$

各温度での周辺分布

$$P(\beta_k) = \sum_x \pi(x, \beta_k) = \frac{Z(\beta_k) e^{g_k}}{Z_{\text{ST}}}$$

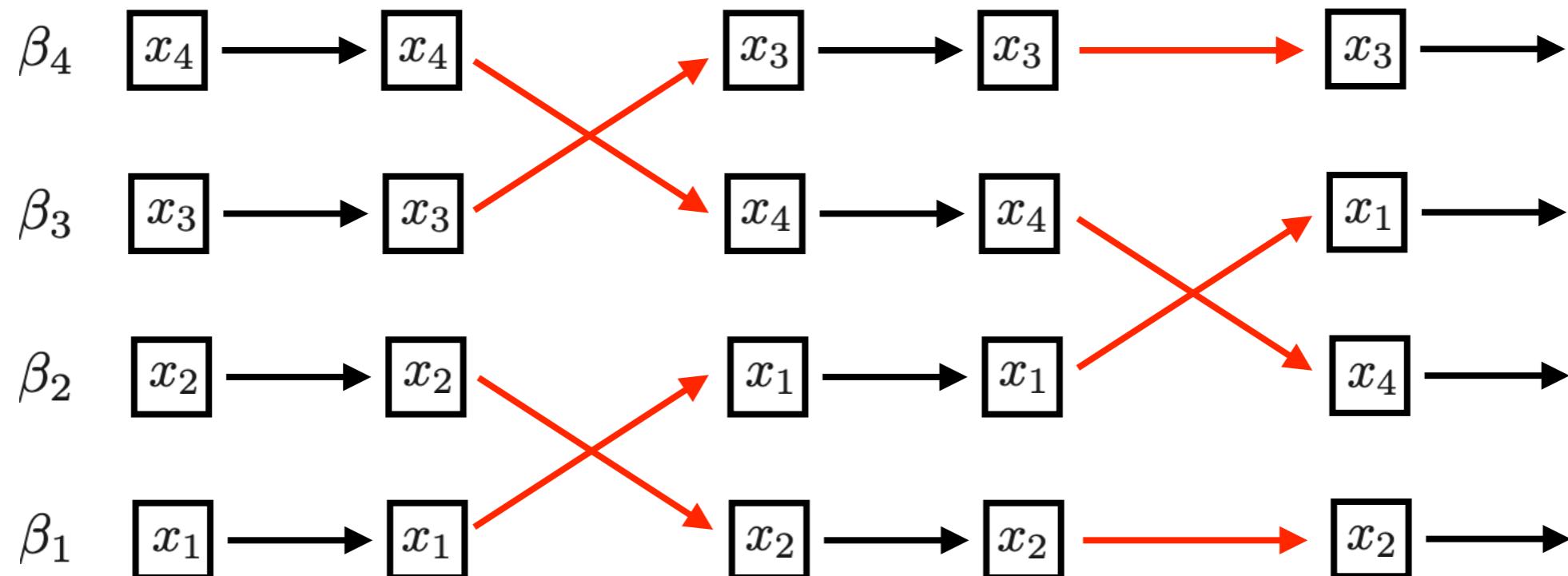
$$g_k = -\ln Z(\beta_k) \text{ とできれば温度に関して一様分布となる}$$

- 短い run を繰り返し行い、 g 関数を少しずつ調整する。

レプリカ交換法（パラレルテンパリング）

- 別々の温度の系（レプリカ）を交換する。
- 規格化定数はキャンセルされるので交換するときは考慮しなくて良い。
- 温度軸上でランダムウォークする。
- 高温と低温をできるだけ速く行ったり来たりしたい（Round trip rate を最大化）
- 並列化が簡単。複数の温度での計算を一度に行える。

$$p_{\text{swap}} = \min \left(1, e^{(\beta_{k+1} - \beta_k)(E(x_{k+1}) - E(x_k))} \right)$$



- C. J. Geyer, Interface Proceedings (1991)
- K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. 65, 1604 (1996)

温度間隔の最適化

- 高温と低温をうまく行き来するには温度間隔の最適化が重要

$$\begin{aligned}\log p_{\text{swap}} &= \Delta\beta(E(x_{k+1}) - E(x_k)) \approx \Delta\beta(\langle E \rangle_{\beta_{k+1}} - \langle E \rangle_{\beta_k}) \\ &\approx \Delta\beta^2 \frac{\partial E}{\partial \beta} \Big|_{\beta=\beta_k} = -\left(\frac{\Delta\beta}{\beta}\right)^2 NC(\beta_k)\end{aligned}$$

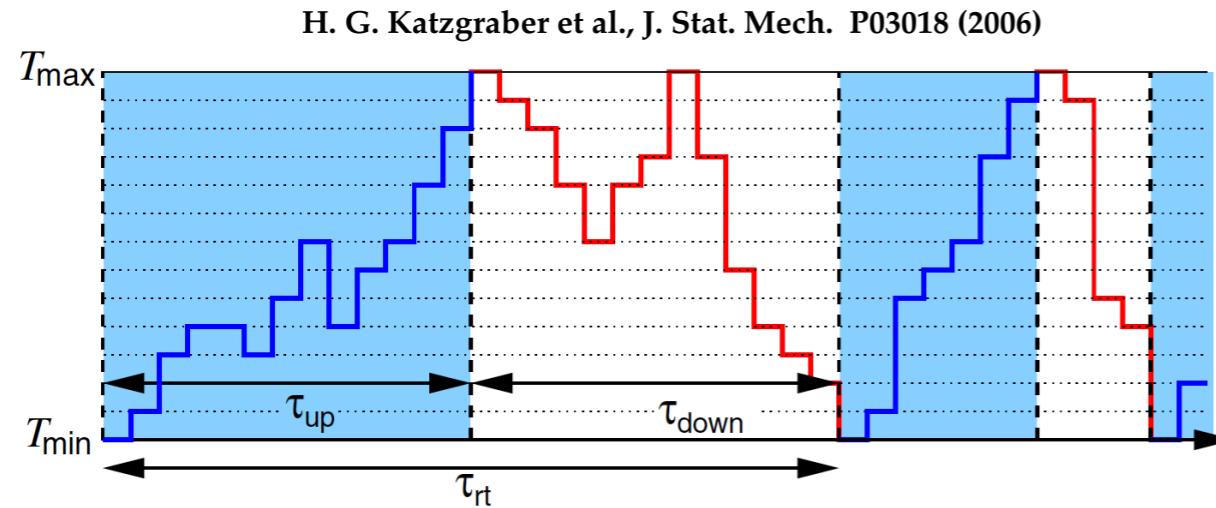
- 一般的に、計算効率を最大化するには交換確率を一定にする

$$\frac{\Delta\beta}{\beta} \sim \frac{1}{\sqrt{NC(\beta)}}$$

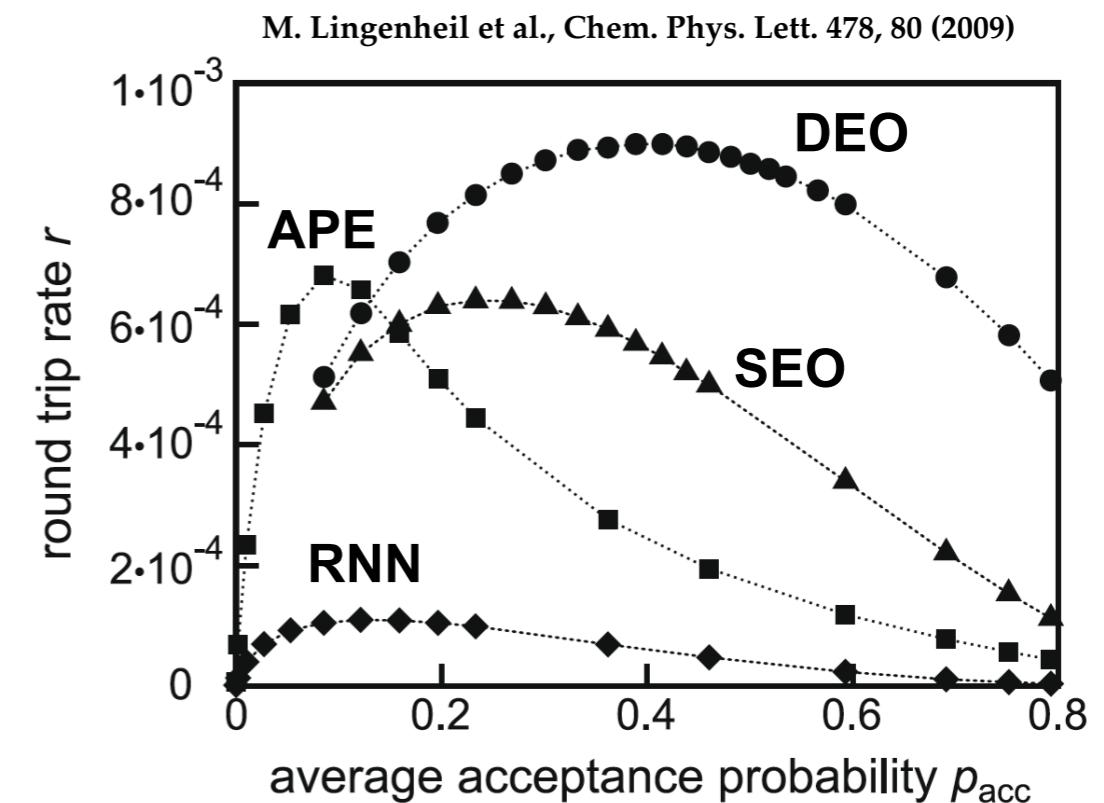
- 系のサイズが大きいほど、温度間隔を狭くする
- 臨界点近くでは温度間隔を小さくする
- テストランで温度間隔を最適化しておく

Round trip rate

- 交換法の計算効率は round trip rate (= round trip time の逆数)で測られる



Oddとevenを交互に繰り返すと効率が高まる
(詳細つりあいの破れ)



幾何級数温度点 \Rightarrow 比熱一定のときに交換確率一定

$$T_i = T_{\min} \alpha^{i-1}$$

$$\alpha(N) = \left(\frac{T_{\max}}{T_{\min}} \right)^{1/(N-1)}$$

エネルギーが正規分布に従うと仮定した場合

$$\pi_i(E) = \frac{1}{\sqrt{2\pi C T_i}} e^{-\frac{(E - C T_i)^2}{2 C T_i^2}}$$

- 温度点を増やすと逆に round trip rate は下がってしまう。

分配関数

- ・ダメな方法

$$\langle e^{\beta E} \rangle = \frac{\sum_x e^{\beta E(x)} e^{-\beta E(x)}}{Z} = \frac{2^N}{Z}$$

$$Z = \frac{2^N}{\langle e^{\beta E} \rangle}$$

- ・ボルツマン重みが指数関数的に小さいときに、指数関数的に大きな値をとる関数の期待値計算 (MCMCの苦手な計算)
- ・Nを大きくすると分散が発散 (ボルツマン分布と一様分布のズレが大きい)

分配関数

- 分布のズレを小さくして、差を集めること（熱力学的積分法）

$$\frac{Z(\beta_2)}{Z(\beta_1)} = \frac{\sum_x e^{-\beta_2 E(x)}}{\sum_x e^{-\beta_1 E(x)}} = \frac{\sum_x e^{(\beta_1 - \beta_2)E(x)} e^{-\beta_1 E(x)}}{\sum_x e^{-\beta_1 E(x)}} = \langle e^{(\beta_2 - \beta_1)E} \rangle_{\beta_1}$$

$$\beta_1 F_{\beta_1} - \beta_2 F_{\beta_2} = \ln \frac{Z(\beta_2)}{Z(\beta_1)}$$

$$\begin{aligned}\beta F_\beta &= (\beta F_\beta - \beta_{K-1} F_{\beta_{K-1}}) + (\beta_{K-1} F_{\beta_{K-1}} - \beta_{K-2} F_{\beta_{K-2}}) + \cdots + (\beta_1 F_{\beta_1} - \beta_0 F_{\beta_0}) + \beta_0 F_{\beta_0} \\ &= \ln \langle e^{(\beta - \beta_{K-1})E} \rangle_{\beta_{K-1}} + \ln \langle e^{(\beta_{K-1} - \beta_{K-2})E} \rangle_{\beta_{K-2}} + \cdots + \ln \langle e^{(\beta_1 - \beta_0)E} \rangle_{\beta_0} + \beta_0 F_{\beta_0} \\ &= \sum_{k=0}^{K-1} \ln \langle e^{(\beta_{k+1} - \beta_k)E} \rangle_{\beta_k}\end{aligned}$$

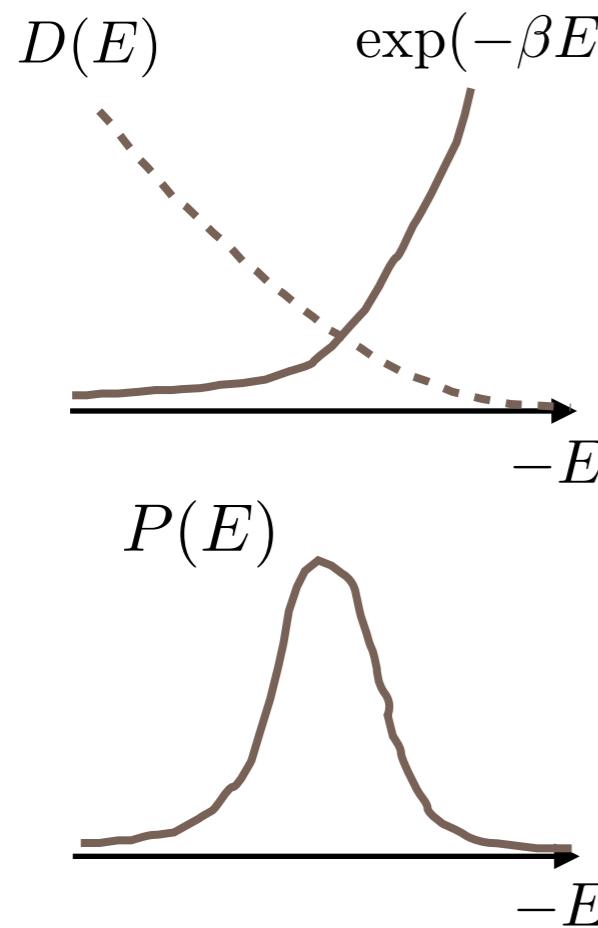
$$0 = \beta_0 < \beta_1 < \cdots < \beta_K = \beta$$

- 分散の大きな推定量に対して一般的に有効

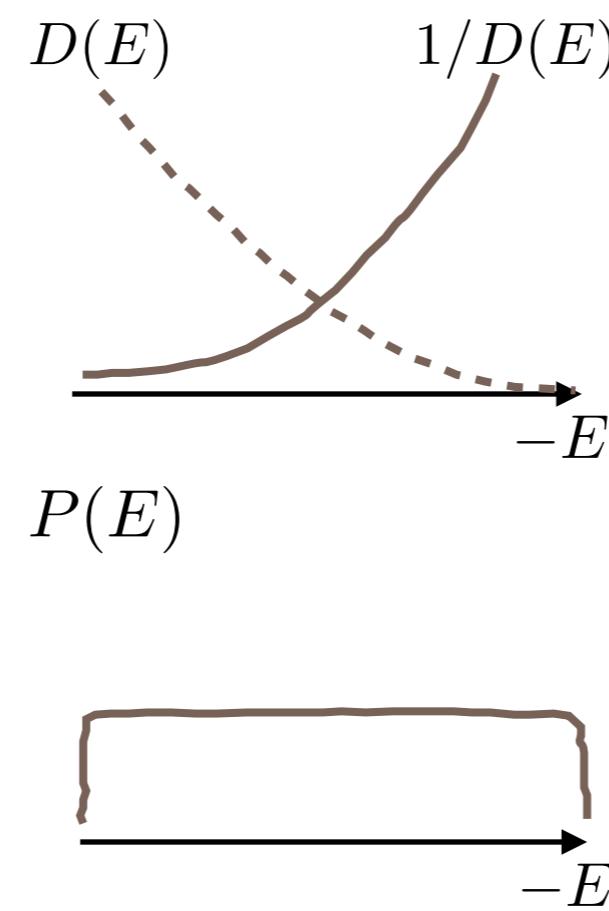
マルチカノニカル法

- ボルツマン分布の代わりに、状態密度の逆数を重みにする。
- エネルギー軸上でランダムウォークさせる。
- 状態密度の良い近似が必要 (Wang-Landauアルゴリズム)。

