モンテカルロ法の前線
—サイコロを振って積分する方法 —

福島 孝治

東京大学 大学院総合文化研究科 広域科学専攻 相関基礎科学系

URL: http://dbs.c.u-tokyo.ac.jp/~fukushima
mailto:hukusima@phys.c.u-tokyo.ac.jp

2003年11月10日

ある日のメーリングリストでの案内

****
*** 確率的情報処理のチュートリアル ***
*** １１月１０日に第３弾!! ***
*** 今回もまたまたわかりやすくお話します． ***
****
****
*** 講義ノートを WEB 上で公開しました． ***
****

開催日時：2003 年 11 月 10 日 (月曜日)

開催場所：ぱるるプラザ京都 (京都駅隣接)

WEB PAGE: http://www.statp.is.tohoku.ac.jp/~kazu/SMAPIP/2003/tutorial/

参加費：無料

若手企業研究者・大学院生を歓迎します. 大学院進学を考えている学部生も大歓迎.

開催主旨: 近年活発化している統計力学からの情報科学への · · ·

Statistical-Mechanical Approach to Probabilistic Information Processing

今日のお品書き

1. はじめに
   - モンテカルロ法と言えば，「πの計算」？
   - なぜモンテカルロ法が必要なのか?
2. メトロポリスのモンテカルロ法
   - マルコフ連鎖モンテカルロ法の原理の概要
   - メトロポリス法，熱浴法，ギブスサンプラー
   - 具体的な例として，イジング模型 ...
3. 注意すべき点
   - なぜその結果を正しいと思えるか？
   - 独立なサンプリング
4. より前線へ
   - アニーリング法とモンテカルロ法の関係
   - 拡張アンサンブルの方法
5. まとめ

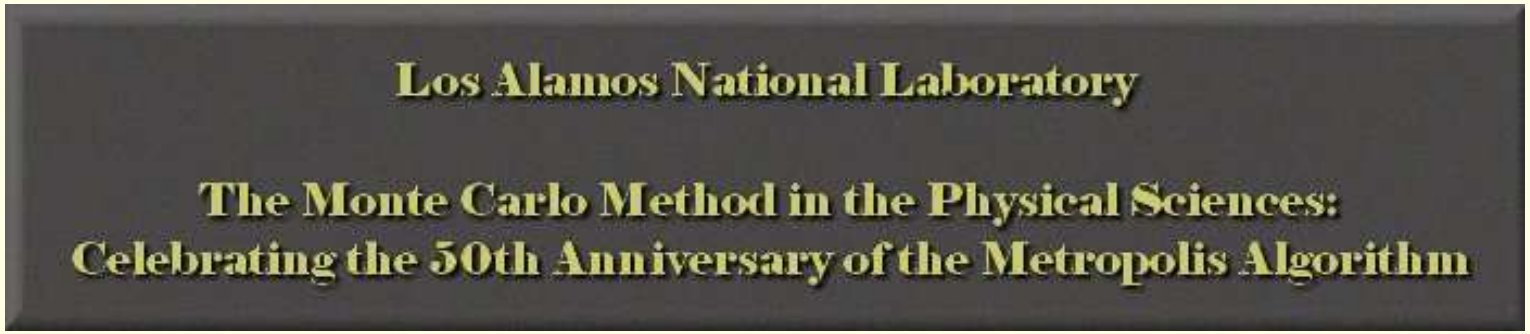
モンテカルロ法50周年

DNA だけじゃない

現在モンテカルロ法と呼ばれている方法の雛型になった論文

the Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller publication in 1953

ちょうど50周年記念ワークショップ



JUNE 9-11, 2003
J. Robert Oppenheimer Study Center
Los Alamos National Laboratory
Los Alamos, New Mexico U.S.A.

http://cnls.lanl.gov/Conferences/MonteCarloMethods/

(マルコフ連鎖)モンテカルロ法の目的

- d次元空間の点$\boldsymbol{X} = (x_1, x_2, \cdots, x_d)$の従う確率分布$P(\boldsymbol{X})$が与えられている．

1. 確率分布$P(\boldsymbol{X})$に従う様に，点$\boldsymbol{X}$をサンプリング,

$$\{\boldsymbol{X}^{(i)}\} = \{\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \cdots, \boldsymbol{X}^{(M)}\}.$$

   - 確率分布$P(\boldsymbol{X})$の逆関数は求めることはできないとする．

2. $\boldsymbol{X}$の関数$A(\boldsymbol{X})$の確率分布$P(\boldsymbol{X})$についての期待値の計算

$$\langle A \rangle = \int_D A(\boldsymbol{X}) P(\boldsymbol{X}) d\boldsymbol{X} = \frac{1}{M} \sum_{i=1}^{M} A(\boldsymbol{X}^{(i)}).$$

   - 一般に多重積分はこの範疇に入ると考えられる．求めたい積分を$I = \int_D dx g(x)$とし，確率分布$P(x)$に従う変数から評価するのは，

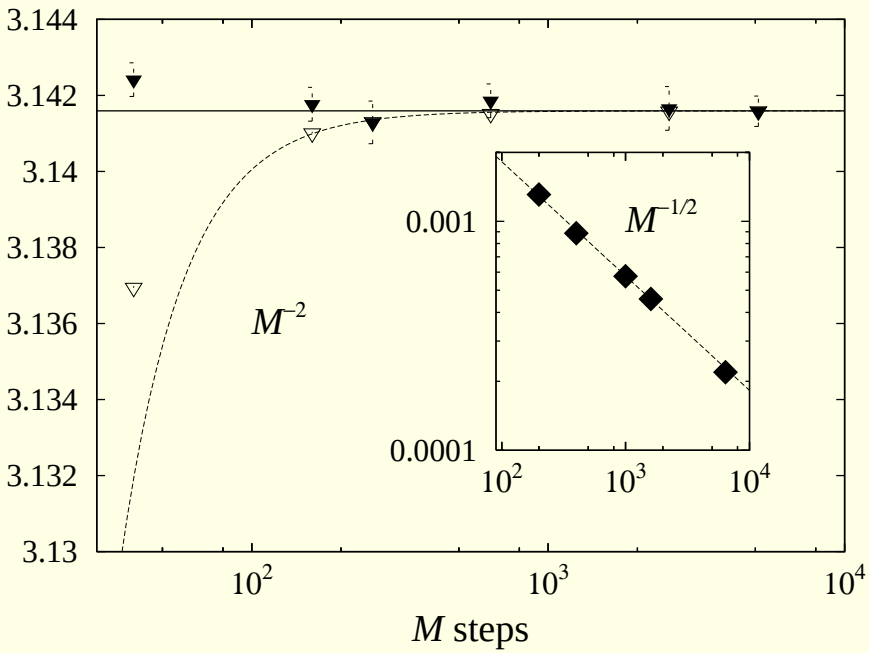
$$I = \int_D dx g(x) = \int_D dx \frac{g(x)}{P(x)} P(x) = \left\langle \frac{g(x)}{P(x)} \right\rangle$$

モンテカルロ積分: π の評価の場合

- 例えば，一変数関数 $g(x) = 4\sqrt{1-x^2}$ を領域 $D \in [0,1]$ で積分．
  - 単純MC積分法：領域 D から一様サンプリングする，$P(x) =$ 一定．

$$\tilde{I} = \frac{1}{M}\left(\frac{g(x^{(1)})}{1/D} + \cdots + \frac{g(x^{(M)})}{1/D}\right)$$

  - モンテカルロ積分の誤差は，$O(M^{1/2})$ 程度
  - 最も簡単な求積法である台形公式の誤差は，$O(M^{-2})$.



