

物性物理におけるモンテカルロ法

川島 直輝 (東京大学物性研究所)

1 モンテカルロ法

1.1 物性物理におけるモンテカルロ法の「守備範囲」

今日物性物理の多くの分野で数値計算のさまざまな手法が使われている。このなかで、個々の構成要素である原子・分子の性質まで含めて全てを計算でカバーするやり方は「第一原理計算」といわれ、工業的な応用を目指す分野においては特に重要な計算手法である。しかし、基礎論的な興味からはしばしば議論されるのは、原子・分子の個別的な性質を捨て去ったあとに残る普遍的な性質、たとえば相図のトポロジカルな性質や、相転移付近で見られる臨界現象であり、これらを議論するには、系のミクロな構造について大胆に単純化したモデルが用いられる。磁性体、合金、液相・気相相転移、神経回路など多くの異なる物理系のモデルであるイジングモデルはその代表例であるが、ミクロなレベルでの基本原理すら異なっている複数の全く異なる系の振る舞いが同一のモデルで記述される、ということ自体、単純化された多体モデルがもつ普遍性を示している。また、基礎論的な興味だけでなく、応用上もモデルパラメータを適当に選びさえすれば十分に正確に現実の物性とあうようにできる場合があり、モデル計算は物性の多くの分野で盛んに行われている。本講義では主にモデル計算に用いられる計算手法について議論する。

モデル計算に用いられている計算手法を大別すると以下のようなになる

- (1) 有限系の厳密解
- (2) 級数展開 (広い意味の摂動論)
- (3) 分子動力学法
- (4) モンテカルロシミュレーション
- (5) 有限要素法 (モデルが変微分方程式の形をとる場合)
- (6) 問題に応じて特殊化した計算手法

無限に大きな系の厳密解が分かるくらいであれば数値計算は必要ないので、ここでは、そのような解を得るのが不可能である場合を考える。無限自由度では不可能であっても、有限自由度であれば計算可能な場合はある。状態空間の次元が有限で十分小さければ、通常は数値的に厳密な取り扱いが可能であり、これが(1)の有限系の厳密解による方法である。 N 個のスピンを含むイジングモデルの場合であれば、 2^N 回程度の計算時間と N 程度のメモリがあればあらゆる物理量の計算が可能である。たとえばおおざっぱに100万回程度足し算や掛け算を実行すれば、20個のスピンの系からなる系の厳密解を計算できる。量子系では古典系に比べてさらに計算に必要な時間やメモリが大きい。たとえば $S = 1/2$ のハイゼンベルクモデ

ルであれば、 N 個のスピンの系を計算するには、 2^N 次元の行列の対角化をしなければならず、これには、 2^N 程度のメモリと 2^{3N} 回程度の計算時間が必要になる。いずれにしても、状態空間の次元が有限とはいえ、その次元は系の構成要素（自由度）の数の関数として指数関数的に増大するので、厳密解を得られるのは、比較的小さい系に関する場合に限られる。問題によっては小さな系の計算によって、問題の本質的な部分が理解できる場合もあるが、臨界現象や相転移などのような協力現象に関しては一般に多数の自由度の系を解く必要があり、有限系厳密解ほ方法では解決しないことが多い。とはいえ、小さい系に関してならどのような情報をも得られるという意味で、厳密解の方法は他の方法にはない長所を持つ非常に強力な方法である。また、指数関数的資源の増大のため、大型のスーパーコンピュータでやろうとパソコンでやろうと計算できる系の大きさには（物理的な観点からは）それほど違いがなく、結果として、研究費が恵まれていない研究者にも使える手法であるとも言える。

(2) の級数展開の代表例は系の分配関数やその他の物理量を温度の逆数で展開したときの係数を求める高温展開である。この方法の長所は高温展開の係数はシステムサイズに依存しないため、係数を求めること自体に熱力学極限における物性を知るという意味がある。しかし、厳密解の方法と同様、例外的な場合を除き、計算する最大次数とそれに必要な計算資源の間には指数関数的な関係があって、計算機の性能が2, 3桁増大しただけでは、物理的観点からいって、さほどの進展は望めない。また、収束半径の問題が付きまとい、収束円内の領域（高温展開では、高々相転移温度の逆数まで）については比較的信頼できる物理量の評価値が得られるが、収束円を超えた領域（転移温度以下の秩序相）に関する情報が得られない（または得にくい）という欠点もある。最近ではこの講義の中心的な話題であるモンテカルロ法を応用して級数展開を行うという方法も提案されているが、その場合得られる係数は統計誤差を含んだものとなり、結局、手段においても得られる結果においても通常のモンテカルロ法と変わらず、通常のモンテカルロ法と同じ長所短所を持つので、ここではそのような手法をモンテカルロ法として扱う。

(3) の分子動力学法は1950年前後、電子計算機の黎明期にロスアラモスで初めて試みられた多体問題計算法であり、剛体球の系で試された。これは、ニュートン方程式やシュレーディンガー方程式など物理的な運動方程式を忠実に（あるいは熱揺らぎを導入する項をいれて）解くことによって時間変化を追う方法であって、熱揺らぎの項が含まれていない場合は基本的にミクロカノニカル分布を生成する方法である。この方法の最大の長所は時間発展に関する情報を得ることができる点である。一方、物理系の時間発展を忠実に再現する方法であるために、遅い緩和現象をもつ物理系に対して応用すると、当然に計算にも時間がかかることになる。例えばたんぱく質のフォールディングの問題については、現象がおきる時間のスケールにくらべて計算上の時間刻みが非常に短いので、計算ステップ数を非常に大きくしなければならないことが、研究上の大きな障害になっている。そのため、時間発展の情報が得られるという長所を犠牲にしてモンテカルロ法と組み合わせて用いるようなやり方が提案されている。

(4) のモンテカルロ法は、本稿の主要テーマで、分子動力学法と同様、電子計算機の歴史の初期にロスアラモスのグループによって考案された。¹ 以下で詳述す

¹N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. "Equations of

講義ノート

るが、この方法の長所は、適用範囲が広く、上述の指数関数的な計算時間、メモリの増大の問題がないために大きな系が扱える、ということである。避けられない欠点は得られる結果に統計誤差がある、ということであるが、これは緩和時間の問題や負符号問題に比べると本質的でない。緩和時間が計算可能な時間の範囲内で、負符号問題がなく、単に統計誤差だけが問題であるならば、計算時間を大きくすることによっていくらかでも誤差を小さくしていけるからである。しかし、一般には臨界点近傍や低温極限で緩和時間が発散したり、準安定状態にはまりこむという問題が生じる。

(5) の有限要素法は、解くべきモデルが変微分方程式の形になっている場合で、流体・弾性体の方程式、場の理論で議論されるモデル、ランダウ・ギンツブルグ方程式、などが代表的な例である。ボーズ凝縮体の時間発展を近似的に表現する Gross-Pitaevski 方程式もこのカテゴリーに分類される。この手法の長所は、繰り込み群など解析的な理論で扱われることが多いモデルであるためにそのような場合には直接的な比較が可能であることである。ただ、格子モデルを調べるほうが数値的には簡便であることが多いので、もし本質的に同一の現象を示すと期待される格子モデルが存在するならば、離散化されたモデルがよく研究される。

このほか、物性で登場する多体問題は実に多岐に渡っているもので、それぞれの問題の特性に応じていろいろな計算手法が考案されている。上で述べたのは比較的広い範囲の問題に適用されている手法であるが、適用範囲は狭いながら、その範囲では強力な計算手法は多い。例えば、1次元量子系に対して非常に強力な手法として、密度行列繰り込み群法などがある。

1.2 単純なモンテカルロ法（棄却法）

正方形に内接するように描かれた円に向かって十分遠くからダーツを投げる。正方形のなかにあたったもののうち、円内に当たったものの割合を調べることで、円周率を求めることができる。これは単純モンテカルロ法であり、いろいろな計算に応用できる。原理的には統計力学のあらゆるモデルの分配関数やカノニカル平均の計算をこの方法で行うことができる。実際には効率が悪すぎて使い物にならないのだが、これを考えることで分かることもあるので、ここから話を始めることにする。

一般にモンテカルロ法とは、乱数を用いて多変数の空間で何らかの与えられた重み付きの平均値を求める問題である。考える変数空間を Ω 、その空間で定義された重みを $W(\Sigma)$ ($\Sigma \in \Omega$)、平均値を求めたい関数を $Q(\Sigma)$ とすると、平均値は

$$\langle Q \rangle \equiv \frac{\sum_{\Sigma \in \Omega} W(\Sigma) Q(\Sigma)}{\sum_{\Sigma \in \Omega} W(\Sigma)} \quad (1)$$

と書ける。 Ω として微視的状态の集合すなわち位相空間、 $W(\Sigma)$ としてボルツマン重み $W(\Sigma) = \exp(-\beta E(\Sigma))$ 、を考えれば、統計力学的なカノニカル平均を求める方法であるとみなせる。

State Calculations by Fast Computing Machines". Journal of Chemical Physics, 21(6):1087-1092, 1953.

もっとも簡単なモンテカルロ法は状態をランダムに発生させたあと、重みに応じてそれを採用したり棄却したりする方法である。重みに上限があれば $W(\Sigma) \leq W_{\max}$ として、(1) を計算する手順として次のようなものが考えられる。

単純モンテカルロ法：

ステップ 1： (乱数を用いて何らかの方法で) Ω の中から 1 つの要素を選ぶ。その際全ての状態は同じ確率で選ばれるようにする。(選ばれた状態を Σ とする.)