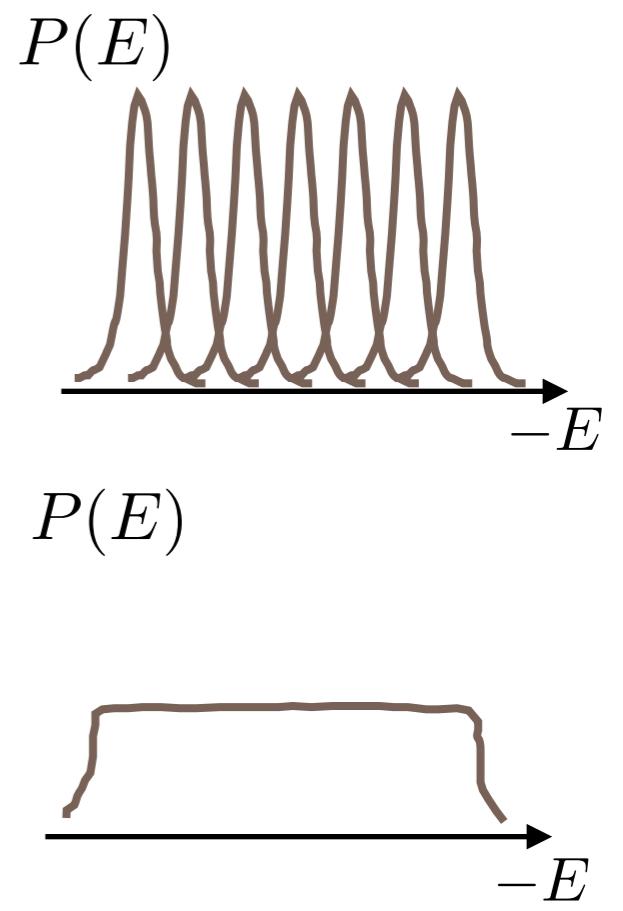
カノニカルアンサンブル



マルチカノニカル法



レプリカ交換法

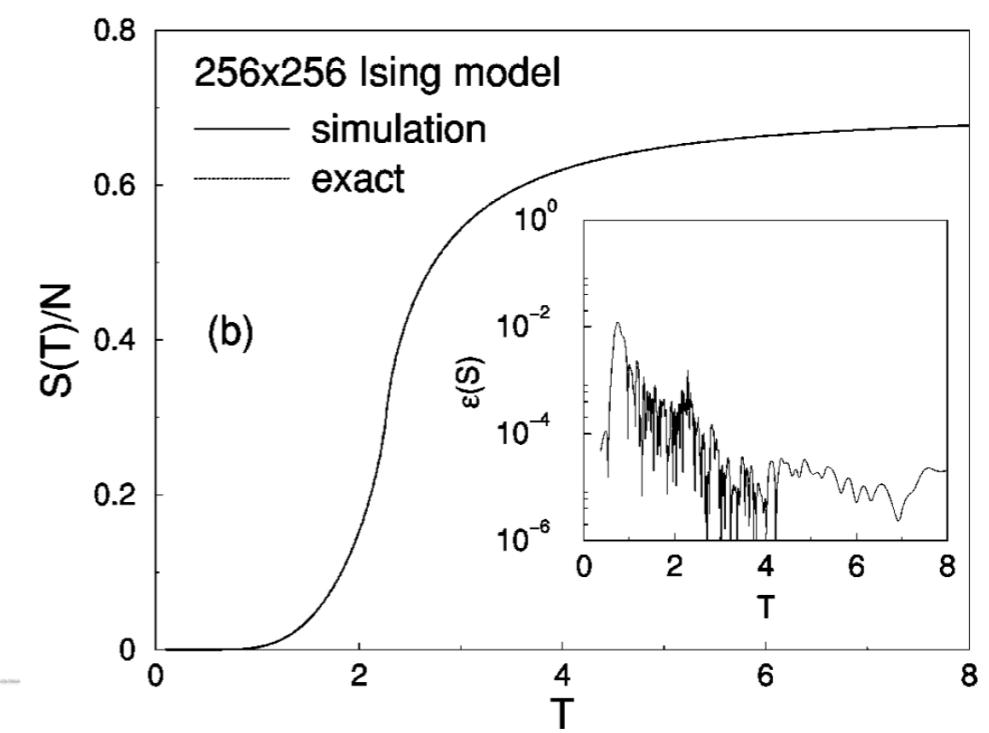
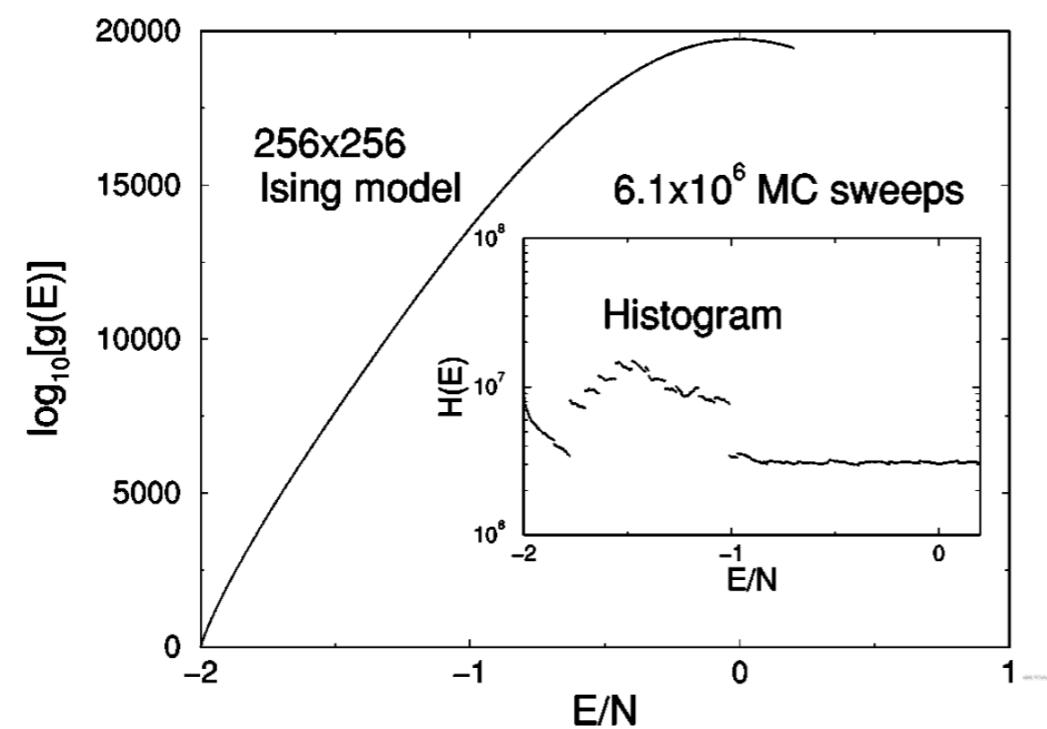


Wang-Landau アルゴリズム

- 状態密度 $g(E)$ を求める手法
- $g(E)=1, H(E)=0$, 修正因子 $f \leftarrow e$ と初期化し、初期状態のエネルギー E を計算
 - 次の状態候補を提案し、エネルギー E' を計算
 - 確率 $p(E \rightarrow E') = \min\left(1, \frac{g(E)}{g(E')}\right)$ で状態遷移（重みは状態密度の逆数）
 - 現在の状態のエネルギー E に対して、 $g(E)$ と $H(E)$ を更新

$$g(E) \leftarrow g(E)f$$

$$H(E) \leftarrow H(E) + 1$$
 - $H(E)$ がほぼフラットになっていたら $H(E)=0$ とリセットし f を更新 $f \leftarrow \sqrt{f}$
 - f が十分 1 に近づいたら終了、そうでなければステップ 1 へ戻る



F. Wang and D. P. Landau,

Phys. Rev. Lett. 86, 2050 (2001), Phys. Rev. E 64, 056101 (2001)

拡張アンサンブル法

- 状態空間を拡張し、緩和の遅い系と速い系をつなぐ手法

	レプリカ交換法	マルチカノニカル法
拡張軸	温度	エネルギー
実装	○	○
並列化	○	△
パラメータ最適化	○ (温度間隔)	△ (状態密度)
分配関数	○	○
連続転移	○	○
一次転移	×	○

MCMCで考えること

- ・ **アンサンブルの決定：状態空間と重み**
拡張アンサンブル法、量子モンテカルロ、リフティング、テンソルネットワーク
- ・ **「良い」状態候補の提案：高い採択率ができるだけ遠くにいく**
クラスターアルゴリズム、ハイブリッド（ハミルトニアン）モンテカルロ、イベント連鎖モンテカルロ、区分決定論的マルコフ過程
- ・ **遷移確率の最適化**
メトロポリス法、熱浴法（ギブスサンプラー）、諏訪・藤堂法、線形計画法
- ・ **確率的イベント発生の効率化**
二分探索、ウォーカーの別名法、イベントドリブン（N-fold way）、thinning
- ・ **推定量の選択**
分配関数、相関長、有限サイズスケーリング、確率近似法、improved estimator

MCMCの条件

- 次の2つの条件を課す

• S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer (1993)
• C. P. Robert and G. Casella, *Monte Carlo statistical methods*, Springer (2004)
• 鎌谷研吾他、モンテカルロ統計計算、講談社

全体つりあい (Global balance) : 欲しい分布が不变分布

$$\pi_i = \sum_j \pi_i p_{ij} = \sum_j \pi_j p_{ji} \quad \forall i$$

エルゴード性 (Ergodicity) : どんな初期分布からでも分布が収束 (離散変数では既約+非周期、ペロン・フロベニウスの定理 (Perron-Frobenius))

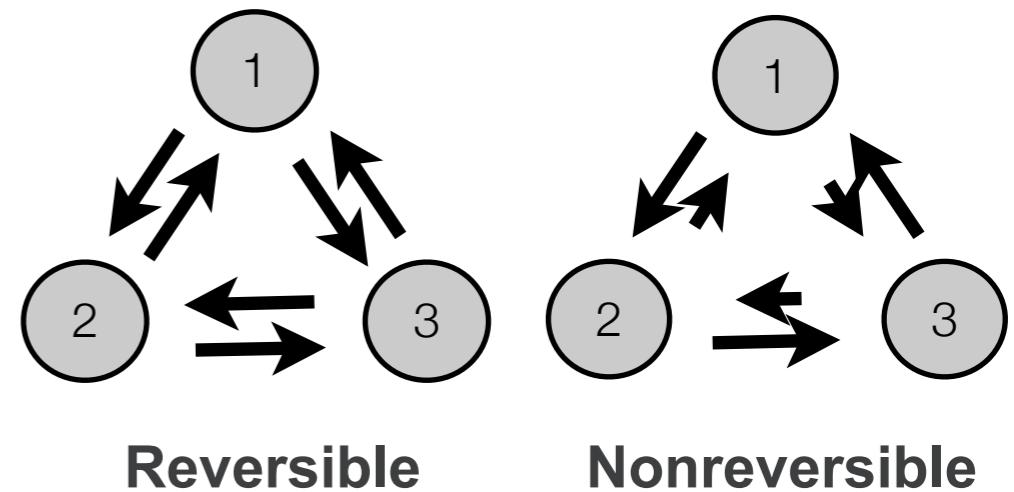
$$\exists k \in \mathbb{N} \text{ s. t. } P^k > 0$$

- 詳細つりあい (Detailed balance, reversibility)

全体つりあいの十分条件 (必要条件ではない)

シンプルかつ実用的

$$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i, j$$



- 詳細つりあいを破って計算効率を改善することが可能 (講義後半)

順序更新では詳細つりあいが破れる

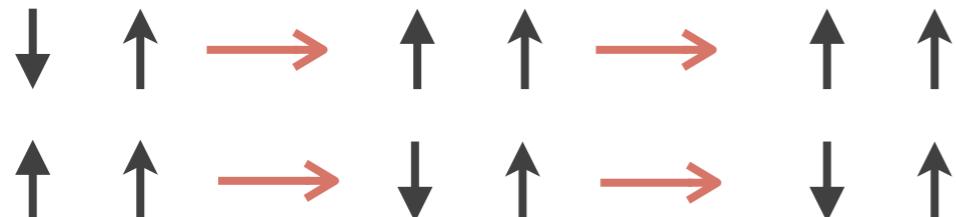
Metropolis, Rosenbluth, Rosenbluth, Teller, Teller, JCP 21, 1087 (1953)

This we do as follows: We place the N particles in any configuration, for example, in a regular lattice. Then we move each of the particles in succession according to the following prescription:

$$\begin{aligned} X &\rightarrow X + \alpha \xi_1 \\ Y &\rightarrow Y + \alpha \xi_2, \end{aligned} \tag{3}$$

2イジングスピニの順序更新の例

毎ステップ左のスピニから順番に更新する



$$p_{i \rightarrow j} = \frac{\pi_j}{\sum_k \pi_k}$$

熱浴法（ギブスサンプラー）

$$(1-p) \times p \times p = p^2(1-p)$$



$$p \times (1-p) \times (1-p) = p(1-p)^2$$

$$\boxed{\pi_i p_{i \rightarrow j} \neq \pi_j p_{j \rightarrow i}}$$

重み

$$\begin{array}{ccc} \uparrow & \uparrow & p \\ \downarrow & \uparrow & 1-p \end{array} \quad p > 1/2$$

メトロポリス法

$$p_{i \rightarrow j} = \frac{1}{n-1} \min \left(1, \frac{\pi_j}{\pi_i} \right)$$

$$(1-p) \times 1 \times \left(1 - \frac{1-p}{p} \right) = \frac{(1-p)(2p-1)}{p}$$

$$\begin{array}{c} \times \\ p \times \frac{1-p}{p} \times 0 = 0 \end{array}$$

全体つりあいは満たされる

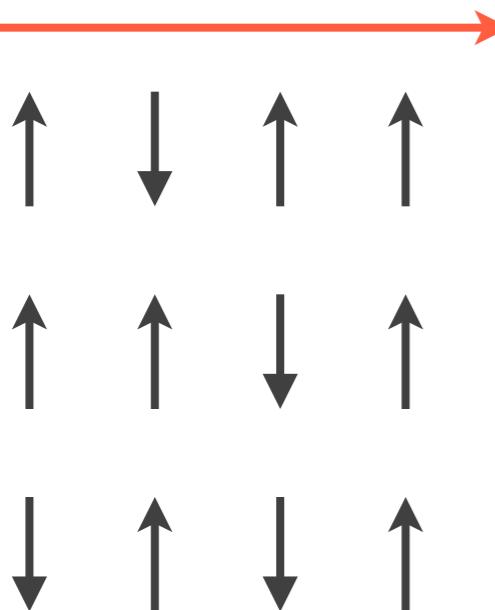
V. I. Manousiouthakis and M. W. Deem,
J. Chem. Phys. 110, 2753 (1999)

局所詳細つりあい

$$\pi_{k_{\ell-1}} P_{k_{\ell-1} k_\ell}^{(\ell)} = \pi_{k_\ell} P_{k_\ell k_{\ell-1}}^{(\ell)}$$

$$j (= k_0) \rightarrow k_1 \rightarrow k_2 \rightarrow \cdots \rightarrow k_{N-1} \rightarrow i (= k_N)$$

毎ステップ同じ順序で更新



$$\begin{aligned} & \sum_{jk_1 \dots k_{N-1}} \underline{\pi_j P_{jk_1}^{(1)}} P_{k_1 k_2}^{(2)} P_{k_2 k_3}^{(3)} \dots P_{k_{N-1} i}^{(N)} \\ &= \sum_{jk_1 \dots k_{N-1}} \underline{P_{k_1 j}^{(1)}} \underline{\pi_{k_1} P_{k_1 k_2}^{(2)}} P_{k_2 k_3}^{(3)} \dots P_{k_{N-1} i}^{(N)} \\ &= \sum_{jk_1 \dots k_{N-1}} \underline{P_{k_1 j}^{(1)}} \underline{P_{k_2 k_1}^{(2)}} \underline{\pi_{k_2} P_{k_2 k_3}^{(3)}} \dots P_{k_{N-1} i}^{(N)} \\ &= \sum_{jk_1 \dots k_{N-1}} \underline{P_{k_1 j}^{(1)}} \underline{P_{k_2 k_1}^{(2)}} \underline{P_{k_3 k_2}^{(3)}} \dots \underline{P_{i k_{N-1}}^{(N)}} \pi_i \\ &= \sum_j \left(P^{(N)} P^{(N-1)} P^{(N-2)} \dots P^{(1)} \right)_{ij} \pi_i \\ &= \pi_i \end{aligned}$$

= 1

状態 i に入ってくる確率流の合計 = 状態 i の重み

詳細つりあいを破ると収束が速まる

- ランジュバンダイナミクスに流れ場を加えると定常分布への収束が速まる

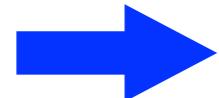
C. R. Hwang et al., Ann. Appl. Probab. 15, 1433 (2005)

M. Ohzeki et al., Phys. Rev. E 92, 012105 (2015)

A. B. Duncan et al., J. Stat. Phys. 163, 457 (2016)

Overdamped Langevin equation:

$$dX_t = \nabla \log \pi(X_t) dt + \sqrt{2} dW_t$$



$$dX_t = (\nabla \log \pi(X_t) + \boxed{\alpha \gamma(X_t)}) dt + \sqrt{2} dW_t$$