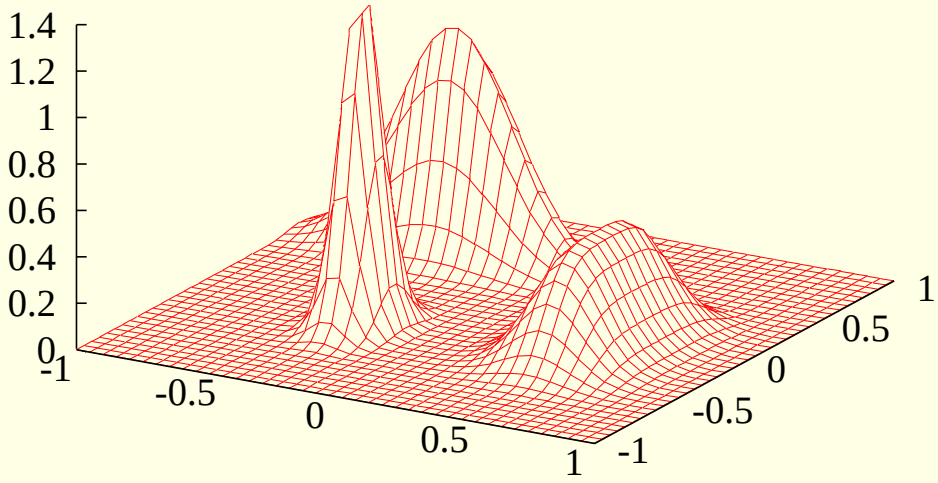
- 実際には，π の計算は，誤差が大きくて，モンテカルロ法ではやらない．
- しかし，モンテカルロ法の誤差は原理的には積分変数の次元に依らない．

$\Longrightarrow$ 次元が大きくなると，モンテカルロ法もやれる?!

モンテカルロ積分をより高次元に

- 一様ランダム・サンプリング積分を高次元へ応用した場合の困難点

1. 高次元空間の中で積分領域を一様にサンプルすることは意外に難しい．
2. ほとんどの点で被積分関数の値が小さく，無駄ばかりが多くなる．
   - 分散が大きくなってしまう．
   - よい評価が困難になる．



二次元分布関数の例：関数の値が有意な値を持つ領域は少なくなる．鞍点評価がよくなるような例だと，ますます辛い．

- 一様ランダム・サンプリングから，重点的サンプリングへ．

$\Longrightarrow$ 分布関数 $P(X)$ に比例するようにサンプルしたい

メトロポリスのモンテカルロ法

- 確率仮定を用いて，逐次的に点列 $\boldsymbol{X}$ を生成する．
  - マルコフ連鎖：直前の点 $\boldsymbol{X}^{(t-1)}$ のみから新しい点 $\boldsymbol{X}^{(t)}$ を決定する．

$$\boldsymbol{X}^{(t-1)} \to \boldsymbol{X}^{(t)} \to \boldsymbol{X}^{(t+1)} \to \cdots \quad : \text{定常分布が} P(\boldsymbol{X}) \text{となるように} \ldots$$

  - モンテカルロ法の原理：遷移確率 $W(\boldsymbol{X} \to \boldsymbol{X}')$ の満たすべき必要条件

(A) 詳細つり合いの条件 遷移確率の関係式

$$P(\boldsymbol{X})W(\boldsymbol{X} \to \boldsymbol{X}') = P(\boldsymbol{X}')W(\boldsymbol{X}' \to \boldsymbol{X})$$

(B) エルゴード条件 任意の2つの点 $\boldsymbol{X}$ と $\boldsymbol{X}'$ の遷移確率がゼロでないか，有限個のゼロでない遷移確率の積で表される．

$\Longrightarrow$ どんな初期条件からでも，唯一つの定常分布に収束する．

遷移確率のいろいろ

- 遷移確率の満たすべき(必要)条件は緩くて，一意的には決定できない．
  - 点$\boldsymbol{X}$の各成分x_iが離散的な多値$(\{\alpha_i^\mu, \mu = 1, \cdots, A_i\})$をとる場合:
  - 遷移は，ある成分x_iの値を更新する状況を考える．
    現在の状態$\boldsymbol{X} = (x_1, \cdots, x_d) \to$ 新しいの状態$\boldsymbol{X}_\mu = (x_1, \cdots, \alpha_i^\mu, \cdots, x_d)$，

  メトロポリス型 : あらかじめ新しい値α_μをどれかに選択．

$$W_{\rm Metropolis}(\boldsymbol{X} \to \boldsymbol{X}_\mu) = \min\left(1, \frac{P(\boldsymbol{X}_\mu)}{P(\boldsymbol{X})}\right) \tag{1}$$

  熱浴法，ギブスサンプラー 取り得る新しい値はA_iとおりある．

$$W_{\rm heatbath}(\boldsymbol{X} \to \boldsymbol{X}_\mu) = \frac{P(\boldsymbol{X}_\mu)}{\sum_\mu P(\boldsymbol{X}_\mu)} \tag{2}$$

- 各成分x_iを一つずつ更新する限り，エルゴート条件は満たされることが多い．

- 連続変数の場合も新しい候補を離散化することで一般化は可能

メトロポリス・アルゴリズム

ステップ **0**： 初期条件を適当に選ぶ： $\boldsymbol{X}^{(0)}$

ステップ **1**： 現在の点 $\boldsymbol{X}^{(t)}$ からランダムに変化させた新しい候補点 $\boldsymbol{X}'$ を生成．

ステップ **2**： 遷移確率 $P(\boldsymbol{X}')/P(\boldsymbol{X}^{(t)})$ を計算

ステップ **3:** 一様乱数 $r \in [0,1]$ を生成し，以下のように次の状態 $X^{(t+1)}$ を決める;

$$X^{(t+1)} = \begin{cases} X' & \text{if } r \leq P(X')/P(X^{(t)}) \\ X^{(t)} & \text{otherwise} \end{cases} \tag{3}$$

ステップ **4**： ステップ 1へ戻り，繰り返す．

ステップ1から4までの試行を1モンテカルロステップと呼ぶことにする．

具体的な例1：イジング模型

- 各格子点にスピン変数 $S_i = \pm 1$

  $\boldsymbol{X} = (S_1, S_2, \cdots, S_N)\,(= (1, -1, 1, \cdots))$

- 確率分布：ギブス関数

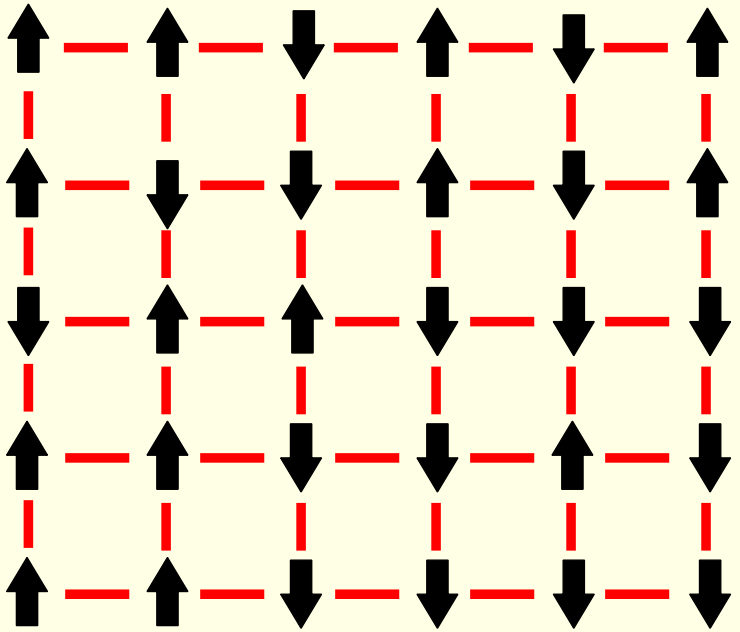
$$P(X) = \frac{\exp(-E(\boldsymbol{X})/T)}{\sum_{\boldsymbol{X}'} \exp(-E(\boldsymbol{X})/T)}$$