ステップ 2： 0 以上 1 未満の一樣乱数を発生し、それが $P_{\text{accept}} = W(\Sigma)/W_{\max}$ 未満なら 1 で選んだ状態を採用する。逆に乱数が P_{accept} 以上なら棄却する。

ステップ 3： 1 に戻る。

この方法で状態 Σ が採用されるたびに $Q(\Sigma)$ を計算し、その値を単純に平均したものがモンテカルロ平均である。

試行回数が多い極限で、モンテカルロ平均は求めたい平均値 (1) に等しくなる。これは以下のように理解できる。ランダム生成された状態が棄却される場合も、値が 0 であったこととしてカウントにいれることにすると、第 i 回目の試行でえられる Q の値、 Q_i の期待値は

$$\overline{Q}_i = \sum_{\Sigma \in \Omega} \frac{1}{|\Omega|} \frac{W(\Sigma)}{W_{\max}} Q(\Sigma) = (|\Omega|W_{\max})^{-1} \sum_{\Sigma \in \Omega} W(\Sigma)Q(\Sigma)$$

となる。試行回数が十分大きければ中心極限定理によって単純平均は期待値に一致するので

$$\frac{1}{N} \sum_{i=1}^N Q_i \approx (|\Omega|W_{\max})^{-1} \sum_{\Sigma \in \Omega} W(\Sigma)Q(\Sigma)$$

特別な場合として、 $Q(\Sigma) = 1$ の場合の単純平均はランダムに発生された状態が採用される回数を全試行回数で割ったものに他ならず、

$$\frac{1}{N} N_{\text{accepted}} \approx (|\Omega|W_{\max})^{-1} \sum_{\Sigma \in \Omega} W(\Sigma)$$

となる。これらを辺々割り算したものが、モンテカルロ平均であるから、

$$\langle Q \rangle_{\text{MC}} \equiv \frac{\sum_{i=1}^N Q_i}{N_{\text{accepted}}} \approx \frac{\sum_{\Sigma \in \Omega} W(\Sigma)Q(\Sigma)}{\sum_{\Sigma \in \Omega} W(\Sigma)} \quad (2)$$

これは、求めたかった重みつき平均 $\langle Q \rangle$ に等しい。

講義ノート

ダーツなげで円周率を求める方法はまさに単純モンテカルロ法の典型である。その場合、 Ω は xy 平面、 $W(x, y)$ は針の落ちた場所の座標が $-1 < x, y < 1$ の範囲にあれば 1, なければ 0 であるような関数。 $Q(x, y)$ は針の落ちた場所の座標が原点を中心とする半径 1 の円内にあれば 1, なければ 0 であるような関数であり、当然 $\langle Q \rangle = \pi/4$ であるから、 $4 \times \langle Q \rangle_{MC}$ を π の近似値とする。

1.3 単純モンテカルロ法の限界

上記のやり方をそのまま統計力学的なモデルに当てはめることは通常は可能であり、自由度が少数であれば実際に計算もできる。たとえば、正方形の 4 隅にスピンを配置した 4 スピンからなるイジングモデルであれば、各スピンの向きをでたらめに設定して（といっても 16 通りのうち 1 通りを選ぶだけだが）、その状態に対するボルツマン重みを計算し、それを重みの最大値で割った確率で状態を採用する。このやり方で全ての物理量を計算することができる。しかし、我々は通常 4 スピンの系にはそれほど関心がなく、自由度数がもっと大きい場合を計算したい。その場合には、この単純なモンテカルロ法は全く機能しない。その理由は、熱力学に寄与すべき状態のほとんど全てについて、状態の採用確率 $W(\Sigma)/W_{\max}$ が小さすぎることである。

これは、熱力学を支配するのがエネルギーでなくエントロピーも考慮した自由エネルギーであるからである。分配関数

$$Z = \sum_{\Sigma \in \Omega} e^{-\beta E(\Sigma)}$$

をエントロピー

$$S(E) \equiv \log \left(\sum_{\Sigma \in \Omega} \Delta(E - E(\Sigma)) \right)$$

を用いて書くと

$$Z = \sum_E e^{-\beta(E - TS(E))}$$

と書ける。ここから、状態和に対して寄与が大きいのは、自由エネルギー

$$F = E - TS(E)$$

の最大値を与えるエネルギーの値 $E^*(T)$ 付近の状態であることがわかる。もちろん、状態単独の重みとしては最低エネルギー値 E_0 をとる基底状態の方が大きいですが、エネルギーが E^* である状態数が莫大であるために、結果として、有限温度の性質に対する基底状態の寄与は無視できるほど小さいものになる。しかし、エネルギーが E^* である状態が単純モンテカルロ法において採用される確率は

$$P_{\text{accept}} \equiv e^{-\beta E^*} / e^{-\beta E_0} = e^{-\beta(E^* - E_0)}$$

であり、エネルギーは系の自由度数 V に比例するので、 $E^* - E_0$ は一般に V に比例して大きくなる。つまり、採用確率は

$$P_{\text{accept}} \propto e^{-aV}$$

の形をとることになって、これは指数関数的に小さい。さらに、ランダムに状態を発生させたときに、エネルギーが E^* である状態が発生される確率はエントロピー $S(E)$ が最大になるエネルギーを E_1 として、

$$P_{\text{generate}} = e^{S(E^*)} / \sum_E e^{S(E)} < e^{-(S(E_1) - S(E^*))}$$

となる。エントロピーもやはり示量性の量であって一般に $S(E_1) - S(E^*)$ は V に比例するから、この確率もやはり e^{-aV} の形になっており指数関数的に小さい。つまり、エネルギーが E^* である状態は、分配関数への寄与が最大であるにも関わらず、単純モンテカルロ法をそのまま適用すると、候補として発生されることからして稀であり、発生されたとしても採用される確率はきわめて小さい、ということが分かる。

以上より、モンテカルロ平均の式

$$\langle Q \rangle_{\text{MC}} \equiv \frac{\sum_{i=1}^N Q_i}{N_{\text{accepted}}} \quad (3)$$

において、分母分子は0であるか非常に小さい数になることが分かる。これが求める期待値になるのは中心極限定理のおかげであるから、採用されるサンプル数が少なければそれは成り立たない。

1.4 マルコフ過程を用いたモンテカルロ法

前節の例から分かることは、棄却法では、和への寄与が大であるような状態を効率よく発生できないことであり、これは系のサイズが小さければ問題ないが、サイズの増大とともに指数関数的に状況が悪化するということである。一般に、積分や平均値に大きく寄与する状態の集団が全状態の集団に比べてきわめて小さいような場合には、全ての状態がいったんは同じ確率で登場するようなやり方をすると使えない状態ばかり登場することになってうまくいかない。そのような場合には、そもそも寄与が大きな状態が最初から優先的に発生されるような方法を考えるしかない。マルコフ過程を用いるモンテカルロ法はそれを狙った方法である。

乱数を用いるあるルールに従って、次々と発生される状態の系列 $\{\Sigma_1, \Sigma_2, \Sigma_3, \dots\}$ を考えよう。とくに、ある時刻での状態が、ルール、直前の状態、そしてその時々発生される乱数の3者にしか依存しないような場合を考える。その場合、現在の状態が Σ' であるときに直後の状態が Σ になる確率は、時刻によらないある関数 $T(\Sigma|\Sigma')$ で表されることになる。次の状態を決めるルールに関する情報は全て T に含まれている。マルコフ過程とは、このように確率的に次々と発生される状態の系列で、しかも未来の状態の出現確率が現在の状態にしか依存せず、過去にどのような履歴であったかには依存しないようなものことである。

マルコフ過程において、 t 番目の状態が Σ である確率を $P^t(\Sigma)$ とすると、

$$P^t(\Sigma) = \sum_{\Sigma'} T(\Sigma|\Sigma') P^{(t-1)}(\Sigma') \quad (4)$$

講義ノート

という漸化式が成立する。ここで、 $P^t(\Sigma)$ をベクトル \mathbf{P}^t の第「 Σ 」成分、同様に、 $T(\Sigma|\Sigma')$ を行列 T の第「 Σ 行 Σ' 列」成分とみなすと、

$$\mathbf{P}^{(t)} = T\mathbf{P}^{(t-1)} \quad (5)$$

と書ける。行列 T はある時刻の状態から次の時刻での状態への遷移を特徴づける確率なので、遷移確率とか、遷移行列とか呼ばれる。

マルコフ過程を用いたモンテカルロ法で、重み W の重みつき平均値を計算するには、 $\lim_{t \rightarrow \infty} P^{(t)}(\Sigma) \propto W(\Sigma)$ となるように遷移行列 T を選ぶ。もし、このように T を選ぶことができれば、十分長いマルコフ過程において、ある時刻から先は状態は $W(\Sigma)$ に比例した確率で出現する。単純モンテカルロ法でのステップ 2 の役割はもともと一様な確率で発生させてしまった状態にフィルターをかけることで、フィルターを通った状態の出現確率が $W(\Sigma)$ に比例するようにすることであった。もともと状態が $W(\Sigma)$ に比例した確率で発生されていれば、ステップ 2 を省略し、全ての状態を「採用」すればよい。

以下ではこのような原理に基づくモンテカルロ法をモンテカルロシミュレーションと呼ぶことにして、それ以外のモンテカルロ法と区別することにする。

1.5 詳細釣り合い

出現確率が望みの重み $W(\Sigma)$ に比例したものに収束するように遷移確率を選ぶことはいつでも可能なのだろうか？可能だとすると、具体的にはどうすればよいのか？

やや天下りだが、与えられた重み W と任意の状態の組 Σ, Σ' に対して、次の詳細釣り合いと呼ばれる条件式を満たす遷移確率 $T(\Sigma|\Sigma')$ を考える。

$$T(\Sigma|\Sigma')W(\Sigma') = T(\Sigma'|\Sigma)W(\Sigma). \quad (6)$$

T は確率なので、当然

$$\sum_{\Sigma'} T(\Sigma'|\Sigma) = 1 \quad (7)$$

も満たさなければならない。このとき、 \mathbf{W} は行列 T の固有値 1 の固有ベクトルである。なぜなら、詳細釣り合いの条件式 (6) から

$$\sum_{\Sigma'} T(\Sigma|\Sigma')W(\Sigma') = \sum_{\Sigma'} T(\Sigma'|\Sigma)W(\Sigma)$$

となるが、右辺は確率としての規格化条件 (7) により $W(\Sigma)$ に等しい。つまり行列、ベクトルの記法で

$$T\mathbf{W} = \mathbf{W}.$$

$T^t\mathbf{P}^{(0)} = \mathbf{P}^{(t)}$ が t を大きくしたときに \mathbf{W} に収束するためには、少なくとも収束したあとの \mathbf{W} に T を作用させたときにそれ以上変化があってはいけないから、上の式は、 T が満たすべき条件が確かに満たされていることを示している。しかし、これだけでは、十分条件であることまでは示したことになる。実際、 T として単位行列を考えれば、明らかに (6) と (7) を満たすが、それは、いつまでたっても最初の状態を生成し続けるだけで役に立たない。

1.6 エルゴード性

遷移行列が単位行列に等しい場合は、時刻0での状態以外は明らかに現れようがない極端な場合であるが、それほど極端でなくても、与えられた初期状態に対して原理的に出現しようがない状態が存在するような遷移確率を用いると、仮にそれが詳細釣り合いを満たしていても、望みの性質は得られない。このような場合を非エルゴード的であるという。逆に、どのような初期状態からはじめたとしても、他の任意の状態がやがては出現しうるような場合をエルゴード的であるということにする。すなわちエルゴード的であるとは、

$$\forall(\Sigma, \Sigma') \left(\exists t \left(T^t(\Sigma'|\Sigma) > 0 \right) \right) \quad (8)$$