Divergence free

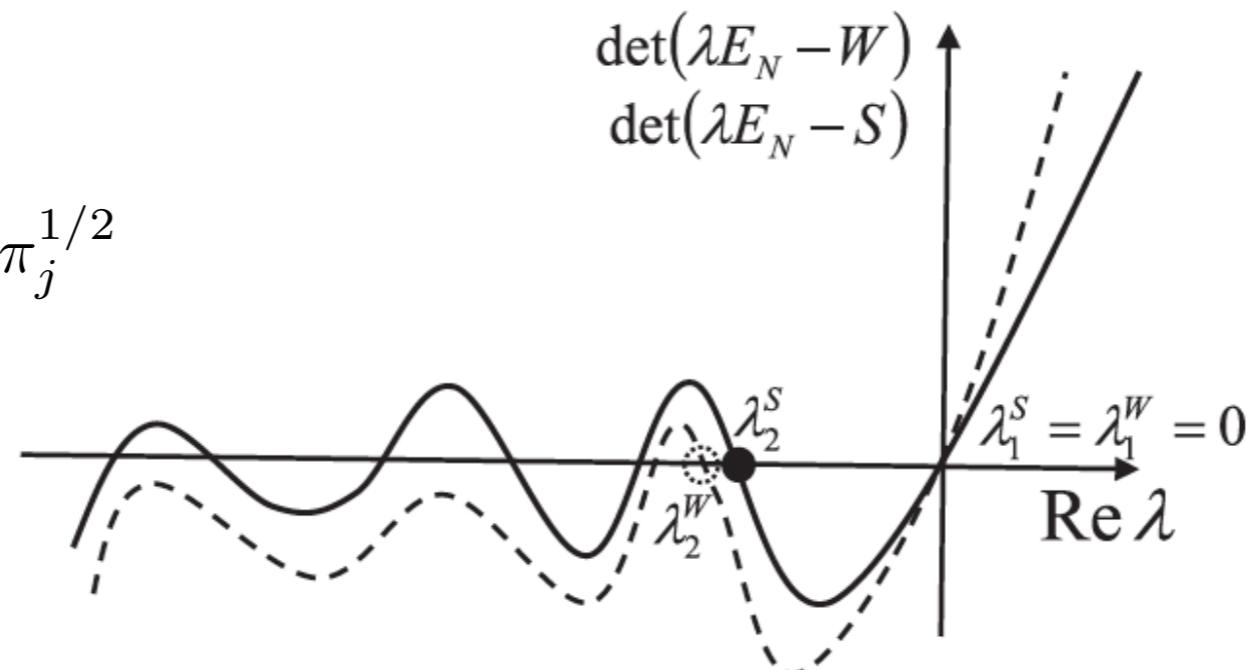
$$\nabla \cdot (\gamma \pi) = 0$$

- 離散変数系で第2固有値は必ず小さくなる

A. Ichiki and M. Ohzeki, Phys. Rev. E 88, 020101(R) (2013)
Y. Sakai and K. Hukushima, Phys. Rev. E 93, 043318 (2016)

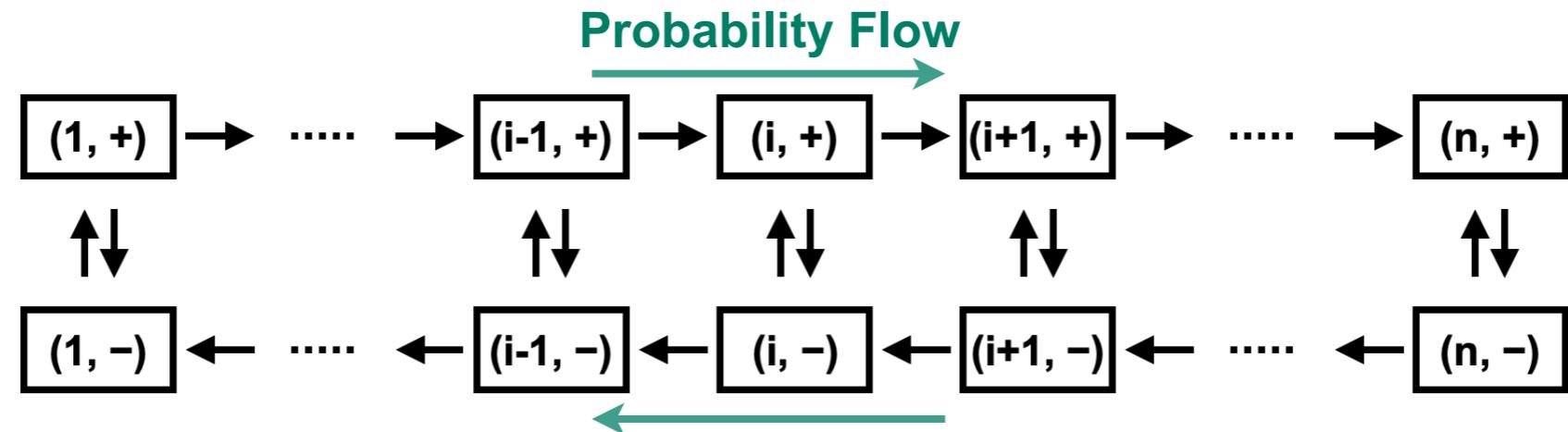
$$W = S + \Gamma$$

$$W_{ij} = \pi_i^{-1/2} q_{ij} \pi_j^{1/2}$$

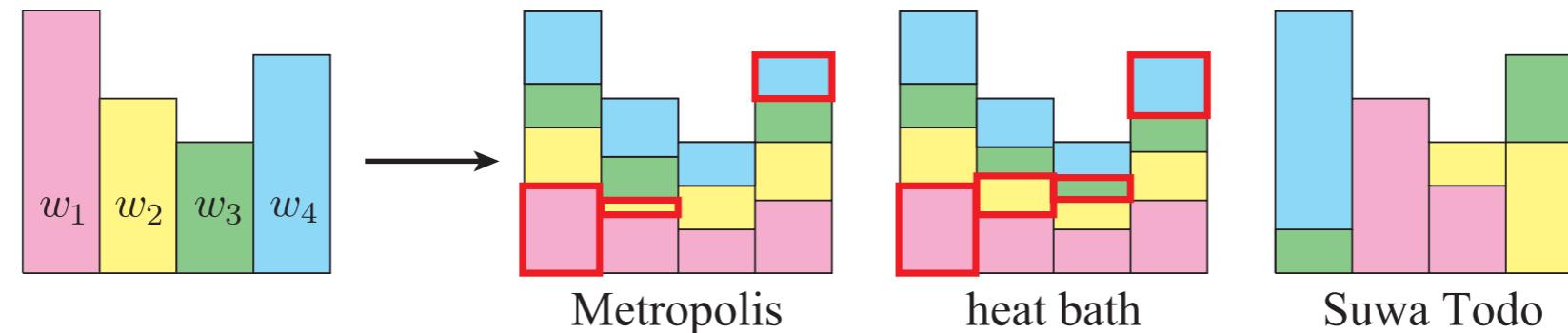


不可逆(Nonreversible)モンテカルロの構成法

- リフティング = 状態空間を拡張し、拡張空間で確率の流れを導入する手法



- 詳細つりあいに縛られない遷移確率の最適化

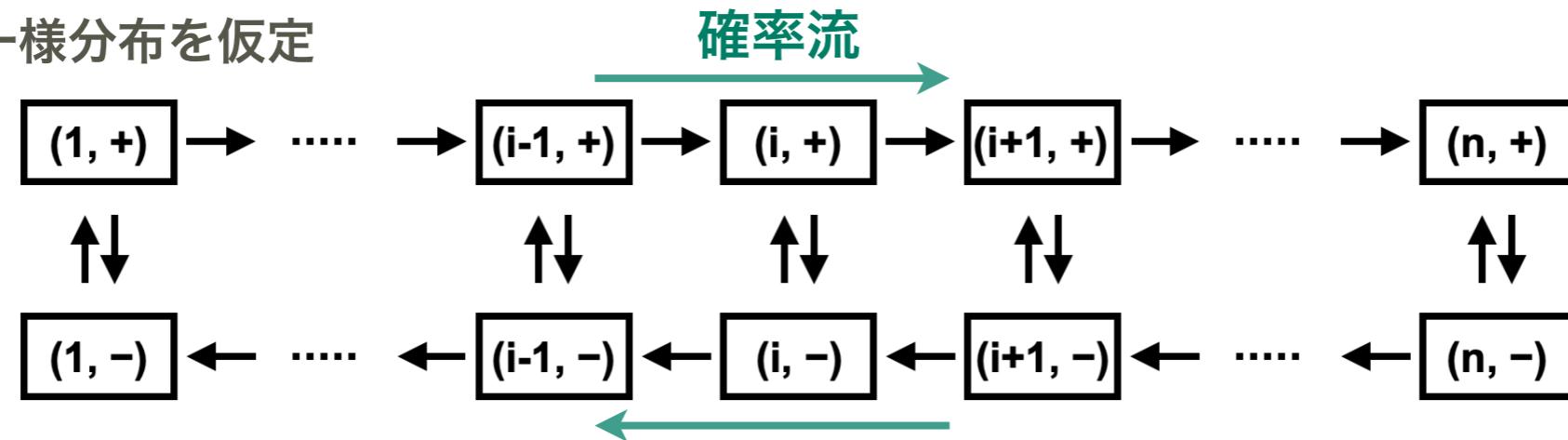


- これらのアプローチを組み合わせることも可能
 - 有向ワームアルゴリズム
 - マルチレプリカ交換モンテカルロ

リフティング (Lifting)

- 拡張空間に確率の流れを導入する

n 状態空間上の一様分布を仮定



P. Diaconis, S. Holmes, and R. M. Neal,
Ann. Appl. Probab. 10, 726 (2000)

F. Chen, L. Lovász, and I. Pak, Proceedings of the 31st Annual
ACM Symposium on Theory of Computing, 275 (1999)

$$t_{\text{mix}} \sim n^2$$

$$t_{\text{mix}} \sim n$$

かき混ぜると速く混ざる



- ハイブリッド（ハミルトニアン）モンテカルロ

Duane–Kennedy–Pendleton–Roweth (1987)

- イベント連鎖モンテカルロ

Bernard–Krauth–Wilson (2009), Nishikawa–Michel–Krauth–Hukushima (2015)

- エネルギー軸でのリフティング

Turitsyn–Chertkov–Vucelja (2011), Fernandes–Weigel (2011)

- 不可逆シミュレーテッドテンパリング

Hukushima–Sakai (2013), Sakai–Hukushima (2016)

- 不可逆レプリカ交換モンテカルロ

Wu (2017), Syed–Côté–Deligiannidis–Doucet (2021)

ねじれ詳細つりあい

$$\pi_{(i,+)} p_{(i,+)} \rightarrow (i+1,+)} = \pi_{(i+1,-)} p_{(i+1,-)} \rightarrow (i,-)}$$

次数の半減

- リフティングによって収束時間の次数を半減させることが可能

平均場（全結合）模型 $H = -\frac{J}{2N} \sum_{ij} \sigma_i \sigma_j$

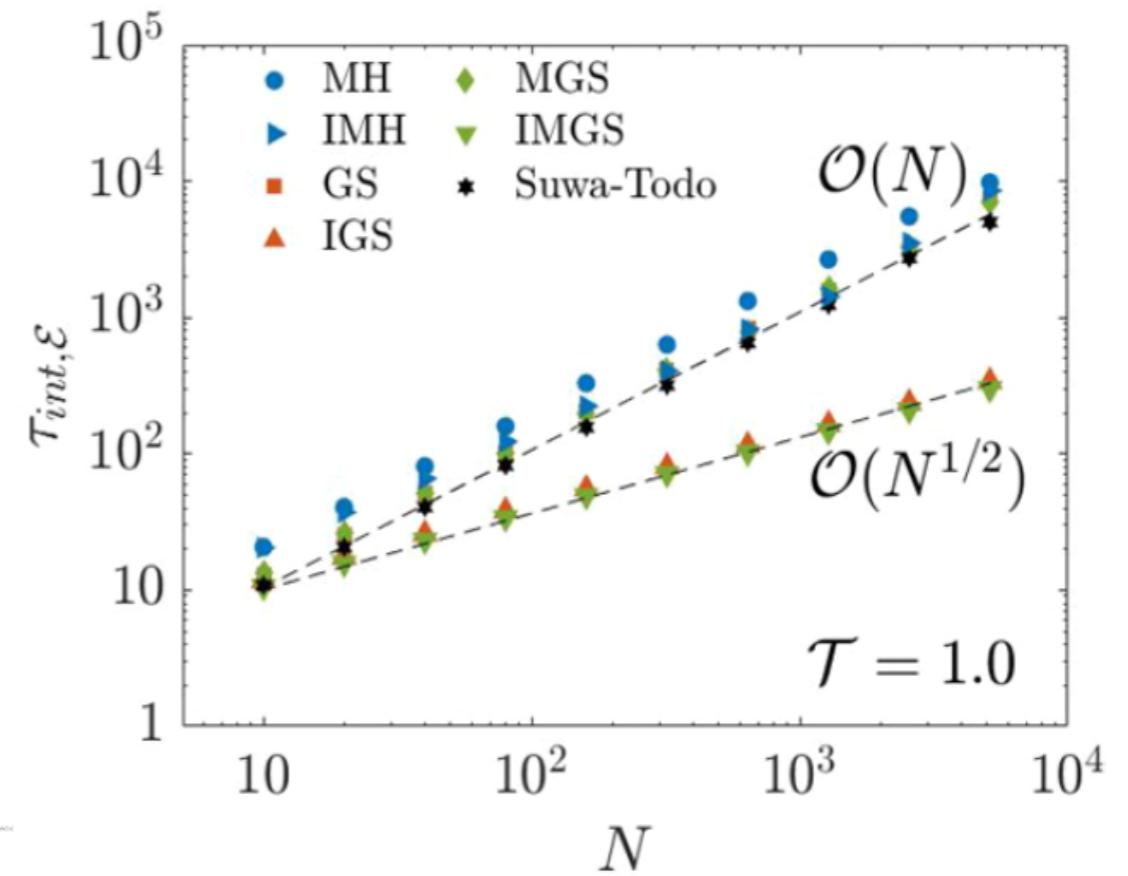
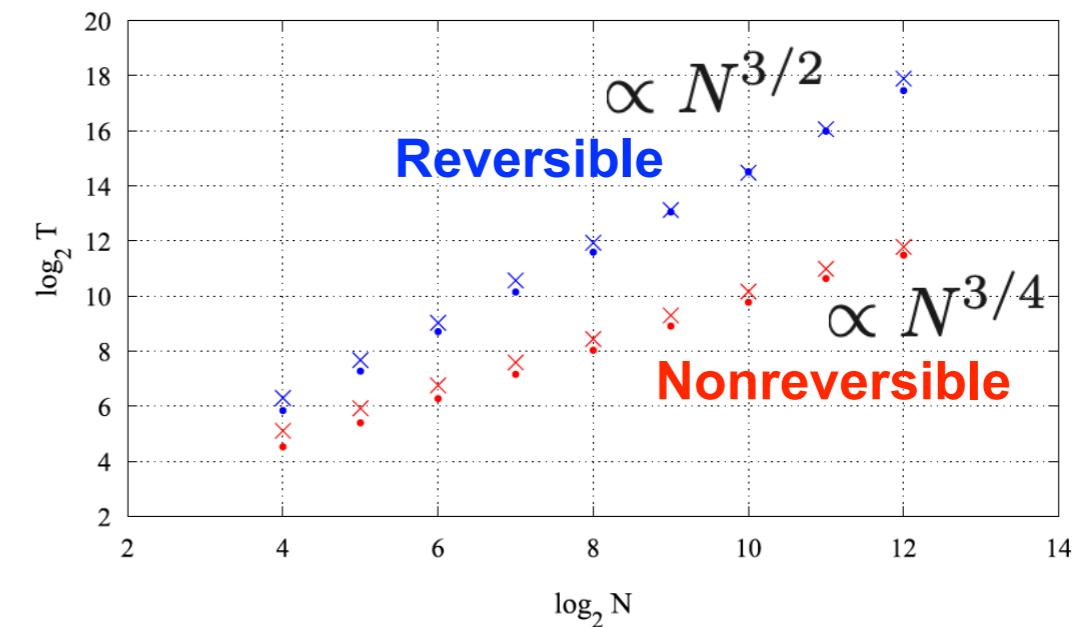
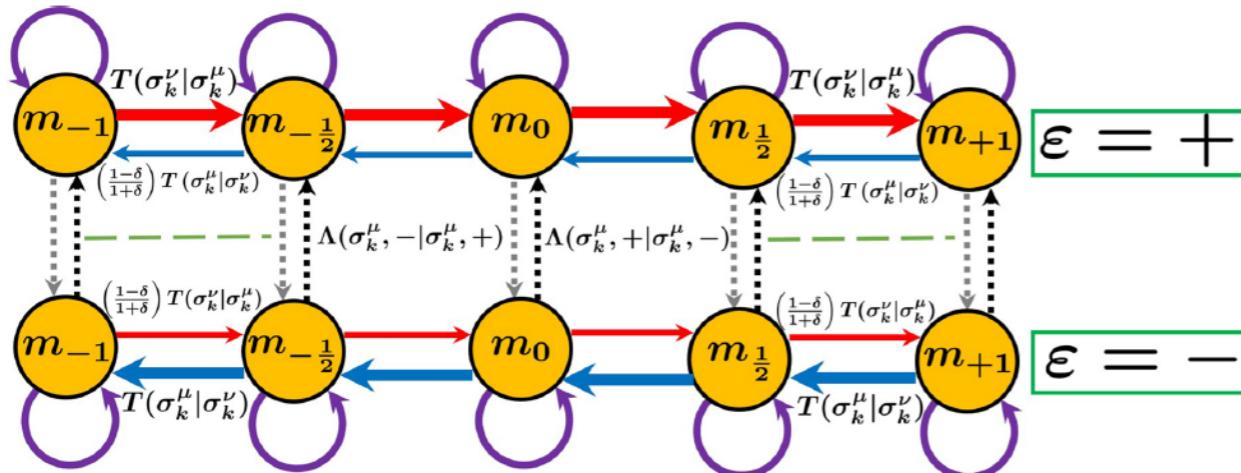
K. S. Turitsyn et al., Physica D 240, 410 (2011)

$$H = - \sum_{\langle i,j \rangle} \delta_{\sigma_i \sigma_j} \quad \sigma_i = 1, 2, \dots, q$$

