

博士学位論文

マルコフ連鎖モンテカルロ法における
詳細つり合い条件の破れの効果とその応用

Theoretical study of violating detailed balance condition
in Markov chain Monte Carlo methods

平成 28 年度

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

酒井 佑士

目次

第 1 章	序論	1
1.1	計算物理学	1
1.2	数値解析	2
1.3	マルコフ連鎖モンテカルロ法	5
1.4	本博士論文の構成	7
第 2 章	マルコフ連鎖モンテカルロ法の理論的性能評価	9
2.1	問題設定	9
2.2	マルコフ連鎖モンテカルロ法の基礎理論	10
2.3	マルコフ連鎖モンテカルロ法の性能と遷移確率行列の固有値	14
第 3 章	詳細つり合い条件を破るマルコフ連鎖モンテカルロ法	20
3.1	Suwa–Todo アルゴリズム	20
3.2	Irreversible Metropolis–Hastings アルゴリズム	23
3.3	詳細つり合い条件の破れと性能評価	29
第 4 章	詳細つり合い条件を破る 1 次元ランダムウォーク	32
4.1	詳細つり合い条件をみたす 1 次元ランダムウォーク	32
4.2	ねじれ詳細つり合い条件をみたす 1 次元ランダムウォーク	33
4.3	本章のまとめ	42
第 5 章	Irreversible simulated tempering アルゴリズム	44
5.1	シミュレーテッドテンパリング	45
5.2	Irreversible simulated tempering アルゴリズム	48
5.3	重み因子の推定手法	51
第 6 章	Irreversible simulated tempering アルゴリズムの性能評価	54
6.1	設定	54

6.2	重み因子の推定	57
6.3	逆温度の緩和ダイナミクス	62
6.4	物理量の緩和と相関	76
6.5	本章のまとめ	89
第 7 章	結論と今後の展望	90
7.1	本博士論文で明らかにしたこと	90
7.2	今後の展望	91
付録 A	行列のスペクトル分解	95
A.1	行列 $J_{\Omega}(a, b)$ のスペクトル分解	95
A.2	行列 $\tilde{J}_{2\Omega}(a, b; c)$ のスペクトル分解	97
付録 B	第 2 固有値の特定	103
B.1	設定 (a) $\delta = \delta'$ の第 2 固有値	103
B.2	設定 (b) $\delta' = 1$ の第 2 固有値	109
付録 C	性能指数の最悪評価	116
C.1	詳細つり合い条件をみたす場合	116
C.2	詳細つり合い条件を破る場合	117
参考文献		121
List of Publications		125
謝辞		126

第 1 章

序論

本博士論文は、数値計算手法のひとつであるマルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo method, MCMC 法) の性能向上を目指したものである。特に「詳細つり合い条件の破れ」がもたらす効果を理解し、その効果を性能改善に活かすというアプローチをとる。本章では現代における数値計算の位置づけやその性能評価方法の例、および MCMC 法の歴史などの背景知識をまとめ、序論とする。

本章の執筆にあたって、文献 [1, 2, 3, 4] を参考にした。

1.1 計算物理学

古代ギリシアから人類は、自然界に起こる物理現象をよりよく理解しようと様々な手法を試みてきた。19 世紀までに確立した研究手法は実験物理学 (experimental physics) と理論物理学 (theoretical physics) に大別される。実験物理学とは実験や観測を通して物理現象をとらえる方法であり、理論物理学は物理現象に対して数理モデルを仮定して理論を構築し、物理現象を説明・予知する方法である。たとえば Tycho Brahe は惑星の運動を精密に観測し、Kepler はそのデータを解析することで「惑星の軌道は太陽を一焦点とする楕円である」ことなど、惑星の運動に法則性を見出した。さらにこの法則は Newton 力学における運動の 3 法則から理論的にも説明されることとなった。一方で Newton 力学は天王星の軌道の不規則性からそれまで知られていなかった惑星の存在を予言し、実際に海王星が発見された。このように実験によって得られた結果を理論が説明し、理論によって予言された未知の現象が観測されるなど、実験物理学と理論物理学というふたつの異なる研究手法は互いに補い合って物理学の発展に貢献してきた。

これらふたつの研究手法に加えて、ここ数十年の間で計算機やアルゴリズムが急速に発達したことを受けて、計算物理学 (computational physics) と呼ばれる新たな研究手法が登場した。理論物理学において物理現象を記述する基礎方程式は解析的に解けないこと

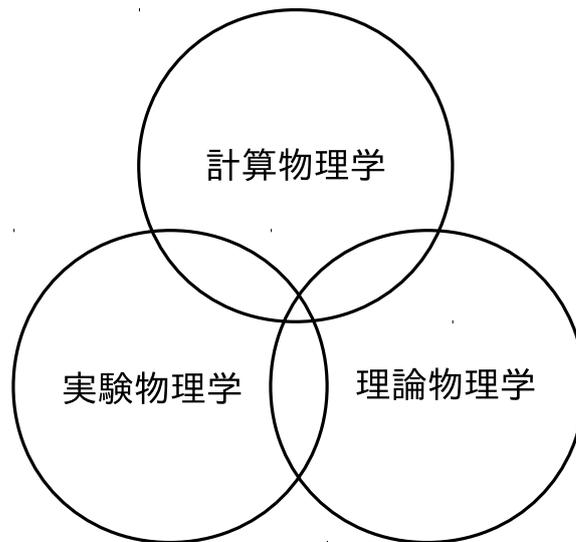


図 1.1 実験物理学・理論物理学・計算物理学はそれぞれが影響を及ぼしあって物理学の発展に貢献している。

がほとんどであり、物理的な仮定や近似を用いて解析が進められてきた。計算物理学では基礎方程式を計算機で「数値的に解く」ことで、基礎方程式が記述する物理現象を理解しようとする。計算機を用いたこの営みは、数値計算またはシミュレーションと呼ばれる。計算物理学は実験物理学や理論物理学と独立しているわけではなく、互いに相補的な関係となっている (図 1.1)。たとえば数値計算では様々な状況を調べることが可能であり、これによって得られた知見から新たな数理モデルが構成されることもある。また大規模な物理実験の前には数値計算を実行し、得られたデータを実験の設計に役立てることがよく行われている。このように現代の物理学は実験物理学・理論物理学・計算物理学がそれぞれの強みを活かし、得られる知見を互いに活用して物理現象の理解を深めている。

1.2 数値解析

1.2.1 数値解析とは

実験物理学では、実験方法の基本原理や設計、測定に伴う不確かさの扱いを十分に理解して実験に臨むことが必要である。計算物理学でも同様に、計算手法の基本原理、得られる数値解やその不確かさの扱いに気を配ることが重要である。特に数値計算で得られる解には、いくつかの要因から必然的に厳密解との誤差が含まれる。したがって、

- なぜその手順にしたがって計算すると「正しい答え」が得られるのか
- 数値的に得られる「正しい答え」はどういう意味で「正しい」のか

といったことを常に意識しておかなければ、自分が何を計算しているのかもわからなくなってしまうたり、得られた結果が説得力を失ってしまう。

計算物理学といっても扱う問題は、非線形方程式を解く・行列を対角化する・微分方程式を解く・積分するなどといった数学的な問題である。これらの数学的な問題を数値計算で扱う手法に関する学問は数値解析 (numerical analysis) と呼ばれる。数値解析では主に、

- ある数学の問題を数値計算で解くアルゴリズムの設計
- 得られる解の誤差評価

に興味がある。以下では数値積分という具体例を通して、数値解析の考え方をみておこう。

1.2.2 数値積分

有界な閉区間 $[a, b]$ 上で定義された実数値 C^0 級関数 f について、定積分

$$I(f) \equiv \int_a^b f(x) dx \quad (1.1)$$

の値を求めることを考えよう。定積分の値が解析的に求められるのはごく一部の関数に対してであり、たとえば $f(x) = \exp(-x^2)$ などの簡単な関数であっても、定積分 $I(f)$ の値を解析的に求めることは不可能であることが知られている。したがって一般には数値計算などから近似的に定積分の値を求めることになる。数値計算によって定積分の値を近似的に求めることを数値積分 (numerical integration) と呼ぶ。

ここでは数値積分の代表例として、複合台形則 (composite trapezoidal rule) を扱う。まず、閉区間 $[a, b]$ 内に $n + 1$ 個の相異なる点

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b \quad (1.2)$$

をとる。一般に各 x_i が等間隔に並んでいる必要はないが、ここでは簡単のため点列 $\{x_i\}_{i=0}^n$ は等間隔に並んでいるとしよう。すなわち点列の間隔を $h \equiv (b - a)/n$ として

$$x_i = a + ih \quad i = 0, 1, \dots, n \quad (1.3)$$

と定める。さらに定積分 $I(f)$ を

$$I(f) = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(x) dx \quad (1.4)$$

と小区間 (x_{i-1}, x_i) での定積分に分割する。ここまでは近似を含んでいない。

次に被積分関数 f を「定積分が容易な関数」に近似することを考える。複合台形則では、小区間 (x_{i-1}, x_i) での関数 f を「2点 $(x_{i-1}, f(x_{i-1}))$, $(x_i, f(x_i))$ を通る1次関数」

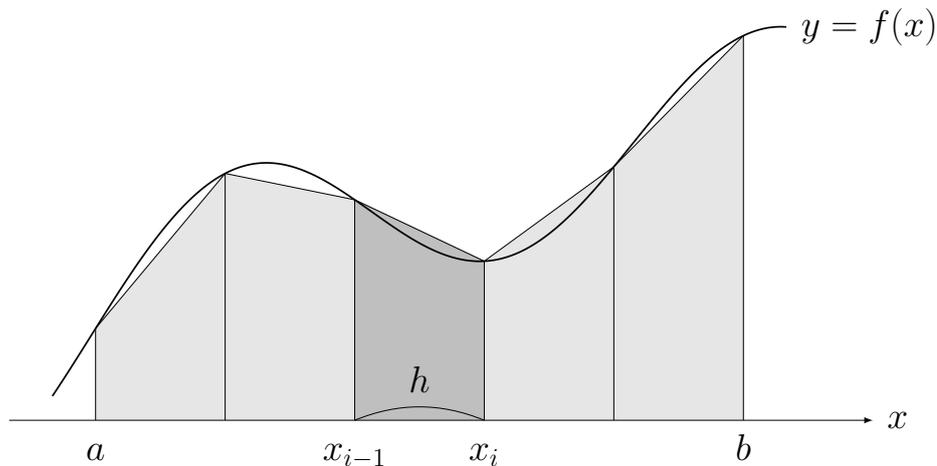


図 1.2 複合台形則による定積分の近似. 区間の分割数は $n = 5$ とした.

に近似する. そして近似後の関数で定積分を実行する. すなわち, 複合台形則では

$$I(f) \simeq \sum_{i=1}^n \frac{f(x_{i-1}) + f(x_i)}{2} h \equiv T_h(f) \quad (1.5)$$

として定積分 $I(f)$ の近似値 $T_h(f)$ を得る (図 1.2).

定積分 $I(f)$ の値を近似して求めたことから, 近似値と厳密な値との誤差が問題となる. ここでは証明を省くが, 以下の関係が成り立つことが示される: 関数 f が閉区間 $[a, b]$ で C^2 級ならば,

$$|T_h(f) - I(f)| \leq \frac{b-a}{12} h^2 \|f''\|_\infty \quad (1.6)$$

が成り立つ. ここで閉区間 $[a, b]$ 上の関数 g に対して $\|g\|_\infty \equiv \max_{x \in [a, b]} |g(x)|$ である. (1.6) 式から近似値 $T_h(f)$ と定積分 $I(f)$ の誤差 $|T_h(f) - I(f)|$ は, 区間内の点間隔 h に対して $O(h^2)$ であることがわかる. 区間の両端の値 a, b を固定したまま区間の分割数 n を増やすと, $b-a = nh$ より点間隔 h は n^{-1} に比例して小さくなる. よって十分多くの点で閉区間 $[a, b]$ を分割することで, 近似値 $T_h(f)$ はいくらでも定積分 $I(f)$ の値に近づき, その誤差は $O(n^{-2})$ であることが示された. このように, 数値解析では

- 計算機で近似解を得るためのアルゴリズムを与えること
- そのアルゴリズムが与える数値解と厳密解の誤差を解析的に評価すること

によって, 数値計算の正当性と精度保証を与える. 計算物理学におけるシミュレーションの結果は, 数値解析によって信頼性のあるものになっているのだ.

数値積分は複合台形則以外にも, 様々な方法が知られている. たとえば複合シンプソン則 (composite Simpson rule) では, 小区間 (x_{i-1}, x_i) における被積分関数 f を 2 次関数

で近似する. すると定積分 $I(f)$ の複合シンプソン則による近似値 $S_h(f)$ について, 以下の誤差評価が成り立つ: 関数 f が閉区間 $[a, b]$ で C^4 級ならば,

$$|S_h(f) - I(f)| \leq \frac{b-a}{2880} h^4 \|f^{(4)}\|_\infty. \quad (1.7)$$

つまり区間の分割数 n を増やすとき, 複合シンプソン則の誤差は複合台形則よりも速やかに小さくなる. このように同じ目的を達成するアルゴリズムであっても, 一般にその性能はアルゴリズムに依存する. 数値計算ではより速く・より正確に計算できるアルゴリズムを性能がよいアルゴリズムと考える. 21 世紀に入って計算機の性能が飛躍的に向上した現在でも, アルゴリズムの性能が悪いと計算資源が無駄になるだけでなく, 現実的な時間でアルゴリズムが終了しないこともある. そのため, 性能がよいアルゴリズムを模索することは現代においても計算物理学の重要な課題となっている.

1.3 マルコフ連鎖モンテカルロ法

前節では 1 次元の定積分における数値積分を扱った. 複合台形則は一般の多重積分についても拡張することができる. d 次元変数 $\mathbf{x} = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ の関数 f と, d 次元空間の部分空間 $D \subset \mathbb{R}^d$ について, 多重積分

$$I(f) = \int_D f(\mathbf{x}) d\mathbf{x} \quad (1.8)$$

の値を計算することを考える. これを複合台形則で扱おうとすると, d 次元積分領域 D を 1 次元あたり $n+1$ 個の点で均一に分割するには $O(n^d)$ の手間がかかる. つまり次元 d が大きくなるほど領域の分割を細かくする手間が増え, 計算精度を上げることが難しくなるのだ. たとえば統計力学で現れる積分などは非常に高次元なことが多く, 複合台形則などでは対応できなくなってしまう. このように, アルゴリズムの有効性は扱う問題によって左右される. ここでは多重積分を題材として, モンテカルロ法 (Monte Carlo method, MC 法) と呼ばれる別の数値積分手法をみる. これをふまえて, 本博士論文のメインとなるアルゴリズムであるマルコフ連鎖モンテカルロ法をとりあげ, その発展の歴史を概観する.

1.3.1 モンテカルロ積分

モンテカルロ法は Ulam らによって 1947 年に提案されたアルゴリズムである [5]. ここではモンテカルロ法の一例として, 多重積分を数値的に実行するモンテカルロ積分を扱う.

モンテカルロ積分では, d 次元積分領域 D から n 個の点 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}$ を一様ラ

ンダムにサンプリングし, 多重積分 (1.8) を

$$I(f) \simeq \frac{1}{n} \sum_{m=1}^n f(\mathbf{x}^{(m)}) \equiv \hat{I}_n(f) \quad (1.9)$$

と近似する. \mathbf{x} が d 次元積分領域 D 上の一様分布にしたがう確率変数であることから, $f(\mathbf{x})$ および $\hat{I}_n(f)$ もまた確率変数である. このとき, 確率変数 $\hat{I}_n(f)$ の期待値は $I(f)$ であり, 分散は $O(n^{-1})$ となることが示せる. すなわち, 近似値 $\hat{I}_n(f)$ と多重積分 $I(f)$ の誤差 $|\hat{I}_n(f) - I(f)|$ は $O(n^{-1/2})$ となる. さらにこの誤差は空間次元 d によらない. このことから空間次元が非常に大きい場合には, 複合台形則よりもモンテカルロ積分が有効であると考えられる.