この条件だけだとやや弱すぎて不便であるので、初期状態によらないある t が存在して、それ以上に時間が経つと、どのような状態も出現確率が0でないようになることを狭義エルゴード的ということにする。形式的には狭義エルゴード的であるとは、

$$\exists t \left(\forall s > t \left(\forall(\Sigma, \Sigma') \left(T^s(\Sigma'|\Sigma) > 0 \right) \right) \right) \quad (9)$$

が成り立つことである。

1.7 イジングモデルのモンテカルロシミュレーション

遷移確率 T のマルコフ過程の極限分布 $\lim_{t \rightarrow \infty} P^{(t)}$ が目的の重み（統計力学ならボルツマン重み） W に比例するものに収束するには、次節で示すように T が詳細釣り合いとエルゴード性を満たすことが十分条件である。よって、望みのマルコフ過程を得るには、これらを指針にして、適当なものを探せばよいことになる。そのようなものの代表例がイジングモデルのシングルスピンフリップによるモンテカルロシミュレーションである。これは以下のような手続きからなる。

シングルスピニアップデート法：

ステップ 0： 適当に初期状態のスピンの配置（たとえばランダムな状態）を用意する。

ステップ 1： 一様ランダムに1つスピンを選ぶ。（以下それを S_0 と呼ぶ）

ステップ 2： S_0 以外の全てのスピンは現在の状態のままにしておいて、 S_0 が上向きである状態 (Σ_{\uparrow}) のエネルギー E_{\uparrow} と下向きに向きである状態 (Σ_{\downarrow}) のエネルギー E_{\downarrow} の差 $\Delta E \equiv E_{\uparrow} - E_{\downarrow}$ を求める。

ステップ 3： 0以上1未満の乱数 r を1つ発生して、それが

$$P_{\uparrow} \equiv \frac{1}{1 + e^{\beta \Delta E}}$$

未満であれば、現在の状態を Σ_{\uparrow} に置き換える。 $r \geq P_{\uparrow}$ であれば、 Σ_{\downarrow} に置き換える。

ステップ 4： ステップ 1に戻る。以下繰り返す。

（注：エネルギー E_{\uparrow} を求める操作には一般にはスピンの総数に比例した計算時間がかかるが、 ΔE を求めるだけであれば、 S_0 と S_0 と直接相互作用しているスピンだけに着目すればよいから、通常は $O(1)$ の計算になる。）

例題 この操作の1サイクルを表す遷移行列を書き下し、それがボルツマン重みを目的の重みとする ($W(\Sigma) = e^{-\beta E(\Sigma)}$) 詳細釣り合いを満たすことを示せ。

1.8 モンテカルロシミュレーションの収束証明

ここでは、状態を Σ でなく i などの文字で表し、状態の重みを $W(\Sigma)$ でなく W_i などと書くことにする。与えられた正の重みベクトル $W_i > 0$ に対して行列 $T = (T_{ij})$ が詳細釣り合いと狭義エルゴード性などをみたせばモンテカルロシミュレーションで出現する状態の確率分布が収束して、しかも、その分布は W に比例することが示せる。つまり、以下が成り立つ。

行列 T_{ij} について

$$T_{ij} \geq 0 \quad (10)$$

$$\sum_i T_{ij} = 1 \quad (11)$$

$$\exists t (\forall s > t (\forall (i, j) ((T^s)_{ij} > 0))) \quad (12)$$

$$T_{ij}W_j = T_{ji}W_i \quad (13)$$

が成り立ち、ベクトル $\mathbf{P}^{(0)}$ が

$$P_i^{(0)} \geq 0 \quad (14)$$

$$\sum_i P_i^{(0)} = 1 \quad (15)$$

であるとき、 $\mathbf{P}^{(t)} \equiv T^t \mathbf{P}^{(0)}$ に対して

$$\lim_{t \rightarrow \infty} \mathbf{P}^{(t)} = \mathbf{P}^*$$

が存在して、

$$\mathbf{P}^* \parallel \mathbf{W}$$

である。

【証明】 T^t の全ての成分が 0 でない正の値をとるようになった時刻 t 以降について証明ができれば十分だから、 T^t をあらためて T とみなしてこれについて証明を行う。つまり、一般性を失わず $T_{ij} > 0$ を仮定してよい。

$P_i^{(t)} = Q_i^{(t)} W_i$ で $Q_i^{(t)}$ を定義する。 $Q_i^{(t)} \rightarrow \text{const}$ を示せばよい。 $U_t = \max_i Q_i^{(t)}$, $L_t = \min_i Q_i^{(t)}$, $D_t = U_t - L_t$ とする。 U_t は単調減少、 L_t は単調増加である。これは以下のように分かる。まず

$$\begin{aligned} U_{t+1} &= \max_i Q_i^{(t+1)} = \max_i \frac{1}{W_i} \sum_j T_{ij} P_j^{(t)} = \max_i \frac{1}{W_i} \sum_j T_{ij} W_j Q_j^{(t)} \\ &= \max_i \sum_j T_{ji} Q_j^{(t)} \\ &= \max_i \sum_j T_{ji} (U_t - (U_t - Q_j^{(t)})) \\ &= U_t - \min_i \sum_j T_{ji} (U_t - Q_j^{(t)}) \end{aligned}$$

であるが、ここで差は

$$\min_i \sum_j T_{ji} (U_t - Q_j^{(t)}) \geq \min_i \max_j T_{ji} (U_t - Q_j^{(t)})$$

講義ノート

$$\begin{aligned} &\geq \max_j T_{\min}(U_t - Q_j^{(t)}) \\ &\geq T_{\min}(U_t - L_t) = T_{\min} D_t \end{aligned}$$

と下限が評価できる. ここで, T_{\min} は

$$T_{\min} \equiv \min_{ij} T_{ij} > 0$$

である. つまり,

$$U_{t+1} \leq U_t - T_{\min} D_t$$

よって U_t は単調減少. 同様にして,

$$L_{t+1} \geq L_t + T_{\min} D_t$$

もわかる. 最後の2式から

$$D_{t+1} \leq D_t - 2T_{\min} D_t = (1 - 2T_{\min}) D_t$$

が得られる. 等号成立は明らかに $\forall i (Q_i^{(t)} = U_t = L_t)$ のとき. よって

$$D_t \leq |1 - 2T_{\min}|^t |D_0| \rightarrow 0 \quad (t \rightarrow \infty) \quad (16)$$

結局 $\min_i Q_i^{(t)}$ と $\max_i Q_i^{(t)}$ の差は $t \rightarrow \infty$ で0になる. つまり, $t \rightarrow \infty$ で $Q_i^{(t)}$ は i によらない.

1.9 「自由エネルギー」の単調増加性 (連続時間の場合)

前節でモンテカルロシミュレーションの平衡分布が正しい分布に一致することが示せた. とくに正しい分布が統計力学におけるギブス分布の場合には, これはモンテカルロシミュレーションの平衡分布が自由エネルギー最小の状態に収束することが言えたことになる. しかし, 途中の過程で, 自由エネルギーが常に単調に減少することまではわからない.

ここではモンテカルロシミュレーションのダイナミクスについて一般的に定義された自由エネルギーが単調減少であることを示す. まず見とおしの良い連続時間の極限をとったダイナミクスについて考える.

連続時間ダイナミクス:

$$\frac{dP_i^{(t)}}{dt} = \frac{1}{\tau} \sum_j (L_{ij} P_j - L_{ji} P_i)$$

ただし, L_{ij} は負でなく

$$L_{ij} W_j = L_{ji} W_i, \quad L_{ii} = 0$$

を満たす.

連続時間ダイナミクスにおいて時間を時間ステップ Δt で離散近時すると、モンテカルロシミュレーションの時間発展方程式となるが、その遷移行列は

$$T_{ij} = \delta_{ij} \left(1 - \frac{\Delta t}{\tau} \sum_k L_{ki} \right) + \frac{\Delta t}{\tau} L_{ij}$$

である。この場合、上の連続時間ダイナミクスの時間発展方程式の右辺は対応するモンテカルロシミュレーションダイナミクスを対角成分と非対角成分に分けて書いただけのものである。

さて、温度を T としてエネルギーを

$$E_i = -T \log W_i$$

その期待値を

$$E = \sum_i P_i E_i = -T \sum_i P_i \log W_i,$$

また、エントロピーは

$$S = \sum_i (-P_i \log P_i)$$

だから、自由エネルギー $F \equiv E - TS$ に対して

$$F \equiv E - TS = T \sum_i P_i \log \frac{P_i}{W_i} \quad (17)$$

連続時間ダイナミクスの時間発展方程式から、

$$\begin{aligned} \frac{dF}{dt} &= T \sum_i \left(\log \frac{P_i}{W_i} + 1 \right) \frac{dP_i}{dt} \\ &= T \sum_i \left(\log \frac{P_i}{W_i} \right) \frac{dP_i}{dt} \\ &= \frac{T}{\tau} \sum_i \log Q_i \sum_j (L_{ij} W_j Q_j - L_{ji} W_i Q_i) \\ &= \frac{T}{\tau} \sum_{ij} L_{ij} W_j \log Q_i (Q_j - Q_i) \\ &= -\frac{T}{2\tau} \sum_{ij} L_{ij} W_j (\log Q_i - \log Q_j) (Q_i - Q_j) \\ &\leq 0 \end{aligned}$$

となり、自由エネルギーが減少し続けることが示せた。

1.10 「自由エネルギー」の単調増加性 (離散時間の場合)

つぎに、もともとの離散時間の場合を考える。 $f(x)$ を任意の上に凸な関数 $f''(x) < 0$ とする。たとえば $f(x) = -x \log x$ とすれば上の場合に対応する。この f を用いて、

講義ノート

自由エネルギーを

$$-F \equiv T \sum_i W_i f\left(\frac{P_i}{W_i}\right) \quad (18)$$

と定義する。(T は熱力学の温度に対応しやすいように入れたが、ここでの議論にはあまり重要でない定数である。) モンテカルロシミュレーションのある離散時刻における量を F, P_i で表し、次の時刻の量をプライムをつけて F', P'_i などと表すことにする。すると以下が成り立つ。

$$\begin{aligned} -F' &\equiv T \sum_i W_i f\left(\frac{P'_i}{W_i}\right) \\ &= T \sum_i W_i f\left(\frac{1}{W_i} \sum_j T_{ij} P_j\right) \\ &= T \sum_i W_i f\left(\sum_j \left(T_{ij} \frac{W_j}{W_i}\right) \frac{P_j}{W_j}\right) \\ &= T \sum_i W_i f\left(\sum_j T_{ji} \frac{P_j}{W_j}\right) \quad \Leftarrow T_{ij} W_j = T_{ji} W_i \end{aligned}$$