1次元強磁場ポツツ模型

F. Faizi et al., J. Chem. Theory Comput. 16, 2124 (2020)

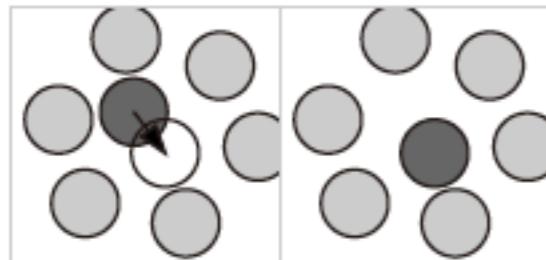
「磁化」軸でのリフティング



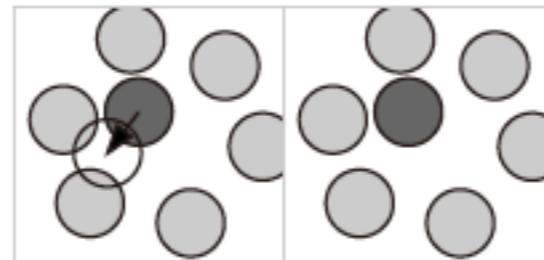
イベント連鎖モンテカルロ

- 剛体球モデルにおける棄却のない状態更新法

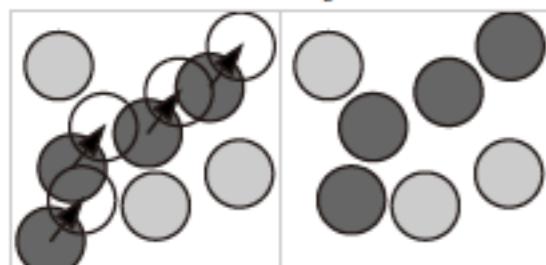
Bernard, Krauth, Wilson (2009)



t_i $t_f = t_i + 1$



t_i $t_f = t_i + 1$



t_i $t_f = t_i + 4$

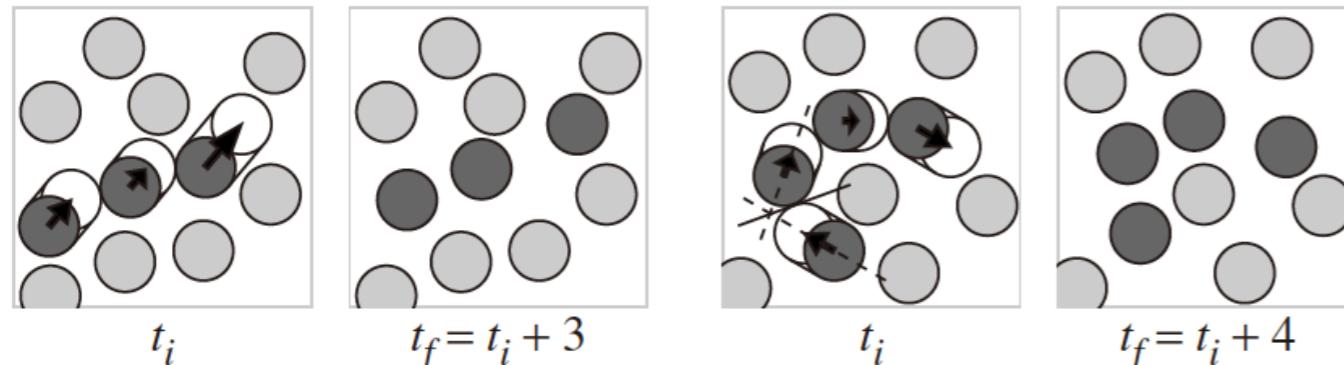


t_i $t_f = t_i + 2$

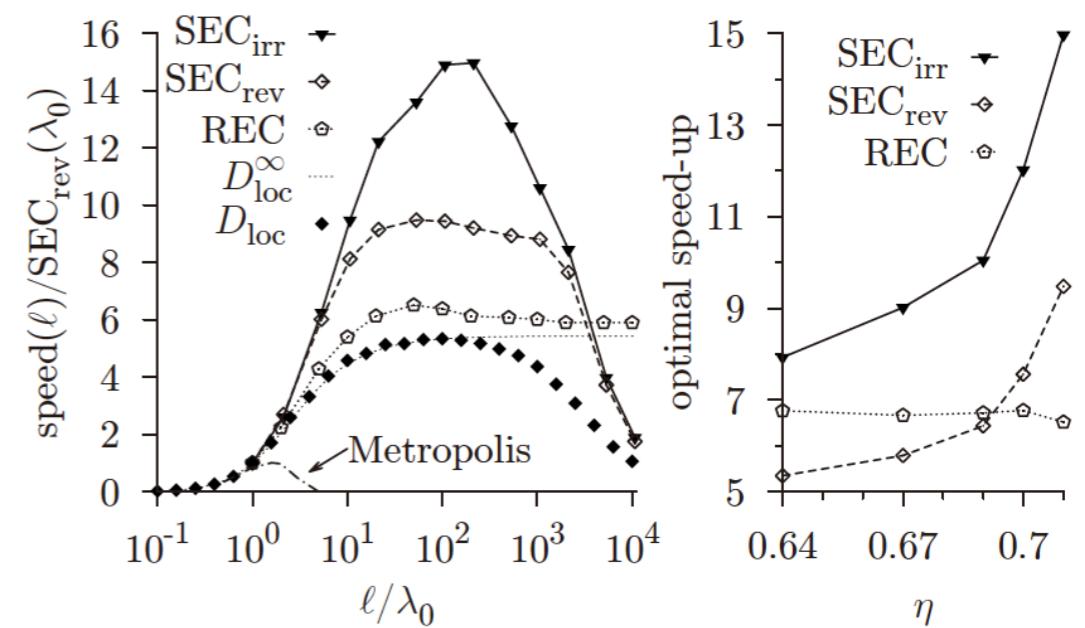
Simple Metropolis

Jaster's chain algorithm

Event-chain algorithm



Generalization: Michel, Kapfer, Krauth (2014)

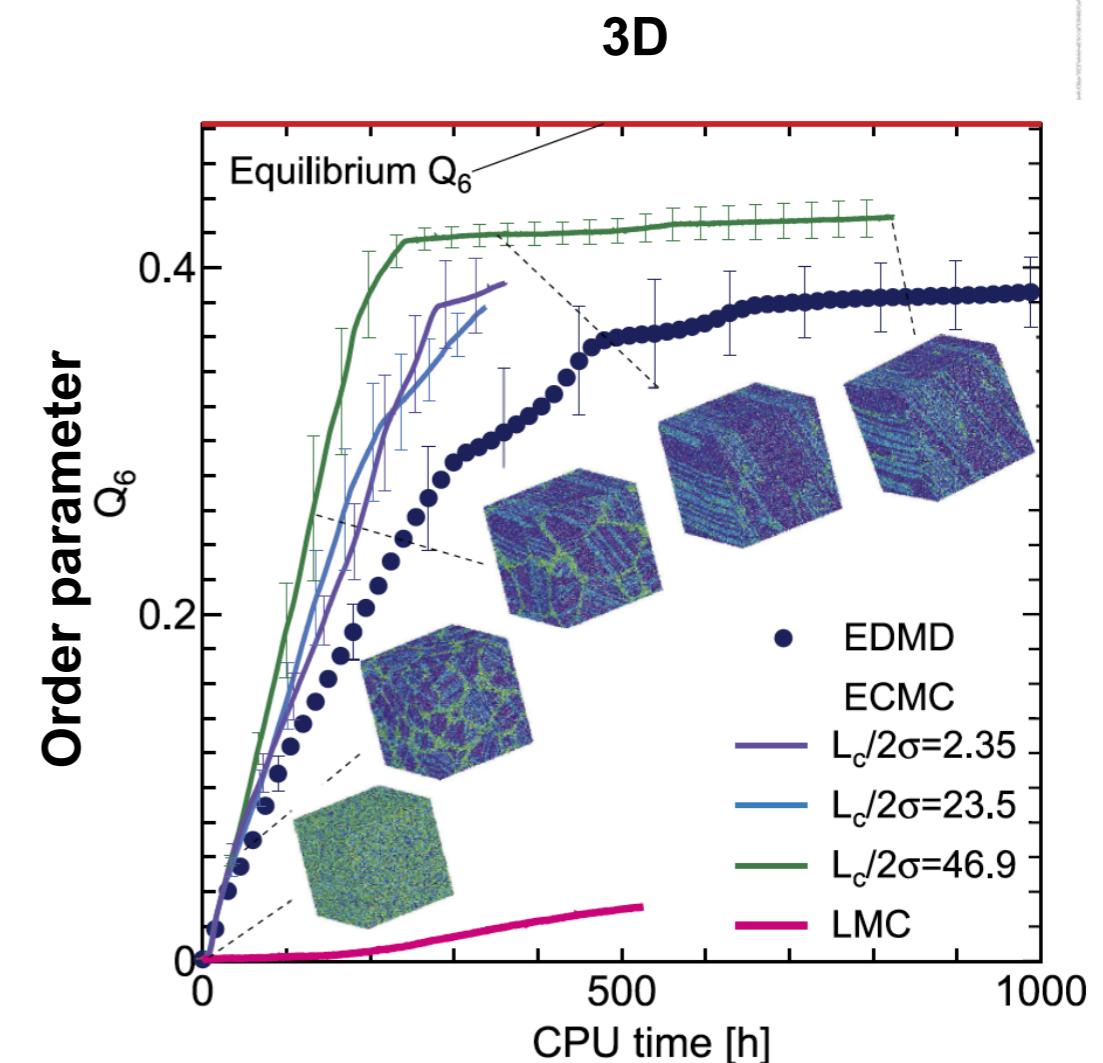
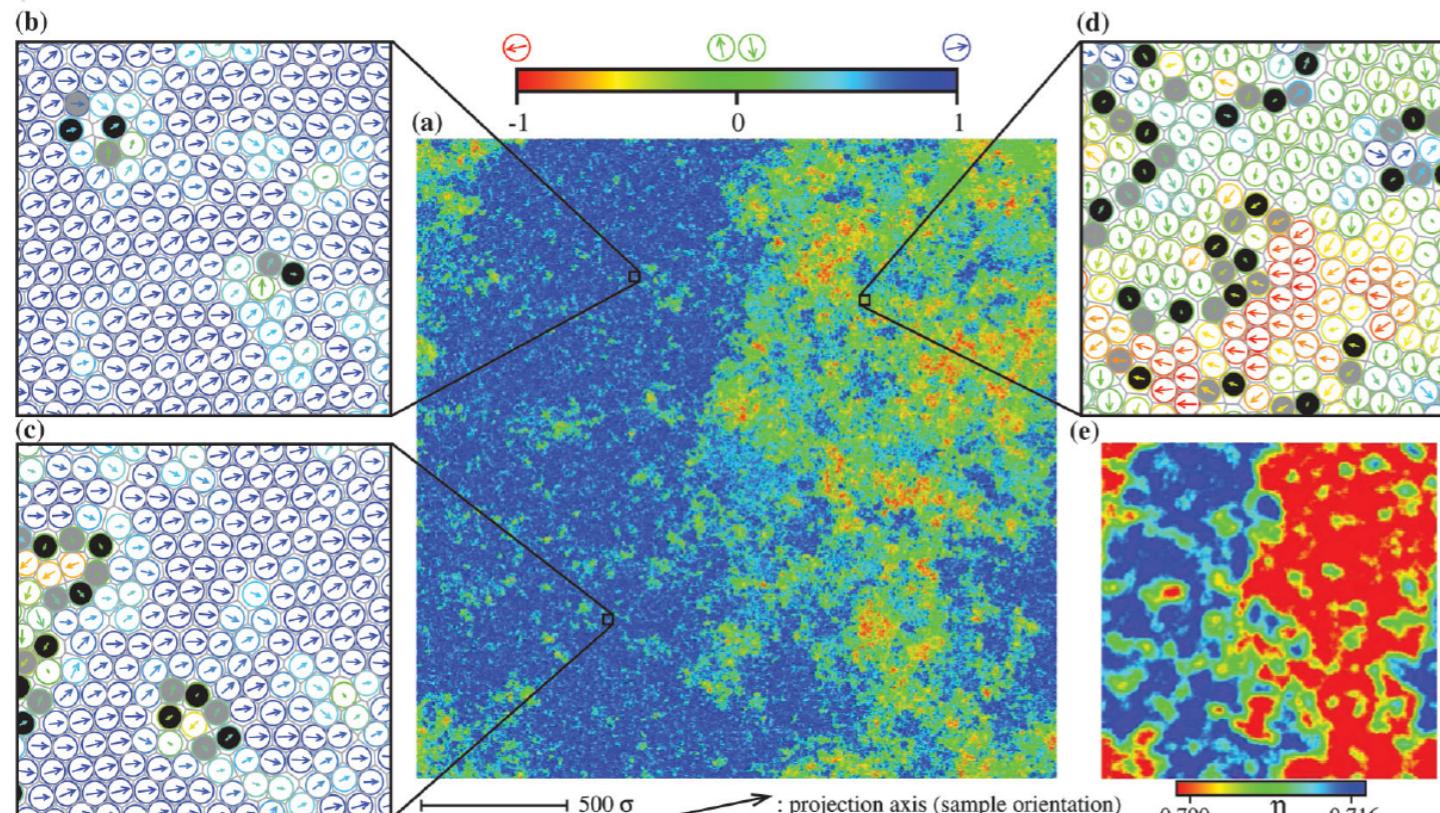


イベント連鎖モンテカルロ

- イベント連鎖モンテカルロはイベントドリブン分子動力学より収束の速いダイナミクスを実現

E. P. Bernard and W. Krauth, Phys. Rev. Lett. 107, 155704 (2011)
M. Isobe and W. Krauth J. Chem. Phys. 143, 084509 (2015)

ヘキサティック状態と液体状態の共存相 (2D)