  $E(\boldsymbol{X})$ はエネルギー関数

$$E(\boldsymbol{X}) = -\sum_{\langle ij \rangle} J_{ij} S_i S_j$$

- 相互作用 $J_{ij} = J$ の場合
  - $S_i S_j = 1$:$-J$ エネルギーが下がる
  - $S_i S_j = -1$:$+J$ エネルギーが上る

- 状態の数は 2^N



- 物理量の例
  - 磁化 $M = \frac{1}{N}\sum_i S_i$

$$\langle |M| \rangle = \left(\sum_{S_i = \pm 1} \sum_{S_2 = \pm 1} \cdots \sum_{S_N} \right) |M| P(\{S\})$$

具体的な例題2：Nクイーン 問題

$N \times N$のボードの上にN個のクイーンがお互いにぶつからないように配置する．

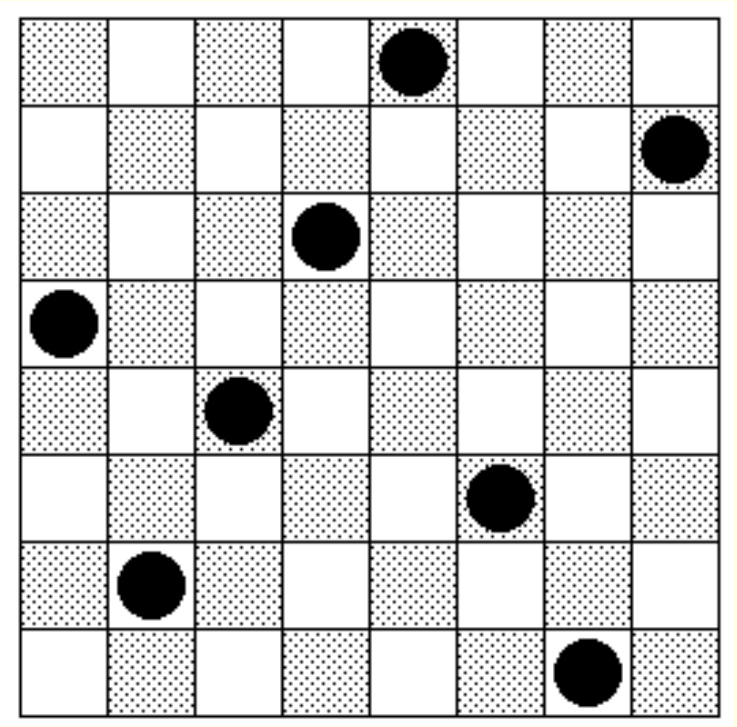
- 各行に1つだけクイーンを配置する．クイーン変数

$$X_i = (1, 2, \cdots, N), (i = 1, \cdots, N)$$

状態の数はN^N ．

- 確率分布はどうとるかは自由だが，ギブス関数的にエネルギーをクイーンがぶつかる件数として，

$$P(\boldsymbol{X}) \propto \exp(-(\text{エネルギー})/T)$$



8クイーン問題の解答例

- クイーン問題はお互いにぶつからない配置，すなわちエネルギーゼロの配置を求めること．温度Tを低くして，確率分布を最大にする配置$\boldsymbol{X}$を求めたい．

イジング模型を例に具体的なメトロポリス・アルゴリズム (1)

ステップ 0： 初期スピン配位

初期点 $\boldsymbol{X}$ をランダムに選ぶ：$X^{(0)} = \left(S_1^{(0)}, S_2^{(0)}, \cdots S_N^{(0)}\right) = (1, -1, \cdots, -1)$

ステップ 1： 次のスピン配位の選択(メトロポリス型)

ランダムに選んだ格子点 k のスピンを反転したスピン配位

$$X' = \left(S_1^{(t)}, S_2^{(t)}, \cdots, -S_k^{(t)}, \cdots, S_N^{(t)}\right)$$

ステップ 2: 遷移確率の計算：遷移コストを評価する．

$$\frac{P\left(\boldsymbol{X}'\right)}{P\left(\boldsymbol{X}^{(t)}\right)} = \frac{\exp\left(\sum_i{}' J S_i^{(t)}(-S_k^{(t)})/T\right)}{\exp\left(\sum_i{}' J S_i^{(t)} S_k^{(t)}/T\right)} = \exp\left(-2\sum_i{}' J S_i^{(t)} S_k^{(t)}/T\right)$$

- 現在のスピン配位だけで表される
- 注目する格子点 k の近接の和だけで計算できる．

イジング模型を例に具体的なメトロポリス・アルゴリズム(2)

ステップ 3： 乱数との比較で更新．

一様乱数rと比較することにより，ルール

$$X^{(t+1)} = \begin{cases} X' & \text{if } r \leq P(X')/P(X^{(t)}) = \exp\left(-2\sum_i{}' JS_i^{(t)}S_k^{(t)}/T\right) \\ X^{(t)} & \text{otherwise} \end{cases}$$

に従って，次の状態$X^{(t+1)}$を決定する．

一つ注意する点は，この更新で実際の状態の変化が無い場合($X^{(t+1)} = X^{(t)}$)でも，状態更新と見なして，平均の計算には考慮する必要がある．

ステップ 4： くりかえし ステップ1に戻って，同じ操作を繰り返す．

- 通常は，全スピン数N回の繰り返しを，1モンテカルロステップ数と言う．
- メトロポリス法で作られるサンプル列$\{X^{(1)}, X^{(2)}, \cdots, X^{(M)}\}$の間には強い相関が残ることが推測される．

イジング模型を例に具体的なメトロポリス・アルゴリズム(3)

ステップ 4 のつづき： 強いMCステップ相関．独立と思えるくらいにサンプル間のMCステップを離す．目安は相関時間τ^*，

$$C(\tau) = \frac{\left\langle A(\boldsymbol{X}^{(t)})A(\boldsymbol{X}^{(t+\tau)})\right\rangle - \left\langle A(\boldsymbol{X}^{(t)}\right\rangle^2}{\left\langle A(\boldsymbol{X}^{(t)})^2\right\rangle - \left\langle A(\boldsymbol{X}^{(t)}\right\rangle^2} \sim \exp(-\tau/\tau^*)$$