ここで、 $\sum_j T_{ji} = 1$ であることから、以下の Jensen の定理が使える。すなわち、一般に

$$a_i \geq 0, \quad \sum_i a_i = 1$$

であれば、上に凸の関数 g に対して、

$$g\left(\sum_i a_i x_i\right) \geq \sum_i a_i g(x_i)$$

であり、かつ、等号成立は $a_i > 0$ であるような全ての i について x_i が同じ値であるときのみである。これを用いると、 f の値を下から抑えることが出来て、

$$-F' \geq T \sum_i W_i \sum_j T_{ji} f\left(\frac{P_j}{W_j}\right) \quad (19)$$

である。以下、

$$\begin{aligned} -F' &\geq T \sum_{i,j} T_{ji} W_i f\left(\frac{P_j}{W_j}\right) = T \sum_{i,j} T_{ij} W_j f\left(\frac{P_j}{W_j}\right) \\ &= T \sum_j W_j f\left(\frac{P_j}{W_j}\right) = -F \end{aligned}$$

となつて、

$$F' \leq F \quad (20)$$

すなわち、自由エネルギーが単調減少であること、言い換えると、モンテカルロシミュレーションにおいて、熱力学第2法則が成り立っていることが分かる。

さらにここから、モンテカルロシミュレーションの収束のもう少し物理的な証明をすることができる。(20)において等号成立は、任意の i について、 i との間に直接遷移のある全ての状態 j が等しい $Q_j = P_j/W_j$ の値を持つことである。たとえば、状態 i_1 と i_2 がともに i との間に0でない直接遷移確率を持っているときに、等号成立状態において、 $Q_{i_1} = Q_{i_2}$ でなければならない。このようなときに、 i_1 と i_2 は「連結」していると呼ぶことにする。

一般には i は i 自身に「連結」しているとは限らないが、以下でこの性質を使うので、 $T_{ii} > 0$ であるとして、全ての状態は自分自身と連結しているとする。そうすることで「連結」という言葉のもつ普通のイメージが正しいことが保障される。また、モンテカルロシミュレーションの実際の適用例の多くでもこの性質はある。

この条件のもとでは、 i から j への直接遷移があることと i と j は直接連結されていることは同じ意味になる。したがって、エルゴード性(広義でよい)は全ての状態が間接的には互いに連結されていることを意味する。つまり、エルゴード性が成り立つなら全ての状態は連結されており、上述の等号が成立する状態では、全ての状態は Q_i の値が等しくなければならないことになる。これは、 $P_i \propto W_i$ が成り立つということである。

F_t には下限があって単調減少だから、 F_t は収束する。これは、どのように小さな正の実数 ϵ を考えても、ある t が存在して、 $s > t$ であるような任意の s に対して、等号を誤差 ϵ 以内で成立させられる、ことを意味する。これには、 $t \rightarrow \infty$ で各 P_i が収束して、 $P_i \propto W_i$ でなければならない。

条件としてエルゴード性だけでなく、 $T_{ii} > 0$ が必要な理由は $T_{ii} = 0$ なら必ずしも直接遷移可能な状態が連結されているとはいえず、上の証明の前提が成り立たないからである。たとえば、状態が 1, 2 の2つしかなく $W_1 = W_2$ であり、遷移確率が $T_{12} = T_{21} = 1$, $T_{11} = T_{22} = 0$ であるとする、この遷移確率は詳細釣り合いを満たし、かつ広義のエルゴード性を満たすが、平衡状態は存在しない。

1.11 モンテカルロシミュレーションの収束の速さ

以上で、モンテカルロシミュレーションが「原理的にはうまくいく」ことがわかった。しかし、実際に役に立つためには、十分効率がよいことが必要であり、それは、十分早く平衡分布に収束するということである。(16)から収束が示されたが、その際に収束因子

$$R \equiv |1 - 2T_{\min}|$$

がでてきた。 T_{\min} は遷移行列の成分のうちもっとも小さいものである。相関時間と収束因子は

$$R < e^{-1/\tau} \sim 1 - \tau^{-1}$$

の関係にあるから

$$\tau < \frac{1}{2T_{\min}}$$

であるといえる。よって、すくなくともこの程度の長さのモンテカルロシミュレーションを行えば、目的の計算を行うことができる。しかし、一般に T_{\min} は系のサイ

講義ノート

ズの関数として指数関数的に小さいことはめずらしくない。このため、上の不等式から得られる「十分な長さ」は一般には実行可能な範囲内に収まらないことがある。

上記の不等式はかなり荒いので、実際の緩和時間とはかけ離れている可能性はあるが、現実には遅い緩和が見られることはある。たとえば、イジングモデルのシングルスピントップデートによるシミュレーションにおいてもこの状況が観察できる。すなわち、高温側では、十分早く平衡状態に達するが、臨界点に近づくにつれて次第に大きなクラスタができ、それらの消長はきわめて緩慢なものになってくる。さらに臨界点以下の秩序相にはいると、自発的に対称性が破れ、正または負の自発磁化が生じる。いったん現れた自発磁化の符号が変わることはほとんどない。自発磁化があらわれること自体はシミュレーションがうまくいっていないことにはならないが、自発磁化が生ずる際にはドメイン壁が形成されることがあり、一度ドメイン壁が形成されるとこれが消えることはほとんどない。このような状況では平衡分布のよいサンプリングを行っているとはいえない。

臨界点近傍において緩和が遅くなる現象、つまり臨界緩和についてはいろいろ調べられている。モンテカルロシミュレーションの緩和時間 τ は温度の関数として

$$\tau \propto \xi^z \propto (T - T_c)^{-z\nu}$$

のような依存性をもって臨界点で発散する。ここで、 ξ は相関長である。臨界指数 z は動的臨界指数と呼ばれ、局所的な状態更新によるアルゴリズムの場合は通常 2 に近い値になることが知られている。

この値が 2 に近いものになることは、直観的には以下のように考えれば理解できる。つまり、臨界点近傍では、直径 ξ 程度のドメインが多数出現する。系が十分に緩和するということは、大雑把に言って現在の状態の情報が消えるということであるから、緩和時間はこれらのドメインがほとんど消滅したり完全に形が変わるかするまでの時間になるであろう。それには、ドメイン壁が ξ 程度の長さだけ移動しなければならない。ドメイン壁の移動は拡散過程に似たものであろうと想像できるので、ドメイン壁が ξ 程度の距離を移動するには、 ξ^2 程度の時間がかかる。したがって、緩和時間は ξ^2 程度の大きさになるであろう。

もちろん、これはかなり乱暴な議論であり、ドメイン壁の移動が完全なランダムウォークでないことやドメインがフラクタル構造をもっていること、などを無視している。実際、動的臨界指数は多くの場合正確には 2 でないことが知られている。

2 クラスタアルゴリズム

前節のランダムウォークに基づく動的臨界指数の荒っぽい評価は局所更新アルゴリズムをどうやって加速するかについてヒントを与えてくれる。ランダムウォークだから ξ^2 程度の時間がかかる, というときに前提になっているのは, 1ステップで1のオーダーの距離だけドメイン壁が移動する, ということである。これは, 1回の状態更新で変更をうける部分が空間的には長さ1のスケールをもっているからである。だからこれを加速するには, もっと大きな単位で状態更新を行うことができる。とくに, 大きさ ξ 程度のクラスタを作って, それを単位にして, 状態更新を行うことができるとよさそうである。これを実現するのがクラスタアルゴリズムである。

2.1 イジングモデルの Fortuin-Kasteleyn 表現

1つの相互作用するスピン対に着目すると, イジングモデルのボルツマン重みは $w(S_i, S_j) = e^{-K S_i S_j}$ ($K \equiv \beta J$) だが, これは形式的に

$$w(S_i, S_j) = v_0 + v_1 \delta_{S_i, S_j} = \sum_{g=0,1} v_g \Delta((S_i, S_j), g)$$

と書ける。ここで,

$$v_0 = e^{-K}, \quad v_1 = e^K - e^{-K}$$

$$\Delta((S_i, S_j), g) = \begin{cases} 1 & (g=0) \\ \delta_{S_i, S_j} & (g=1) \end{cases}$$

である。これを利用して Fortuin と Kasteleyn はイジングモデルの分配関数を次のように変形した。²

$$\begin{aligned} Z &= \sum_{\Sigma} W(\Sigma) \\ &= \sum_{\Sigma} \prod_b w(\Sigma_b) \quad \Leftarrow b \equiv (i, j), \Sigma_b \equiv (S_i, S_j) \\ &= \sum_{\Sigma} \prod_b \sum_{\Gamma_b=0,1} v(\Gamma_b) \Delta(\Sigma_b, \Gamma_b) \\ &= \sum_{\Sigma} \sum_{\Gamma} \prod_b v(\Gamma_b) \Delta(\Sigma_b, \Gamma_b) \quad \Leftarrow \Gamma \equiv \{\Gamma_b\} \\ &= \sum_{\Sigma} \sum_{\Gamma} V(\Gamma) \Delta(\Sigma, \Gamma) \end{aligned} \tag{21}$$

$$\begin{aligned} &\Leftarrow V(\Gamma) \equiv \prod_b v(\Gamma_b), \Delta(\Sigma, \Gamma) \equiv \prod_b \Delta(\Sigma_b, \Gamma_b) \\ &= \sum_{\Sigma} \sum_{\Gamma} W(\Sigma, \Gamma) \quad \Leftarrow W(\Sigma, \Gamma) \equiv V(\Gamma) \Delta(\Sigma, \Gamma) \end{aligned} \tag{22}$$

²P. W. Kasteleyn and C. M. Fortuin, J. Phys. Soc. Jpn. Suppl. 26s (1969) 11; C. M. Fortuin and P. W. Kasteleyn, Physica (Utrecht) 57 (1972) 536.

講義ノート

$\Gamma_b = 1$ であるようなボンド b には線 (edge) を描き, そうでなければなにも描かないことにすると, Γ は格子に描かれたグラフに対応する. グラフ中でつながっているスピンの集団をクラスタと呼ぶことにすると, $\Delta(\Sigma, \Gamma)$ は, 各クラスタごとにそれに属するスピンが同一の方向を向いているときに 1 で, それ以外の場合に 0 であるような関数である. したがって, グラフ Γ が与えられたとき, 和に寄与するような (これを「 Γ のもとで許される」と表現することにする) Σ は各クラスタの個数分だけの自由度しかもっていないことになる. クラスタの総数を $N_c(\Gamma)$ として, そのような状態は $2^{N_c(\Gamma)}$ 個ある. これを用いて (21) を以下のように書きなおせる.

$$Z = \sum_{\Gamma} V(\Gamma) 2^{N_c(\Gamma)} \propto \sum_{\Gamma} p^{N_b} (1-p)^{N_p - N_b} q^{N_c(\Gamma)}$$

ここで, $q = 2$, N_p は最近接スピン対の総数, N_b はボンドの総数, $p = w_1 / (w_0 + w_1) = 1 - e^{-2K}$ である. q を 1 にした極限がパーコレーションモデルの分配関数を与え, これによってイジングモデルとパーコレーションの関係が明らかになる. それだけでなくこの書き換えはモンテカルロシミュレーションの新しいアルゴリズムを導く.