一般化イベント連鎖モンテカルロ

- 一般的な2体相互作用系に拡張可能

メトロポリスフィルター

$$p^{\text{Met}}(x \rightarrow x') = \min \left[1, \prod_{i < j}^N e^{-\Delta V_{ij}} \right]$$

$$\min(1, e^{-\Delta V}) = e^{-[\Delta V]^+} \rightarrow 1 - [dV]^+$$

$$[\Delta V]^+ = \max(0, \Delta V)$$

$$p(s)ds = e^{-\int_0^{E^*} [dV_{ij}]^+ ds} dV_{ij}$$

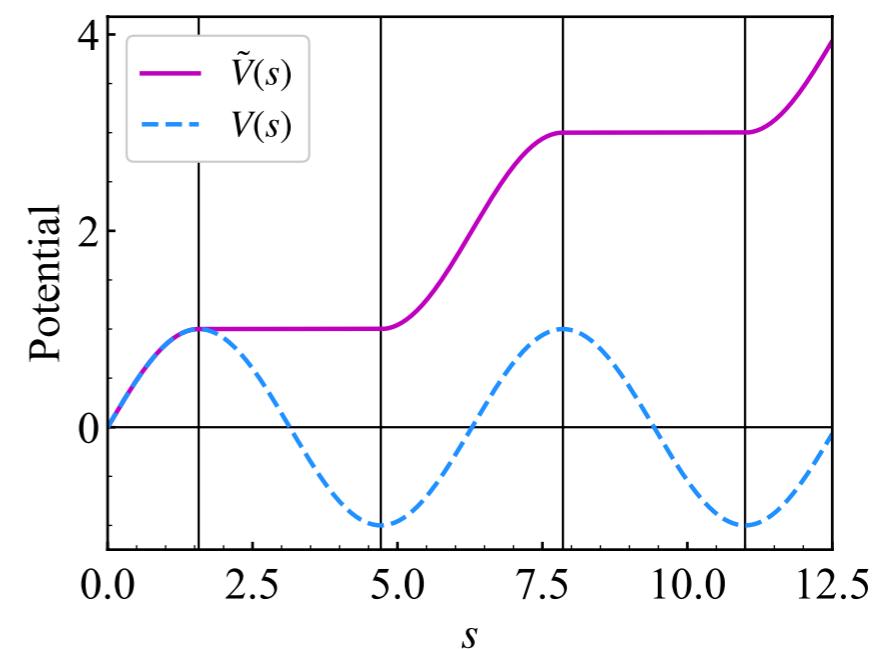
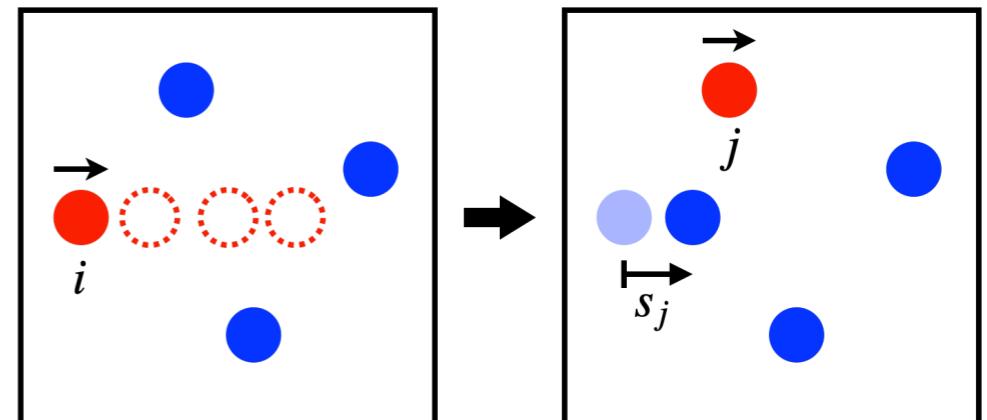
$$E^* = \int_0^{E^*} [dV_{ij}]^+ = \int_0^{s_j} \left[\frac{\partial V_{ij}}{\partial s} \right]^+ ds =: \tilde{V}(s_j)$$

$$E^* \sim e^{-E^*}$$

E. A. J. F. Peters and G. de With, Phys. Rev. E 85, 026703 (2012)
M. Michel et al., J. Chem. Phys. 140, 054116 (2014).

因子化メトロポリスフィルター

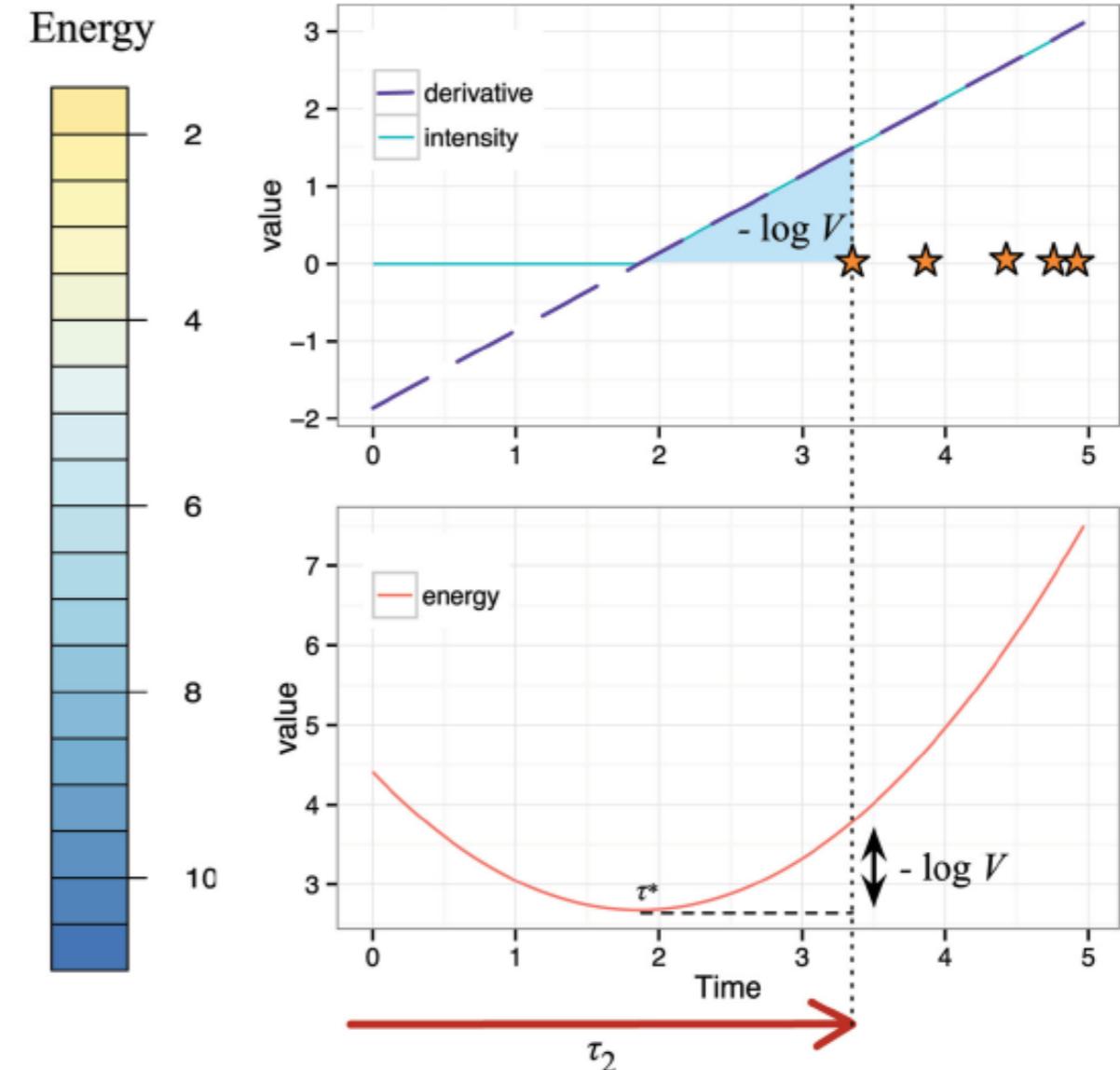
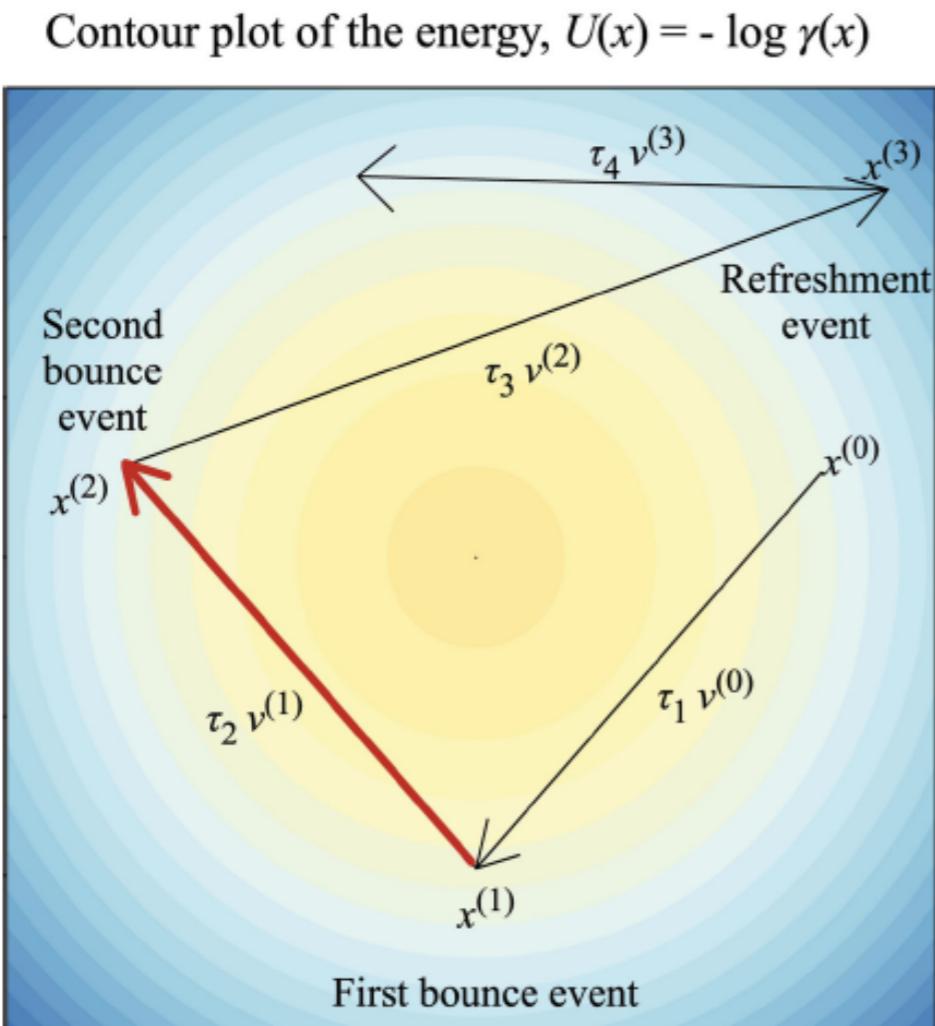
$$p^{\text{Fact}}(x \rightarrow x') = \prod_{i < j}^N \min \left[1, e^{-\Delta V_{ij}} \right]$$



跳ね返り粒子サンプリング

- 連続時間の不可逆区分決定論的マルコフ過程で時間発展
- メトロポリス法とは全く異なるサンプリング法

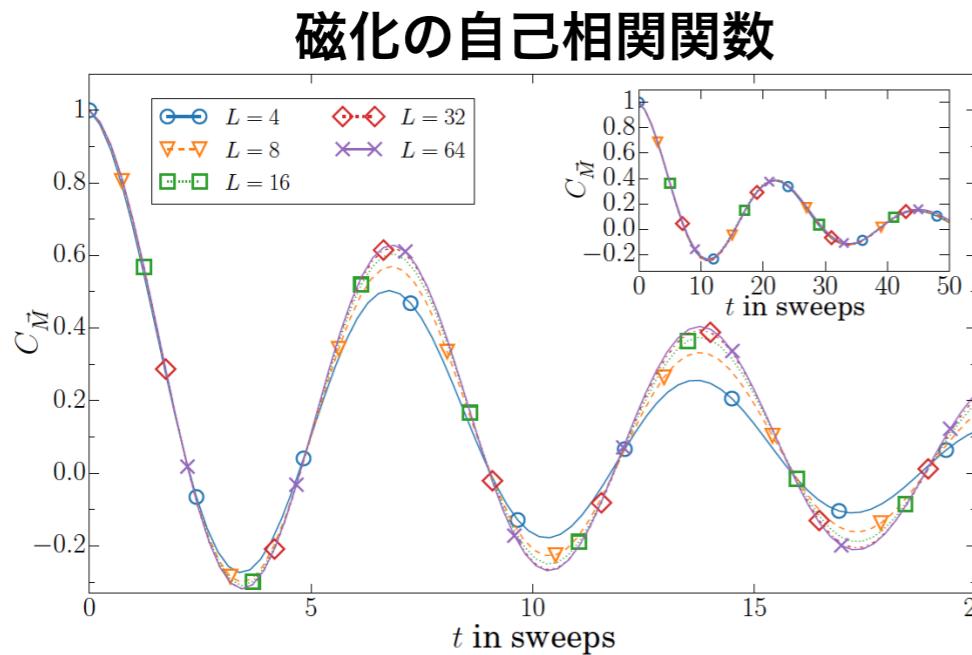
• E. A. J. F. Peters and G. de With, Phys. Rev. E 85, 026703 (2012)
• A. Bouchard-Côté, J. Am. Stat. Assoc. 113, 855 (2018)



- イベント時間の逆引きを効率的に行えれば、ハイブリッド（ハミルトニアン）モンテカルロより効率的になりうる

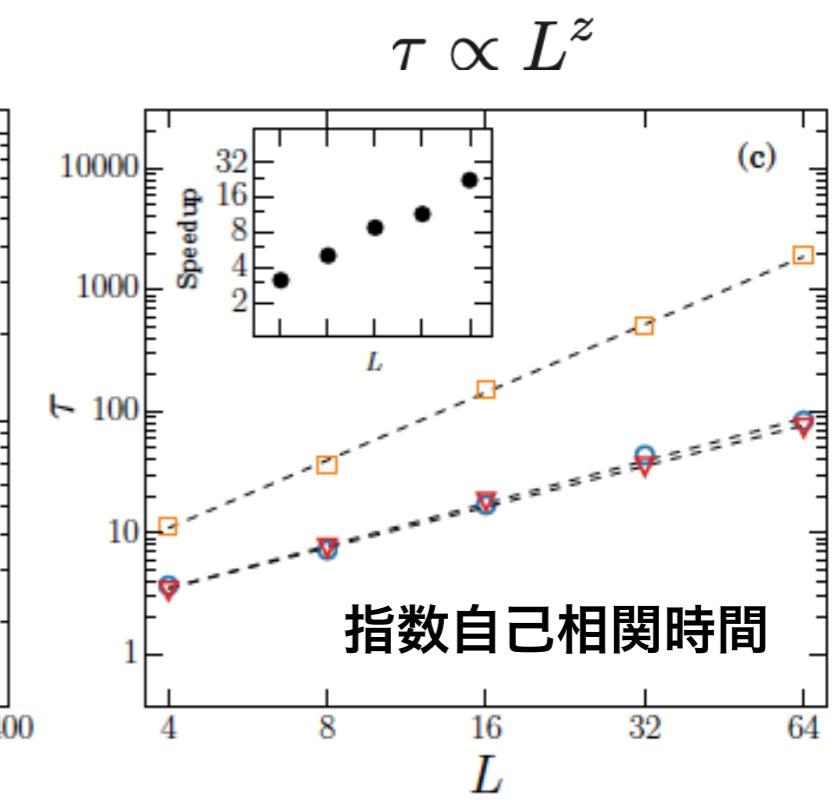
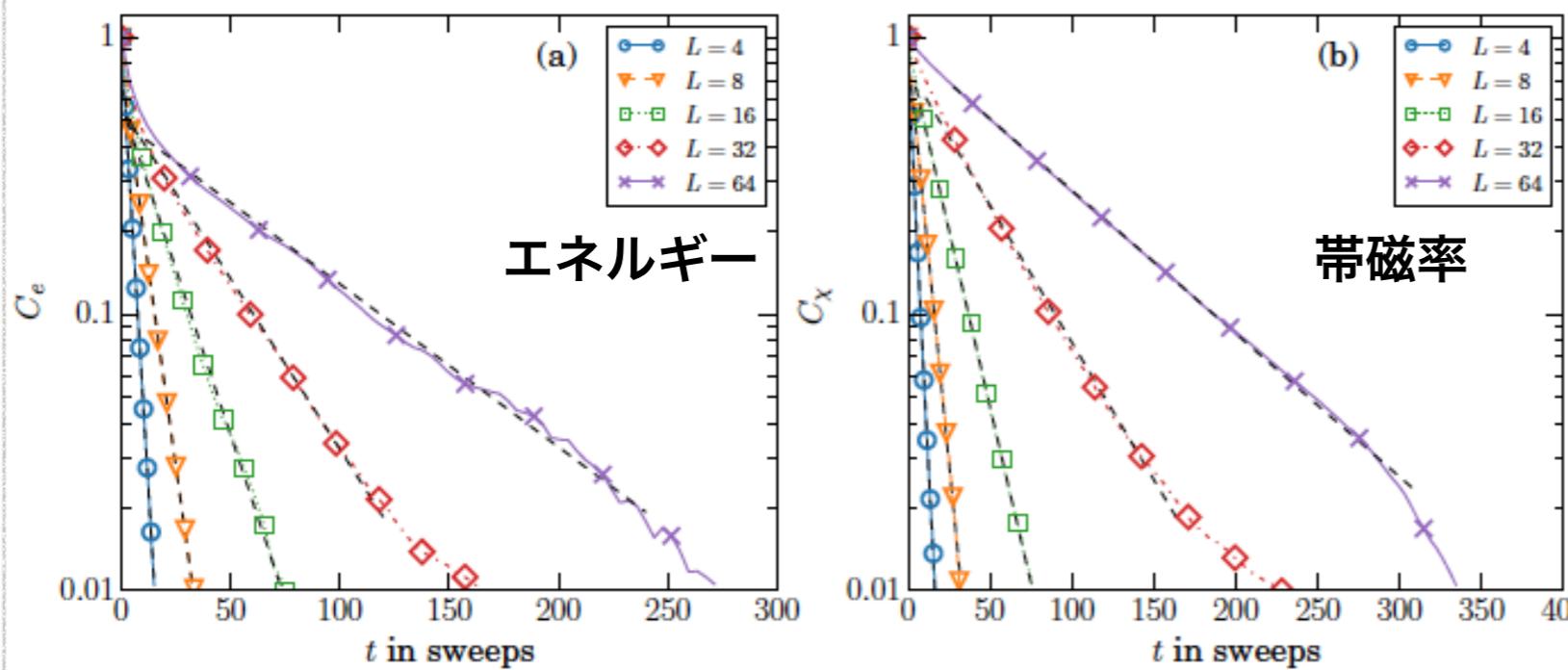
スピニ系でのイベント連鎖