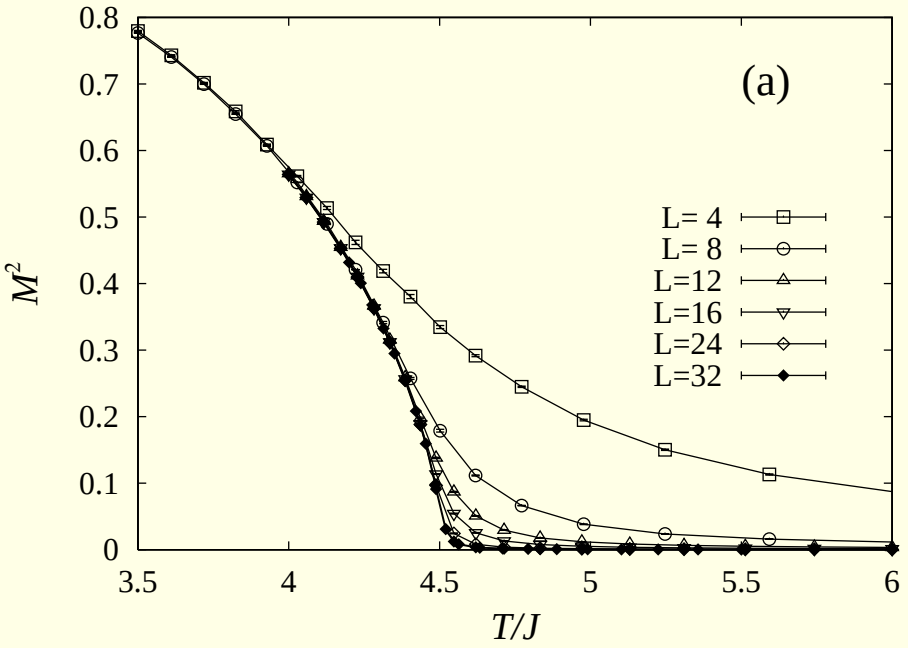
を見積もって，τ^*くらいは離してサンプリングしたい．

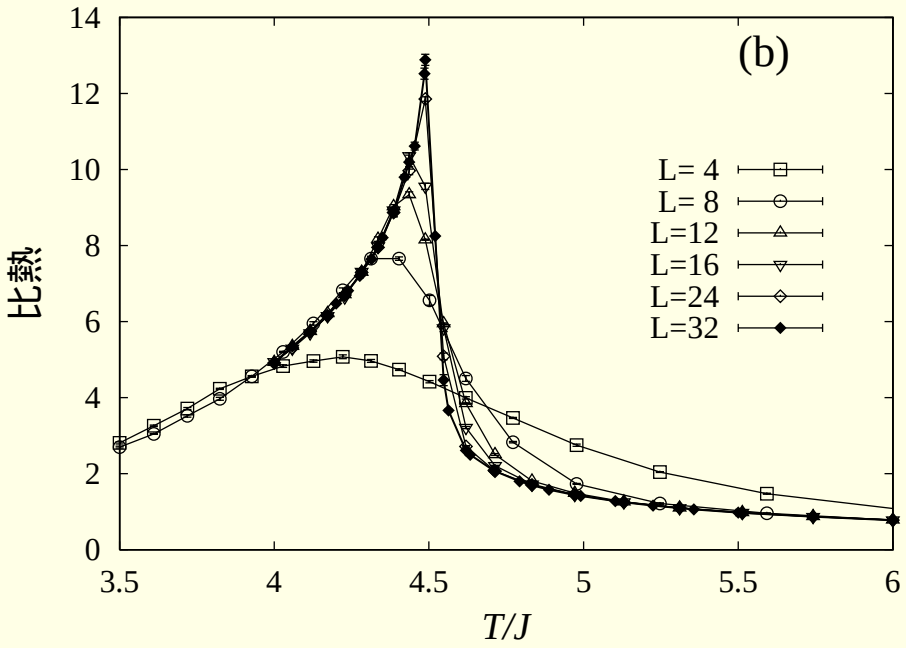
$$\{\boldsymbol{X}^{(t)}, \boldsymbol{X}^{(t+\tau^*)}, \boldsymbol{X}^{(t+2\tau^*)}, \cdots, \boldsymbol{X}^{(t+M\tau^*)} \ldots\}$$

これは，単純モンテカルロ積分から重点サンプリングに移行する際の代償か．

イジング模型の結果

3 次元イジング模型の結果：磁化 (左) と比熱 (右) の温度依存性





- ある温度で特異な振るまいが見られる．相転移の存在．
- スピン数は，$N = 4^3 \sim 32^3$．$N = 32^3$ のときの全状態数は 2^{32768}．
- この程度の計算が一晩で，できる時代になった．

注意すべき点

- ある修士論文審査会にて，

  審査の先生：「あなたのモンテカルロ計算はなぜ正しいといえるのか?」

注意すべき点

- ある修士論文審査会にて，

  審査の先生：「あなたのモンテカルロ計算はなぜ正しいといえるのか?」

  学生：「私を信じて下さい!!」

注意すべき点

- ある修士論文審査会にて，

  審査の先生：「あなたのモンテカルロ計算はなぜ正しいといえるのか?」

  学生：「私を信じて下さい!!」

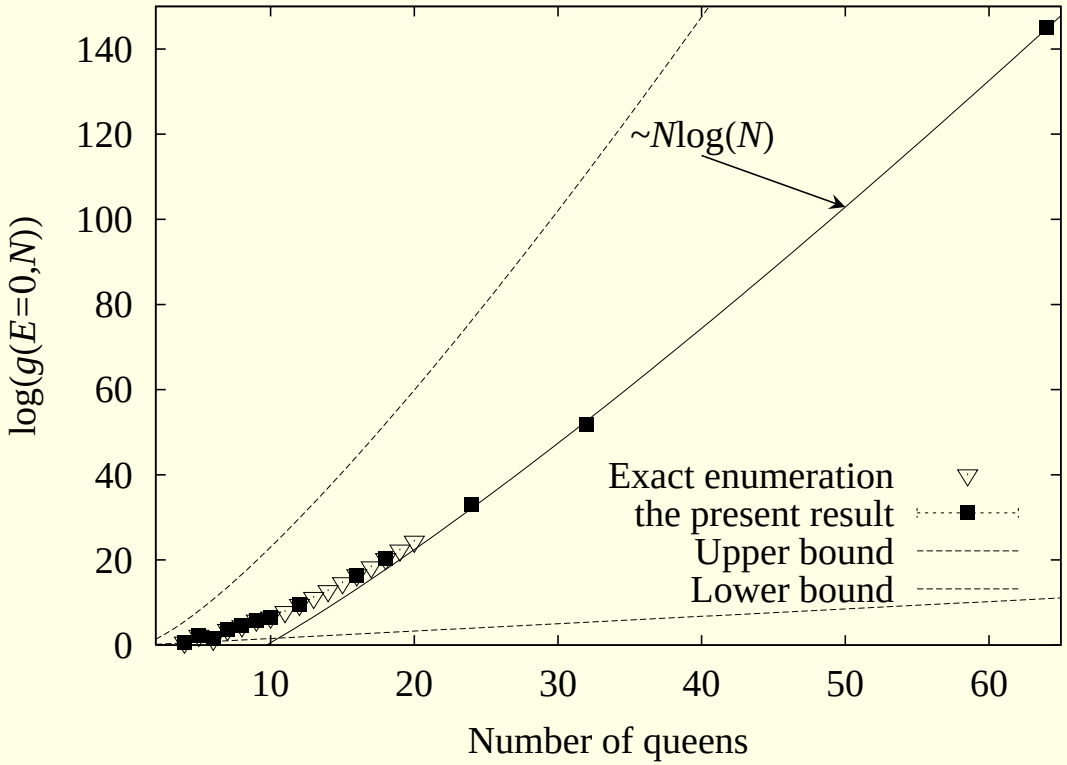
  - プログラムのバグ:いろいろありうる．いつもある変数だけ更新を忘れているとか ...
  - 乱数の性質： 周期性の問題． 64bit化が進む中で，どうなるか．
  - エルゴード性の問題
    - 本質的なダメなアルゴリズム：広義のバグ?
    - 実質的にエルゴード性が満たされない場合：原理的には正しくても ...
  - 有効サンプリングの問題：統計誤差の評価，サンプル点数 M は幾つ必要か?