2.2 Swendsen-Wang アルゴリズム

Fortuin-Kasteleyn の表示にもとづいて, Swendsen と Wang は以下のアルゴリズムをイジングモデルに対して考案した.³

Swendsen-Wang アルゴリズム :

ステップ 0 : 適当に初期状態のスピン配置 (たとえばランダムな状態) を用意する.

ステップ 1 : 各最近接スピン対ごとに 2 つのスピンをつなぐ線を描くかどうかを決める. すなわち, もし 2 つのスピンが同じ向きを向いているばあい, 0 以上 1 未満の一様乱数を発生してそれが $1 - e^{-2K}$ 以下であれば線を描く. それ以外のときは, 描かない (すでに描いてあれば線を消す).

ステップ 2 : ステップ 1 でできたグラフの各クラスタごとに 0 以上 1 未満の乱数を発生して, それが $1/2$ 以下であればクラスタに属するすべてのスピンを -1 とし, そうでなければ $+1$ にする.

ステップ 3 : 1 に戻る.

なぜこのアルゴリズムで正しいサンプリングができるかを考える. そのためには, (22) から出発する.

$$Z = \sum_{\Sigma} \sum_{\Gamma} W(\Sigma, \Gamma)$$

³R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. 58 (1987) 86.

重みが Σ だけでなく Γ の関数になっているので、マルコフ過程も Σ の空間のものでなく、スピン状態とグラフの組 (Σ, Γ) のなす空間内のマルコフ過程を考えるのが自然である。Swendsen-Wang アルゴリズムはこの拡大された状態空間内のマルコフ過程を定義するものであり、上記の手続きを遷移行列の形であらわすと、ステップ 1 とステップ 2 それぞれに対応する遷移行列を用いて

$$T = T_2 T_1$$

と書ける。 T_1 はステップ 1 の、 T_2 はステップ 2 の遷移行列である。ステップ 1 では、スピン状態は変化しないので、

$$T_1(\Sigma', \Gamma' | \Sigma, \Gamma) = \Delta(\Sigma' | \Sigma) \times \hat{T}_1(\Gamma' | \Sigma)$$

の形をしている。逆にステップ 2 ではグラフは変化しないので、

$$T_2(\Sigma', \Gamma' | \Sigma, \Gamma) = \hat{T}_2(\Sigma' | \Gamma) \times \Delta(\Gamma' | \Gamma)$$

少し考えると明らかになることだが、実は Swendsen-Wang アルゴリズムの $\hat{T}_1(\Gamma | \Sigma)$ と $\hat{T}_2(\Sigma | \Gamma)$ は Fortuin-Kasteleyn 表現 (22) の $W(\Sigma, \Gamma)$ を用いて

$$\hat{T}_1(\Gamma | \Sigma) = \frac{W(\Sigma, \Gamma)}{W(\Sigma)}, \quad \hat{T}_2(\Sigma | \Gamma) = \frac{W(\Sigma, \Gamma)}{W(\Gamma)} \quad (23)$$

と表せる。ここで、

$$W(\Gamma) \equiv \sum_{\Sigma'} W(\Sigma', \Gamma), \quad W(\Sigma) \equiv \sum_{\Gamma'} W(\Sigma, \Gamma'), \quad (24)$$

である。(とくに $W(\Sigma)$ は普通の意味のボルツマン重みに他ならない.)

(23) を認めると、

$$T_1(\Sigma', \Gamma' | \Sigma, \Gamma) W(\Sigma, \Gamma) = \Delta(\Sigma', \Sigma) \frac{W(\Sigma, \Gamma')}{W(\Sigma)} W(\Sigma, \Gamma) = \frac{\Delta(\Sigma', \Sigma) W(\Sigma, \Gamma') W(\Sigma', \Gamma)}{W(\Sigma)}$$

となるが、最後の式は明らかに $(\Sigma, \Gamma) \leftrightarrow (\Sigma', \Gamma')$ の入れ替えに対して不変だから、 T_1 が詳細釣り合い条件を満たすことがわかる。同様に T_2 も詳細釣り合い条件を満たす。エルゴード性については、 $T_1 T_2$ ともそれぞれ単独では明らかに満たさないが $T = T_2 T_1$ はエルゴード性をみたす。(たとえば、きわめて小さいが 0 ではない確率で、ステップ 1 で 1 本もボンドが描かれないことが起こりえる。このとき、ステップ 2 ではいかなるスピン状態も等確率であらわれることになる。) しかし残念ながら T は必ずしも詳細釣り合い条件を満たさないで、すでに行った証明の適用範囲外になってしまう。

クラスタアルゴリズムの収束を証明するには (20) を示した過程を思い出せばよい。この式自体は詳細釣り合いさえ満たせば成立する。つまり、 T_1 を作用させても、 T_2 を作用させても、自由エネルギーは単調減少する。状態空間が有限集合なら自由エネルギーに下限があるから、やがて収束する。十分収束した状態では、 T_1 を作

講義ノート

用させても T_2 を作用させても自由エネルギーが減少しないはずであるが、それは、(20) の等号成立条件から T_1 でつながったパス上の状態 i について $Q_i = P_i/W_i$ が等しいと同時に T_2 でつながったパスについても Q_i が等しいことを意味する。 $T_2 T_1$ がエルゴード的であることは、任意の2状態が T_1 の定めるパスと T_2 の定めるパスのどちらかかまたは両方を使えばつながっていることを意味しているので、結局 Q_i は全ての i について同じ値を持たなければならない。すなわち、Swendsen-Wang アルゴリズムのように、個別にはエルゴード性を満たさない遷移行列でも、組み合わせでエルゴード性をみたせば、全体としては、モンテカルロシミュレーションは正しい分布に収束する。

2.3 クラスタアルゴリズムにおける物理量の計算

Γ は計算の都合上現れた人工的なものであるから、通常われわれが計算したい量は Σ にのみ依存する物理量 $Q(\Sigma)$ である。組み合わせ状態 $\tilde{\Sigma} = (\Sigma, \Gamma)$ が重み $W(\Sigma, \Gamma)$ に比例する状態で出現していれば、第 t モンテカルロステップで実現する状態を $\tilde{\Sigma}_t \equiv (\Sigma_t, \Gamma_t)$ として、

$$\overline{Q}_t \equiv \overline{Q(\tilde{\Sigma}_t)} = \frac{\sum_{\Sigma, \Gamma} W(\Sigma, \Gamma) Q(\Sigma)}{\sum_{\Sigma, \Gamma} W(\Sigma, \Gamma)}$$

となるが、

$$\sum_{\Gamma} W(\Sigma, \Gamma) = W(\Sigma)$$

であることを使って Γ に関する和を先にとってしまうと、

$$\overline{Q}_t \equiv \overline{Q(\tilde{\Sigma}_t)} = \frac{\sum_{\Sigma} W(\Sigma) Q(\Sigma)}{\sum_{\Sigma} W(\Sigma)} = \langle Q \rangle$$

となり、正しい平均値に等しいことが分かる。だから、

$$\langle Q \rangle_{\text{MC}} \equiv \frac{1}{T} \sum_{t=1}^T Q_t$$

は中心極限定理により長いシミュレーションの極限で正しい期待値に一致することが分かる。

以上から、クラスタアルゴリズムにおいても、シミュレーション中に登場するスピン状態 Σ にのみ注目すれば、通常のモンテカルロシミュレーションと同じ様に物理量の計算ができることがわかった。しかし、実際によく用いられる方法として、グラフ Γ を積極的に物理量計算に活用する方法がある。ここでは、2点相関関数と帯磁率について考える。

2点相関関数 $\langle S_i S_j \rangle$ はスピン状態に依存しない量

$$\Delta_{ij}(\Gamma) = \begin{cases} 1 & (i, j \text{ が同じクラスタに属する}) \\ 0 & (i, j \text{ が異なるクラスタに属する}) \end{cases}$$

を用いて次のように表せる.

$$\begin{aligned} \langle S_i S_j \rangle &= \frac{\sum_{\Sigma} W(\Sigma) S_i S_j}{\sum_{\Sigma} W(\Sigma)} = \frac{\sum_{\Sigma, \Gamma} W(\Sigma, \Gamma) S_i S_j}{\sum_{\Sigma} W(\Sigma, \Gamma)} = \frac{\sum_{\Sigma, \Gamma} V(\Gamma) \Delta(\Sigma, \Gamma) S_i S_j}{\sum_{\Sigma, \Gamma} V(\Gamma) \Delta(\Sigma, \Gamma)} \\ &= \frac{\sum_{\Gamma} V(\Gamma) 2^{N_c(\Gamma)} \Delta_{ij}(\Gamma)}{\sum_{\Gamma} V(\Gamma) 2^{N_c(\Gamma)}} = \frac{\sum_{\Gamma} V(\Gamma) \Delta_{ij}(\Gamma)}{\sum_{\Gamma} V(\Gamma)} = \langle \Delta_{ij}(\Gamma) \rangle \end{aligned}$$

つまり, スピン状態のみに依存する量 $S_i S_j$ の期待値はグラフのみに依存する量 $\Delta_{ij}(\Gamma)$ の期待値に等しいので, どちらを計算しても統計誤差を別にすれば同じ結果が得られるはずである. 実際には $\Delta_{ij}(\Gamma)$ を計算するほうが統計誤差を小さく抑えられる. このため, このようなグラフによる観測量は improved estimator と呼ばれる.

関係式 $\langle S_i S_j \rangle = \langle \Delta_{ij}(\Gamma) \rangle$ は単に計算に便利であるというだけでなく, クラスタアルゴリズムがなぜ効率がよいかを示している. なぜならば, この関係式は, クラスタアルゴリズムで生成されるクラスタの大きさが, 相関の及ぶ空間的な範囲の大きさに等しいことを示しているからである.