- 3次元強磁性ハイゼンベルグ模型において動的臨界指数（自己相関時間の次数）がほぼ半減



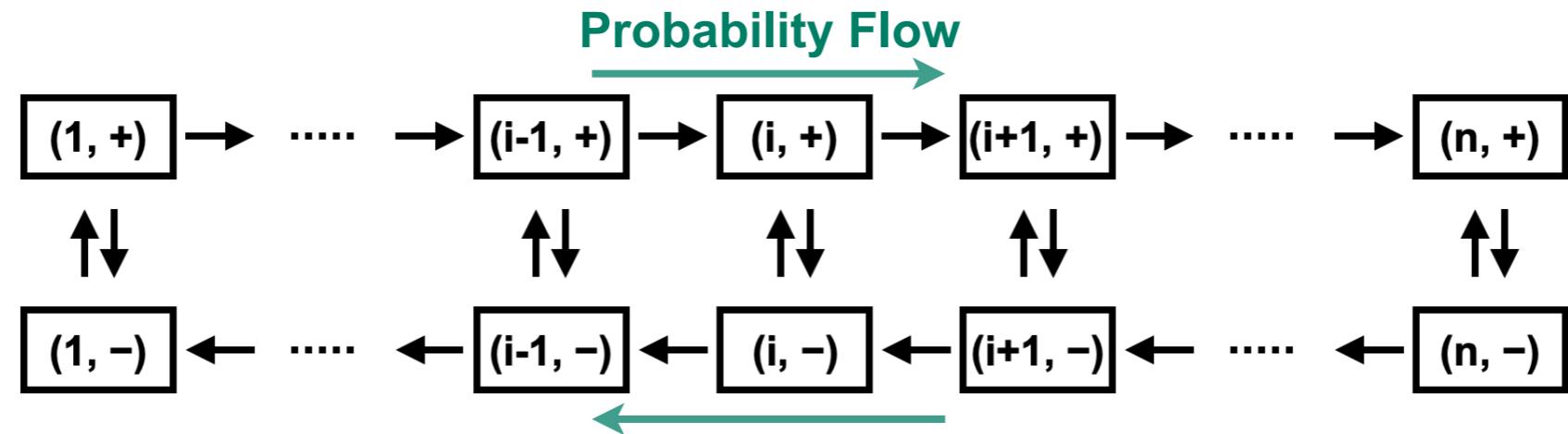
- M. Michel et al, EPL 112, 20003 (2015)
- Y. Nishikawa et al., Phys. Rev. E 92, 063306 (2015)

- スピニ波を効率的にサンプリング
- XY模型ではボルテックス位置の更新が難しい

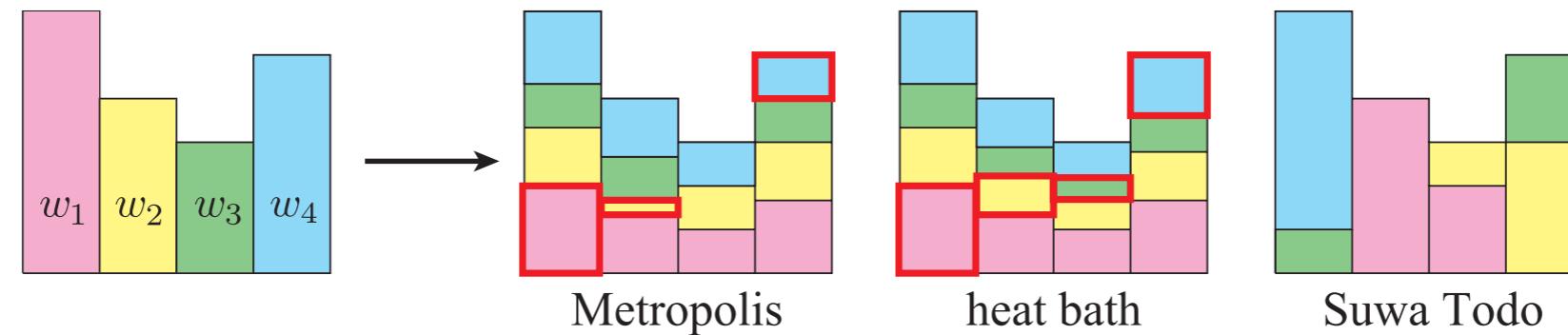


不可逆(Nonreversible)モンテカルロの構成法

- リフティング = 状態空間を拡張し、拡張空間で確率の流れを導入する手法



- 詳細つりあいに縛られない遷移確率の最適化



- これらのアプローチを組み合わせることも可能

- 有向ワームアルゴリズム
- マルチレプリカ交換モンテカルロ

Peskunの定理

2つの遷移行列を比べて、全ての非対角要素が大きいか等しいとき $P_2 \geq P_1$ と定める。2つの遷移行列が詳細つりあいを満たすとき、任意の関数に対して

$$v(f, P_1) \geq v(f, P_2)$$

$$\sigma^2 \simeq \frac{v}{M}$$

つまり P_2 を用いた方が必ず効率が良い。

- P. H. Peskun. Biometrika 60, 607 (1973)
- R. M. Neal and J. S. Rosenthal, arXiv:2305.18268

簡単な改良の例：

$$P_{ij}^{MG} = \min(\pi_i / (1 - \pi_j), \pi_i / (1 - \pi_i))$$

“Metropolized Gibbs Sampler” = 現在の状態を除いたGibbs sampler + Metropolis

$$\pi_i = \frac{w_i}{\sum_i w_i}$$
$$P^{MG} = \begin{bmatrix} 0 & \frac{\pi_1}{1-\pi_1} & \frac{\pi_1}{1-\pi_1} & \cdots & \frac{\pi_1}{1-\pi_1} \\ \frac{\pi_2}{1-\pi_1} & 1 - \dots & \frac{\pi_2}{1-\pi_2} & \cdots & \frac{\pi_2}{1-\pi_2} \\ \frac{\pi_3}{1-\pi_1} & \frac{\pi_3}{1-\pi_2} & 1 - \dots & \cdots & \frac{\pi_3}{1-\pi_3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\pi_n}{1-\pi_1} & \frac{\pi_n}{1-\pi_2} & \frac{\pi_n}{1-\pi_3} & \cdots & 1 - \dots \end{bmatrix}, \quad P^{MG} \geq P^G$$

Liu (1996),
Loison et al. (2004)

$$\pi_1 \leq \pi_2 \leq \cdots \leq \pi_n$$

状態空間における確率流

目標：重みから最適な遷移確率を見つける

$$\{\pi_i\} \rightarrow \{p_{i \rightarrow j}\}$$

確率流

$$v_{ij} := \pi_i p_{i \rightarrow j}$$

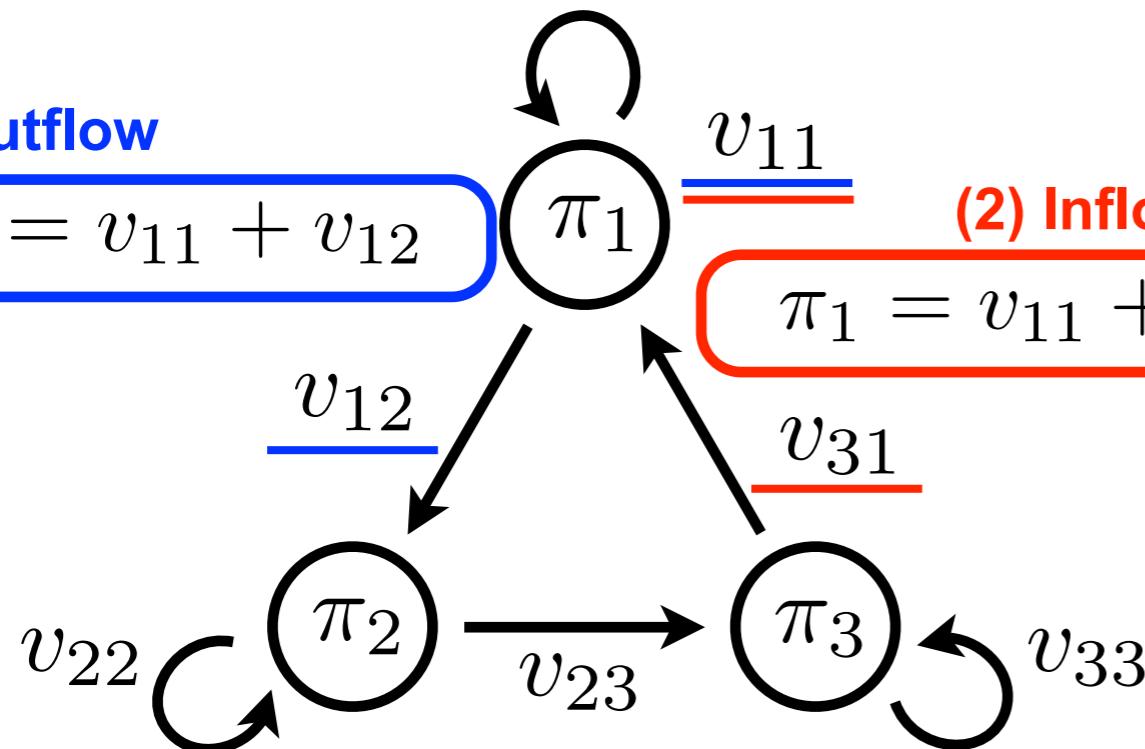
n: 状態候補の数

全体つりあい条件

$$\pi_i = \sum_{j=1}^n \pi_i p_{i \rightarrow j} = \sum_{j=1}^n \pi_j p_{j \rightarrow i}$$

(1) Outflow

$$\pi_1 = v_{11} + v_{12}$$



(2) Inflow

$$\pi_1 = v_{11} + v_{31}$$

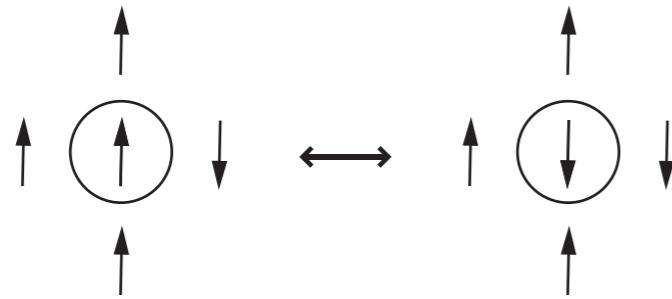
確率の保存

$$\pi_i = \sum_{j=1}^n v_{ij} \quad \forall i \quad (1)$$

全体つりあい

$$\pi_j = \sum_{i=1}^n v_{ij} \quad \forall j \quad (2)$$

幾何学的割当て



Metropolis:

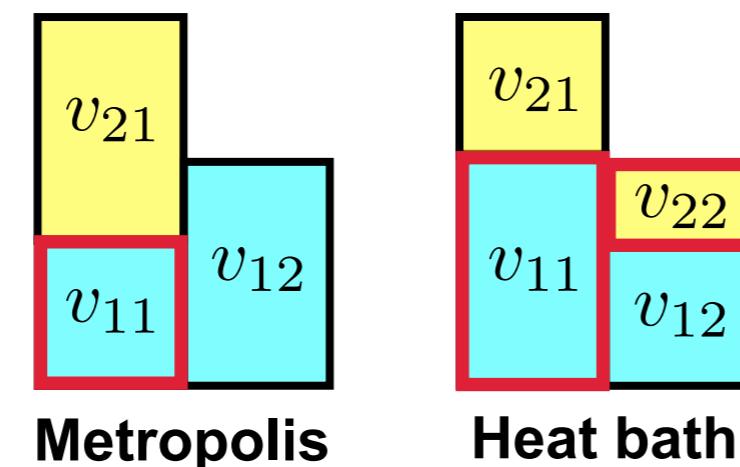
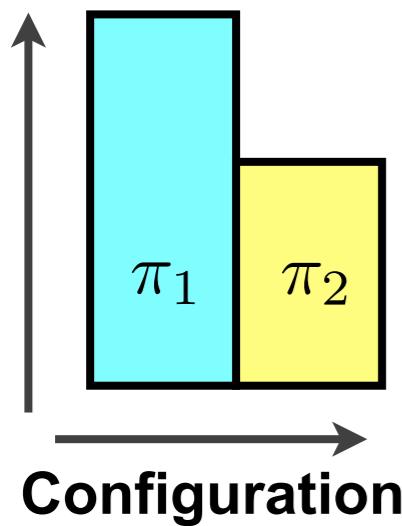
$$v_{ij} = \frac{1}{n-1} \min(\pi_i, \pi_j) \quad i \neq j$$

Heat bath:

$$v_{ij} = \frac{\pi_i \pi_j}{\sum_{k=1}^n \pi_k} \quad \forall i, j$$

Reversibility: $v_{ij} = v_{ji}$

Weight



Configuration

棄却率 \propto 赤い四角で囲まれた面積

イジングスピンの局所更新の例

確率の保存

$$\pi_i = \sum_{j=1}^n v_{ij} \quad \forall i \quad (1)$$

全体つりあい

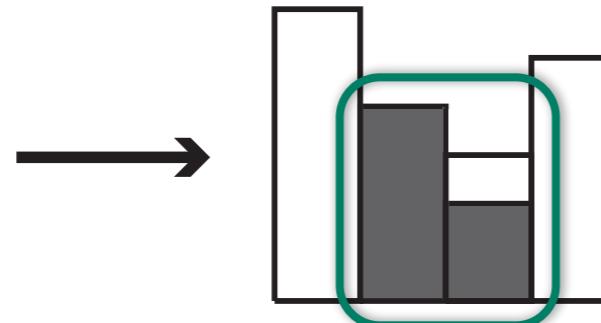
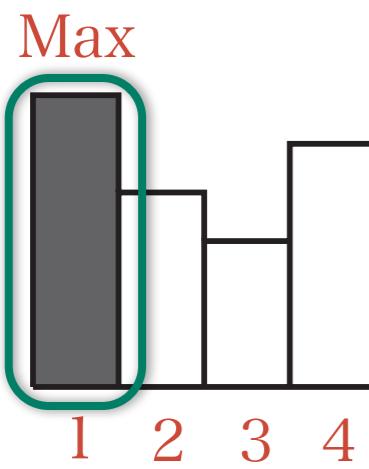
$$\pi_j = \sum_{i=1}^n v_{ij} \quad \forall j \quad (2)$$



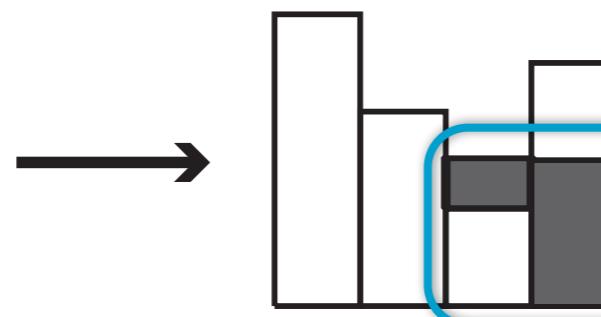
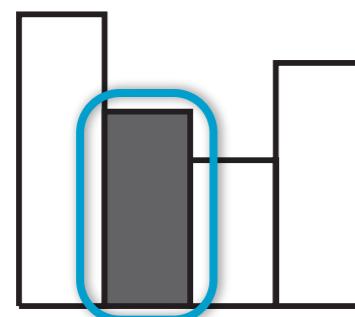
幾何学的割当て:
全体の箱の形を保ったまま
色（確率流）を割り当てる

割り当てアルゴリズム

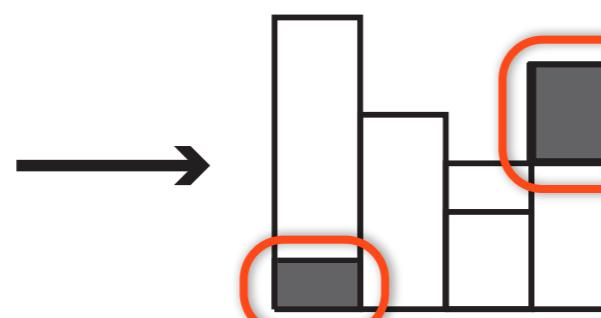
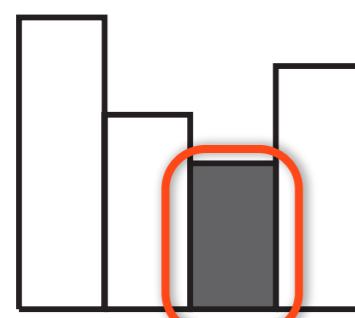
HS and S. Todo,
Phys. Rev. Lett. 105, 120603 (2010)