モンテカルロ法では d 次元積分領域 D からの一様サンプリングが必要となる. しかし高次元空間での一様サンプリングはそれなりに難しいだけでなく, 被積分関数 f において d 次元積分領域 D 内で多重積分 $I(f)$ への寄与が大きい領域は一般にごくわずかである. つまり一様なサンプル $\mathbf{x}^{(m)}$ をたくさん用いてもそのほとんどは無駄となり, 結果的に分散が大きくなってしまう. この困難を解決する方法として, 積分領域から点を一様にサンプリングするかわりにマルコフ連鎖を用いるのがマルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo method, MCMC 法) である. MCMC 法の構成法や理論的性能評価は次章で詳しく議論することにして, 以下では MCMC 法の性能改善の歴史を概観する.

1.3.2 マルコフ連鎖モンテカルロ法の性能改善の歴史

MCMC 法は高次元確率分布からのサンプリングやその分布の下での期待値 (多重積分) を, 乱数を用いて評価する汎用的な数値計算手法である. 1953 年に液体のシミュレーション方法として提唱されて以来 [6], MCMC 法は物理学の枠を超えて, 化学や生物学, 情報科学, 経済学などの様々な分野に幅広く応用され, それぞれの分野の発展に貢献してきた [7, 8, 9]. 特に統計力学の分野では必然的に高次元確率分布を扱うニーズがあり, 数値計算による相転移と臨界現象へのアプローチには MCMC 法が主要な役割を果たしてきた. その中で, MCMC 法の性能改善を試みる研究が多く行われ, 従来よりも効率的な MCMC 法が提案されてきた.

アルゴリズムの性能は速さと正確性で評価できると 1.2.2 節で述べた. MCMC 法の性能におけるこれらの観点は, 定常分布への収束の速さと推定値の分散の大きさを評価することで議論できる. つまり推定値が真の値へ速やかに収束し, かつ分散が十分に小さいならば, その MCMC 法は効率がよいといえる. ひとつ目の観点である収束の速さは, 緩和時間の定数倍程度の改善をもたらす定量的改善と, 動的臨界指数の変化をもたらす定性的改善に大別される. シミュレーテッドテンパリング [10, 11] やマルチカノニカル法 [12], レプリカ交換法 [13] に代表される拡張アンサンブル法は, 定常分布を目的の分布から拡張

することにより、定常分布への収束性を定量的に改善している。一方で Swendsen–Wang アルゴリズム [14] や Wolff アルゴリズム [15] といったクラスターアルゴリズムは、スピン系の問題において大域的な状態更新を行うことによって、定常分布への収束性を定性的に改善している。ただし、スピングラス系で機能するクラスターアルゴリズムは必ずしも成功していない。ふたつ目の観点については、棄却率（現在の状態にとどまる確率）が小さい MCMC 法の遷移確率を構成することで推定値の分散（漸近分散）を小さくできることが Peskun の定理として知られている [16]。この定理の意味で、Metropolis–Hastings アルゴリズム [17] は熱浴法 (Gibbs sampler) [18] よりも性能がよいことが論じられている。

MCMC 法を用いる際には、定常分布の設定と遷移確率の構成が必要である。特に設定した定常分布への収束を保証するために、遷移確率はつり合い条件をみたすように構成しなくてはならない。しかしつり合い条件は比較的ゆるい条件であり、つり合い条件をみたす遷移確率を見出すのは一般には困難である。そのためつり合い条件の十分条件として詳細つり合い条件が広く用いられている。前述した MCMC 法の改善や性能評価はすべて詳細つり合い条件の枠組みの中でなされてきたものである。しかしながら、あくまでも詳細つり合い条件はつり合い条件の十分条件でしかない。MCMC 法の効率を考えると、詳細つり合い条件を課すことが最適である保証はこれまでなされていない。

近年、詳細つり合い条件を破りつつも、つり合い条件はみたす MCMC 法がいくつか提案されており、その有用性が数値的に示されている [19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38]。さらに詳細つり合い条件の破れの有用性は様々な文脈から議論されている [39, 40, 41, 42, 43]。また一般に詳細つり合い条件の破れが MCMC 法の性能に及ぼす影響について、部分的にはあるが理論的に示されつつある [24, 44]。しかしながら、これらの理論的結果は詳細つり合い条件の破れた MCMC 法に対して比較対象となっている MCMC 法が必ずしも実際に比較したいものとなっていないなど、不十分な点が残されている。このように、効率の観点からみた詳細つり合い条件の位置づけや、詳細つり合い条件を破る効果は十分に理解されているとはいえない。これらを明らかにすることは、より効率的な MCMC 法の構築において不可欠である。

1.4 本博士論文の構成

以上の背景をふまえ、本博士論文は詳細つり合い条件の破れがもたらす効果のより深い理解、およびこれまで以上に効率的な MCMC の開発を目的とした。

本博士論文の構成は以下のとおりである。第 2 章と第 3 章は先行研究で得られた知見の整理に充てられている。第 2 章では MCMC 法の基本原理を述べた上で、MCMC 法の理論的な性能評価について詳細つり合い条件の枠組みの中で示されていることを中心にまとめる。第 3 章では、これまでに提案されてきた詳細つり合い条件を破る MCMC 法

を概観し、詳細つり合い条件を破ったときに成り立つこととしてこれまでに知られている結果をまとめる。詳細つり合い条件を破る MCMC 法の理論的な性能評価の一環として、第4章では最もシンプルなモデルである1次元ランダムウォークに詳細つり合い条件の破れを導入し、その効果を議論する。特にねじれ詳細つり合い条件を用いた irreversible Metropolis-Hastings アルゴリズムを適用したときの対応する遷移確率行列を解析的に対角化し、得られた固有値、固有ベクトルから緩和時間や漸近分散を議論する。以上で得られた知見から、詳細つり合い条件の破れの効果を活かした応用を第5章および第6章で議論する。第7章で本博士論文をまとめ、今後の課題と展望を述べる。

第2章

マルコフ連鎖モンテカルロ法の 理論的性能評価

詳細つり合い条件の破れとその効果を論じるにあたって、本章ではマルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo method, MCMC 法) の数理的基礎を概観する。特に MCMC 法が「正しく」動作することを保証する理論的背景と、性能の理論的評価に焦点を当てる。特に、詳細つり合い条件の範囲内で成り立つことと詳細つり合い条件の有無によらず成り立つこととの区別を強調する。

なお本論文で用いるベクトルはすべて行ベクトルとして定義される。また本章を執筆するにあたって、文献 [29, 45, 46, 47, 48, 49, 50] を参考にした。

2.1 問題設定

以下では簡単のため、有限離散自由度系に対する MCMC 法を扱う。系の状態数を $\Omega \in \mathbb{N}$, 状態空間を $I = \{1, 2, \dots, \Omega\}$ と書く。状態空間 I 上に確率分布 $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_\Omega) \in \mathbb{R}^\Omega$ がひとつ与えられているとする。第 i 成分 π_i は状態 $i \in I$ が実現する確率を表す。また確率分布 $\boldsymbol{\pi}$ は規格化条件

$$\sum_{i=1}^{\Omega} \pi_i = 1 \quad (2.1)$$

をみたす。ここでは簡単のため、すべての状態 $i \in I$ に対して $\pi_i > 0$ を仮定する。この確率分布 $\boldsymbol{\pi}$ を目的分布と呼ぶ。

ここでの目的は、目的分布 $\boldsymbol{\pi}$ にしたがう状態列を状態空間 I からサンプリングするこ

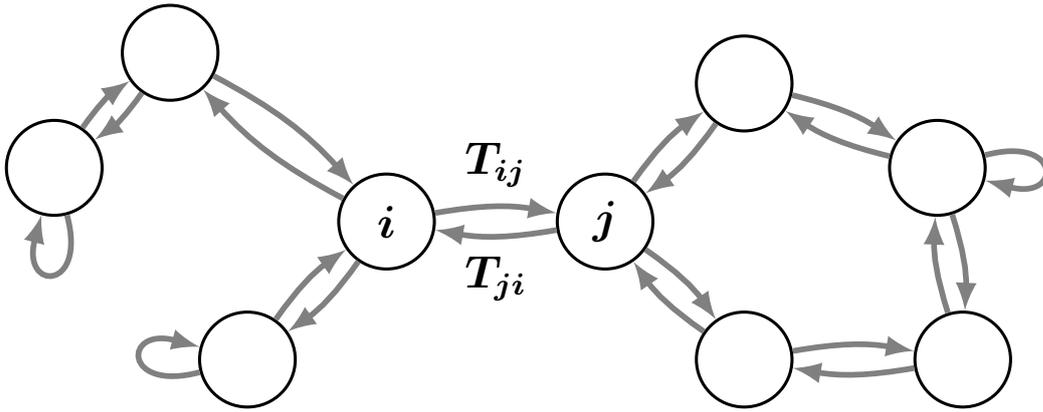


図 2.1 状態遷移図の例.

とや、状態空間 I 上に定義された物理量 f の、目的分布 π に関する期待値

$$\langle f \rangle_{\pi} \equiv \sum_{i=1}^{\Omega} \pi_i f_i = \pi \mathbf{f}^{\top} \quad (2.2)$$

を評価することである。ここで実数 f_i は状態 $i \in I$ における物理量 f の値を表す。また $\mathbf{f} \equiv (f_1, f_2, \dots, f_{\Omega}) \in \mathbb{R}^{\Omega}$ である。たとえばスピン数 N のイジング模型の状態数は $\Omega = 2^N$ となり、スピン数の増加にともなって期待値 $\langle f \rangle_{\pi}$ を直接計算することは計算量的に困難となる。このような困難は次元の呪いと呼ばれている。次元の呪いを回避しつつこれらの目的を数値的に実行する手段として、MCMC 法がしばしば用いられる。

2.2 マルコフ連鎖モンテカルロ法の基礎理論

MCMC 法は、定常分布が目的分布 π である遷移確率行列を用いてマルコフ連鎖を生成する数値計算手法である。本節では MCMC 法における遷移確率行列の性質やその構成法を述べる。

2.2.1 遷移確率行列とマスター方程式

離散時間のマルコフ連鎖は最も基本的な確率過程のひとつであり、現在の状態 $i \in I$ から次の状態 $j \in I$ への遷移確率が系のこれまでたどった履歴に一切依存せず、 i と j のみに依存して決まる。いま状態 $i \in I$ から状態 $j \in I$ へのマルコフ連鎖の単位時間あたりの遷移確率を T_{ij} と書き、 T_{ij} を (i, j) 成分にもつ正方行列 $\mathbf{T} = (T_{ij})_{i, j \in I} \in \mathbb{R}^{\Omega \times \Omega}$ を考える。 \mathbf{T} を遷移確率行列 (transition probability matrix) と呼ぶ。また状態 i をノード、非零の遷移確率 T_{ij} を状態 i から状態 j への有向エッジとして表現した離散グラフを、遷移確率行列が定める状態遷移図と呼ぶ (図 2.1)。特に遷移確率行列 \mathbf{T} は確率行列であり、す

すべての $i, j \in I$ に対して $T_{ij} \geq 0$, かつすべての $i \in I$ に対して

$$\sum_{j=1}^{\Omega} T_{ij} = 1 \quad (2.3)$$

をみます.

マルコフ連鎖における n ステップ後の確率分布を $\mathbf{p}^{(n)} = (p_1^{(n)}, p_2^{(n)}, \dots, p_{\Omega}^{(n)}) \in \mathbb{R}^{\Omega}$ とする. 第 i 成分 $p_i^{(n)}$ は n ステップ後にマルコフ連鎖の状態が $i \in I$ である確率を表す. このとき, 確率分布 $\mathbf{p}^{(n)}$ の時間発展はマスター方程式

$$\mathbf{p}^{(n+1)} = \mathbf{p}^{(n)} \mathbf{T} \quad (2.4)$$

で記述される. マスター方程式 (2.4) は形式的に解けて,

$$\mathbf{p}^{(n)} = \mathbf{p}^{(0)} \mathbf{T}^n \quad (2.5)$$

を得る. すなわち, 任意の時刻 $n \in \mathbb{N}$ における系の確率分布 $\mathbf{p}^{(n)}$ は, 初期分布 $\mathbf{p}^{(0)}$ と遷移確率行列 \mathbf{T} によって完全に特徴づけられる.

マスター方程式の解 (2.5) をもう少し詳細に検討してみよう. 以下では簡単のため, 遷移確率行列 \mathbf{T} は対角化可能であると仮定する. 遷移確率行列 \mathbf{T} の相異なる固有値を $\lambda_1, \lambda_2, \dots, \lambda_K \in \mathbb{C}$ とする. 固有値には縮退があってもよいとし, 相異なる固有値の数 K は一般に Ω 以下である. また一般性を失うことなく $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_K|$ としてよい. 固有値 λ_k の重複度を $m_k \in \mathbb{N}$ とする. さらに固有値 λ_k に属する左固有ベクトル, 右固有ベクトルをそれぞれ $\mathbf{u}_{k,\sigma}, \mathbf{v}_{k,\sigma} \in \mathbb{C}^{\Omega}$ とする. $\sigma = 1, 2, \dots, m_k$ は縮退を区別する添字である. このとき固有ベクトルの定義からそれぞれ

$$\mathbf{u}_{k,\sigma} \mathbf{T} = \lambda_k \mathbf{u}_{k,\sigma}, \quad \mathbf{T} \mathbf{v}_{k,\sigma}^{\top} = \lambda_k \mathbf{v}_{k,\sigma}^{\top} \quad (2.6)$$

が成り立つ. ここでベクトル \mathbf{a} に対して \mathbf{a}^{\top} はその転置を表す. またクロネッカーのデルタを

$$\delta_{\alpha\beta} = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases} \quad (2.7)$$

と表すとき, 対角化可能性の下で一般性を失うことなく,

$$\mathbf{u}_{k,\sigma} \mathbf{v}_{l,\rho}^{\top} = \delta_{kl} \delta_{\sigma\rho} \quad (2.8)$$

が成り立つようにできる^{*1}. このとき遷移確率行列 \mathbf{T} のスペクトル分解は

$$\mathbf{T} = \sum_{k=1}^K \lambda_k \left[\sum_{\sigma=1}^{m_k} \mathbf{v}_{k,\sigma}^{\top} \mathbf{u}_{k,\sigma} \right] \quad (2.9)$$

^{*1} (i) $k \neq l$ のとき, $\mathbf{u}_{k,\sigma} \mathbf{v}_{l,\rho}^{\top} = 0$ であること: 固有ベクトルの定義より, $\lambda_k \mathbf{u}_{k,\sigma} \mathbf{v}_{l,\rho}^{\top} = \mathbf{u}_{k,\sigma} \mathbf{T} \mathbf{v}_{l,\rho}^{\top} = \lambda_l \mathbf{u}_{k,\sigma} \mathbf{v}_{l,\rho}^{\top}$ が成り立つ. よって $(\lambda_k - \lambda_l) \mathbf{u}_{k,\sigma} \mathbf{v}_{l,\rho}^{\top} = 0$. $k \neq l$ より $\lambda_k \neq \lambda_l$ ゆえ, これより

で与えられる。これより遷移確率行列 T の n 乗はただちに

$$T^n = \sum_{k=1}^K (\lambda_k)^n \left[\sum_{\sigma=1}^{m_k} \mathbf{v}_{k,\sigma}^\top \mathbf{u}_{k,\sigma} \right] \quad (2.10)$$

であることがわかる。これを用いると、マルコフ連鎖における n ステップ後の確率分布 $\mathbf{p}^{(n)}$ は (2.5) 式から

$$\mathbf{p}^{(n)} = \sum_{k=1}^K (\lambda_k)^n \left[\sum_{\sigma=1}^{m_k} (\mathbf{p}^{(0)}) \mathbf{v}_{k,\sigma}^\top \mathbf{u}_{k,\sigma} \right] \quad (2.11)$$

と表せる。以上から、マスター方程式の解として与えられる確率分布 $\mathbf{p}^{(n)}$ は初期分布 $\mathbf{p}^{(0)}$ および遷移確率行列 T の固有値 $\{\lambda_k\}$, 固有ベクトル $\{\mathbf{u}_{k,\sigma}\}$, $\{\mathbf{v}_{k,\sigma}\}$ によって記述できることがわかった。