- 一般的な対策はない： 必要条件の積み重ね

  - 厳密にできるところをチェック．例えばサイズの小さいところで，全状態数え上げ．
  - 満たすべき恒等式や不等式のチェック

N クイーン問題の例

N クイーン問題の「解の個数」のクイーン数 N 依存性



- 小さいクイーン数 N では，全状態をバックトラック等で探索して答えがわかる．
- 下限，上限の不等式の評価もある．

モンテカルロ法における定常分布への遅い緩和の問題

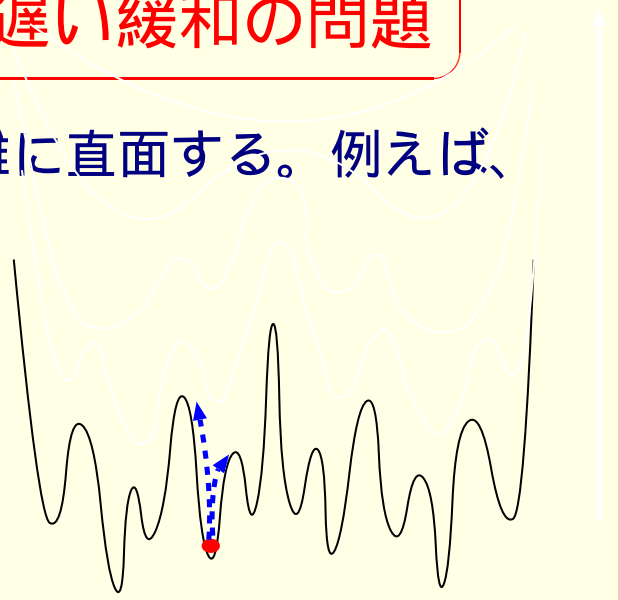
ある状況では、マルコフ連鎖の状態の遷移確率が著しく小さくなり、実質的に遷移が起こらない。 $\Longrightarrow$ 独立なサンプリングが少なくなり、期待値の評価に問題が生じる

広い状態空間のある局所的な部分に留まり続けるというイミで

実質的なエルゴード条件の破れ $\simeq$ 遅い緩和の問題

と呼ばれていて、 物理的に興味のある系でこの困難に直面する。例えば、

- 複雑な系での分布の多峰性
  拘束条件の強い系で広く一般的に見られる．
- ある相転移に関係する **slowing down**
  - 2次相転移のcritical slowing down
  - 1次転移に関する核生成

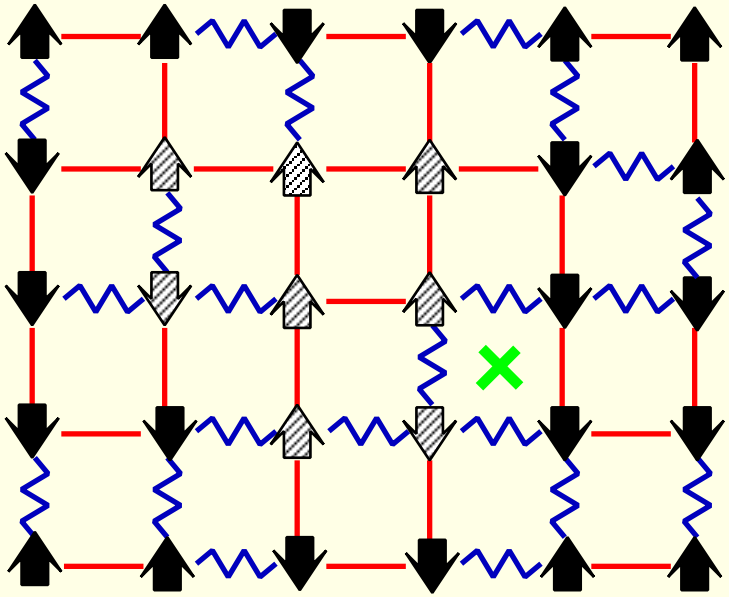


(自由)エネルギー表面の多数の局所的な安定状態 $=$ 分布の多峰性

拘束条件の強い例

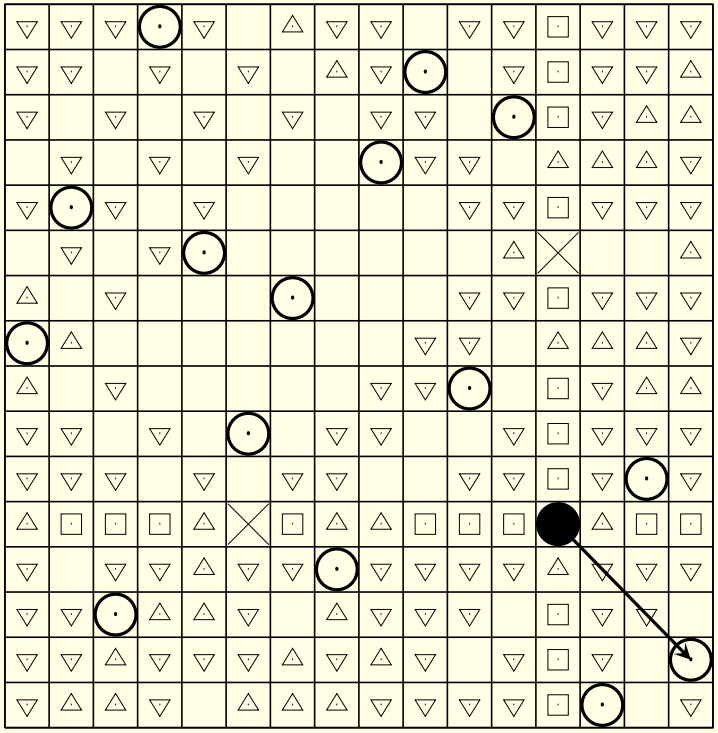
イジング模型の応用：スピングラス

- 相互作用 J_{ij} を正負ランダムにすると ...
  波線は -1, 直線は $+1$
- 緑× の四角形にはうまくスピンを置けない．
- エネルギーが低い状態が沢山ある．
- 薄色のスピンを全部ひっくり返しても良い．
- が，一つずつひっくり返す途中にエネルギーの高い点がある．



N クイーン問題の例：

- この配位は，一つのクイーンがぶつかっている．
- 一つずつ動かすと，一度は2つぶつかっている配位を通らなければならない．



Simulated Annealing: 最適化技法としての観点

S. Kirkpatrick, et al, Science **220** 671, (1983).

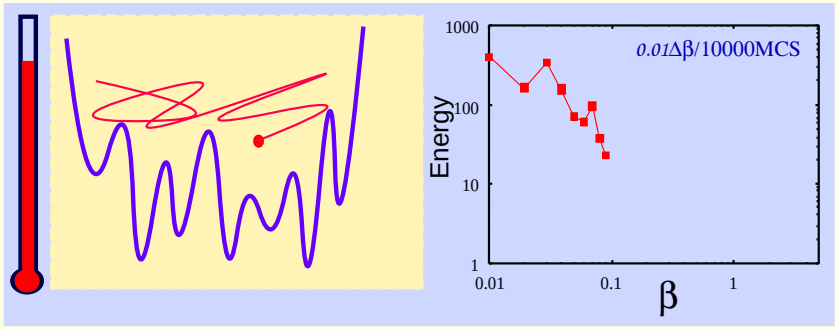
- 一般の確率分布 $P(X)$ に対して、温度パラメータ T の導入

$$P_T(X) = \frac{P^{1/T}(X)}{\sum_X P^{1/T}(X)} \qquad \left(= \frac{\mathrm{e}^{-E(X)/T}}{\sum_X \mathrm{e}^{-E(X)/T}} \quad \text{for} \quad \text{Gibbs} \quad \text{dist.} \right)$$