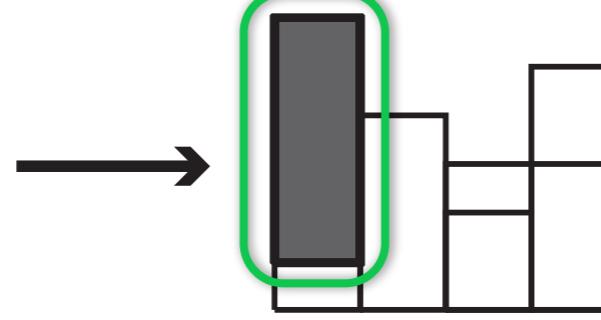
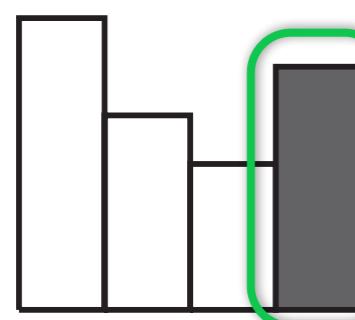
- 候補の中で最大の重みを持つ状態を見つける (π_1)。



- 最大の重み π_1 を次の箱に割り当てる ($i=2$)、残りをその次に割り当てる ($i=3$)。



- 次の重み π_2 を2の続きの場所に割り当てる。



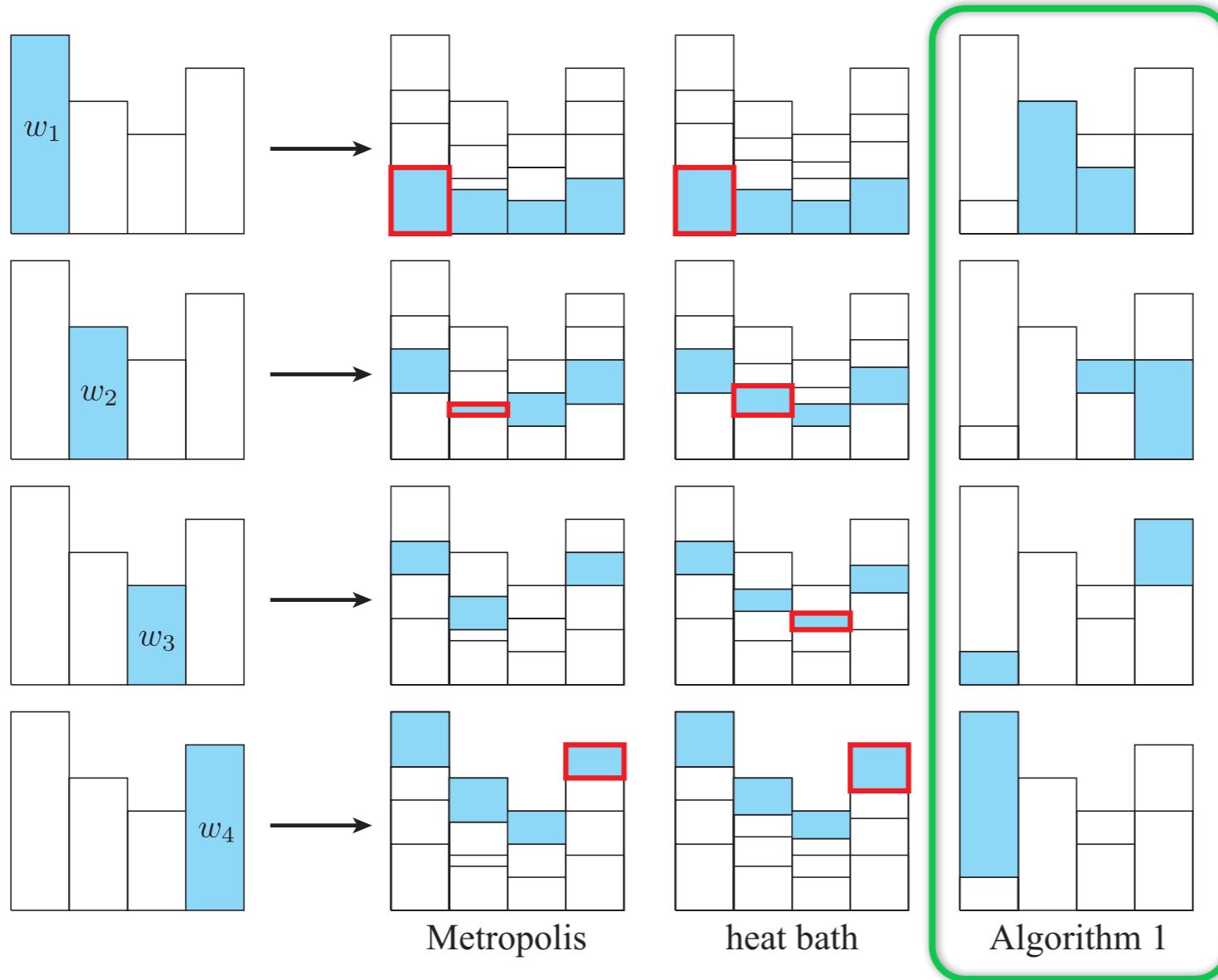
- 同様に続きの場所に重みを割り当てる。

- 遷移確率を次の式から決める。

$$p_{i \rightarrow j} = v_{ij} / \pi_i$$

自分自身に割り当てた重みがない \Rightarrow 棄却のない状態遷移

アルゴリズムの比較



棄却率がゼロになる条件

$$\pi_1 \leq \sum_{i=2}^n \pi_i$$

平均棄却率

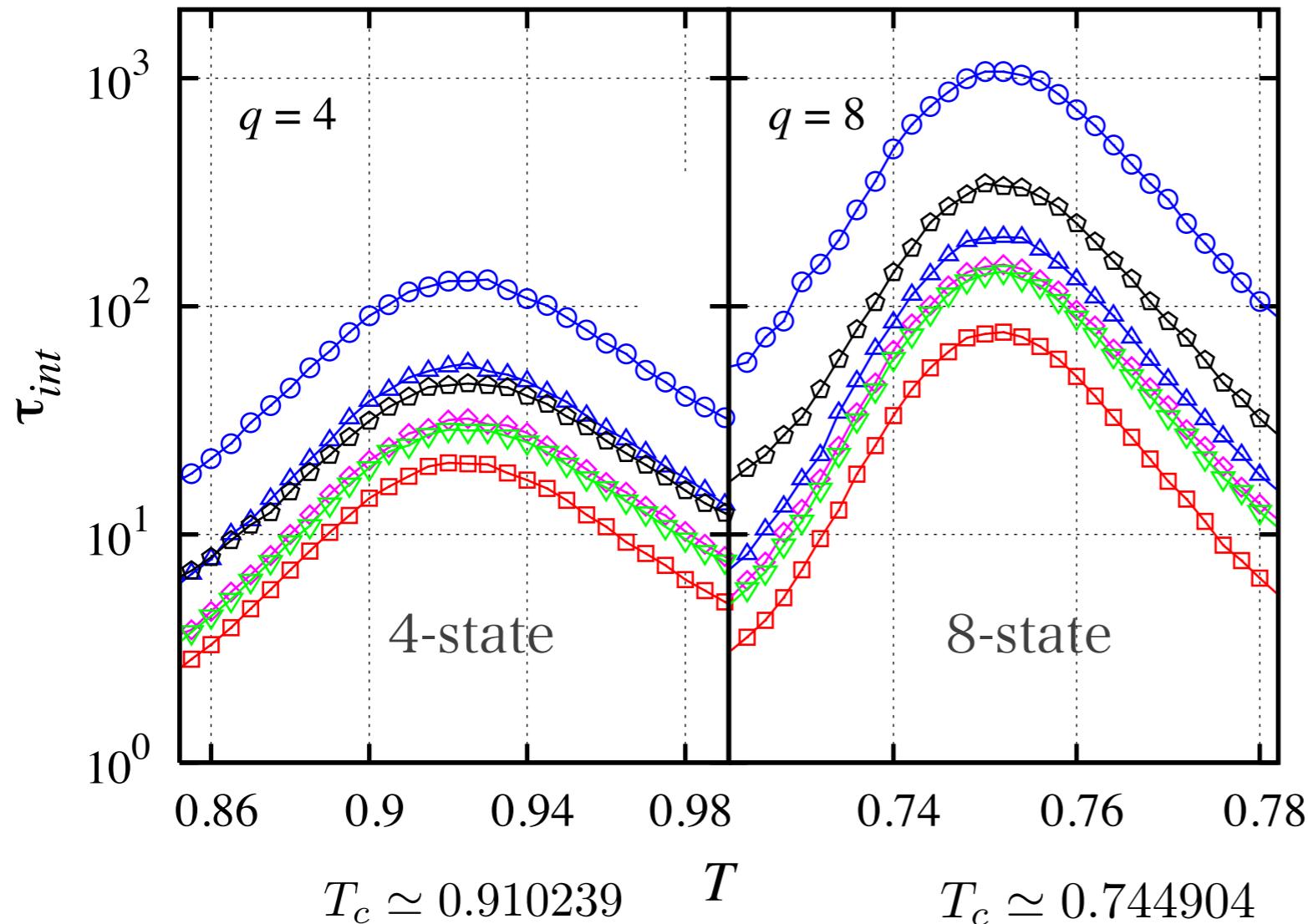
$$\frac{\sum_i v_{ii}}{\sum_j \pi_j}$$

- 遷移確率は事前に計算可能で、シミュレーション中の余分な計算コストなし
- 一般的に詳細つりあいを破る初めてのアルゴリズム
- 状態候補が与えられたときに、常に棄却率を最小化

ポツツ模型での改善

正方格子 $L = 16$

構造因子（秩序変数2乗）の積分自己相関時間



Metropolis

$\times 6.4$

Heat bath

$\times 2.7$

$\times 14$

$\times 2.6$

強磁性 q -状態ポツツ模型

$$\sigma_i = 1, 2, \dots, q$$

$$H = - \sum_{\langle i,j \rangle} \delta_{\sigma_i \sigma_j}$$

- Metropolis —○—
- Heat Bath —△—
- Metropolized Gibbs —◇—
- Iterative MG —▽—
- Optimal Average —◇—
- Optimal Rejection —□—

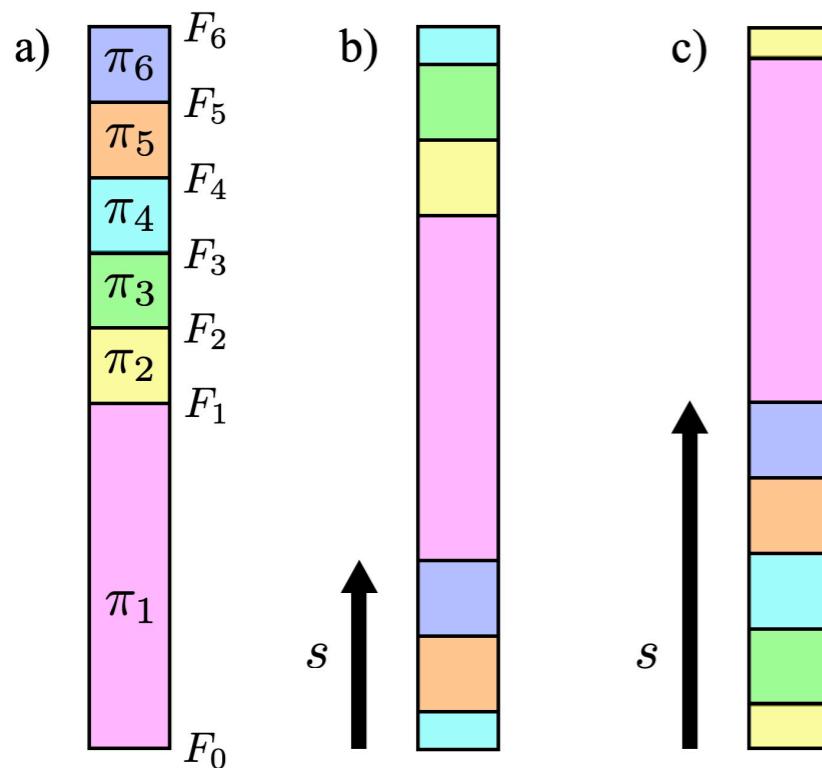
転移温度

$$T_c = 1 / \ln(1 + \sqrt{q})$$

計算効率の大幅な改善

棄却率制御アルゴリズム

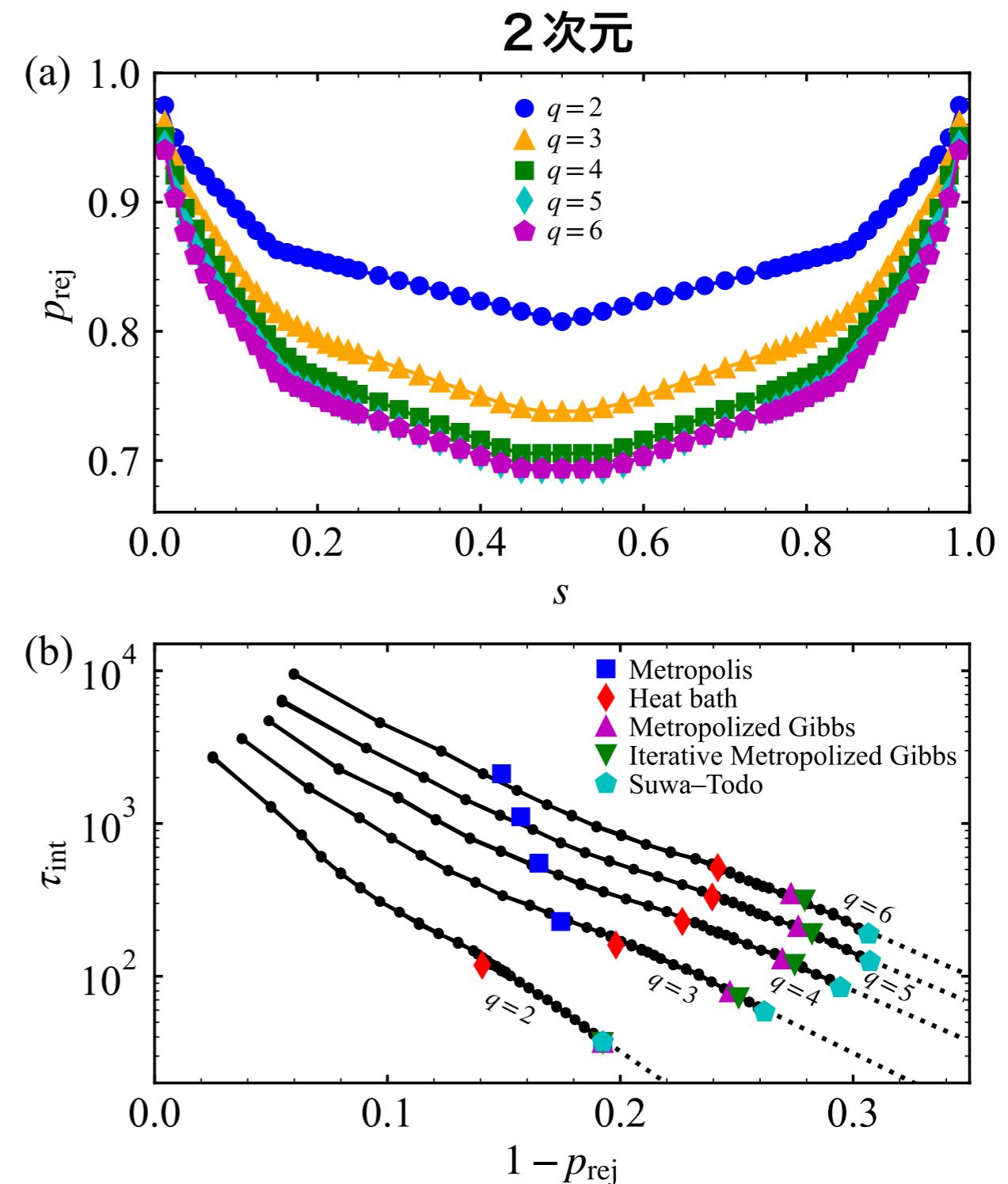
- ・棄却率を下げるとき自己相関時間が指数関数的に減少（強磁性ポツツ模型）
- ・解析解が得られ簡単に実装可能



累積分布関数における重みシフト

$$F_i = \sum_{j=1}^i \pi_j \quad (1 \leq i \leq n)$$

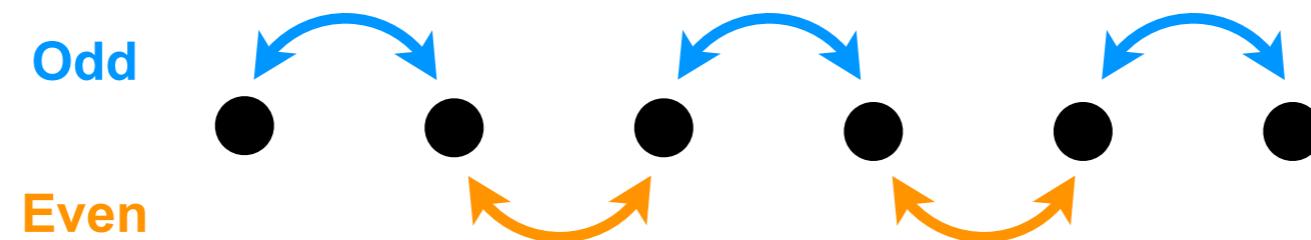
HS, Physica A 633, 129368
(2024)