更に帯磁率を考えると, V_c をクラスタ c の体積として,

$$\begin{aligned} \chi &= \frac{\beta}{V} \left\langle \sum_{ij} S_i S_j \right\rangle = \frac{\beta}{V} \left\langle \sum_{ij} \Delta_{ij}(\Gamma) \right\rangle = \frac{\beta}{V} \left\langle \sum_c \sum_{i,j \in c} 1 \right\rangle \\ &= \frac{\beta}{V} \left\langle \sum_c V_c^2 \right\rangle = \beta \langle \bar{V}_c \rangle \end{aligned}$$

となる. ここで, \bar{V}_c は $\bar{V}_c \equiv \sum_c V_c^2 / \sum_c V_c$ で定義される平均クラスタサイズである.

2.4 クラスタアルゴリズムの収束の速さ

局所更新アルゴリズムの動的臨界指数が 2 に近いことはドメイン壁のランダムウォークに起因することを述べたが, 前節で分かった様に, Swendsen-Wang アルゴリズムでは, クラスタがドメインと同程度の大きさであるために, 動的臨界指数が大幅に小さくなることが期待される.

実際, Swendsen-Wang アルゴリズムについては, 動的臨界指数が $z \sim 0$ であることが知られている.

$$\tau \propto \log(T - T_c)$$

であるとする研究者もいるが, いずれにしても, 臨界点付近では局所更新アルゴリズムに比べてきわめて速く平衡状態に近づく.

2.5 Wolff による拡張

Swendsen-Wang アルゴリズムは連続スピン系に適用できるように Wolff によって拡張された。⁴ 彼のアイデアは2つのそれぞれ独立にも応用のできるアイデアからなっている。

1. スピン空間内の鏡像面などを考え、その鏡像面に対する鏡像変換をするかしないか、という離散自由度に着目して、もとの連続自由度系を離散自由度系とみなすこと。
2. 全系をクラスタに分割するのではなく、ランダムに選んだ1点からスタートして、その点を含むクラスタのみを生成すること。

1 番目のアイデアは鏡像変換で符号が変わるようなカイラリティーや渦度などの秩序変数が重要である場合にシミュレーションの効率化に大きな影響を与える。2 番目のアイデアは量子系のワームアルゴリズムや向きつきループアルゴリズムなどを考える際に重要である。

⁴U. Wolff, "Collective Monte Carlo Updating for Spin Systems", Phys. Rev. Lett., 62 (1989) 361.

3 拡張アンサンブル法

クラスタアルゴリズムで緩和が劇的に速くなることがあることが分かったが、それには、相関長とクラスタサイズが同程度であることが必要である。Swendsen-Wang アルゴリズムの場合にはそれが成り立ったが、一般にもそれが成り立つ保障はない。拡張アンサンブルの方法はクラスタアルゴリズムとはまったく異なるアプローチで遅い緩和の問題を解決する試みであり、多くの場合に成功している。

3.1 一般論

拡張アンサンブル法には多くの種類や流儀があるが、基本的な考え方は、モンテカルロシミュレーションで用いる重み、すなわち、詳細釣り合い条件 (6) に登場する重みを平均値の定義式 (1) に現れる重み $W(\Sigma)$ とは異なるものにとることによって、計算全体の効率化を図るというものである。たとえば、モンテカルロシミュレーションを重み $W_0 (\neq W)$ に関して詳細釣り合いを満たすような遷移行列を用いて行なったとする。このとき、重み W による平均値は $R \equiv W/W_0$ として、

$$\langle Q \rangle = \frac{\sum_{\Sigma} W(\Sigma) Q(\Sigma)}{\sum_{\Sigma} W(\Sigma)} = \frac{\sum_{\Sigma} W_0(\Sigma) R(\Sigma) Q(\Sigma)}{\sum_{\Sigma} W_0(\Sigma) R(\Sigma)} = \frac{\langle R(\Sigma) Q(\Sigma) \rangle_0}{\langle R(\Sigma) \rangle_0} \quad (25)$$

最後の表式の $\langle \dots \rangle_0$ は重み W_0 に関する期待値であるが、これはモンテカルロ法の平均に置き換えられるから、結局

$$\langle Q \rangle = \frac{\langle R(\Sigma) Q(\Sigma) \rangle_{MC}}{\langle R(\Sigma) \rangle_{MC}}$$

のように、 RQ と R という2つの量のモンテカルロ平均を計算することで求められるということが分かる。

よくある応用は、 W がパラメータ (たとえば温度など) を含んでいて、1度のシミュレーションで複数のパラメータの値における重みつき平均値を求めるというものである。パラメータを X 、それに対応する重みを W_X とし、さらにそれに関して平均値を求めたいと考えている X の値を X_1, X_2, \dots, X_r とすると、モンテカルロシミュレーションでは、各モンテカルロステップごとに

$$R_{X_k}(\Sigma_t) \equiv \frac{W_{X_k}(\Sigma_t)}{W_0(\Sigma_t)} \quad (k = 1, 2, \dots, r)$$

と $Q(\Sigma_t)$ を計算することで、 $R_k Q$ と R_k のモンテカルロ平均をすべての k について同時に求めることができる。

拡張アンサンブル法のなかでとく便利でよく用いられる方法は、 W_X と W_0 がともに共通の1つないし少数の量の関数になっている場合である。それらを $A_i(\Sigma)$ $i = 1, 2, \dots, p$ とすると、 R も A だけの関数になるから、各モンテカルロステップで $R_k(\Sigma_t)$ を求めるかわりに $A_t \equiv A(\Sigma_t)$ を求めておき、その値を $Q_t \equiv Q(\Sigma_t)$ の

講義ノート

値とともに、すべての t についてファイルに出力して保存しておく。(通常のシミュレーションにおいて、物理量計測の回数は 10^8 回程度かそれ以下であるから、 10^9 バイト程度のディスクスペースがあればこれは実行可能である。) なぜなら、そうしておくとして以下の式を用いて、任意の X の値についての平均値をあとから求められるからである。

$$\langle Q \rangle_X = \frac{\langle R_X(\mathbf{A}_t) Q_t \rangle_{MC}}{\langle R_X(\mathbf{A}_t) \rangle_{MC}} \quad (26)$$

3.2 再重み付け法

上記の拡張アンサンブル法のもっとも簡単な応用例は、 X が逆温度 β であり、 W_β が逆温度 β におけるカノニカル分布関数、 W_0 がある特定の逆温度 β_0 におけるカノニカル分布関数の場合であり、これは再重み付け法 (reweighting method)⁵ と呼ばれる。 W_β , W_0 はともにエネルギーを通してしか状態に依存しないので、重みの比は

$$R_\beta(E) = e^{-(\beta-\beta_0)E}$$

で簡単に計算できる。この例では、ある特定の逆温度で1度普通のシミュレーションを行い、その際に、モンテカルロステップごとのエネルギーの値 E_t と、計算したい物理量の値 Q_t をファイルに出力しておき、あとから

$$\langle Q \rangle_\beta \approx \frac{\sum_t e^{-(\beta-\beta_0)E_t} Q_t}{\sum_t e^{-(\beta-\beta_0)E_t}}$$

によって任意の逆温度におけるカノニカル平均値を計算できる。

この方法は簡単だが実用的で非常に有力な方法である。第1の利点は、通常のモンテカルロシミュレーションのプログラムがあれば、ほとんど手を加える必要なく計算ができることである。第2の利点は、多くの問題においては臨界現象に興味があり、その場合臨界点近傍の比較的狭い範囲を重点的に調べたいが、再重み付け法はそれに適していることである。

3.3 再重み付け法の問題点

一方、再重み付け法の欠点は、第1に実際に行うシミュレーションが普通のシミュレーションであるために、普通のシミュレーションが緩和が遅いものである場合、その困難がそのまま現れることであり、第2に以下で考察するように再重み付け法で精度よく計算できるのは実際にシミュレーションを行った温度の近傍に限られ、広い範囲の温度依存性を調べるのには向かないことである。

再重み付け法によって十分な精度で物理量を求められるパラメータの範囲を考えてみる。話を具体的にするために、最もよく使われる前節の例を考える。すなわち、

⁵A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. 61 (1988) 2635.

シミュレーションはある温度におけるボルツマン重み $e^{-\beta_0 E(\Sigma)}$ で行い, パラメータ X は逆温度 β , A はエネルギー E である場合である. 再重み付けの式を

$$\langle Q \rangle_\beta = \frac{\langle e^{-(\beta-\beta_0)E(\Sigma)} Q(\Sigma) \rangle_0}{\langle e^{-(\beta-\beta_0)E(\Sigma)} \rangle_0} \quad (27)$$

この式の分母分子が精度よくとまることが実用的な計算が可能である条件である. 分子を考えると

$$\begin{aligned} \langle e^{-(\beta-\beta_0)E(\Sigma)} Q(\Sigma) \rangle_0 &= \frac{1}{Z_0} \sum_E W_0(E) e^{-(\beta-\beta_0)E} \sum_{\substack{\Sigma \\ E(\Sigma)=E}} Q(\Sigma) \\ &= \frac{1}{Z_0} \sum_E W_0(E) e^{-(\beta-\beta_0)E} \Omega(E) Q(E) \\ &= \frac{1}{Z_0} \sum_E e^{-\beta_0 F_0(E)} e^{-(\beta-\beta_0)E} Q(E) \end{aligned} \quad (28)$$

ここで $\Omega(E)$ は巨視的な量の組 E を決めたときにそれに対応する微視的な状態の総数, すなわち状態密度であり,

$$Q(E) \equiv \frac{1}{\Omega(E)} \sum_{\substack{\Sigma \\ E(\Sigma)=E}} Q(\Sigma)$$

は E で指定されるミクロカノニカルアンサンブルにおける物理量 Q のミクロカノニカル平均,

$$F_0(E) \equiv -T_0 \log(W_0(E)\Omega(E)) = E - ST_0$$

は自由エネルギーである. 同様に (27) の分子は

$$\langle e^{-(\beta-\beta_0)E(\Sigma)} \rangle_0 = \frac{1}{Z_0} \sum_E e^{-\beta_0 F_0(E)} e^{-(\beta-\beta_0)E} \quad (29)$$

式 (29) において, 和をとられる関数

$$e^{-\beta_0 F_0(E)} e^{-(\beta-\beta_0)E} = e^{-\beta F(E)}$$

を仔細に眺めると, まず, これが逆温度 β におけるエネルギー分布関数に等しいことが分かる. したがって, この関数は逆温度 β におけるエネルギーの期待値 $E(\beta)$ を中心とした, ガウス分布で近似されるような形をしている. このような関数の和を精度よく計算するためには, シミュレーションにおいて, $E(\beta)$ 付近のエネルギーの状態が十分に多い回数出現しなくてはならない. 一方, このモンテカルロシミュレーションでエネルギーが実際に出現する頻度は逆温度 β_0 でのエネルギー分布関数 $e^{-\beta_0 F_0(E)}$ である. したがってこれは, $E(\beta_0)$ のまわりのガウス分布である. もし, β と β_0 が大きく異なっているとすると, 2つの分布が重なりあわないほど離れていることになり, 和に本来寄与するはずの状態がシミュレーションではほとんど出現しないことになる. この場合計算は破綻する.