2.2.2 Perron–Frobenius の定理

我々の目標は、マルコフ連鎖を用いて目的分布 $\boldsymbol{\pi}$ にしたがうサンプリングや期待値評価を行うことであった。 $n \rightarrow \infty$ の極限で確率分布 $\mathbf{p}^{(n)}$ が目的分布 $\boldsymbol{\pi}$ に収束するマルコフ連鎖を構成できれば、マルコフ連鎖の十分なステップ後に得られる状態列は近似的に目的分布にしたがうサンプルとみなせるだろう。以下では、初期分布によらず $\lim_{n \rightarrow \infty} \mathbf{p}^{(n)} = \boldsymbol{\pi}$ となるために遷移確率行列 T が満たすべき条件を明らかにする。

まず一般に、遷移確率行列は固有値 1 を少なくともひとつもつ。またベクトル $\mathbf{1} \equiv (1, 1, \dots, 1)$ は遷移確率行列の固有値 1 に属する右固有ベクトルのひとつである。これらは遷移確率行列の性質 (2.3) を書き直した式 $T\mathbf{1}^\top = \mathbf{1}^\top$ から明らかである。また 1 以外の固有値を $\lambda \in \mathbb{C}$ とするとき、不等式 $|\lambda| \leq 1$ が成り立つ*2。よって (2.11) 式の形から、 $\lim_{n \rightarrow \infty} \mathbf{p}^{(n)} = \boldsymbol{\pi}$ となるためには

- 残りの固有値 λ がすべて不等式 $|\lambda| < 1$ をみたし
- 目的分布 $\boldsymbol{\pi}$ が固有値 1 に属する左固有ベクトルである

$\mathbf{u}_{k,\sigma} \mathbf{v}_{l,\rho}^\top = 0$ がしたがう。(ii) 固有値 λ_k に属する固有ベクトル $\{\mathbf{u}_{k,\sigma}\}_{\sigma=1}^{m_k}$, $\{\mathbf{v}_{k,\sigma}\}_{\sigma=1}^{m_k}$ は、適当に線形変換することで $\mathbf{u}_{k,\sigma} \mathbf{v}_{k,\rho}^\top = \delta_{\sigma\rho}$ となるように取り直せること：以下添字 k を省略する。行列 $UV^\top \equiv (\mathbf{u}_\sigma \mathbf{v}_\rho^\top)_{\sigma,\rho=1,2,\dots,m} \in \mathbb{C}^{m \times m}$ に対して $\text{rank}(UV^\top) = m$ であることが必要十分である。行列 UV^\top の第 σ 行目を $\mathbf{a}_\sigma \equiv (\mathbf{u}_\sigma \mathbf{v}_1^\top, \mathbf{u}_\sigma \mathbf{v}_2^\top, \dots, \mathbf{u}_\sigma \mathbf{v}_m^\top) \in \mathbb{C}^m$ と書くとき、ベクトル $\{\mathbf{a}_\sigma\}_{\sigma=1}^m$ が線形独立であればよい。 $c_1, c_2, \dots, c_m \in \mathbb{C}$ として $c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_m \mathbf{a}_m = \mathbf{0}$ とおくと、この第 ρ 列目は $(c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_m \mathbf{u}_m) \mathbf{v}_\rho^\top = 0$ 。ここで固有ベクトル $\{\mathbf{u}_{k,\sigma}\}_{\sigma=1}^{m_k}$, $\{\mathbf{v}_{k,\sigma}\}_{\sigma=1}^{m_k}$ はそれぞれ線形独立であり、同じ固有空間を張ることから、 $c_1 = c_2 = \dots = c_m = 0$ がしたがう。 ■

*2 以下の証明は文献 [48] による。遷移確率行列 $T \in \mathbb{R}^{\Omega \times \Omega}$ の固有値 $\lambda \in \mathbb{C}$ に属する左固有ベクトルのひとつを $\mathbf{u} \in \mathbb{C}^\Omega$ とするとき、 $\mathbf{u}T = \lambda \mathbf{u}$ が成り立つ。 $\mathbf{u} \neq \mathbf{0}$ よりベクトル \mathbf{u} の成分の絶対値が最大値をとる列を第 j 列目とすると、 $|u_j| > 0$ かつ $|\lambda| |u_j| = \left| \sum_{i=1}^{\Omega} u_i T_{ij} \right| \leq |u_j| \left| \sum_{i=1}^{\Omega} T_{ij} \right| = |u_j|$ が成り立つ。よって $|u_j| > 0$ より $|\lambda| \leq 1$ がしたがう。 ■

ことが必要十分である。しかし一般に成り立つ不等式 $|\lambda| \leq 1$ だけでは、 $\lambda \neq 1$ だが $|\lambda| = 1$ である固有値が存在する可能性を排除できない*3。いかなるときに残りの固有値 λ がすべて不等式 $|\lambda| < 1$ をみたすのかについて、遷移確率行列に対する Perron–Frobenius の定理がある。定理の主張は以下のとおりである。

遷移確率行列に対する Perron–Frobenius の定理

エルゴード的な遷移確率行列 T に対して、次が成り立つ。

- (i) T の最大固有値 1 は単根である。
- (ii) T の最大固有値 1 に対する固有ベクトルとして、すべての成分が正であるベクトルをとることができる。
- (iii) T の 1 以外の固有値はすべて絶対値が 1 より小さい。

ある正の整数 n に対して遷移確率行列 T の n 乗 T^n のすべての (i, j) 成分が正となるとき、遷移確率行列 T はエルゴード的 (ergodic) であるという。またこの性質をエルゴード性 (ergodicity) とかエルゴード条件と呼ぶ。エルゴード的な遷移確率行列 T では、有限回の遷移で任意の状態 $i \in I$ から任意の状態 $j \in I$ へと遷移できる。 $\mathbf{v}_{1,1} = \mathbf{1}$ および Perron–Frobenius の定理から、遷移確率行列にエルゴード条件を課すことで、

$$\lim_{n \rightarrow \infty} \mathbf{p}^{(n)} = \mathbf{u}_{1,1} \equiv \boldsymbol{\nu} \quad (2.12)$$

が任意の初期分布 $\mathbf{p}^{(0)}$ に対して成り立つことが示される。さらに Perron–Frobenius の定理の主張 (ii) から、収束先のベクトル $\boldsymbol{\nu}$ を適切に規格化することで確率分布とみなすことができる。規格化後のベクトル $\boldsymbol{\nu}$ を定常分布 (stationary distribution) と呼ぶ。

エルゴード条件に加えて、定常分布 $\boldsymbol{\nu}$ が我々の目的分布 $\boldsymbol{\pi}$ に一致する条件

$$\boldsymbol{\pi} = \boldsymbol{\pi}T \quad (2.13)$$

を課すことで、 $\lim_{n \rightarrow \infty} \mathbf{p}^{(n)} = \boldsymbol{\pi}$ が保証される。(2.13) 式をつり合い条件 (balance condition, BC) と呼ぶ。以上のことから、エルゴード条件およびつり合い条件をみたす遷移確率行列を用いることによって、原理的には我々の目標を達成できることが示された。

2.2.3 詳細つり合い条件

以下では特に、つり合い条件をみたす遷移確率行列を具体的に構成するための処方箋である詳細つり合い条件をみていこう。

*3 $\lambda \neq 1$ だが $|\lambda| = 1$ なる固有値の存在は、マルコフ連鎖の周期性と密接な関係がある。本論文では非周期的なマルコフ連鎖しか扱わないため、これ以上は議論しない。文献 [49] などを参照のこと。

遷移確率行列の性質 (2.3) を考慮してつり合い条件 (2.13) を成分ごとに書き下すと, 任意の状態 $i \in I$ について

$$\sum_{j \neq i} \pi_i T_{ij} = \sum_{j \neq i} \pi_j T_{ji} \quad (2.14)$$

となる. つまりつり合い条件は, 状態 i から流出する確率流 $\pi_i T_{ij}$ と, 他の状態 j から流入する確率流 $\pi_j T_{ji}$ が状態空間全体としてつり合うことを要請するものである. ひとつの目的分布 π に対してつり合い条件をみたす遷移確率行列は数多く存在する. しかしつり合い条件は緩すぎて, (2.13) 式だけからつり合い条件をみたす遷移確率行列を実際に見出すことは一般に難しい. そこでより厳しい条件として, 任意の状態 $i \in I$ と $j \neq i$ の組 (i, j) に対して

$$\pi_i T_{ij} = \pi_j T_{ji} \quad (2.15)$$

を課す. (2.15) 式を詳細つり合い条件 (detailed balance condition, DBC) と呼ぶ. 詳細つり合い条件をみたす MCMC 法を可逆な MCMC 法 (reversible MCMC) といい, 詳細つり合い条件を破る MCMC 法は非可逆な MCMC 法 (irreversible MCMC) という. Metropolis–Hastings アルゴリズム [17] や熱浴法 [18] などのこれまで広く用いられてきた MCMC 法は, すべて詳細つり合い条件をみたすように構成されている. しかしあくまでも詳細つり合い条件はつり合い条件の十分条件であり, 詳細つり合い条件を破りつつもつり合い条件をみたす遷移確率行列を構成することは原理的に可能である. 近年では Suwa–Todo アルゴリズム [20] に代表されるように, 詳細つり合い条件を破ることで従来の MCMC 法よりも性能が向上する例がいくつか報告されている. 非可逆な MCMC 法については, 第3章で詳しく扱う.

2.3 マルコフ連鎖モンテカルロ法の性能と遷移確率行列の固有値

MCMC 法の性能はふたつの観点から論じられることが多い. ひとつは目的分布への収束の速さであり, もうひとつは得られるサンプル間の相関の大きさである. 本節ではこれらの MCMC 法の性能と, 対応する遷移確率行列の固有値の関係を概説する. 加えて, MCMC 法の性能と詳細つり合い条件との関係についてこれまでに知られている結果をまとめる.

以下ではエルゴード条件および目的分布 π に対してつり合い条件をみたす遷移確率行列 T を考える. このとき (2.9) 式および (2.11) 式から, 遷移確率行列 T とマルコフ連鎖

における n ステップ後の確率分布 $\mathbf{p}^{(n)}$ はそれぞれ

$$\mathbf{T} = \mathbf{A} + \sum_{k=2}^K (\lambda_k)^n \left[\sum_{\sigma=1}^{m_k} \mathbf{v}_{k,\sigma}^\top \mathbf{u}_{k,\sigma} \right], \quad (2.16)$$

$$\mathbf{p}^{(n)} = \boldsymbol{\pi} + \sum_{k=2}^K (\lambda_k)^n \left[\sum_{\sigma=1}^{m_k} (\mathbf{p}^{(0)} \mathbf{v}_{k,\sigma}^\top) \mathbf{u}_{k,\sigma} \right] \quad (2.17)$$

で与えられる. ここで行列 $\mathbf{A} \equiv \mathbf{1}^\top \boldsymbol{\pi} \in \mathbb{R}^{\Omega \times \Omega}$ は極限行列 (limiting matrix) と呼ばれる.

2.3.1 収束の速さと第 2 固有値

マルコフ連鎖を用いて目的分布からのサンプリングを実行する際には, n ステップ後の確率分布 $\mathbf{p}^{(n)}$ が速やかに目的分布 $\boldsymbol{\pi}$ へと収束することが望ましい. Perron–Frobenius の定理から, $\lambda_1 = 1$ 以外の固有値 λ_k ($k = 2, 3, \dots, K$) はすべて不等式 $|\lambda_k| < 1$ をみたす. これと (2.17) 式より, 確率分布 $\mathbf{p}^{(n)}$ は $n \rightarrow \infty$ の極限で初期分布 $\mathbf{p}^{(0)}$ によらず目的分布 $\boldsymbol{\pi}$ へと収束することが保証されている. さらに (2.17) 式より, 目的分布への収束の速さは絶対値の意味での第 2 固有値 (second-largest eigenvalue modulus, SLEM)

$$\eta \equiv \max_{2 \leq k \leq K} |\lambda_k| \quad (2.18)$$

で特徴づけられることがわかる. つまり,

$$\tau_{\text{relax}} = -\frac{1}{\log \eta} \quad (2.19)$$

で定義される緩和時間 τ_{relax} が収束の速さを特徴づける時定数となる. マルコフ連鎖を始めてからおおむね τ_{relax} ステップ程度たつて得られる状態列は, 目的分布にしたがうサンプルとみなしてよいだろうというわけである.

また, n ステップ後の物理量 f の期待値 $\langle f \rangle_n$ は

$$\langle f \rangle_n \equiv \sum_{i=1}^{\Omega} p_i^{(n)} f_i = \mathbf{p}^{(n)} \mathbf{f}^\top \quad (2.20)$$

と定義される. (2.17) 式から, これは

$$\langle f \rangle_n = \langle f \rangle_{\boldsymbol{\pi}} + \sum_{k=2}^K (\lambda_k)^n \left[\sum_{\sigma=1}^{m_k} (\mathbf{p}^{(0)} \mathbf{v}_{k,\sigma}^\top) (\mathbf{u}_{k,\sigma} \mathbf{f}^\top) \right] \quad (2.21)$$

と書き換えられる. これより, 期待値 $\langle f \rangle_n$ の収束の速さもまた SLEM η によって決まることがわかる. さらに, ある整数 $k \in \{2, 3, \dots, K\}$ についてすべての $\sigma = 1, 2, \dots, m_k$

で $\mathbf{u}_{k,\sigma} \mathbf{f}^\top = 0$ が成り立つならば, 固有値 λ_k は期待値 $\langle f \rangle_n$ の緩和ダイナミクスに影響を及ぼさないこともわかる. 特にベクトル \mathbf{f} が SLEM η に対応する左固有ベクトルのすべてと直交するならば, SLEM η は期待値 $\langle f \rangle_n$ の緩和に影響を及ぼさない. 同様に, 初期分布 $\mathbf{p}^{(0)}$ が SLEM η に対応する右固有ベクトルのすべてと直交するならば, SLEM η はどの物理量の期待値の緩和にも影響を及ぼさない. このことから, ある特定の物理量の期待値や特別な初期分布における物理量の期待値を測定し, その緩和だけから分布全体の緩和を議論する際には注意を要する.

2.3.2 相関と漸近分散

分布の収束の速さと並んで MCMC 法の性能を特徴づけるのが, 得られるサンプル間の相関の大きさである. 目的分布から独立に状態をサンプリングすることが理想である. しかしマルコフ連鎖を用いることで, 得られる状態の間に相関が生じることは避けられない. サンプル間の相関が小さい MCMC 法は性能がよいといえる.

MCMC 法では, 得られるマルコフ連鎖 $(X^{(n)})_{n=0,1,2,\dots}$ を用いて物理量 f の期待値 $\langle f \rangle_\pi$ をサンプル平均

$$\hat{I}_n(f) \equiv \frac{1}{n} \sum_{m=1}^n f(X^{(m)}) \quad (2.22)$$

で評価する. ここで n は評価に用いたサンプル数であり, $f(X^{(m)})$ は状態 $X^{(m)}$ における物理量 f の値を表す. 特に, いくつかの条件の下で中心極限定理

$$\sqrt{n} \left(\hat{I}_n(f) - \langle f \rangle_\pi \right) \Rightarrow \mathcal{N}(0, v(f, \pi, T)) \quad (2.23)$$

が成り立つ [51]. ここで $\mathcal{N}(\mu, \nu)$ は平均 μ , 分散 ν の正規分布を表し, 矢印 \Rightarrow は分布収束を表す. また $v(f, \pi, T)$ は漸近分散 (asymptotic variance) と呼ばれる量であり,

$$v(f, \pi, T) \equiv \lim_{n \rightarrow \infty} n \operatorname{var} \left[\hat{I}_n(f) \right] \quad (2.24)$$

で定義される. (2.24) 式から, サンプル平均 $\hat{I}_n(f)$ の分散は十分大きな n に対しておおむね $v(f, \pi, T)/n$ で与えられることがわかる. すなわち, 漸近分散が小さいほど期待値の推定精度が高くなる.

漸近分散 $v(f, \pi, T)$ はマルコフ連鎖におけるサンプル間の相関に依存している. 目的分布 π から独立にサンプリングされた状態列であれば, 漸近分散 $v(f, \pi, T)$ は目的分布 π による分散 $\operatorname{var}_\pi[f] \equiv \langle f^2 \rangle_\pi - \langle f \rangle_\pi^2$ に一致する. しかしマルコフ連鎖では状態間の相関によって, 一般に漸近分散 $v(f, \pi, T)$ は独立サンプリングでの分散 $\operatorname{var}_\pi[f]$ より大きくなる. つまり, MCMC 法における相関は漸近分散によって特徴づけることができる.