不可逆レプリカ交換

- 交換するレプリカペアの取り方を偶奇で繰り返すと詳細つりあいは破れる
- リフティングによる状態空間の拡張 (+/- モード)

T. Okabe et al., Chem. Phys. Lett. 335, 435 (2001)
S. Syed et al., J. R. Stat. Soc. Series B. 84, 321 (2022)

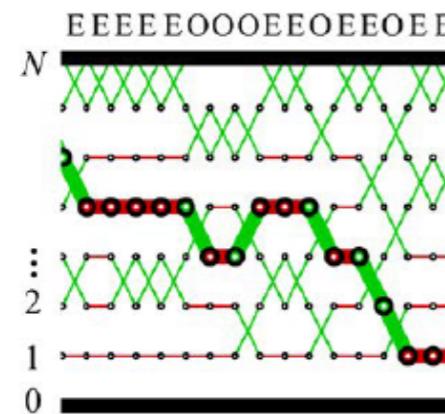


ねじれ詳細つりあい

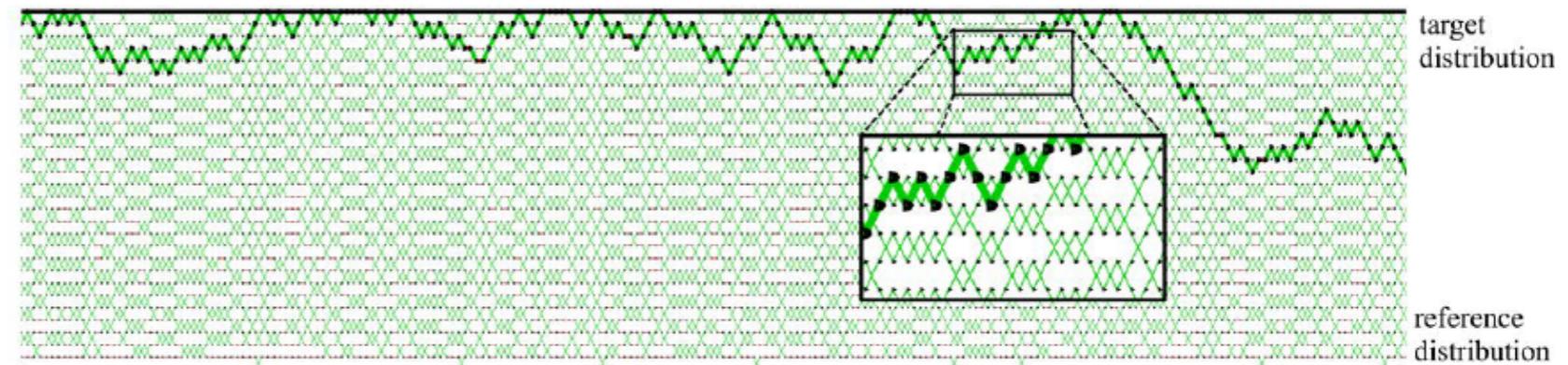
$$K^{\text{DEO}}((i, \varepsilon), (i', \varepsilon')) = K^{\text{DEO}}((i', -\varepsilon'), (i, -\varepsilon)),$$

$N = 8$

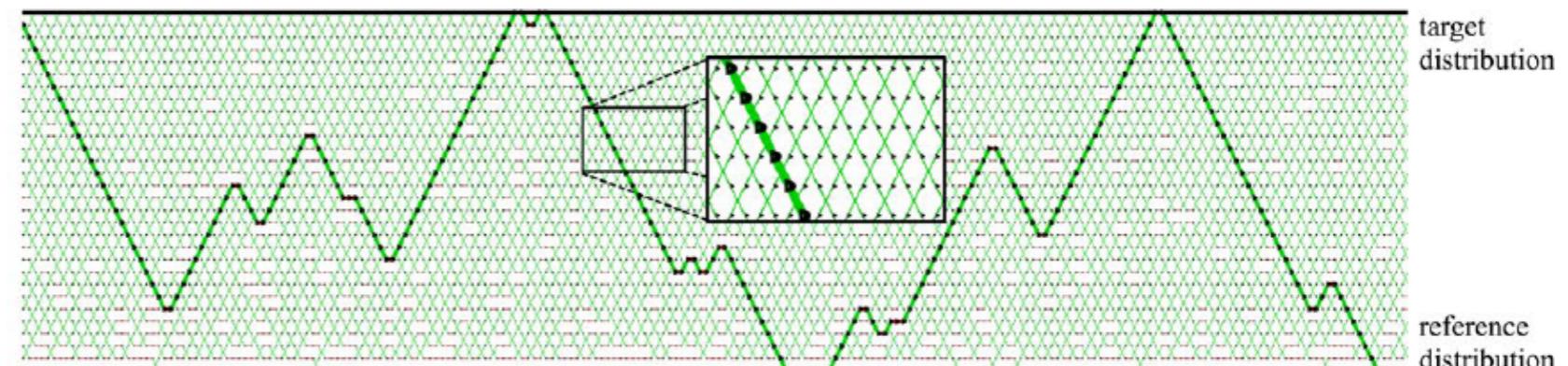
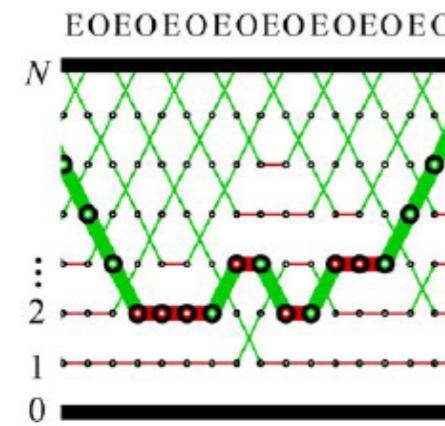
Reversible
(SEO)



$N = 30$

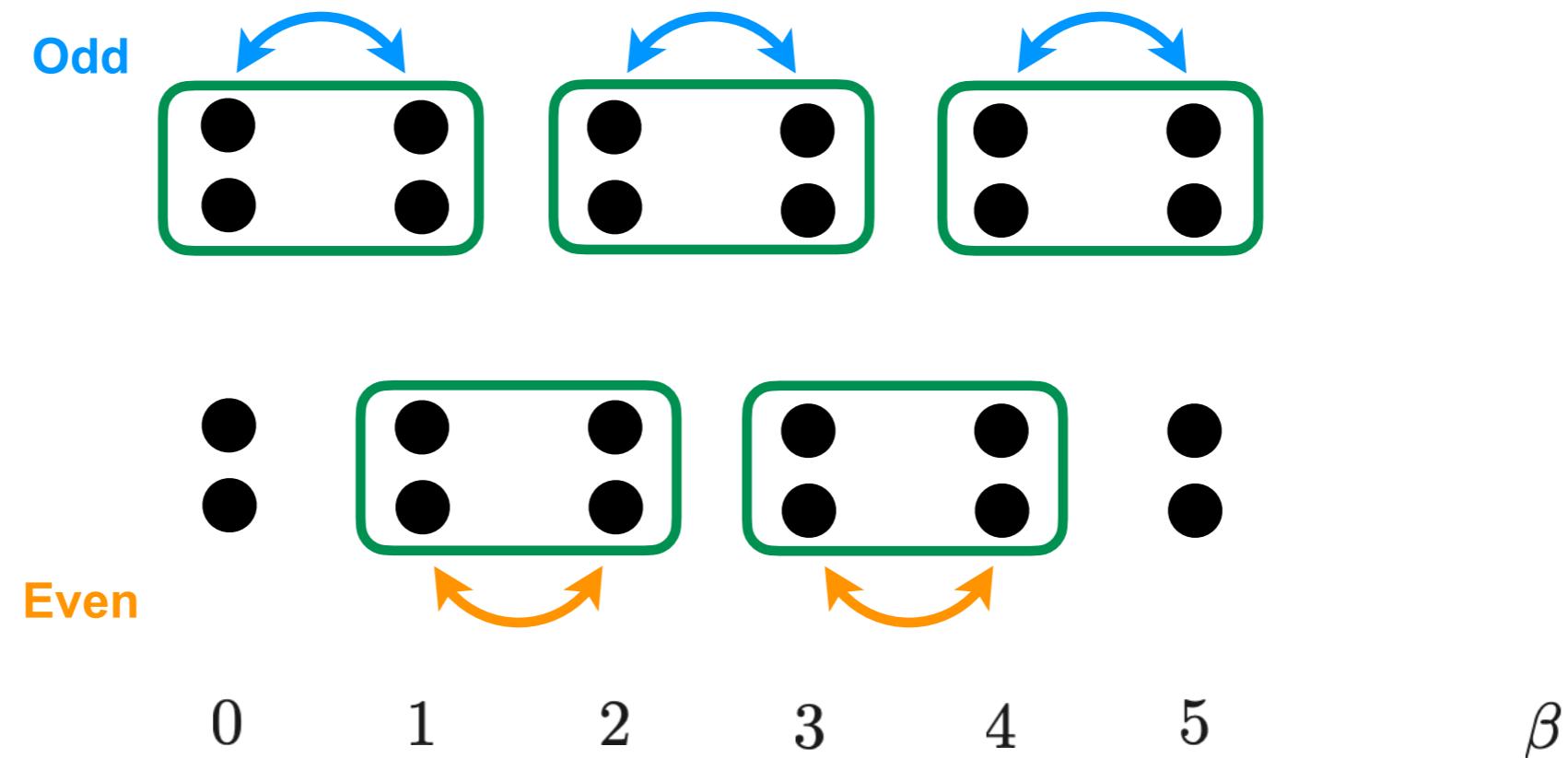


Non-reversible
(DEO)



マルチレプリカ交換

- マルチレプリカ交換によりレプリカ流密度は増加可能
- メトロポリス法を超えた確率最適化



m workers for each temperature

$$q(m) = \# \text{ Possible states} = \binom{2m}{m}$$

m	1	2	3	4
$q(m)$	2	6	20	70

線型計画法

- レプリカ流密度（交換するレプリカの割合）を最大化する遷移確率を求める

Maximize $\sum_{ij} \pi_i p_{ij} d_{ij}$

d_{ij} : Distance between i and j
(Number of swapped workers)

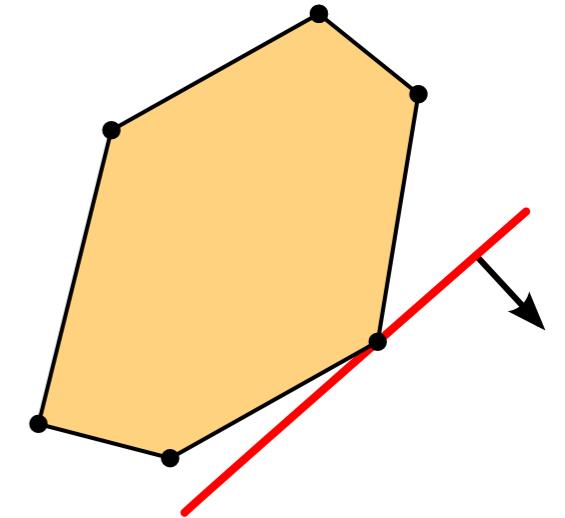
Constraints $p_{ij} \geq 0$

$$\sum_j p_{ij} = 1 \quad \forall i \quad \text{(Probability conservation)}$$

$$\sum_j \pi_j p_{ji} = \pi_i \quad \forall i \quad \text{(Global balance)}$$

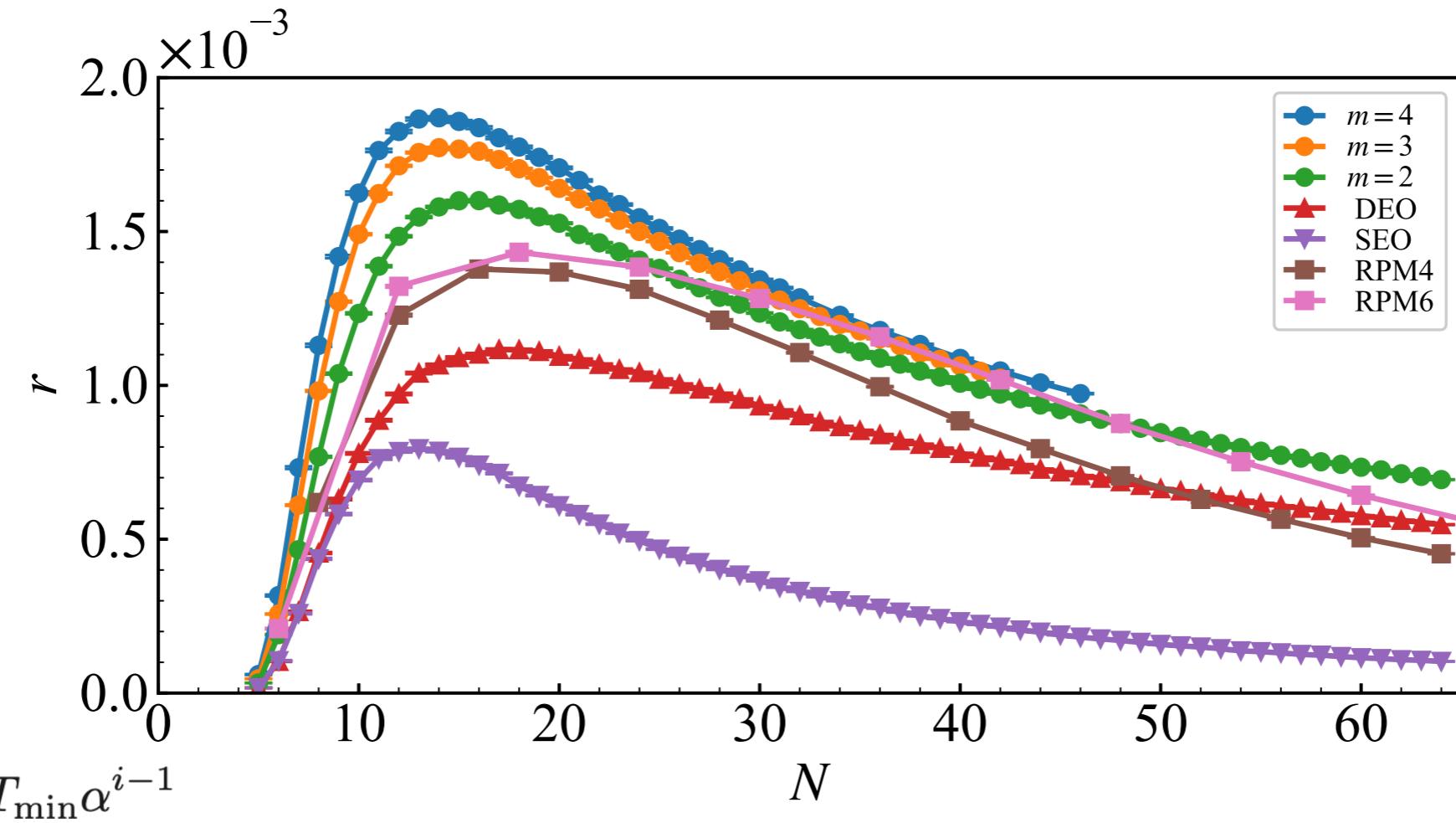
or

$$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i, j \quad \text{(Detailed balance)}$$



マルチレプリカ交換による改善

- 計算効率 (round trip rate) が改善される



$$T_i = T_{\min} \alpha^{i-1}$$

$$\alpha(N) = \left(\frac{T_{\max}}{T_{\min}} \right)^{1/(N-1)}$$

$$\pi_i(E) = \frac{1}{\sqrt{2\pi C T_i}} e^{-\frac{(E - CT_i)^2}{2CT_i^2}}$$

$$T_{\min} = 0.5$$

$$T_{\max} = 2.0$$

$$C = 200$$

- 局所的にも詳細つりあいを課さない方が条件数が少なくなり速く解が見つかる

まとめ

- ・マルコフ連鎖モンテカルロ法は任意の分布からのサンプリングを可能にする。
 - 多体問題、ベイズ推定の強力な計算手法
- ・次元の呪いを解く代わりに自己相関が問題となりうる。
 - 統計誤差の解析が重要 (binning 解析)
- ・自己相関時間を短くするアプローチ
 - クラスター更新 (Swendsen-Wang アルゴリズム)
 - ハイブリッド (ハミルトニアン) モンテカルロ
 - 拡張アンサンブル法 (レプリカ交換法、マルチカノニカル法)
- ・確率の流れを操作する不可逆モンテカルロ法で効率化が可能
 - リフティング
 - 遷移確率の最適化