- 高温から低温に徐々に温度を下げる：徐冷法

高温極限： 無作為抽出法

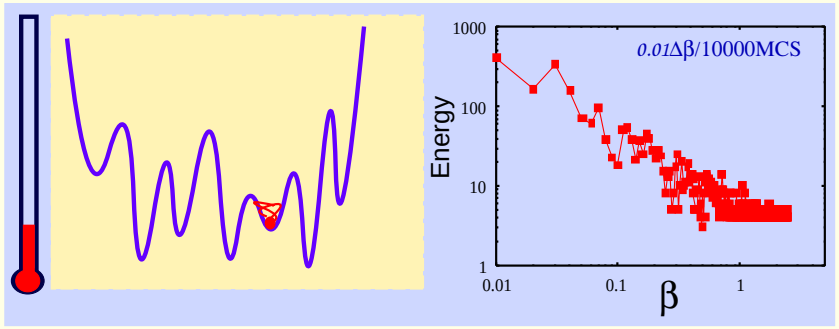
どんな状態も一様にサンプリング
状態空間を広く探索できるかも
エネルギーの低い状態は探せない
準安定状態に捕まる可能性は低い



→⇒⟹

低温極限： 急冷法

分布 P_T を最大にするような状態 X^* が実現されやすい。
常にエネルギーの低い状態を探す
準安定状態に捕まる可能性は高い



SA法の限界....ジレンマ

1. 温度のスケジュールはどうすればよいか?
   - 速く温度を下げすぎると、修正が難しい。
   - 遅すぎると、無駄な計算。
     (対数的にゆっくりと冷やせばよいが、現実的ではない。)
2. 多数に縮退した解はどのように求めるのか？
   - 何度も繰り返すのか？
3. 途中の温度の情報は？
   - 温度を下げる過程で詳細釣り合いの条件を破っているので、定常分布の情報は得られない。サンプリングとしての保障がない。

詳細釣り合いを満たして、定常分布を得たい

低温で準安定状態に捕まってしまう

$\Longleftrightarrow$

温度を上げ下げして準安定状態から脱出したい

定常分布(相対的な重み)わからない

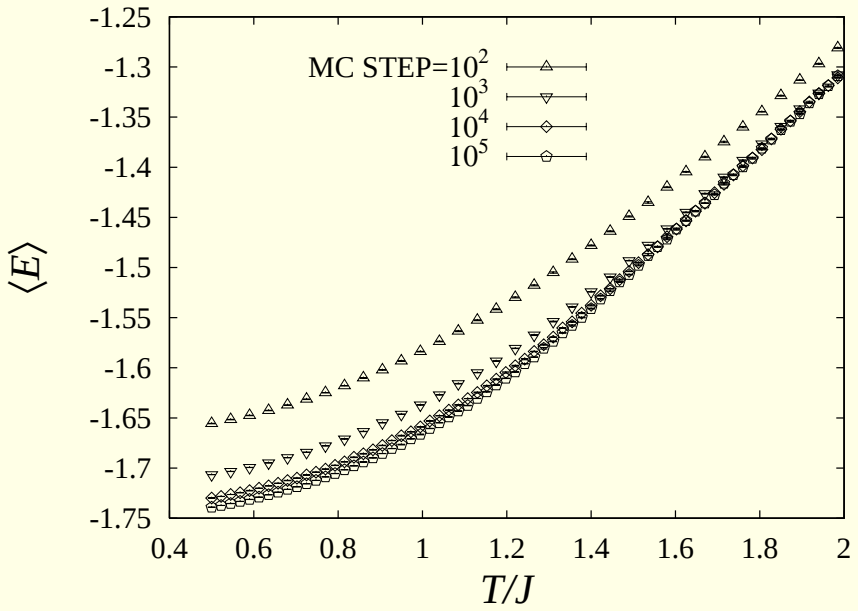
SAの例.....サンプリング手法ではない ...

スピングラス系へのSimulated Annealingの適用結果：

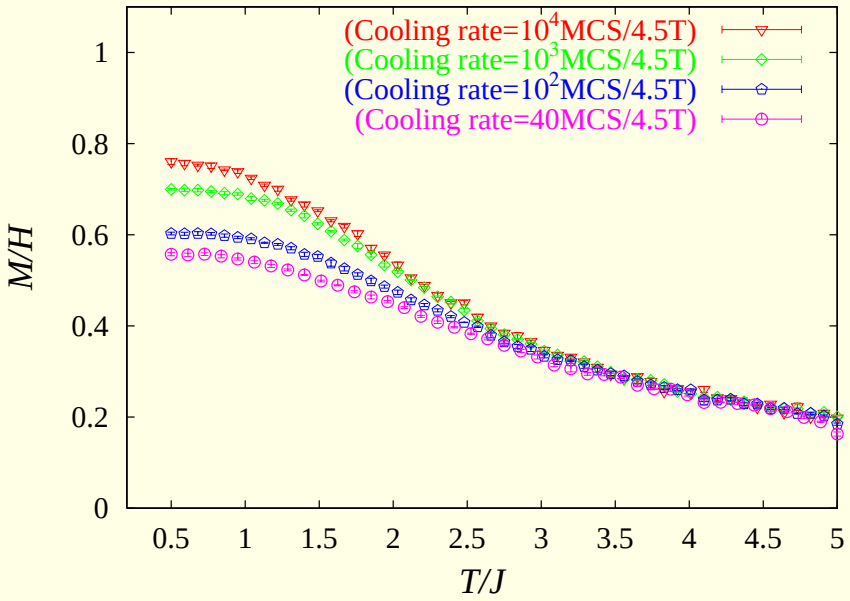
冷却速度を変えたときの物理量の期待値

定常状態が保証されていないので，アニーリング中の温度の情報はない．

エネルギー



磁場中磁化



冷却速度に依存した結果

$\Longrightarrow$ SA法は最適化手法であって，サンプリング手法ではない．

最近のモンテカルロ法の発展：前線へ

1. 非局所的状態更新 (クラスターアルゴリズム)
   - Swendsen-Wang (1987), Wolf (88): スピン系で固まりでフリップする方法
   - 高分子の非局所ムーブとして，pivot algorithm
2. 拡張アンサンブル法

   確率分布関数を拡張したり、合体したアンサンブルを考えてみる。

   - マルチカノニカル法：Berg-Neuhaus, (1991)
     - entropic sampling：Lee
     - Broad histogram MC：Oliveira (1998)
     - Flat histogram MC，Transition Matrix MC：Wang (1999)
     - Wang-Landau method....
   - Simulated tempering：Marinari-Parisi, (1992)
     - Expanded ensemble method：Lyubartsev et. al. (1992)
   - 交換法：Hukushima-Nemoto, (1996)
     - Multiple coupled Markov chain MC，Parallel tempering

Exchange Monte Carlo Method (1)

レプリカ系：ある調べたい系 $E(X)$ に対して，M 個のレプリカ系を考える。

$$E_{\rm eff}(\{X\}) = \sum_{m=1}^{M} \beta_m E(X_m),$$

拡張された状態は $\{X\} = \{X_1, X_2, \cdots, X_M\}$ で表される。

拡張された確率分布関数

$$P_{\rm eq}(\{X\};\{\beta\}) = \prod_{m=1}^{M} P_{\rm eq}(X_m;\beta_m) = \prod_{m=1}^{M} \frac{1}{Z(\beta_m)} \exp(-\beta_m E(X_m))$$

モンテカルロ・ステップ

1. (通常ステップ) それぞれのレプリカについて、状態更新

$$X_m \Longrightarrow X_m'$$

2. 2つのレプリカ系 X_m と X_n について状態の交換

$$\{X_m, X_n\} \Longrightarrow \{X_n, X_m\}$$

Exchange Monte Carlo Method (2)