漸近分散は数値計算では評価しにくいいため、積分自己相関時間 $\tau_{\text{int},f}$ によってマルコフ連鎖における相関を評価することが多い。積分自己相関時間は

$$\tau_{\text{int},f} = \sum_{n=1}^{\infty} C_f^{(n)} \quad (2.25)$$

で定義される。ここで

$$C_f^{(n)} \equiv \frac{\langle f_{\pi} f \rangle_n - \langle f \rangle_{\pi}^2}{\text{var}_{\pi}[f]} \quad (2.26)$$

は自己相関関数である。また、対角行列 $\mathbf{B} \equiv \text{diag}(\pi_i)_{i \in I} \in \mathbb{R}^{\Omega \times \Omega}$ を用いて $\langle f_{\pi} f \rangle_n \equiv \mathbf{f} \mathbf{B} \mathbf{T}^n \mathbf{f}^{\top}$ で与えられる。漸近分散 $v(f, \pi, \mathbf{T})$ と積分自己相関時間 $\tau_{\text{int},f}$ は

$$v(f, \pi, \mathbf{T}) = (1 + 2\tau_{\text{int},f}) \text{var}_{\pi}[f] \quad (2.27)$$

なる関係式で結ばれている。これより相関が強いとき推定値の分散が大きくなり、推定精度が悪くなるのが積分自己相関時間によって定量化できることがわかる。

漸近分散と遷移確率行列の関係をみてみよう。まず基本行列 (fundamental matrix) \mathbf{Z} を導入する:

$$\mathbf{Z} \equiv (\mathbf{I} - \mathbf{T} + \mathbf{A})^{-1} = \mathbf{I} + \sum_{n=1}^{\infty} (\mathbf{T}^n - \mathbf{A}). \quad (2.28)$$

基本行列 \mathbf{Z} を用いると、漸近分散 $v(f, \pi, \mathbf{T})$ は以下のとおり書き換えられる [45]:

$$\begin{aligned} v(f, \pi, \mathbf{T}) &= \mathbf{f} [\mathbf{B} \mathbf{Z} + (\mathbf{B} \mathbf{Z})^{\top} - \mathbf{B} - \mathbf{B} \mathbf{A}] \mathbf{f}^{\top} \\ &= \sum_{k=2}^K \frac{1 + \lambda_k}{1 - \lambda_k} \left[\sum_{\sigma=1}^{m_k} (\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^{\top}) (\mathbf{u}_{k,\sigma} \mathbf{f}^{\top}) \right]. \end{aligned} \quad (2.29)$$

これより、漸近分散 $v(f, \pi, \mathbf{T})$ は遷移確率行列 \mathbf{T} のすべての固有値に依存することがわかる。このことは、緩和時間 τ_{relax} が SLEM のみで特徴づけられたことと対照的である。

2.3.3 詳細つり合い条件と漸近分散

これまでは詳細つり合い条件の有無によらない一般的な議論であった。ここでは詳細つり合い条件を課すことで、漸近分散についてふたつの強い主張が成り立つことをみてみよう。ひとつはある最悪評価についての不等式、もうひとつは2通りの遷移確率行列が与えられたとき、それぞれの漸近分散の間に成り立つ不等式である。

目的分布 π に対して詳細つり合い条件をみたす遷移確率行列 \mathbf{T} を考える。このとき、対角行列 $\mathbf{B} = \text{diag}(\pi_i)_{i \in I}$ に対して $\mathbf{B}^{\pm 1/2} \equiv \text{diag}(\pi_i^{\pm 1/2})_{i \in I}$ と定め、

$$\mathbf{S} \equiv \mathbf{B}^{1/2} \mathbf{T} \mathbf{B}^{-1/2} \quad (2.30)$$

なる相似変換を考えると、詳細つり合い条件から行列 S は実対称行列となることがわかる。実対称行列は直交行列によって対角化可能であり、その固有値はすべて実数である。また相似変換で固有値は重複度も含めて保存する。これらのことから、詳細つり合い条件をみたま遷移確率行列 T は対角化可能であり、その固有値はすべて実数である。実対称行列 S の固有値を $\lambda_k \in \mathbb{R}$ ($k = 1, 2, \dots, K$) とするとき、一般性を失うことなく $1 = \lambda_1 > \lambda_2 > \dots > \lambda_K \geq -1$ とできる。また固有値 λ_k に属する正規直交化された固有ベクトルを $\mathbf{x}_{k,\sigma} \in \mathbb{R}^\Omega$ とするとき*4、遷移確率行列 T の対応する左固有ベクトル $\mathbf{u}_{k,\sigma}$ 、右固有ベクトル $\mathbf{v}_{k,\sigma}$ はそれぞれ $\mathbf{u}_{k,\sigma} = \mathbf{x}_{k,\sigma} \mathbf{B}^{1/2}$ 、 $\mathbf{v}_{k,\sigma} = \mathbf{x}_{k,\sigma} \mathbf{B}^{-1/2}$ で与えられる。するとこのとき、漸近分散の表式 (2.29) に現れる項 $(\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^\top)(\mathbf{u}_{k,\sigma} \mathbf{f}^\top)$ について

$$(\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^\top)(\mathbf{u}_{k,\sigma} \mathbf{f}^\top) = (\mathbf{f} \mathbf{B}^{1/2} \mathbf{x}_{k,\sigma}^\top)^2 \geq 0 \quad (2.31)$$

が成り立つ。これを用いると、漸近分散と分散の比 (performance index [45], 性能指数) について、その上限は絶対値をとらない意味での第2固有値のみによって与えられることが示される*5。すなわち、性能指数 $\Gamma(f, \pi, T)$ を

$$\Gamma(f, \pi, T) \equiv \frac{v(f, \pi, T)}{\text{var}_\pi[f]} \quad (2.32)$$

と定義するとき、

$$\Gamma(f, \pi, T) \leq \frac{1 + \lambda_2}{1 - \lambda_2} \quad (2.33)$$

なる不等式が成り立つ。ここで等号は、ベクトル \mathbf{f} が第2固有値 λ_2 に対応する右固有ベクトル $\{\mathbf{v}_{2,\sigma}\}_{\sigma=1}^{m_2}$ の線形結合で与えられるときに成立する。つまり詳細つり合い条件をみたま場合は、第2固有値が小さい遷移確率行列を構成することで性能指数の最悪評価を改善できる。

もうひとつの主張は Peskun の定理と呼ばれるものである [16]。定理の主張は以下のとおりである。

Peskun の定理 [16]

同じ目的分布 π に対して詳細つり合い条件をみたまふたつの遷移確率行列 $T = (T_{ij})_{i,j \in I}$ 、 $T' = (T'_{ij})_{i,j \in I}$ について、任意の状態 $i \in I$ と $j \neq i$ の組 (i, j) に対して不等式 $T_{ij} \geq T'_{ij}$ が成り立つならば、任意の物理量 f に対して不等式

$$v(f, \pi, T) \leq v(f, \pi, T') \quad (2.34)$$

が成り立つ。

*4 実対称行列では左固有ベクトルと右固有ベクトルは同じになる。

*5 証明は付録 C に譲る。

簡単にいえば、遷移確率行列の非対角成分が大きいほど漸近分散が小さくなるということである。Peskun の定理を指導原理として、Metropolis–Hastings アルゴリズムや熱浴法よりも漸近分散の意味で効率を向上させようという先行研究もある [52, 53, 54, 55]。しかし一方で、詳細つり合い条件が破れた遷移確率行列では必ずしも Peskun の定理と同様の主張が成り立つとは限らない。詳細つり合い条件の破れと MCMC 法の性能の関係については次章の 3.3 節で議論する。

第3章

詳細つり合い条件を破る マルコフ連鎖モンテカルロ法

本章では詳細つり合い条件を破る MCMC 法の具体的な構成例を概観する。詳細つり合い条件の破り方にはいくつかのパターンがある。ここでは「元の状態空間をそのまま用いているか/状態空間を拡張しているか」という基準で先行研究を分類する。それぞれのパターンの代表として、Suwa-Todo アルゴリズムと irreversible Metropolis-Hastings アルゴリズムを解説する。最後に詳細つり合い条件の破れと MCMC 法の性能の関係について、これまでに示されていることとまだわかっていないことを列挙し、次章への橋渡しとする。

3.1 Suwa-Todo アルゴリズム

一般の系について、詳細つり合い条件を破りつつもつり合い条件はみたま MCMC 法が最初に提示された例として Suwa-Todo アルゴリズムが知られている [20, 21, 22]。Suwa-Todo アルゴリズムの提案以前には 1 次元ランダムウォークなど特定のモデルに対するアルゴリズムは存在していたものの、その一般化にはいくつか困難な点が残っていた。元の状態空間をそのまま用いる例として、Suwa-Todo アルゴリズムの考え方をみてみよう。なお Suwa-Todo アルゴリズムは離散系のみならず連続系でも適用可能な形で提案されているが、ここでは簡単のため離散系のみを扱う。また状態空間を拡げずに詳細つり合い条件の破れを導入する手法として、Suwa-Todo アルゴリズム以外ではたとえば文献 [23, 24, 25] などがある。特に文献 [24] では一般に元の状態遷移図にループがある場合、遷移確率行列に詳細つり合い条件の破れを導入できることが議論されている。また文献 [25] ではこの議論をふまえて、vorticity matrix と呼ばれる反対称行列を用いることで詳細つり合い条件をみたま遷移確率行列に詳細つり合い条件の破れを

導入した. さらに従来の Metropolis-Hastings アルゴリズムを拡張した Non-reversible Metropolis-Hastings アルゴリズムを構成した.

Suwa-Todo アルゴリズムの構成における基本的な考え方は, 「重みの幾何学的埋め立てによる棄却率最小化」である. 棄却率とは, マルコフ連鎖において現在の状態から他の状態に遷移せずにとどまる遷移確率であり, 遷移確率行列の対角成分に相当する. Suwa-Todo アルゴリズムは, 棄却率が最小となるように遷移確率を構成することで MCMC 法の効率が向上するという思想に基づいている. 以下では「重みの幾何学的埋め立て」について詳しくみることにする.

現在の状態を i とし, 次の状態として n 通りの候補 $j = 1, 2, \dots, n$ がある状況を考える. この候補には現在の状態 i も含まれているとする. いま状態 i が出現する重みを w_i とし^{*1}, 状態 i から状態 j への遷移確率を $p_{i \rightarrow j}$ とする. さらに確率流 $v_{ij} \equiv w_i p_{i \rightarrow j}$ なる量を導入する. このとき, 確率保存則およびつり合い条件はそれぞれ

$$w_i = \sum_{j=1}^n v_{ij} \quad \forall i, \quad (3.1)$$

$$w_j = \sum_{i=1}^n v_{ij} \quad \forall j \quad (3.2)$$

と表される. これより, 確率流は「重みの幾何学的埋め立て」と解釈できることがわかる. つまり重み w_i に比例する高さの箱を w_1, w_2, \dots, w_n と並べたとき, その面積は確率保存則から確率流 v_{ij} の添字 j による和に等しい. またつり合い条件は, すべての状態 i から確率流 v_{ij} を状態 j に埋め立てるとちょうど w_j になることを表している. 言い換えれば, つり合い条件をみたす遷移確率行列を構成することは重みを過不足なく埋め立てることができる確率流の組 $\{v_{ij}\}$ を見出すことにほかならない. Suwa-Todo アルゴリズムは, この「重みの幾何学的埋め立て」において棄却率に対応する確率流 v_{ii} が最小となるように構成されている. Suwa-Todo アルゴリズムの具体的な構成法はアルゴリズム 3.1 のとおりである.

Suwa-Todo アルゴリズムにしたがって重みを埋め立てたとき, 確率流 v_{ij} は「重み w_i を j 番目の箱に埋め立てた量」として定義される. また詳細つり合い条件は確率流の対称性 $v_{ij} = v_{ji}$ と同値である. 例として図 3.1 には, $n = 4$ での Metropolis-Hastings アルゴリズム, 熱浴法, Suwa-Todo アルゴリズムにおける重みの幾何学的埋め立てが描かれている. 図 3.1 において, Suwa-Todo アルゴリズムでは状態 1 から状態 2 への確率流 v_{12} が存在する一方で, その逆向きの確率流 v_{21} はゼロである. すなわち $v_{12} \neq v_{21}$ であり詳細つり合い条件が破れている. また図 3.1 において赤太線で囲まれた部分の総和は棄却率の期待値に対応する量 $\sum_{i=1}^n v_{ii}$ を表している. Metropolis-Hastings アルゴリズム

*1 目的分布は重み w_i を規格化して得られる.

アルゴリズム 3.1: Suwa-Todo アルゴリズム

(1) [最大重みの特定]

最大の重みをひとつ選ぶ. 複数の重みが等しい値で最大となっているとき, その中から任意にひとつ選ぶ. 一般性を失うことなく最大の重みを w_1 とする. 残りの重みの順序は問わない.

(2) [最大重み w_1 の埋め立て]

最大の重み w_1 を次の箱 ($i = 2$) に埋め立てる. 次の箱が一杯になり重みが余った場合は, さらにその次の箱 ($i = 3$) に埋め立てる. 重み w_1 がすべて他の箱に埋め立てられるまで埋め立てを行う.

(3) [次の重み w_2 の埋め立て]

最初に埋め立てられた箱の重み w_2 を, 手順 (2) で最後に埋め立てられてまだ空きがある箱に埋め立てる. 手順 (2) と同様に重み w_2 がすべて他の箱に埋め立てられるまで埋め立てを行う.

(4) [残った重みの埋め立て]

手順 (3) を残りの重み w_3, w_4, \dots, w_n についても同様に順番に行う. $i \geq 2$ の箱がすべて埋まった時点で余っている重みはすべて $i = 1$ の箱に埋め立てる.

や熱浴法では棄却率がゼロとならないのに対して, Suwa-Todo アルゴリズムでは棄却なし (rejection-free) が実現している. 一般に Suwa-Todo アルゴリズムを用いると, 棄却率に対応する確率流 v_{ii} は

$$v_{ii} = \begin{cases} \max\left(0, w_1 - \sum_{i=2}^n w_i\right) & i = 1 \text{ のとき,} \\ 0 & \text{それ以外} \end{cases} \quad (3.3)$$

となることが示される. これより

$$w_1 \leq \frac{1}{2} \sum_{i=1}^n w_k \quad (3.4)$$

のとき rejection-free が実現することがわかる.

文献 [20] では Suwa-Todo アルゴリズムを 2 次元正方格子上的 q 状態ポッツ模型に適用し, 転移温度近傍での自己相関時間を評価している. その結果, 2 次転移を示す $q = 4$, および 1 次転移を示す $q = 8$ のいずれにおいても Metropolis-Hastings アルゴリズムや熱浴法などと比較して自己相関時間に数倍程度の改善がみられることが数値的に示され

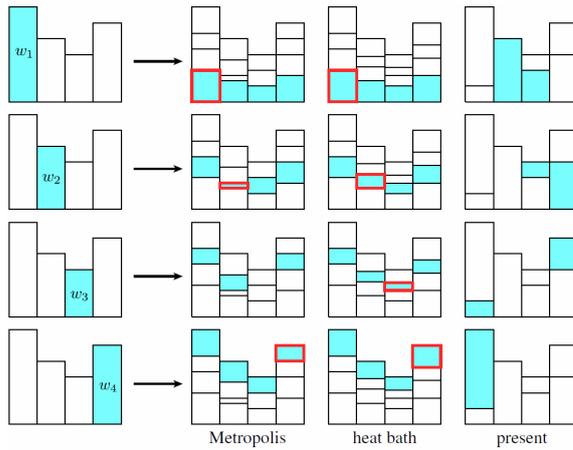


図 3.1 Metropolis–Hastings アルゴリズム (Metropolis), 熱浴法 (heat bath), および Suwa–Todo アルゴリズム (present) における確率流 v_{ij} の幾何学的構成法. 赤太線で囲まれた部分が棄却率に対応する確率流 v_{ii} を表す. 文献 [20] より引用.