講義ノート

実際に β と β_0 はどの程度まで離れていても効率的な計算ができるであろうか？
これには分布の幅を考えればよい。エネルギー分布の幅 ΔE の2乗は

$$C = \beta^2 \langle E^2 \rangle - \langle E \rangle^2 = \beta^2 (\Delta E)^2$$

で熱容量に比例しているので、分布の幅は $\Delta E \sim T\sqrt{C}$ と表せる。一方、2つの分布の中心間の距離は

$$|E(\beta) - E(\beta_0)| \sim \frac{dE}{d\beta} \Delta\beta = T^2 C \Delta\beta$$

よって、分布が重なる条件は

$$T\sqrt{C} > T^2 C \Delta\beta$$

つまり、

$$\Delta\beta < \frac{1}{T\sqrt{C}}$$

となる。ここで、分母にある熱容量は系の大きさを規格化するまえの示量性の量であるから、系の体積に比例しておおきくなる。したがって、大きなシステムほど、この方法でカバーできる温度の範囲が狭いことが分かる。また、比熱は転移点近傍で発散する（有限系ではサイズのべき乗で抑えられるが）ので、転移点近傍では再重み付け法を行うにしても、正しい転移点に近い温度でのシミュレーションを行う必要があることになる。

3.4 マルチカノニカル法

再重み付け法では、 W_0 として、あるパラメータ値における通常のボルツマン重みがとられたために、シミュレーション中に実際に出現するエネルギー値の分布は非常に幅の狭いガウス分布になり、その結果再重み付けによって実用的計算が可能な温度範囲は狭いものになった。

しかし、一般論から明らかのように、あとで再重み付け平均をとるのであれば、シミュレーションで使う重みがもとのボルツマン重みと同じ形をしている必要はない。ハミルトニアンを e の肩に載せた形をしている必要は全くないのである。マルチカノニカル法⁶では、この自由度を積極的に利用して、物理量の組 \mathbf{A} の分布関数が広い範囲でほぼ一定値になるように W_0 を調節する。結果として、通常マルチカノニカル法で用いられる W_0 はどのようなモデルのボルツマン重みにも似ていないようなものになる。物理量の組 \mathbf{A} として最もよく用いられるのは、やはりエネルギーであるが、緩和を早くしたりするために、エネルギーのほかにもいろいろな物理量の組み合わせが用いられることも多い。だから、ここでは一般的記号である \mathbf{A} のまま議論をすることにする。

⁶B. A. Berg and T. Neuhaus, "Multicanonical ensemble: A new approach to simulate first-order phase transitions" Phys. Rev. Lett. (1992) 68; B. A. Berg, "Multicanonical simulations step by step", Comp. Phys. Commun. 153 (2003) 397.

マルチカノニカル法ではモンテカルロシミュレーションをいくつかのフェーズに分けて行う。各フェーズでシミュレーションにつかう重み関数はフェーズ内では一定で、各フェーズのシミュレーションの結果に基づいて、次のフェーズで用いる重み関数を決定する。その際に目的とするのは \mathbf{A} の出現頻度分布ができるだけ均一になることである。ここで「均一」という点はそれほど厳密である必要はない。 \mathbf{A} について極端にサンプリングされにくい場所がなくなればそれでよいのである。

各フェーズではモンテカルロステップごとに得られた \mathbf{A} の値から \mathbf{A} の出現頻度のヒストグラム $H(\mathbf{A})$ を作る。出現頻度が多いところはその部分の重みが大きすぎたということだから、次のセットでは、重み関数として、前回使った重み関数をヒストグラムで割ったものを使う。第 k フェーズで用いる重み関数を $W_0^{(k)}(\mathbf{A})$ 、得られたヒストグラムを $H(\mathbf{A})$ とすると、

$$W_0^{(k+1)}(\mathbf{A}) = \frac{W_0^{(k)}(\mathbf{A})}{H(\mathbf{A}) + 1} \quad (30)$$

とするわけである。(右辺分母の 1 は一度も訪れなかった場所に関してゼロ割がおこることを防ぐため。)

前節までの部分で明らかのように、シミュレーションで用いる重み関数 W_0 がどんなものであれ、その結果得られる巨視的物理量の組 \mathbf{A} の出現頻度分布関数が、平均値を求めるときの重み関数 W に対応する分布関数と十分重なりを持っていさえすれば、再重み付けの考え方をを用いて正しい平均値を求めることができる。したがって、マルチカノニカル法の最終フェーズでもし \mathbf{A} の出現頻度が十分に均一になっていれば、(均一な分布関数はどんな関数とでも大きな重なりをもっているはずだから) 再重み付けをして、任意のパラメータの値で平均値を求めることができるはずだ、というのが基本的な考え方である。

この方法はもうひとつの利点を持っている。それは、通常重みと全く異なる重み関数でシミュレーションを行うために、緩和の仕方も当然通常のシミュレーションとは異なり、通常のシミュレーションよりも早く緩和する場合があることである。もし、臨界点をもつ系に対して $\mathbf{A} = E$ としてこの方法を適用すると、シミュレーションの最中、エネルギーは臨界温度における期待値を含めて広い範囲を上下することになる。これは、温度を自律的に変更しているようなものであり、低温(低エネルギー)で一度秩序化した状態もしばらくまつと、高温(高エネルギー)となり、自然と無秩序状態に戻る。再び低温になると再度秩序化するが、秩序の出方は最初のとときとは無関係である。通常エネルギーや温度一定のシミュレーションでは、すでに秩序化している状態からスタートするといくら待っても別の秩序状態に変化することはない。たとえば、転移温度よりも低温でのイジングモデルのシミュレーションで大半のスピンの向きが上向きに整列した状態が一度出現すると、その後下向きに秩序化した状態は(系が十分に大きければ)出現しない。これに対してマルチカノニカル法のシミュレーションでは両者が比較的頻繁に移り変わることが観測できる。

3.5 Wang-Landau 法

しかし、実際にマルチカノニカル法をイジングモデルなどに適応してみると、いろいろな問題点に気付く。もっとも本質的なのは、分布の幅が広がっていく速度が十

講義ノート

分に速くないことである。たとえば，最初のフェーズで用いる重みとしてもっとも簡単に $W^{(1)} = \text{const}$ を採用したとする。これだと，エネルギーの分布は高温極限における分布，すなわち $E = 0$ のまわりの幅 \sqrt{V} 程度のガウス分布になる。マルチカノニカル法の手順によってこの範囲の状態は次のフェーズでは重みを下げられるので，次のフェーズでは， $E = 0$ 付近のエネルギーをとりにくくなり，エネルギー空間でより広い範囲のランダムウォークが実現する。といっても，前回まったく足を踏み入れなかったエネルギー領域については，情報がないうままであり，当然この領域では重みをうまく調節できていないから，その領域に足を踏み入れたとたん進みが遅くなる。結果として，前のフェーズで実現した分布よりも広いことは広いがそれほど目立った広がりにはならず，最終的に全エネルギー領域をカバーする分布が実現するまでに非常に多くのフェーズを経なければならないということになる。

この欠点を劇的に解消したのが Wang と Landau によるマルチカノニカル法の改良である。⁷ 普通のマルチカノニカル法との違いは，普通のマルチカノニカル法においては各フェーズ内では重みが一定に保たれているのに対して，改良法では，フェーズ内でも刻々と重みが増減することである。WLでは，フェーズの終わりではなく，各モンテカルロスイープが終わるごとに以下の処理を行う。

$$W_0^{(k)}(\mathbf{A}) \leftarrow e^{-\eta^{(k)}} W_0^{(k)}(\mathbf{A}) \quad (31)$$

ここで， $\eta^{(k)}$ はフェーズ内では一定に保たれる正の定数である。各フェーズの終わりに

$$\eta^{(k+1)} = \eta^{(k)} / 2 \quad (32)$$

などによって，次のフェーズで用いられる値を決める。通常は $\eta^{(1)} \sim 0.1$ 程度から初めて， $\eta^{(k)} \sim 10^{-8}$ 程度になるまで，フェーズを繰り返す。 η の値が大きい初期のフェーズでは，同じエネルギーの値を続けてとると，急激にそのエネルギー値に対するペナルティがつく（そのエネルギー値に対する重みが減少する）ので系は「強制的」に他のエネルギーの状態をとることになる。普通のマルチカノニカル法でも同じことは起きるが，フェーズ内ではペナルティが重みに反映されないため，改良法に比べるとはるかにマイルドである。実際に改良法ではとくに η が大きい初期のフェーズでは分布関数の広がり方が普通のマルチカノニカル法に比べてはるかに早いことが観測される。

ただし，モンテカルロ法に使われる重みが同一フェーズ内では一定でなく，前のステップの結果に依存するということは，厳密にはマルコフ過程でない，ということの意味している。それはすなわち， $\eta = 0$ でない限りWL法で得られる分布は厳密にはどのような重みに対する平衡分布にもなっていない，ということである。したがって，一般の物理量の期待値を計算するためには， $\eta = 0$ かまたは十分に小さい値で最終フェーズを行って，そのフェーズで得られた情報のみに基づいて計算が行われる。

⁷F.-G. Wang, D. P. Landau, "Efficient, multiple-range random walk algorithm to calculate the density of states", Phys. Rev. Lett. 86 (2001) 2050.