詳細釣り合いの条件

$$P(\{\cdots,X,\cdots,X',\cdots\};\{\cdots,\beta_m,\cdots,\beta_n,\cdots\})\times W(X,X';\beta_m,\beta_n)$$
$$=P(\{\cdots,X',\cdots,X,\cdots\};\{\cdots,\beta_m,\cdots,\beta_n,\cdots\})W(X',X;\beta_m,\beta_n).$$

$$\frac{W(X,X';\beta_m,\beta_n)}{W(X',X;\beta_m,\beta_n)}=\exp(-\Delta),$$

交換のコスト関数は、

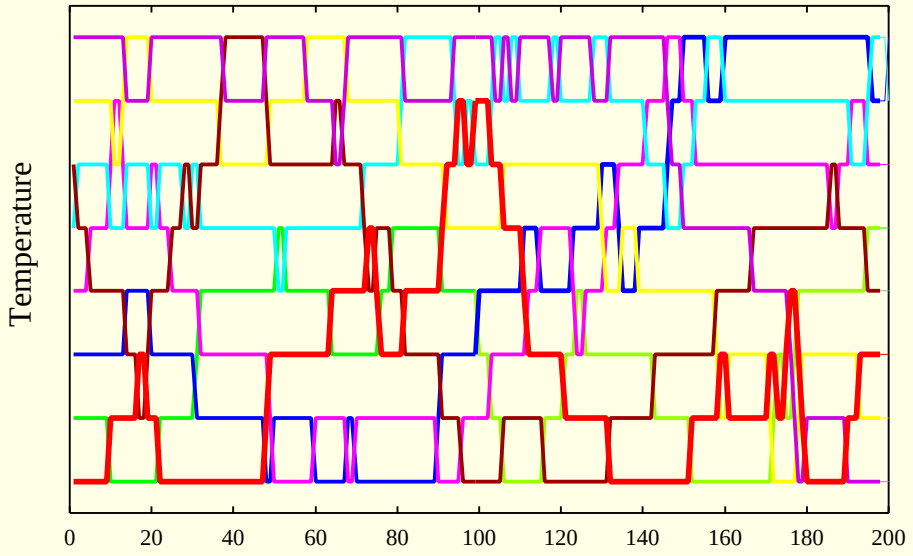
$$\Delta(X,X';\beta_m,\beta_n)=(\beta_n-\beta_m)(E(X)-E(X')).$$

交換のための遷移確率

$$W(X,X';\beta_m,\beta_n)=\begin{cases}\min[1,\exp(-\Delta)], & \text{for Metropolis type,}\\ \frac{\exp(-\Delta/2)}{\exp(-\Delta/2)+\exp(\Delta/2)} & \text{for heat bath type.}\end{cases}$$

Exchange Monte Carlo Method (3)

Monte Carlo Procedure



1. 通常のMCの手続きに従って、それぞれのレプリカは独立に同時に状態更新する

$$\{X_m\} \Longrightarrow \{X_m^{'}\}$$

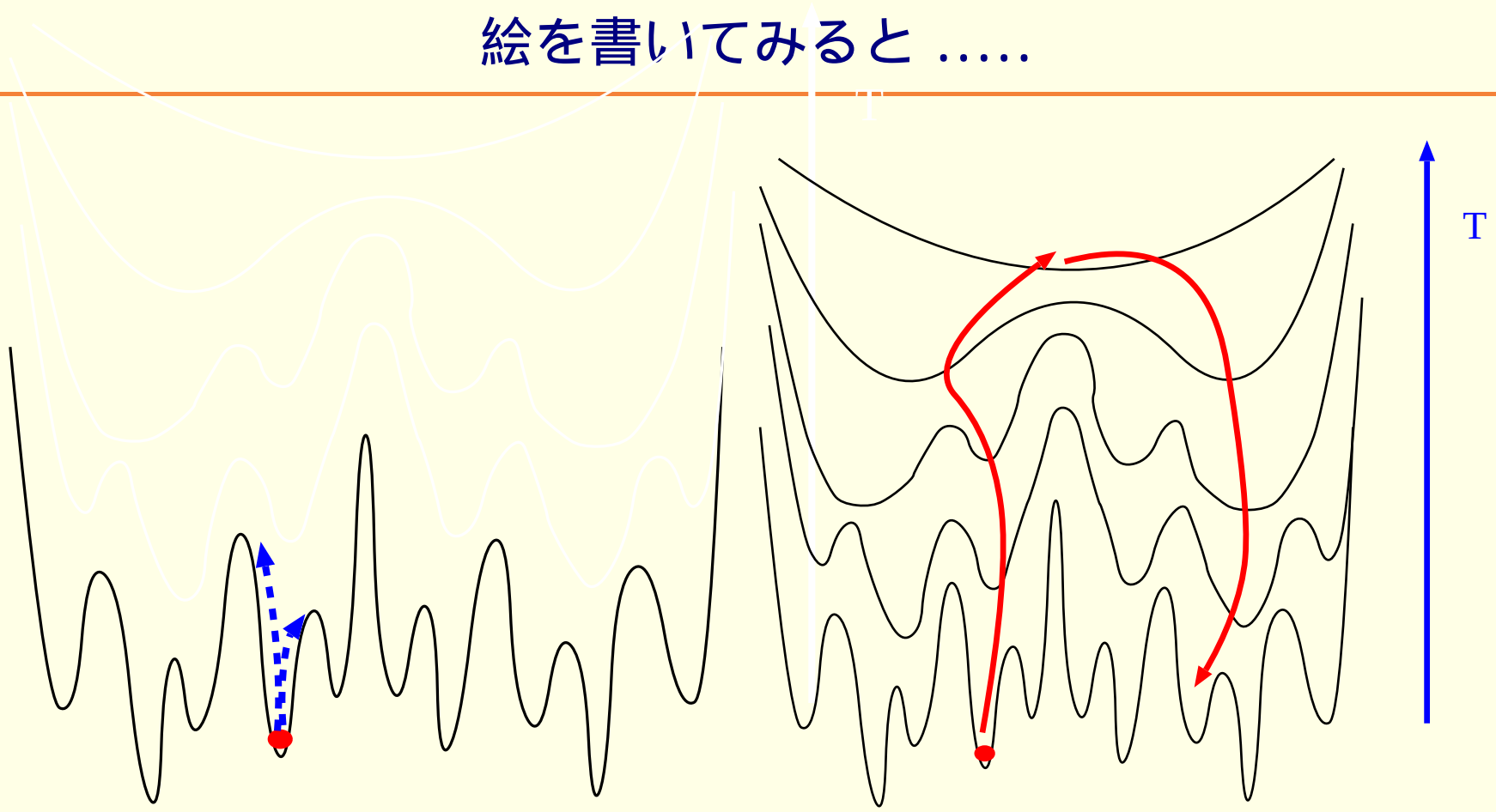
2. 2つのレプリカの状態 X_m と X_{m+1} を遷移確率 $W(X_m, X_{m+1})$ に従って交換する

$$\{X_m, X_{m+1}\} \Longrightarrow \{X_{m+1}, X_m\}$$

- それぞれのレプリカ状態は温度軸上をランダムウォークする
  - 自分自身で自動的アニーリング(ヒーティング)
  - 温度上昇もするので、準安定状態から脱出する手段を備えている