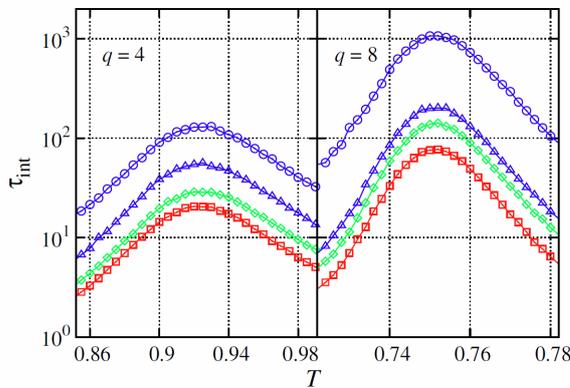


図 3.2 2次元 q 状態ポッツ模型の転移温度近傍における自己相関時間 τ_{int} . 左が $q = 4$, 右が $q = 8$ での結果である. 丸, 三角, ひし形, 四角の各プロットはそれぞれ Metropolis–Hastings アルゴリズム, 熱浴法, locally optimal update [54], Suwa–Todo アルゴリズムでの結果を示している. 文献 [20] より引用.

ている (図 3.2). しかしながら $q = 4$ のとき自己相関時間のシステムサイズ依存性から見積もった動的臨界指数は, Suwa–Todo アルゴリズムを用いても Metropolis–Hastings アルゴリズムや熱浴法の場合から改善がみられないことが報告されている [22]. つまり Suwa–Todo アルゴリズムにおける詳細つり合い条件の破れは, 系の緩和を定量的に加速させるものの, 定性的な変化はもたらさないことを示唆している.

3.2 Irreversible Metropolis–Hastings アルゴリズム

元の状態空間を拡張した上で詳細つり合い条件の破れを導入する手法を総称して lifting と呼ぶ. また lifting によって構成された MCMC 法を lifted MCMC 法と呼ぶ. ここではこれまでに提案されている lifted MCMC 法のうち, ねじれ詳細つり合い条件を用いた手法である irreversible Metropolis–Hastings (IMH) アルゴリズム [26, 29] に注目し, その構成方法と性能評価をまとめる. またこれ以外の lifted MCMC 法は, たとえば文献 [30, 31, 32] などで議論されている. 文献 [30, 31] ではいずれも離散自由度系の状態空間を 2 倍に拡張した上で, それぞれの状態空間では詳細つり合い条件を破りつつも拡張系全体でつり合い条件をみたす遷移確率行列の構成法を提案している. ま

た粒子系に対して提案された event-chain Monte Carlo (ECMC) アルゴリズム [32] は、その後より一般の連続系に適用できる形に拡張され、連続スピン系などに応用されている [33, 34, 35, 36, 37, 38].

IMH アルゴリズムでは新たな確率変数を導入し状態空間を拡張した上で、詳細つり合い条件にかわる新たな条件としてねじれ詳細つり合い条件を課すことで遷移確率を構成する。ねじれ詳細つり合い条件はその概念が提案された当初、全結合イジング模型に対してのみ適用可能なアルゴリズムにおいて提案されていた。その後筆者らは、より一般の離散自由度系へ適用可能な形にアルゴリズムを拡張した [29]. これらをまとめて IMH アルゴリズムと呼んでいる。次章でこのアルゴリズムを具体的な模型に適用し、その性能評価を行うことから、ここでは IMH アルゴリズムを詳しく解説する。

状態空間の拡張

状態空間が $I = \{1, 2, \dots, \Omega\}$ で与えられる有限離散自由度系を考える。また状態空間 I 上の目的分布を $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_\Omega) \in \mathbb{R}^\Omega$ とする。以下ではこの系を着目系と呼ぶ。

まず、着目系に lifting parameter と呼ばれる新たな確率変数 ε を導入する。lifting parameter ε は $+$ か $-$ の 2 値をとる。これによって状態空間は I から $\tilde{I} \equiv I \times \{+, -\}$ へと 2 倍に拡張される。また拡張系全体の状態は、着目系の状態 $i \in I$ と lifting parameter $\varepsilon \in \{+, -\}$ の組 $(i, \varepsilon) \in \tilde{I}$ で記述されることになる。lifting parameter ε の導入にともなって、着目系の状態空間 I 上の目的分布 $\boldsymbol{\pi}$ を拡張系の状態空間 \tilde{I} 上の確率分布に拡張する必要がある。状態 $(i, \varepsilon) \in \tilde{I}$ をとる確率を $\pi_{(i, \varepsilon)}$ と書くとき、拡張系の目的分布 $\tilde{\boldsymbol{\pi}} \in \mathbb{R}^{2\Omega}$ を

$$\begin{aligned} \tilde{\boldsymbol{\pi}} &\equiv (\pi_{(1,+)}, \pi_{(2,+)}, \dots, \pi_{(\Omega,+)}, \pi_{(1,-)}, \pi_{(2,-)}, \dots, \pi_{(\Omega,-)}) \\ &= \frac{1}{2}(\boldsymbol{\pi}, \boldsymbol{\pi}) \end{aligned} \quad (3.5)$$

とする。拡張系の目的分布 $\tilde{\boldsymbol{\pi}}$ は、lifting parameter ε を周辺化して得られる状態 $i \in I$ の確率が着目系の目的分布 $\boldsymbol{\pi}$ における確率 π_i と一致し、また lifting parameter ε の周辺分布が一様分布となるように設定されている。

拡張系の遷移確率行列

次に拡張系の状態空間 \tilde{I} における遷移確率行列を構成する。特に本手法では、マルコフ連鎖の 1 ステップで着目系の状態 $i \in I$ と lifting parameter $\varepsilon \in \{+, -\}$ を同時に更新する遷移は考えない。すなわち状態 $(i, \varepsilon) \in \tilde{I}$ から状態 $(j, \varepsilon') \in \tilde{I}$ への遷移確率を $\tilde{T}(j, \varepsilon' | i, \varepsilon)$ と書くとき、 $i \neq j$ かつ $\varepsilon \neq \varepsilon'$ のとき $\tilde{T}(j, \varepsilon' | i, \varepsilon) = 0$ とする。いま、これ以外

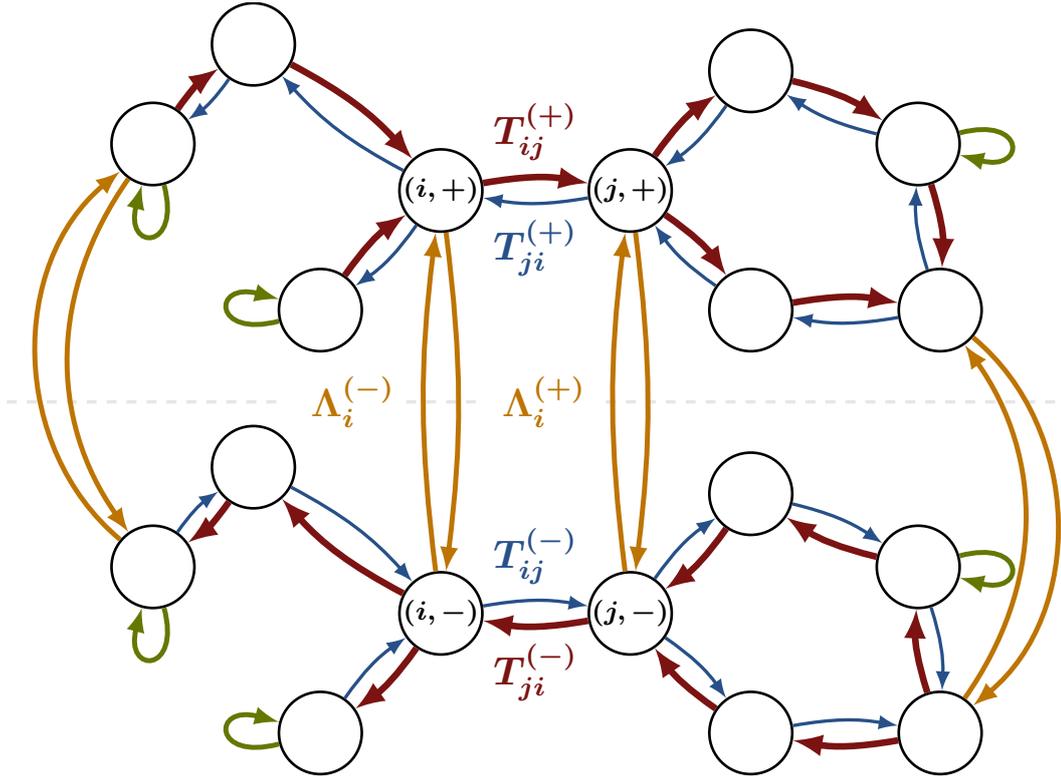


図 3.3 lifting parameter の導入によって拡張された状態空間における状態遷移図の例.

の遷移確率を

$$T_{ij}^{(\varepsilon)} \equiv \tilde{T}(j, \varepsilon | i, \varepsilon), \quad (3.6)$$

$$\Lambda_i^{(\varepsilon)} \equiv \tilde{T}(i, -\varepsilon | i, \varepsilon) \quad (3.7)$$

と表すことにする. ここで $T_{ij}^{(\varepsilon)}$, $\Lambda_i^{(\varepsilon)}$ はそれぞれ, lifting parameter ε を固定した下での着目系の遷移確率, および着目系の状態を固定した下で lifting parameter ε を flip する遷移確率である. これより拡張系の状態空間 \tilde{I} 上の状態遷移図は一般に図 3.3 で与えられる. これらをまとめた行列をそれぞれ $\mathsf{T}^{(\varepsilon)} \equiv (T_{ij}^{(\varepsilon)})_{i,j \in I} \in \mathbb{R}^{\Omega \times \Omega}$, $\Lambda^{(\varepsilon)} \equiv \text{diag}(\Lambda_i^{(\varepsilon)})_{i \in I} \in \mathbb{R}^{\Omega \times \Omega}$ とするとき, 拡張系の状態空間 \tilde{I} における遷移確率行列 $\tilde{\mathsf{T}} \in \mathbb{R}^{2\Omega \times 2\Omega}$ として

$$\tilde{\mathsf{T}} = \begin{pmatrix} \mathsf{T}^{(+)} & \Lambda^{(+)} \\ \Lambda^{(-)} & \mathsf{T}^{(-)} \end{pmatrix} \quad (3.8)$$

なる形で与えられるものを考える. 行列 $\tilde{\mathsf{T}}$ は拡張系の状態空間 \tilde{I} における遷移確率行列であるので, すべての $(i, \varepsilon) \in \tilde{I}$ に対して確率保存則

$$\sum_{j=1}^{\Omega} T_{ij}^{(\varepsilon)} + \Lambda_i^{(\varepsilon)} = 1 \quad (3.9)$$

をみtas.

ねじれ詳細つり合い条件

拡張系の状態空間 \tilde{I} における遷移確率行列 \tilde{T} が目的分布 $\tilde{\pi}$ を定常分布にもつための条件として, つり合い条件 $\tilde{\pi} = \tilde{\pi}\tilde{T}$ を課す. 従来ならばつり合い条件の十分条件として詳細つり合い条件を課し遷移確率を構成していた. しかしここではそのかわりとして, ねじれ詳細つり合い条件 (skew detailed balance condition, SDBC) と呼ばれる条件を課す:

$$\pi_i T_{ij}^{(+)} = \pi_j T_{ji}^{(-)}. \quad (3.10)$$

ねじれ詳細つり合い条件とは, $\varepsilon = +$ の空間における確率流 $\pi_i T_{ij}^{(+)}$ と, $\varepsilon = -$ の空間における逆向きの確率流 $\pi_j T_{ji}^{(-)}$ がつり合うことを要請するものである. これによって詳細つり合い条件は一般に破れることになる.

ねじれ詳細つり合い条件だけでは, まだつり合い条件の十分条件とはなっていない. つり合い条件の式に確率保存則 (3.9) およびねじれ詳細つり合い条件 (3.10) を用いて変形すると, lifting parameter ε を flip する遷移確率 $\Lambda_i^{(\varepsilon)}$ について

$$\Lambda_i^{(+)} - \Lambda_i^{(-)} = \sum_{j \neq i} \left(T_{ij}^{(-)} - T_{ij}^{(+)} \right) \quad (3.11)$$

なる条件を得る. 言い換えれば, 遷移確率行列 \tilde{T} にねじれ詳細つり合い条件 (3.10) と $\Lambda_i^{(\varepsilon)}$ の条件 (3.11) を課すことによって, 目的分布 $\tilde{\pi}$ への収束が保証される*2.

Irreversible Metropolis–Hastings アルゴリズム

以下ではねじれ詳細つり合い条件をみだす遷移確率行列 \tilde{T} を実際に構成し, 計算機上でねじれ詳細つり合い条件をみだすマルコフ連鎖を実現するための irreversible Metropolis–Hastings アルゴリズムを提示する.

まず詳細つり合い条件をみだす従来の Metropolis–Hastings アルゴリズムと同様に, 着目系の遷移確率 $T_{ij}^{(\varepsilon)}$ を提案確率 $q_{ij}^{(\varepsilon)}$ と採択確率 $w_{ij}^{(\varepsilon)}$ の積に分解する:

$$T_{ij}^{(\varepsilon)} = q_{ij}^{(\varepsilon)} w_{ij}^{(\varepsilon)}. \quad (3.12)$$

このとき, IMH アルゴリズムにおける 1 ステップの手順はアルゴリズム 3.2 のとおりである.

拡張系の目的分布 $\tilde{\pi}$ は lifting parameter ε について周辺化すると着目系の目的分布 π と等価になることから, IMH アルゴリズムで得られるマルコフ連鎖 $\{(i^{(n)}, \varepsilon^{(n)})\}_{n=1,2,\dots}$ から物理量の期待値を計算するには, 各ステップでの lifting parameter の値 $\varepsilon^{(n)}$ を気にせずに着目系の状態列 $(i^{(n)})_{n=1,2,\dots}$ だけを用いればよい.

*2 ここではエルゴード性を仮定している.

アルゴリズム 3.2: irreversible Metropolis–Hastings アルゴリズム

(1) [着目系の状態更新]

現在の状態が $(i, \varepsilon) \in \tilde{I}$ であるとする. このとき提案確率 $q_{ij}^{(\varepsilon)}$ で次の状態候補 $j \in I$ を提案する. さらに採択確率 $w_{ij}^{(\varepsilon)}$ で提案された候補 j を採択し, 状態を (j, ε) へと更新する.

(2) [lifting parameter の状態更新]

手順 (1) で状態更新が棄却された場合のみ, この手順を行う. 確率 $\lambda_i^{(\varepsilon)}$ で lifting parameter ε を flip し, 状態を $(i, -\varepsilon)$ へと更新する. ここで確率 $\lambda_i^{(\varepsilon)}$ は

$$\lambda_i^{(\varepsilon)} \equiv \frac{\Lambda_i^{(\varepsilon)}}{1 - \sum_{j \neq i} q_{ij}^{(\varepsilon)} w_{ij}^{(\varepsilon)}} \quad (3.13)$$

で与えられる.

(3) 手順 (1), (2) で状態更新がともに棄却されたとき, 現在の状態 (i, ε) を次の状態として採択する.