3.6 マルチカノニカル法における期待値の計算

重みがエネルギーだけの関数である場合、物理量 Q の期待値は状態密度 $\Omega(E)$ と Q のマイクロカノニカル平均 $Q(E)$ を用いて

$$\langle Q \rangle = \frac{\sum_E \Omega(E) W(E) Q(E)}{\sum_E \Omega(E) W(E)} \quad (33)$$

と書ける。すなわち、 $\Omega(E)$ と $Q(E)$ が正確にわかっているならば、 $W(E)$ がどのようなものであろうと、正確な期待値を計算することができる。よくある応用は $Q(E)$ が加法的な量である場合だが、その場合は中心極限定理により、系自体のサイズが大きければ相対的に小さな分散しかもたない。だから、エネルギーが E である状態が実際にはそれほど多数回出現しなくても、かなり正確な期待値を知ることができる。これに対して、 $\Omega(E)$ はどうであろうか。

上の期待値を計算するには、 $\Omega(E)$ の絶対値を正確に知る必要はなく、 E_0 を適当なエネルギーの基準値として、 $\Omega(E)/\Omega(E_0)$ が分かれば十分である。マルチカノニカル法でこれを最も素直に求めるには、 W_0 を用いたシミュレーションで観測されるエネルギー値の出現状態分布つまりヒストグラム $H(E)$ を使う。特定のエネルギー値の出現頻度は、重みと状態数の積 ($H \propto \Omega \times W_0$) であるから

$$\tilde{\Omega}(E) \equiv \frac{\Omega(E)}{\Omega(E_0)} \approx \frac{H(E)W_0(E_0)}{H(E_0)W_0(E)}$$

が成り立つ。長いシミュレーションの極限では上の近似は厳密になる。しかし、実際の有限ステップのシミュレーションでは、個々のエネルギー値が実現される回数は少ないことが多く、そのときに上の近似式を用いると統計誤差が非常に大きいものになる。具体的にはシミュレーションが全部で N ステップであり、相関時間が $\tau \propto V^2$ であるとする、 $H(E)$ に関する独立なサンプル数は

$$\frac{N}{\tau} \propto \frac{N}{V^2}$$

である。このとき、上の近似式で得られる $\tilde{\Omega}(E)$ に関する統計誤差は

$$\Delta \tilde{\Omega} \propto \frac{V}{\sqrt{N}}$$

となる。

これに対して、次に述べる方法を用いると、より精度のよい評価が得られることがある。出発点は詳細釣り合いと同じ形をした式

$$T(\Sigma'|\Sigma)W_0(\Sigma) = T(\Sigma|\Sigma')W_0(\Sigma')$$

である。ここで、 T はこの式を満たす任意の関数でよく、実際にシミュレーションで用いられる遷移確率であってもよいが、そうでなくても以下は成立する。これに

講義ノート

対して、 W_0 は実際にモンテカルロシミュレーションで用いられる重みである。任意にエネルギーの値 E, E' を選び、 $E(\Sigma) = E$, $E(\Sigma') = E'$ となるような全ての Σ, Σ' について上の式の両辺の和をとる。すると

$$T(E'|E)W(E)\Omega(E) = T(E|E')W(E')\Omega(E') \quad (34)$$

が得られる。ここで、 $T(E'|E)$ は、

$$T(E'|E) \equiv \frac{1}{\Omega(E)} \sum_{\substack{\Sigma \\ (E(\Sigma)=E)}} \sum_{\substack{\Sigma' \\ (E(\Sigma')=E')}} T(\Sigma'|\Sigma)$$

で定義され、もし $T(\Sigma'|\Sigma)$ が遷移行列であれば、 $T(E'|E)$ は、「現在の状態のエネルギーが E であるときに次の状態のエネルギーが E' である確率」を意味する。もし、ある定数 J に対して $T(E \pm J|E)$ が常に 0 でないなら、(34) から

$$\tilde{\Omega}(E) = \frac{\Omega(E)}{\Omega(E_0)} = \prod_{k=0}^{K-1} \frac{T(E_{k+1}|E_k) W(E_k)}{T(E_k|E_{k+1}) W(E_{k+1})} = \left(\prod_{k=0}^{K-1} \frac{T(E_{k+1}|E_k)}{T(E_k|E_{k+1})} \right) \times \frac{W(E)}{W(E_0)}$$

(ただし、 $E_k \equiv E_0 + kJ$, $E \equiv E_K$) が得られる。ここで $T(E'|E)$ は「次の状態が E' である確率」のミクロカノニカル平均

$$T(E'|E) = \langle T(E'|\Sigma) \rangle_E$$

とも書くことができるので、モンテカルロシミュレーションで評価することが可能である。すると、結果として、 $\tilde{\Omega}$ の評価が得られる。重要なのは $T(E'|E)$ はマクロな物理量と同様に系のサイズが大きいたときには比較的少ないサンプル数によって精度のよい評価を得ることができる場合があることである。⁸

例として、イジングモデルの拡張アンサンブルシミュレーションにおいて、 $T(\Sigma'|\Sigma)$ がシングルスピントリップのメトロポリスアルゴリズムのような遷移確率である場合を考える。 Σ_i を第 i 番目のスピン S_i のみにおいて Σ と異なる状態、 $t(E', E)$ を $t(E', E)W(E) = t(E, E')W(E')$ を満たす関数であるとして、

$$T(\Sigma'|\Sigma) = t(E(\Sigma'), E(\Sigma)) \sum_{i=1}^V \Delta(\Sigma', \Sigma_i)$$

とかけるから、

$$T(E'|E) = t(E'|E) \times \left\langle \sum_{\substack{\Sigma' \\ (E(\Sigma')=E')}} \sum_i \Delta(\Sigma', \Sigma_i) \right\rangle_E = t(E'|E) \times \langle M(E') \rangle_E$$

とかける。ここで、 $M(E')$ はそれを反転することによってエネルギーが E' になるようなスピンの個数である。すなわち、この場合、 $\tilde{\Omega}$ を評価することは、 $T(E'|E)$ を評価することに帰着し、 $T(E'|E)$ を評価することは、反転後のエネルギーが E' に

⁸J. S. Wang and R. H. Swendsen, J. Stat. Phys. 106 (2002) 245.

なるスピンの個数の期待値を求めることに帰着する。これは、明らかに示量性の量であり、精度よくもとめられる。また、この例の場合には、(34) はよりシンプルに

$$\langle M(E') \rangle_E \Omega(E) = \langle M(E) \rangle_{E'} \Omega(E') \quad (35)$$

とかけて、これはブロードヒストグラム関係式と呼ばれている。この場合の $\tilde{\Omega}(E)$ を求めるための式は

$$\tilde{\Omega}(E) = \prod_{k=0}^{K-1} \frac{\langle M(E_{k+1}) \rangle_{E_k}}{\langle M(E_k) \rangle_{E_{k+1}}}$$

である。

3.7 レプリカ交換法

マルチカノニカル分布では、分布がフラットになるように重み $W(E)$ を調節したが、ある意味では分布が最初からフラットになるように固定された方法がレプリカ交換法である。⁹ この方法は、普通のモンテカルロシミュレーションのプログラムがあれば、非常に簡単な修正を行うだけで使えるようになることと、大規模並列化に向いている、という大きな長所をもっており、拡張アンサンブル法のなかでもとくに重要な方法である。

パラメータ (たとえば逆温度) X を含む重み $W_X(\Sigma)$ を考える。レプリカ交換法では L 個の状態の組

$$\tilde{\Sigma} \equiv (\Sigma_1, \Sigma_2, \dots, \Sigma_L)$$

と、 L 個のパラメータ値の組

$$\tilde{X} \equiv (X_1, X_2, \dots, X_L)$$

を考える。通常これらのパラメータはとなりあうものが近い、つまり、 $X_i \sim X_{i+1}$ であるようにしておく。また、 W_{X_i} のことを簡単に W_i と書くことにする。

⁹K. Hukushima and K. Nemoto, "Exchange Monte Carlo method and application to spin glass simulations", J. Phys. Soc. Jpn. 65, (1996) 1604.

講義ノート

レプリカ交換法：

ステップ0：ランダムに初期スピン状態の組 $\tilde{\Sigma}$ を選ぶ。

ステップ1：各 μ について，第 μ レプリカの状態 Σ_μ を更新する．このときに用いるのは重み W_μ に従う何らかのモンテカルロシミュレーションの1スイープである．

ステップ2：隣接したペア $(\mu, \mu+1)$ について，確率

$$p_{\text{swap}} \equiv \min \left(1, \frac{W_\mu(\Sigma_{\mu+1})W_{\mu+1}(\Sigma_\mu)}{W_\mu(\Sigma_\mu)W_{\mu+1}(\Sigma_{\mu+1})} \right)$$

で， Σ_μ と $\Sigma_{\mu+1}$ を入れ替える．この操作を $\mu = 1, 2, \dots, L-1$ について行う．

ステップ3：ステップ1に戻る．

レプリカ交換法における「状態」とは L 個のレプリカ状態と 1 から L までの L 個の数の順列 P の組み合わせ

$$(\tilde{\Sigma}, P) \equiv (\Sigma_1, \Sigma_2, \dots, \Sigma_L, P)$$

である．レプリカ交換法によるモンテカルロシミュレーションは，この拡張された状態空間内のマルコフ過程である．重みは単純な重みの積

$$\tilde{W}(\tilde{\Sigma}, P) \equiv \prod_{\mu=1}^L W_\mu(\Sigma_{P(\mu)})$$

である．この分布関数で，量 $Q(\Sigma_{P(\mu)})$ の期待値を考えると，

$$\overline{Q(\Sigma_{P(\mu)})} = \frac{\sum_{\tilde{\Sigma}, P} \tilde{W}(\tilde{\Sigma}, P) Q(\Sigma_{P(\mu)})}{\sum_{\tilde{\Sigma}, P} \tilde{W}(\tilde{\Sigma}, P)}$$

分子だけ取り出すと

$$\begin{aligned} \sum_{\tilde{\Sigma}, P} \tilde{W}(\tilde{\Sigma}, P) Q(\Sigma_{P(\mu)}) &= \sum_P \sum_{\tilde{\Sigma}} \left(\prod_{\alpha} W_\alpha(\Sigma_{P(\alpha)}) \right) Q(\Sigma_{P(\mu)}) \\ &= \sum_P \left(\prod_{\alpha} Z_\alpha \right) \sum_{\Sigma_{P(\mu)}} Z_\mu^{-1} W_\mu(\Sigma_{P(\mu)}) Q(\Sigma_{P(\mu)}) \\ &= \sum_P \left(\prod_{\alpha} Z_\alpha \right) \langle Q \rangle_\mu = L! \left(\prod_{\alpha} Z_\alpha \right) \langle Q \rangle_\mu \end{aligned}$$

すなわち, 任意の Q に対して

$$\overline{Q(\Sigma_{P(\mu)})} = \langle Q \rangle_{\mu}$$

が成立する. つまり, 平衡状態で $\Sigma_{P(\mu)}$ の従う分布は t によらず W_{μ} に比例することを意味している.

補足

本講義では講義テキストとして本書き下ろしノートのほかにレビュー論文を用いた. 本ノートは講義1日目と2日目(最初の4コマ(1コマ=90分))の内容に対応し, レビュー論文は3日目(最後の2コマ)で用いられた. このレビュー論文は

Naoki Kawashima and Kenji Harada: J. Phys. Soc. Jpn. **73** (2004) 1379.

として公表されており, 通常の論文と同様に JPSJ ウェブサイトからダウンロード可能である.