- それぞれの温度では定常分布が再現される：平均値の計算ができる．

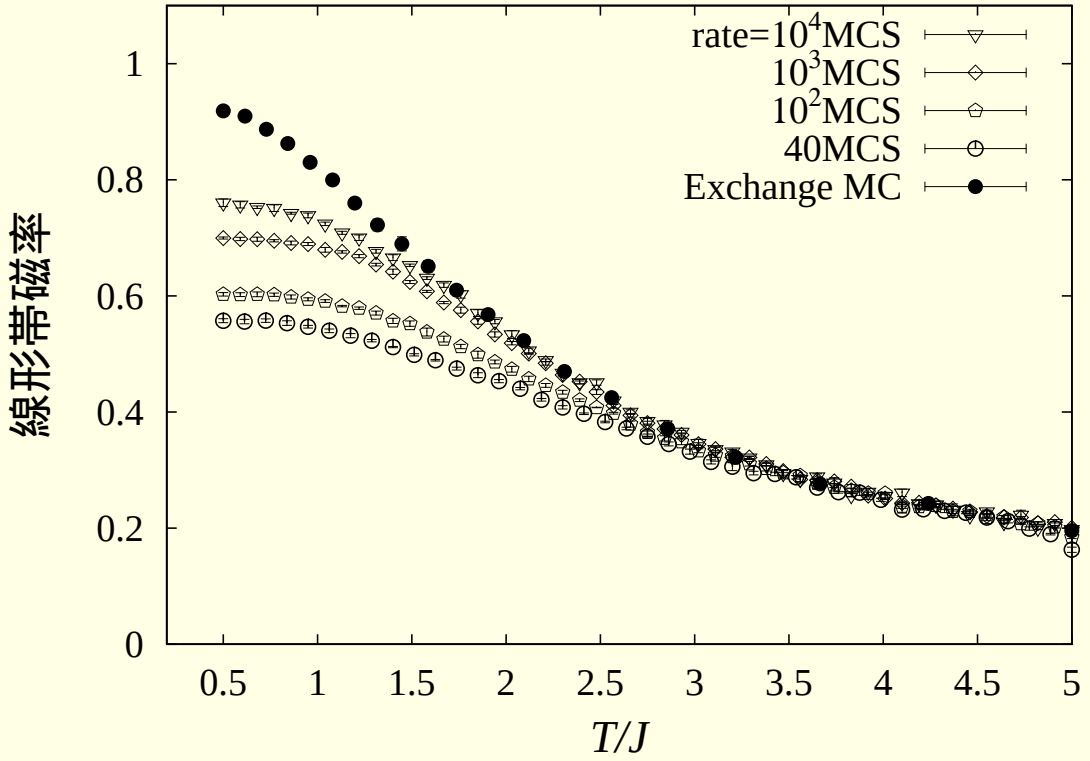
絵を書いてみると .....



深みにはまりこんだら，這い上がるのではなく，回り込むのがよい．

交換法の応用例

先のスピングラスの例



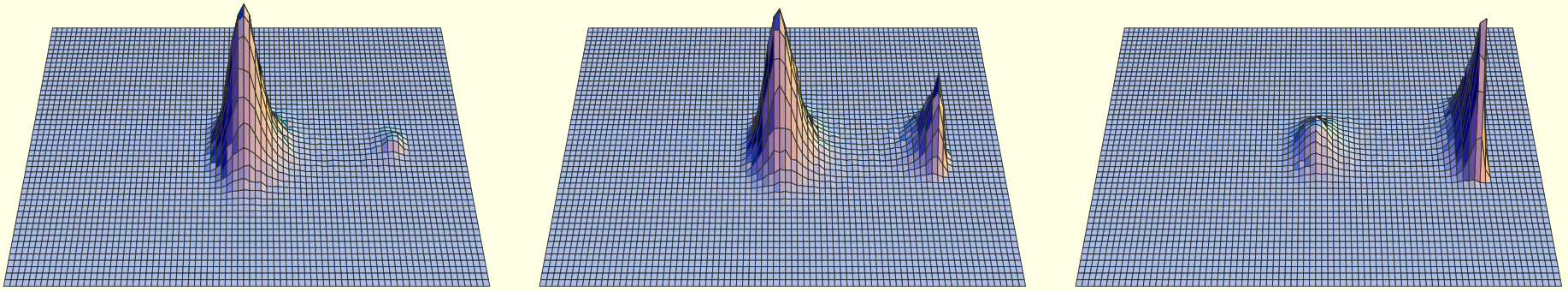
- 今日お見せしたイジング模型の例，N クイーン問題の例も交換法の結果

交換モンテカルロ法の応用例の一部

- **Polymer**
  - A. Irbäck and E. Sandelin, J. Chem. Phys. 110 (1999) 12256.
- **Protein folding simulation** ：
  - A. Mitsutake, Y. Sugita, Y. Okamoto, Biopolymer 60 (2001) 96.
- **Glass transition, Supercooled liquid**
  - R. Yamamoto and W. Kob, PRE 61 (2000) 5473.
- **Strongly correlated system**
  - P.Sengupta, A.W.Sandvik and D.K.Campbell, PRB **65** 155113, (2002).
- **Optimization Problems, SAT Problems**
  - K. Pinn and C. Wieczerkowski(1998)
  - K. Hukushima, Comp. Phys. Comm **147** 77, (2002)
- **Spin Glasses, Random spin systems**
  - K. Hukushima, ...many other groups....

Sourlas符号への応用： IBIS2003のポスター発表 (理研 井上さん)

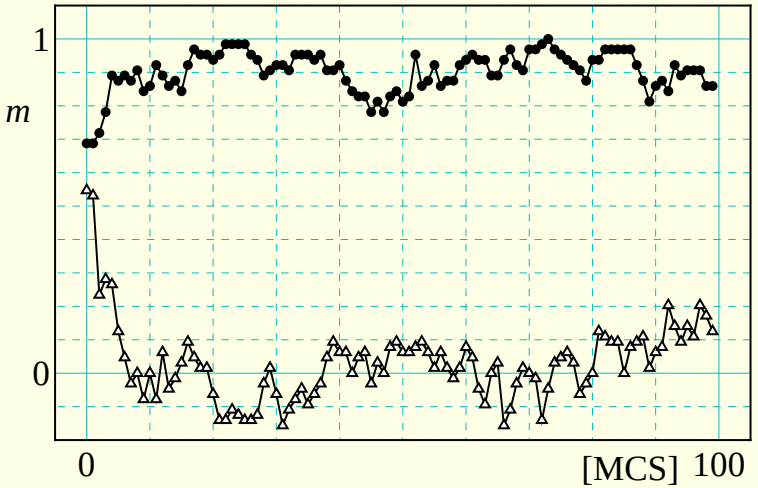
ある量のヒストグラム： 温度$T = 1.20, 1.15, 1.10$



交換法を用いることで，復号成功成分 (右の山) と失敗成分 (左) を同時に確認できる．

共存温度での復号ダイナミクス

右の温度で，いろいろな初期条件から通常 MC での復号を試みると，ある初期条件を境界に失敗相に入り込んでしまう．



まとめ

- (マルコフ連鎖) モンテカルロ法の基本として，メトロポリス法の解説
  - メトロポリス法，熱浴法
  - メトロポリス・アルゴリズムのイジング模型への具体的な応用
- 最近の進展のひとつとして，交換法について詳しく解説
  - 具体的な応用例については，それほど紹介できていない．
  - マルチカノニカル法，あるいはその亜種の応用例は交換法よりも多いくらい．
  - ヒストグラム法等のテクニカルな発展は他にもいろいろある．

参考図書として追加!

http://www.ism.ac.jp/ iba/iwanami.html

- 私は物理屋なので，物理の話題が中心になってしまった．
- テキストの参考文献には挙げなかったが，ベイズ統計と統計物理に関する入門書で，かつ，モンテカルロ法についても触れているのためにここで紹介したい．
- 本当は，共同研究者の近刊の本を引用していなくて，しかられただけ．