提案確率 $q_{ij}^{(\varepsilon)}$ および採択確率 $w_{ij}^{(\varepsilon)}$ の具体的な選び方を考えよう. まず着目系の状態空間 I 上で目的分布 π に対して詳細つり合い条件をみたす遷移確率行列 $T = (T_{ij})_{i,j \in I}$ をひとつ用意する. 遷移確率 T_{ij} は提案確率 q_{ij} と採択確率 w_{ij} の積で与えられているとする. これを用いて, ねじれ詳細つり合い条件をみたす遷移確率 $T_{ij}^{(\varepsilon)}$ を

$$T_{ij}^{(\varepsilon)} = \frac{1 + \Delta_{ij}^{(\varepsilon)}}{2} q_{ij} w_{ij} \quad (3.14)$$

と構成できる. ここで $\Delta_{ij}^{(\varepsilon)}$ は状態 $i, j \in I$ および lifting parameter ε の関数であり, $|\Delta_{ij}^{(\varepsilon)}| \leq 1$ かつ $\Delta_{ij}^{(+)} = \Delta_{ji}^{(-)}$ をみたすようにこちらで定める. あとは (3.14) 式が (3.12) 式と整合するように, 因子 $(1 + \Delta_{ij}^{(\varepsilon)})/2$ を IMH アルゴリズムの提案確率 $q_{ij}^{(\varepsilon)}$ と採択確率 $w_{ij}^{(\varepsilon)}$ のどちらかに割り当てればよい. たとえば文献 [27, 28] では, イジング模型に対して IMH アルゴリズムを構成している. そこではねじれ詳細つり合い条件からくる因子 $(1 + \Delta_{ij}^{(\varepsilon)})/2$ をすべて採択確率に割り当てている. つまり詳細つり合い条件をみたす従来の Metropolis–Hastings アルゴリズムと同様に, イジングスピンをランダムにひとつ選ぶという提案確率を用いている. その上で詳細つり合い条件の破れた採択確率 $w_{ij}^{(\varepsilon)}$ を用いて状態を更新し, 全体としてねじれ詳細つり合い条件をみたしている. 一方, 文献 [26] ではねじれ詳細つり合い条件からくる因子 $(1 + \Delta_{ij}^{(\varepsilon)})/2$ をすべて提案確率に割り当てて

いる. 全結合イジング模型に対する IMH アルゴリズムにおいて, lifting parameter が $\varepsilon = +$ のときは上向きスピンの中からひとつ選び, $\varepsilon = -$ のときは下向きスピンの中からひとつ選ぶという提案確率である. その上で採択確率は従来の Metropolis–Hastings アルゴリズムで用いられているものを使用している. このように提案確率と採択確率の積 $q_{ij}^{(\varepsilon)} w_{ij}^{(\varepsilon)}$ がねじれ詳細つり合い条件をみたしていればよく, 提案確率か採択確率のいずれで詳細つり合い条件を破ってもよい.

次に lifting parameter ε を flip する遷移確率 $\Lambda_i^{(\varepsilon)}$ の選び方を考えよう. 条件 (3.11) および確率保存則 (3.9) から要請される不等式

$$0 \leq \Lambda_i^{(\varepsilon)} \leq 1 - \sum_{j \neq i} T_{ij}^{(\varepsilon)} \quad (3.15)$$

をみたす遷移確率 $\Lambda_i^{(\varepsilon)}$ として, これまでのところ SH₁ 型, SH₂ 型, SH₃ 型, TCV 型と呼ばれる 4 種類の構成法が知られている:

$$\text{SH}_1 \text{ 型} \quad \Lambda_{i,\text{SH}_1}^{(\varepsilon)} = \sum_{j \neq i} T_{ij}^{(-\varepsilon)}, \quad (3.16)$$

$$\text{SH}_2 \text{ 型} \quad \Lambda_{i,\text{SH}_2}^{(\varepsilon)} = \frac{1}{2} \left[1 - \varepsilon \sum_{j \neq i} (T_{ij}^{(+)} - T_{ij}^{(-)}) \right], \quad (3.17)$$

$$\text{SH}_3 \text{ 型} \quad \Lambda_{i,\text{SH}_3}^{(\varepsilon)} = \sum_{j \neq i} \frac{1 - \Delta_{ij}^{(\varepsilon)}}{2} T_{ij}, \quad (3.18)$$

$$\text{TCV 型} \quad \Lambda_{i,\text{TCV}}^{(\varepsilon)} = \max \left[0, -\varepsilon \sum_{j \neq i} (T_{ij}^{(+)} - T_{ij}^{(-)}) \right]. \quad (3.19)$$

以上のとおり構成した遷移確率を用いることで, IMH アルゴリズムを実行できる.

Irreversible Metropolis–Hastings アルゴリズムの性能評価

IMH アルゴリズムはこれまでいくつかの系に対して適用され, 数値的に性能が評価されている. はじめてねじれ詳細つり合い条件の考え方が提案された研究では, 全結合イジング模型に対して IMH アルゴリズムが実装された [26]. この例では lifting parameter を flip する遷移確率に TCV 型を用い, イジングスピンの自己相関関数を測定した. さらにこれを用いて自己相関時間を見積もった. その結果詳細つり合い条件を破ることで, 相転移温度直上における自己相関時間のイジングスピン数 N 依存性がおおむね $O(N^{3/2})$ から $O(N^{3/4})$ へと質的に改善することが報告されている. さらにこの結果は理論的にも説明されている [56]

また筆者らは 1 次元イジング模型に対して IMH アルゴリズムを適用し, 磁化密度の時間発展を解析的・数値的に調べた [27]. ねじれ詳細つり合い条件によって導入された関

数 $\Delta_{ij}^{(\epsilon)}$ の選び方から, ここでは詳細つり合い条件の破れ度合いがパラメタ $\delta \in [-1, 1]$ によって表現された. δ はその絶対値が大きいほど詳細つり合い条件の破れが大きく, $\delta = 0$ とすることで詳細つり合い条件が回復するパラメタとなっている. SH₁ 型, SH₂ 型, TCV 型の遷移確率を用いた数値計算によって, 詳細つり合い条件の破れが磁化密度の緩和を加速することが示された. またマスター方程式の解析によって, その加速度合いが定量的に評価された. その結果, SH₁ 型では磁化密度の緩和時間が定数倍の改善にとどまる一方, SH₂ 型では動的臨界指数 z を $z = 2$ から $z = 1$ へと減少させ, 緩和ダイナミクスが質的に改善していることが明らかとなった. TCV 型については数値的に観測した磁化密度の緩和ダイナミクスに減衰振動のような振る舞いがみられ, 遷移確率行列に複素固有値の存在が示唆された.

さらに IMH アルゴリズムは 2 次元・3 次元の強磁性イジング模型にも適用され, その動的臨界指数が評価されている [28]. その結果, 高温領域での磁化密度の緩和ダイナミクスには改善がみられるものの, 相転移温度直上ではその改善がみえなくなり, 動的臨界指数にも詳細つり合い条件の有無による変化は認められなかった.

3.3 詳細つり合い条件の破れと性能評価

以上でみたとおり, 詳細つり合い条件を破ることで従来の MCMC 法よりも性能が向上する例がこれまでいくつか報告されている. しかしこれらの結果はいずれも数値的に得られたものであり, 理論的に説明されているものはわずかである. ここではまず, 詳細つり合い条件の破れが MCMC 法の性能に及ぼす影響について一般に示されている重要な結果をふたつ紹介する. これらはいずれも, 詳細つり合い条件の破れた MCMC 法の性能と, それに対応する可逆な MCMC 法の性能の間に成り立つ不等式を与える. しかし一方で, これらの不等式だけでは詳細つり合い条件の破れの効果を十分に理解したとはいえないことを指摘する. これをふまえて今後議論すべき論点を明らかにする.

以下ではエルゴード的な遷移確率行列 T を考える. 遷移確率行列 T は目的分布 π に対してつり合い条件をみたすが, 詳細つり合い条件をみたすとは限らないとする. このとき, 遷移確率行列 T に対応する可逆な遷移確率行列 T_0 を以下のとおり定義する:

$$T_0 \equiv \frac{1}{2}(T + B^{-1}T^TB). \quad (3.20)$$

ここで $B \equiv \text{diag}(\pi_i)_{i \in I}$ は目的分布 π から定まる対角行列である. 行列 T_0 は遷移確率行列 T の additive reversibilization と呼ばれる. 遷移確率行列 T のつり合い条件から $T_0 \mathbf{1}^T = \mathbf{1}^T$ が成り立つので, additive reversibilization T_0 は遷移確率行列になっている. また additive reversibilization T_0 の (i, j) 成分 $T_{0,ij}$ は $i \neq j$ に対して

$T_{0,ij} = (T_{ij} + \pi_j T_{ji} \pi_i^{-1})/2$ で与えられることから,

$$\pi_i T_{0,ij} = \frac{1}{2}(\pi_i T_{ij} + \pi_j T_{ji}) = \pi_j T_{0,ji} \quad (3.21)$$

が成り立つ. すなわち, additive reversibilization T_0 は目的分布 π に対して詳細つり合い条件をみたす遷移確率行列である. 特に遷移確率行列 T が目的分布 π に対して詳細つり合い条件をみたすならば, 対応する additive reversibilization T_0 は元の遷移確率行列 T に一致する.

ある遷移確率行列とその additive reversibilization の間には, その固有値について以下の関係が成り立つことが示されている [44]. 原論文では有限離散状態・連続時間マスター方程式で議論されているが, そこで得られる結果は適切な離散化を介して有限離散状態・離散時間マスター方程式でも成り立つ.

Ichiki–Ohzeki (2013) [44]

目的分布 π に対してつり合い条件をみたすエルゴード的な遷移確率行列 T およびその additive reversibilization T_0 の固有値を, 重複も許してそれぞれ $\lambda_k^T \in \mathbb{C}$, $\lambda_k^{T_0} \in \mathbb{R}$ ($k = 1, 2, \dots, \Omega$) とする. 一般性を失うことなく, $1 = \lambda_1^T > \operatorname{Re}\lambda_2^T \geq \dots \geq \operatorname{Re}\lambda_\Omega^T > -1$, $1 = \lambda_1^{T_0} > \lambda_2^{T_0} \geq \dots \geq \lambda_\Omega^{T_0} > -1$ としてよい. このとき

$$\operatorname{Re}\lambda_2^T \leq \lambda_2^{T_0} \quad (3.22)$$

が成り立つ.

また同じ設定の下, 漸近分散についても以下の関係が示されている [24].

Sun et al. (2010) [24]

目的分布 π に対してつり合い条件をみたすエルゴード的な遷移確率行列を T とし, その additive reversibilization を T_0 とする. このとき

- (i) ゼロでない任意の物理量 f に対して, 不等式

$$v(f, \pi, T) \leq v(f, \pi, T_0) \quad (3.23)$$

が成り立つ.

- (ii) 遷移確率行列 T が詳細つり合い条件を破っているならば, 不等式

$$v(f, \pi, T) < v(f, \pi, T_0) \quad (3.24)$$

をみたす物理量 f が少なくともひとつ存在する.

以上から, 緩和の意味でも相関の意味でも詳細つり合い条件の破れた MCMC 法は

additive reversibilization に対応する可逆な MCMC 法よりも効率がよいことが一般に示されたことになる。ただし、これらの結果をもって「詳細つり合い条件は破ったほうがよい」と結論づけるのは早計である。一般に詳細つり合い条件を破る MCMC 法をひとつ構成したとしよう。このときその additive reversibilization に対応する MCMC 法は、従来用いられている MCMC 法に対応物がない架空のものになっていることが普通である。実際これまで提案されてきた詳細つり合い条件を破る MCMC 法のほとんどは additive reversibilization に対応する MCMC 法が不明である。特に従来の可逆な MCMC 法を拡張して詳細つり合い条件を破った irreversible Metropolis–Hastings (IMH) アルゴリズム [50] でも、その additive reversibilization は元の遷移確率行列に一致しない。そのため、これらの結果は「詳細つり合い条件を破れば、詳細つり合い条件をみたすが素性はよくわからないアルゴリズムと比較して性能がよい」という主張にとどまっている。実際に性能を比較したい相手と比較できていないのである。

特定のアルゴリズムに絞った議論でも、詳細つり合い条件の破れの効果は理論的に十分に理解されているとは言いがたい。たとえば Suwa–Todo アルゴリズムの設計指針は「棄却率最小化」であり、その根底には Peskun の定理がある。たしかに棄却率が小さいほど他の状態への遷移が頻繁に起こり、直感的にはサンプル間の相関が小さくなりそうである。しかし Peskun の定理は詳細つり合い条件を課した範囲でしか証明されておらず、詳細つり合い条件を破ってもなお同様の主張が成り立つのかは明らかではない。また IMH アルゴリズムではいくつかの特定の模型で理論的に性能が評価されており、それらはいずれも詳細つり合い条件の破れが性能を向上させると結論づけている。しかしより一般の系に対しても IMH アルゴリズムを用いることがよいのかどうかは議論の余地がある。さらに IMH アルゴリズムを用いた数値計算では複素固有値の存在が示唆されていた。前章でみたとおり詳細つり合い条件をみたす遷移確率行列には実固有値しか現れないため、複素固有値の出現は詳細つり合い条件を破ったことによるものである。しかし複素固有値の出現が MCMC 法の性能向上に結びつくかどうかは明らかではない。このように、詳細つり合い条件の破れが MCMC 法の性能にもたらす影響は理論的に明らかでない点が多く残されている。本博士論文では次章以降、IMH アルゴリズムの性能評価を通してこれらの論点を吟味し、いくつかの論点を否定的に解決した。

第 4 章

詳細つり合い条件を破る 1 次元ランダムウォーク

本章では MCMC 法の性能評価のひとつとして, 1 次元リング上のランダムウォークを解析する. はじめに詳細つり合い条件をみたすランダムウォークの結果をまとめる. その後詳細つり合い条件の破れを導入した先行研究を概観し, 特にねじれ詳細つり合い条件を用いた lifting によって詳細つり合い条件の破れを導入した場合の解析結果を詳細に述べる. 解析では対応する遷移確率行列を厳密に対角化し, すべての固有値, 固有ベクトルを導出する. これらに基づいて MCMC 法の性能を議論する.

本章の主要な結果は文献 [50] として出版されている.

4.1 詳細つり合い条件をみたす 1 次元ランダムウォーク

状態空間を $I = \{1, 2, \dots, \Omega\}$ とし, 1 次元リング上に状態 $i = 1, 2, \dots, \Omega$ が周期境界条件の下で並んでいるとする. この状態空間 I 上に以下の遷移確率行列 T を定義する:

$$T = (1 - \alpha)I_{\Omega} + J_{\Omega} \left(\frac{\alpha}{2}, \frac{\alpha}{2} \right). \quad (4.1)$$

ここで $0 < \alpha \leq 1$ は遷移レート, 行列 $I_{\Omega} \in \mathbb{R}^{\Omega \times \Omega}$ は Ω 次の単位行列を表す. また, 実数 a, b に対して

$$J_{\Omega}(a, b) \equiv \begin{pmatrix} 0 & a & & & b \\ b & 0 & a & & \\ & \ddots & \ddots & \ddots & \\ & & b & 0 & a \\ a & & & b & 0 \end{pmatrix} \in \mathbb{R}^{\Omega \times \Omega} \quad (4.2)$$

と定義する. この遷移確率行列 T が定める状態遷移図を図 4.1 に示す. 遷移確率行列 T は 2 重確率行列であり, その定常分布 π はすべての状態 $i \in I$ について $\pi_i = 1/\Omega$ なる一

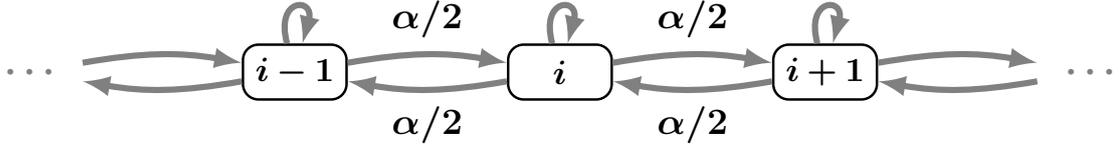


図 4.1 1次元リング上で詳細つり合い条件をみたすランダムウォークの状態遷移図.

様分布で与えられる. さらに遷移確率行列 T はこの一様分布に対して詳細つり合い条件をみたしている.

(4.1) 式で与えられる遷移確率行列 T は厳密対角化が可能である. 付録 A.1 の結果を用いると, 遷移確率行列 T の固有値 λ_k は $k = 1, 2, \dots, [\Omega/2] + 1$ に対して

$$\lambda_k = 1 - \alpha + \alpha \cos \theta_k \quad (4.3)$$

で与えられる. ここで $[x]$ は実数 x を超えない最大の整数を表す. また $\theta_k \equiv 2\pi(k-1)/\Omega$ である. さらに遷移確率行列 T における固有値 λ_k の重複度 m_k や左固有ベクトル $\mathbf{u}_{k,\sigma}$, 右固有ベクトル $\mathbf{v}_{k,\sigma}$ はすべて行列 $J_\Omega(\alpha/2, \alpha/2)$ のそれらと等しい. これらに基づいて, 緩和時間 τ_{relax} や漸近分散 $v(f, \pi, T)$ を解析的に評価できる.

4.2 ねじれ詳細つり合い条件をみたす 1次元ランダムウォーク

lifting を用いたランダムウォークにはいくつかの先行研究がある. Diaconis らは lifting の考えに基づいて 1次元ランダムウォークを拡張し, 詳細つり合い条件の破れたランダムウォークを構成した [57]. さらに全変動距離 (total variation distance) と χ^2 距離 (χ^2 distance) の収束を解析し, いずれも詳細つり合い条件をみたす 1次元ランダムウォークと比較して収束の速さが定性的に改善することを示した. また Chen らは lifting によってランダムウォークの mixing time が状態数 Ω に対して $O(\Omega^2)$ から $O(\Omega)$ まで小さくできることを示した [58]. Vucelja は 2次元トーラス上のランダムウォークに対してねじれ詳細つり合い条件の手法に基づく lifting を用い, 位置の自己相関関数を数値的に測定した [59]. その結果詳細つり合い条件の破れによって自己相関関数の緩和が加速するとともに, 減衰振動的な振る舞いがみられ, 複素固有値の出現が示唆された. いずれの結果も詳細つり合い条件の破れが MCMC 法の性能改善に寄与することを支持するものである.

本節では, 1次元リング上のランダムウォークに対してねじれ詳細つり合い条件を用いた lifting によって詳細つり合い条件の破れを導入する. 本博士論文のアプローチは先行研究のものとは異なる. 本博士論文では拡張された遷移確率行列の固有値, 固有ベクトル

をすべて解析的に導出し、それらを用いて緩和時間や漸近分散を計算する。その結果として、詳細つり合い条件の破れによって緩和時間や性能指数の最悪評価が改善できることを示す。これらの結果はすべて先行研究の結果と整合するものである。しかし一方で、パラメタの値によっては詳細つり合い条件の破れが必ずしも MCMC 法の性能を改善させるわけではないことも明らかとなる。

4.2.1 遷移確率行列の拡張

ねじれ詳細つり合い条件の手法を1次元ランダムウォークに適用しよう。まず、前節で議論した可逆な1次元ランダムウォークを着目系とする。この着目系に lifting parameter $\varepsilon \in \{+, -\}$ を導入することで、状態空間は2倍に拡張される。拡張系の状態はランダムウォークの状態 i と lifting parameter ε の組 (i, ε) で記述される。これにともなって、目的分布もまた拡張される。拡張系の目的分布 $\tilde{\pi}$ は lifting parameter ε も加えた一様分布へと拡張される。すなわち、拡張系が状態 (i, ε) をとる確率 $\pi_{(i, \varepsilon)}$ は $\pi_{(i, \varepsilon)} = 1/(2\Omega)$ で与えられる。

拡張系の状態空間における遷移確率行列 \tilde{T} を、

$$\tilde{T} = \left[1 - \left(1 + \frac{\delta - \delta'}{2} \right) \frac{\alpha}{2} - \gamma \right] I_{2\Omega} + \tilde{J}_{2\Omega} \left(\frac{1 + \delta}{2} \frac{\alpha}{2}, \frac{1 - \delta'}{2} \frac{\alpha}{2}; \gamma \right) \quad (4.4)$$

と定義する。ここで実数 a, b, c に対して

$$\tilde{J}_{2\Omega}(a, b; c) \equiv \begin{pmatrix} J_{\Omega}(a, b) & cI_{\Omega} \\ cI_{\Omega} & J_{\Omega}(b, a) \end{pmatrix} \in \mathbb{R}^{2\Omega \times 2\Omega} \quad (4.5)$$

である。 δ, δ' は詳細つり合い条件の破れを特徴づけるパラメタであり、不等式 $|\delta| \leq 1$ および $|\delta'| \leq 1$ をみたす。絶対値 $|\delta|$ や $|\delta'|$ が大きいほど詳細つり合い条件の破れが大きく、特に $\delta = \delta' = 0$ のとき詳細つり合い条件をみたすランダムウォークに帰着する。また γ は lifting parameter ε を flip する遷移確率である。第3章で述べた IMH アルゴリズムの一般論では、 ε -flip の遷移確率は SH₁ 型, SH₂ 型, SH₃ 型, および TCV 型の4通りしか構成されていなかった。しかし今回の問題では、 ε -flip の遷移確率 $\Lambda_i^{(\varepsilon)}$ に対する条件 (3.11) を書き下すと $\Lambda_i^{(+)} = \Lambda_i^{(-)}$ が得られる。そのためこれらの値をともに γ で表すとき、遷移確率 γ は

$$0 \leq \gamma \leq 1 - \left(1 + \frac{\delta - \delta'}{2} \right) \frac{\alpha}{2} \quad (4.6)$$

の範囲で任意の値をとることができる。この範囲には、SH₁ 型, SH₂ 型, SH₃ 型, および TCV 型に対応する遷移確率 γ の値

$$\gamma_{\text{SH}_1} = \left(1 + \frac{\delta - \delta'}{2} \right) \frac{\alpha}{2}, \quad \gamma_{\text{SH}_2} = \frac{1}{2}, \quad \gamma_{\text{SH}_3} = \left(1 - \frac{\delta - \delta'}{2} \right) \frac{\alpha}{2}, \quad \gamma_{\text{TCV}} = 0 \quad (4.7)$$

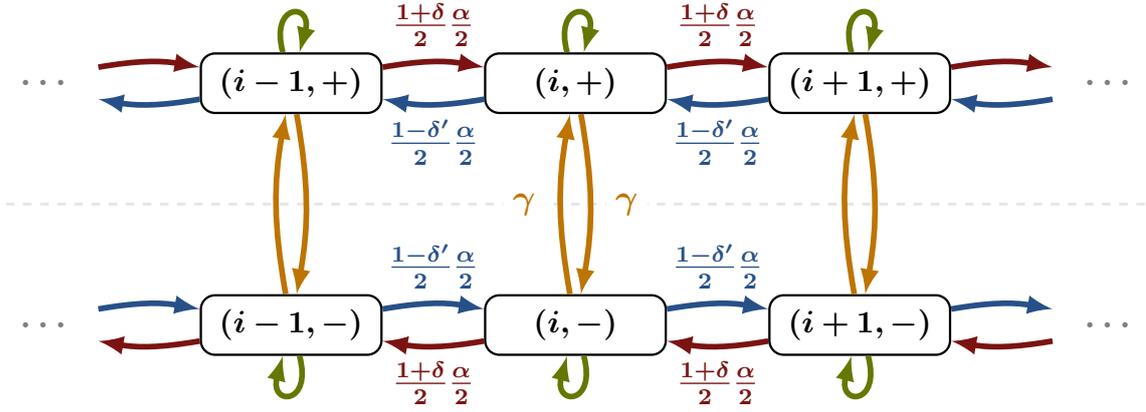


図 4.2 ねじれ詳細つり合い条件を用いた lifting によって拡張された, 1次元ランダムウォークの状態遷移図.

がすべて含まれている. 拡張系の遷移確率行列 \tilde{T} が定める状態遷移図は図 4.2 で与えられる.

拡張系の遷移確率行列 \tilde{T} の固有値および固有ベクトルはすべて解析的に導出できる. 付録 A.2 の結果を用いると, 遷移確率行列 \tilde{T} の固有値 $\tilde{\lambda}_k^\pm$ は $k = 1, 2, \dots, [\Omega/2] + 1$ に対して

$$\tilde{\lambda}_k^\pm = 1 - \left(1 + \frac{\delta - \delta'}{2}\right) \alpha \sin^2 \frac{\theta_k}{2} - \gamma \pm \sqrt{\gamma^2 - \left(\frac{\delta + \delta'}{2}\right)^2 \left(\frac{\alpha}{2} \sin \theta_k\right)^2} \quad (4.8)$$

で与えられる. 付録 A.2 で詳しく述べられているとおり, パラメタ δ, δ' , および γ の値によってはいくつかの固有値が複素数になりうる. さらにある整数 $q \in \{2, 3, \dots, [(\Omega + 1)/2]\}$ に対して $\gamma = |\delta + \delta'|(\alpha/4) \sin \theta_q$ が成り立つとき, 遷移確率行列 \tilde{T} は対角化可能でない. また固有値 $\tilde{\lambda}_k^\pm$ の重複度 m_k^\pm および固有ベクトル $\tilde{\mathbf{u}}_{k,\sigma}^\pm, \tilde{\mathbf{v}}_{k,\sigma}^\pm$ はすべて行列 $\tilde{J}_{2\Omega}(a, b; c)$ において $a = (1 + \delta)\alpha/4, b = (1 - \delta')\alpha/4, c = \gamma$ としたときのそれらと等しい. 以上の結果を用いることで, 拡張されたランダムウォークの緩和時間や漸近分散を解析的に評価できる.

4.2.2 性能の比較

まずねじれ詳細つり合い条件をみたす遷移確率行列 \tilde{T} の additive reversibilization をみておこう. 遷移確率行列 \tilde{T} の additive reversibilization \tilde{T}_0 は

$$\tilde{T}_0 = \left[1 - \left(1 + \frac{\delta - \delta'}{2}\right) \frac{\alpha}{2} - \gamma\right] I_{2\Omega} + \tilde{J}_{2\Omega} \left[\left(1 + \frac{\delta - \delta'}{2}\right) \frac{\alpha}{4}, \left(1 + \frac{\delta - \delta'}{2}\right) \frac{\alpha}{4}; \gamma \right] \quad (4.9)$$

で与えられる. これより本節で議論しているランダムウォークの問題では, 拡張系の遷移確率行列 \tilde{T} で $\delta = \delta' = 0$ としたものが, $\delta = \delta' (\neq 0)$ としたときの遷移確率行列 \tilde{T} の additive reversibilization と等しい. つまり遷移確率行列のパラメタ δ, δ' 依存性をあらわに書けば

$$\tilde{T}(\delta = 0, \delta' = 0) = \tilde{T}_0(\delta = \delta') \quad (4.10)$$

である. したがって第2章で述べたとおり, $\delta = \delta' (\neq 0)$ のときの遷移確率行列 \tilde{T} は Ichiki–Ohzeki の結果 [44] から第2固有値の実部が, Sun らの結果 [24] から漸近分散がそれぞれ $\delta = \delta' = 0$ としたときの遷移確率行列 \tilde{T} のそれと比較して改善する. しかし, 拡張系の緩和時間や漸近分散が着目系と比較して改善するかどうかは明らかではない. また $\delta \neq \delta'$ のときは additive reversibilization にパラメタ δ, δ' 依存性が残るなど, 性能の比較がやや煩雑になる. そのためここでは2通りの設定

- 設定 (a) $\delta = \delta'$
- 設定 (b) $\delta' = 1$

を考察し, 性能の比較を行う. 以下では簡単のため $\alpha = 1/2$ とし, 状態数 Ω は十分大きな偶数とする. 前節までに得られた遷移確率行列 T および \tilde{T} の解析結果を用いて, これらの遷移確率行列に対応する MCMC 法の性能を確率分布の緩和時間, 物理量の期待値の緩和時間, および漸近分散の観点から比較する.

確率分布の緩和時間

ランダムウォークにおける確率分布の緩和時間を考える. 第2章の一般論から, 確率分布の収束の速さは絶対値の意味での第2固有値 (second-largest eigenvalue modulus, SLEM) によって特徴づけられる. 着目系の遷移確率行列 T の SLEM η は

$$\eta = \cos^2 \frac{\pi}{\Omega} \quad (4.11)$$

で与えられる. よってこのとき着目系の緩和時間 τ_{relax} は $\Omega \gg 1$ の下で

$$\tau_{\text{relax}} = -\frac{1}{\ln(\cos^2 \frac{\pi}{\Omega})} \simeq \frac{\Omega^2}{\pi^2} \quad (4.12)$$

で与えられる. これに対して, 拡張系の遷移確率行列 \tilde{T} の SLEM $\tilde{\eta}$ を与える固有値はパラメタの値に応じて変化する. 詳細は付録 B に譲ることにする. いずれにおいても拡張系の緩和時間 $\tilde{\tau}_{\text{relax}}$ は SLEM $\tilde{\eta}$ を用いて

$$\tilde{\tau}_{\text{relax}} = -\frac{1}{\ln \tilde{\eta}} \quad (4.13)$$

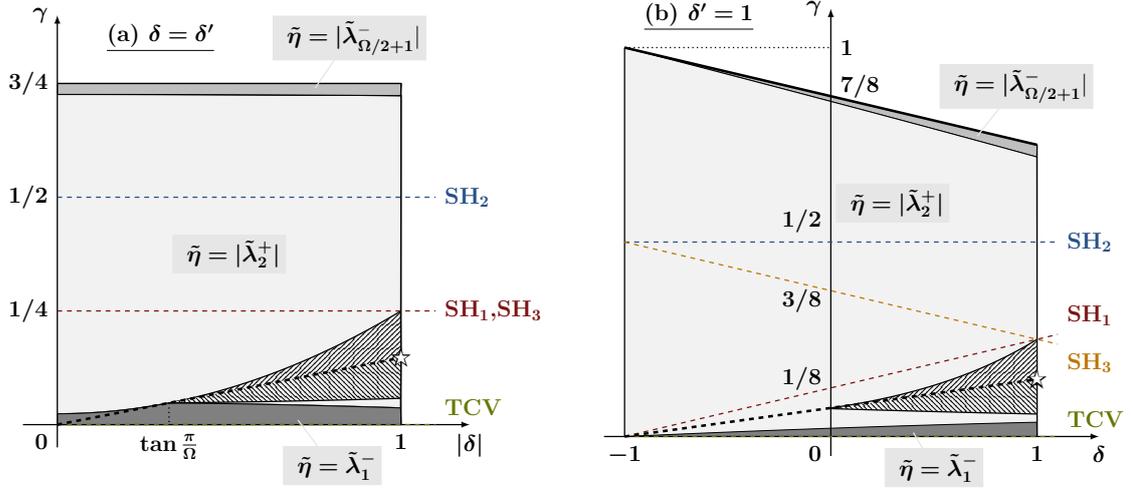


図 4.3 拡張系の遷移確率行列 \tilde{T} の SLEM $\tilde{\eta}$ をパラメタ (δ, γ) の値で分類した図. 設定 (a) $\delta = \delta'$, 設定 (b) $\delta' = 1$ のそれぞれにおいて, 斜線部の領域では着目系の遷移確率行列 T の SLEM η に対して不等式 $\tilde{\eta} \leq \eta$ が成り立っている. 太い破線はパラメタ δ を固定したときに拡張系の緩和時間 $\tilde{\tau}_{\text{relax}}$ が最小となるパラメタ γ の値を示している. また星印 (☆) はパラメタ (δ, γ) 空間全体で拡張系の緩和時間 $\tilde{\tau}_{\text{relax}}$ が最小となる点を表す. 太い破線より下の領域では固有値 $\tilde{\lambda}_2^+$ が複素数となる.

で与えられる.

拡張系の緩和時間 $\tilde{\tau}_{\text{relax}}$ を着目系の緩和時間 τ_{relax} と比較することで, 緩和時間が改善するパラメタ領域を明らかにする. 緩和時間が改善する, すなわち不等式 $\tilde{\tau}_{\text{relax}} \leq \tau_{\text{relax}}$ が成り立つパラメタ領域 (δ, γ) は設定 (a) $\delta = \delta'$, 設定 (b) $\delta' = 1$ のそれぞれに対して

$$\frac{1}{2} \sin^2 \frac{\pi}{\Omega} + \frac{1}{16} \frac{\sin^2 \frac{2\pi}{\Omega}}{1 + \cos^2 \frac{\pi}{\Omega}} \left(\delta^2 - \tan^2 \frac{\pi}{\Omega} \right) \leq \gamma \leq \frac{1}{4} \left(\sin^2 \frac{\pi}{\Omega} + \delta^2 \cos^2 \frac{\pi}{\Omega} \right), \quad (4.14)$$

および

$$\frac{4 \sin^2 \frac{\pi}{\Omega}}{4 - (1 + \delta) \sin^2 \frac{\pi}{\Omega}} \left[\cos^2 \frac{\pi}{\Omega} + \left(\frac{3 - \delta}{4} \right)^2 \right] \leq \gamma \leq \frac{3 - \delta}{8} \left[\sin^2 \frac{\pi}{\Omega} + \left(\frac{1 + \delta}{3 - \delta} \right)^2 \cos^2 \frac{\pi}{\Omega} \right] \quad (4.15)$$

と得られる. これらの領域は図 4.3 の斜線部で示されている. 特に, 拡張系の緩和時間 $\tilde{\tau}_{\text{relax}}$ は設定 (a) $\delta = \delta'$ および設定 (b) $\delta' = 1$ のそれぞれに対して, とともに $|\delta| = 1$, $\gamma = (1/4) \sin(2\pi/\Omega)$ と選んだとき最小となる. つまり緩和時間の観点では, これらのパラメタの値が最良である. このとき拡張系の SLEM $\tilde{\eta}$ は

$$\tilde{\eta} = 1 - \frac{1}{2} \sin^2 \frac{\pi}{\Omega} - \frac{1}{4} \sin \frac{2\pi}{\Omega} \quad (4.16)$$

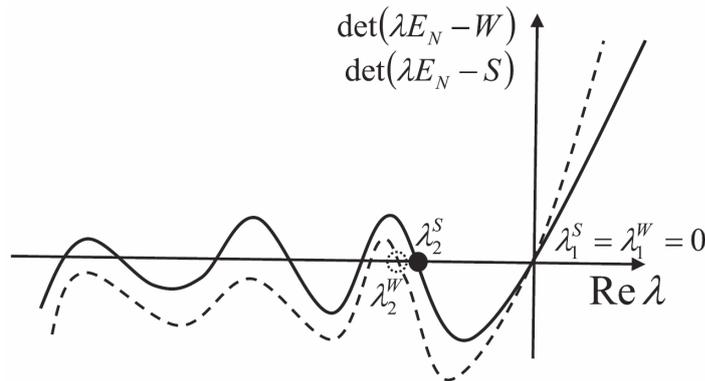


図 4.4 文献 [44] における, 詳細つり合い条件の破れた遷移確率行列 W の特性多項式のグラフ (破線), およびその additive reversibilization S の特性多項式のグラフ (実線). それぞれのグラフと横軸 $\text{Re}\lambda$ の交点が実固有値を与え, その交点は右から左に向かって最大固有値, 第 2 固有値, ... に対応している. 連続時間マスター方程式による議論であるため, 遷移確率行列 W およびその additive reversibilization S の最大固有値は $\lambda_1^S = \lambda_1^W = 0$ であることに注意. 白抜き破線丸印および塗りつぶし丸印はそれぞれ遷移確率行列 W およびその additive reversibilization S の第 2 固有値 λ_2^W, λ_2^S を表す. またこの図は遷移確率行列 W の第 2 固有値 λ_2^W および第 3 固有値 λ_3^W が実数である状況を描いている. 文献 [44] より引用.

となることから, 拡張系の緩和時間 $\tilde{\tau}_{\text{relax}}$ は状態数 Ω が十分大きいとき, 漸近的に $\tilde{\tau}_{\text{relax}} \simeq O(\Omega)$ と振る舞う. すなわち, 緩和時間の状態数 Ω 依存性が $O(\Omega^2)$ から $O(\Omega)$ へと質的に改善している. これは詳細つり合い条件の破れによってランダムウォークの拡散的な緩和ダイナミクスが弾道的なダイナミクスへと質的に変化したことを示している. 一方で, ε -flip の遷移確率 γ を SH_1 型, SH_2 型, SH_3 型, および TCV 型のいずれにとっても, 拡張系の緩和時間 $\tilde{\tau}_{\text{relax}}$ は前述の最小値よりも大きい. また着目系の緩和時間 τ_{relax} よりも大きくなってしまふ. この結果は, 詳細つり合い条件の破れがいつでも MCMC 法の性能を改善するとは限らないことを示唆している.

詳細つり合い条件の破れたマルコフ連鎖では一般に複素固有値が出現しうる. 今回の問題でも, 図 4.3 に示すとおり複素固有値が現れるパラメタ領域が存在する. SLEM を与える固有値が実数であっても緩和時間が改善することもあれば, それが複素数であっても緩和時間はむしろ悪化することもある. このことは, 複素固有値の出現が必ずしも緩和時間の改善に寄与するわけではないことを示している. また今回の場合, 拡張系の緩和時間の最小を与えるパラメタ (δ, γ) は SLEM を与える固有値が実数から複素数となる境界に存在することは興味深い. 実際このとき拡張系の遷移確率行列 \tilde{T} は対角化可能でなく, 2 次の Jordan 細胞をふたつもつ Jordan 標準形へと変形される. このように固有値が縮退し対角化可能性が破れるところで緩和時間が最小となることは, 詳細つり合い条件の破れと第 2 固有値の関係を示した Ichiki–Ohzeki による議論 [44] から示唆される. 文献 [44]

では詳細つり合い条件の破れによって第 2 固有値 λ_2 が小さくなる方向にシフトする一方で、第 3 固有値 λ_3 は大きくなる方向にシフトする (図 4.4). さらに詳細つり合い条件の破れを「大きく」すると、ふたつの固有値 λ_2, λ_3 は等しくなった後、共役な複素固有値となる。つまり第 2 固有値が実数のうちは、 $\lambda_2 = \lambda_3$ と縮退するとき第 2 固有値が最小となる。今回のランダムウォーク問題における最適なパラメタの値は、ちょうどこの状況を実現していると考えられる。

期待値の緩和時間

直前の小節で議論した緩和時間は、確率分布が収束する速さを定量的に表したものである。ここでは物理量の期待値の緩和ダイナミクスを議論する。

lifting parameter ε の導入により状態空間を拡張したので、物理量もまた拡張系の状態空間に対して定義する必要がある。目的分布からのサンプリングや期待値の計算など、通常の目的を考えれば lifting parameter ε の値に依存する物理量を考える必要はない。すなわち着目系の状態空間 I 上で定義される物理量を f とするとき、拡張系の状態空間 \tilde{I} 上で定義される物理量のうち $\hat{f}': (i, \varepsilon) \mapsto f_i$ なる形のものをだけ考慮すればよい。物理量 f' に対応するベクトル $\mathbf{f}' \in \mathbb{R}^{2\Omega}$ は、物理量 f に対応するベクトルを $\mathbf{f} \in \mathbb{R}^{\Omega}$ として $\mathbf{f}' = (\mathbf{f}, \mathbf{f})$ で与えられる。さらに簡単のため、以下では拡張系の初期分布 $\tilde{\mathbf{p}}^{(0)}$ は lifting parameter ε について一様とする。すなわち着目系の初期分布 $\mathbf{p}^{(0)} \in \mathbb{R}^{\Omega}$ に対して、拡張系の初期分布 $\tilde{\mathbf{p}}^{(0)} \in \mathbb{R}^{2\Omega}$ を

$$\tilde{\mathbf{p}}^{(0)} = \frac{1}{2}(\mathbf{p}^{(0)}, \mathbf{p}^{(0)}) \quad (4.17)$$

と与える。このとき、物理量 \hat{f}' の n ステップ後の期待値 $\langle f' \rangle_n$ は

$$\begin{aligned} \langle f' \rangle_n &\equiv \tilde{\mathbf{p}}^{(n)} \mathbf{f}'^{\top} \\ &= \langle f' \rangle_{\tilde{\pi}} + \sum_{k=2}^{[\Omega/2]+1} \left\{ \sum_{\varepsilon=\pm} (\tilde{\lambda}_k^{\varepsilon})^n \left[\sum_{\sigma=1}^{m_k} (\tilde{\mathbf{p}}^{(0)} \tilde{\mathbf{v}}_{k,\sigma}^{\varepsilon\top}) (\tilde{\mathbf{u}}_{k,\sigma}^{\varepsilon} \mathbf{f}'^{\top}) \right] \right\} \\ &= \langle f \rangle_{\pi} + \sum_{k=2}^{[\Omega/2]+1} \left[\sum_{\varepsilon=\pm} \frac{(\tilde{\lambda}_k^{\varepsilon})^n}{1 - (A_k^{\varepsilon})^2} \right] \left[\sum_{\sigma=1}^{m_k} (\mathbf{p}^{(0)} \mathbf{v}_{k,\sigma}^{\top}) (\mathbf{u}_{k,\sigma} \mathbf{f}^{\top}) \right] \end{aligned} \quad (4.18)$$

と与えられる。これより、期待値 $\langle f' \rangle_n$ は指数関数的に目的分布での期待値 $\langle f \rangle_{\pi}$ へと収束すること、およびその典型的な収束の速さは前小節で議論した拡張系の緩和時間 $\tilde{\tau}_{\text{relax}}$ で与えられることがわかる。なお、固有値 $\tilde{\lambda}_1^-$ に属する左固有ベクトル $\tilde{\mathbf{u}}_{1,1}$ は常に物理量 f' に対応するベクトル $\mathbf{f}' = (\mathbf{f}, \mathbf{f})$ と直交するため、物理量の緩和には現れない。よって図 4.3 において固有値 $\tilde{\lambda}_1^-$ が SLEM $\tilde{\eta}$ を与える領域でも、lifting parameter ε の値によらない物理量の緩和は確率分布の緩和より速い。また本章で議論しているランダムウォーク

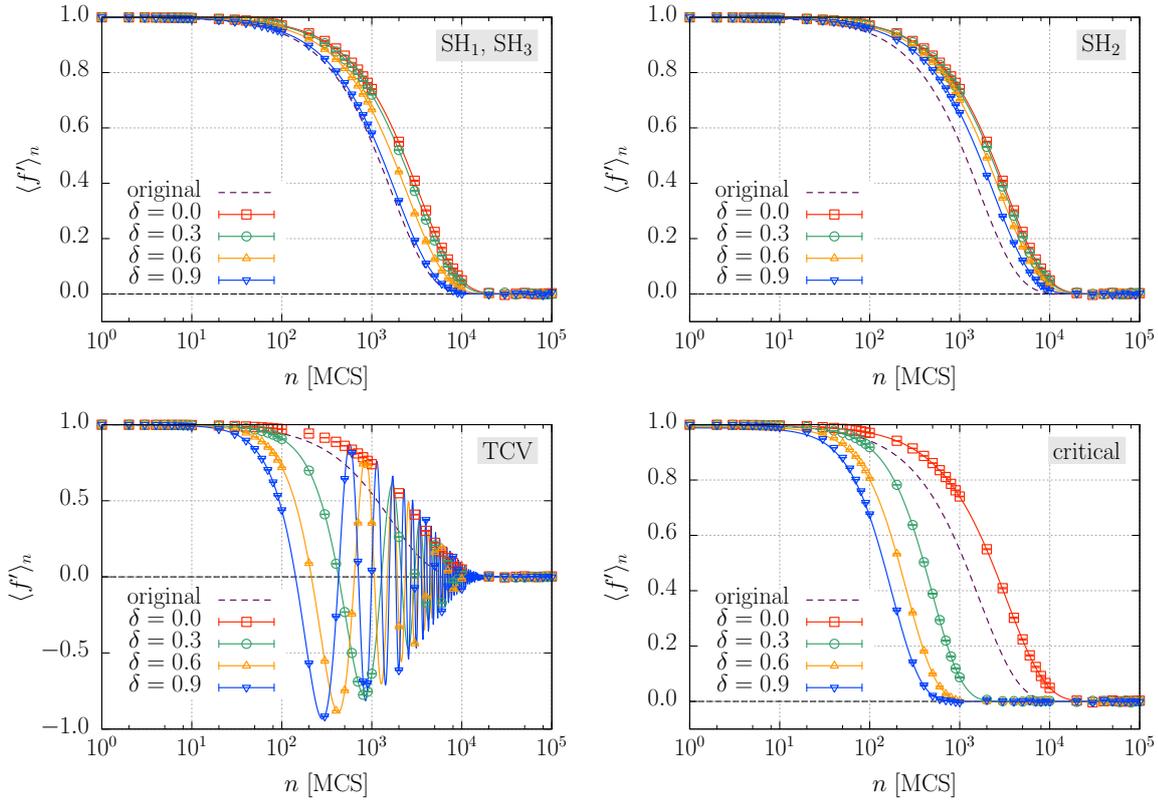


図 4.5 設定 (a) $\delta = \delta'$, および $\alpha = 1/2$ の下での, いろいろなパラメタ δ, γ の値に対する物理量の期待値 $\langle f' \rangle_n$ の時間発展. 数値計算には irreversible Metropolis–Hastings アルゴリズム [50] を用いた. また $\Omega = 128$ とした. 遷移確率 γ の値はそれぞれ $\gamma_{\text{SH}_1} = \gamma_{\text{SH}_3} = \alpha/2$, $\gamma_{\text{SH}_2} = 1/2$, $\gamma_{\text{TCV}} = 0$, $\gamma_{\text{critical}} = |\delta|(\alpha/2) \sin(2\pi/\Omega)$ である. 初期状態 $(i^{(0)}, \varepsilon^{(0)})$ は $i^{(0)} = 1$ とし, $\varepsilon^{(0)}$ は一様ランダムに ± 1 とした. 各プロットおよび不確かさは独立な 10^4 回の試行から評価した. original (紫破線) は着目系の遷移確率行列 \mathbf{T} による物理量の期待値 $\langle f' \rangle_n$ の時間発展 (解析解) を表す.

クの問題では一般に, 着目系における期待値 $\langle f \rangle_n$ の緩和に固有値 λ_k が寄与しないとき, 拡張系の期待値 $\langle f' \rangle_n$ の緩和に固有値 $\tilde{\lambda}_k^\pm$ は寄与しない.

設定 (a) $\delta = \delta'$ における物理量 $f': (i, \varepsilon) \mapsto \cos[2\pi(i-1)/\Omega]$ の期待値 $\langle f' \rangle_n$ の時間発展を図 4.5 に示す. この例では $\langle f' \rangle_\pi = 0$ へと収束していく様子がわかる. SH₁ 型, SH₂ 型および SH₃ 型の遷移確率ではいずれもパラメタ δ の値を大きくすることで緩和が加速しているものの, 着目系での緩和より遅くなっている. このように, lifting を用いたからといっていつでも緩和が加速するわけではない. TCV 型では $\delta \neq 0$ のとき減衰振動的な緩和がみられる. これは拡張系の固有値 $\tilde{\lambda}_2^\pm$ が複素固有値となっていることを反映している. しかしこのときでも減衰振動の包絡線から定まる拡張系の緩和時間 $\tilde{\tau}_{\text{relax}}$ は着目系の緩和時間 τ_{relax} より大きくなっており, 複素固有値の出現がむしろ緩和を遅くしている例となっている. これらはいずれもさきほどの SLEM の議論と整合している. 一方で解析

解から導かれた遷移確率 γ の最適値 $\gamma_{\text{critical}} = |\delta|(\alpha/2) \sin(2\pi/\Omega)$ を用いた緩和ダイナミクスは, $\delta \neq 0$ とすることで収束が着目系よりも 10 倍以上加速していることがみてとれる. このとき拡張系の遷移確率行列 $\tilde{\mathbf{T}}$ は対角化可能でなくなっている. 一般にある整数 $q \in \{2, 3, \dots, [(\Omega + 1)/2]\}$ を用いて遷移確率 γ が $\gamma_q \equiv |\delta|(\alpha/2) \sin \theta_q$ で与えられるとき, 拡張系の遷移確率行列 $\tilde{\mathbf{T}}$ は対角化可能でない. このとき付録 A の結果を用いると, (4.18) 式は

$$\begin{aligned} \langle f' \rangle_n = \langle f \rangle_{\boldsymbol{\pi}} + \sum_{k \neq 1, q} \left[\sum_{\varepsilon = \pm} \frac{(\tilde{\lambda}_k^\varepsilon)^n}{1 - (A_k^\varepsilon)^2} \right] \left[\sum_{\sigma=1}^{m_k} (\mathbf{p}^{(0)} \mathbf{v}_{k,\sigma}^\top) (\mathbf{u}_{k,\sigma} \mathbf{f}^\top) \right] \\ + (\tilde{\lambda}_q^n + \gamma_q n \tilde{\lambda}_q^{n-1}) \left[\sum_{\sigma=1,2} (\mathbf{p}^{(0)} \mathbf{v}_{q,\sigma}^\top) (\mathbf{u}_{q,\sigma} \mathbf{f}^\top) \right] \end{aligned} \quad (4.19)$$

と変更される. ここで $\tilde{\lambda}_q = \tilde{\lambda}_q^\pm$ である. 固有値 $\tilde{\lambda}_q$ の緩和モードに指数関数の前に n の 1 次の項が加わっている. 図 4.5 の例は $q = 2$ に対応しており, SLEM に対応する緩和モードが臨界緩和を示している. このときでも緩和時間は前小節で議論したとおり SLEM で決まり, 緩和ダイナミクスが質的に改善している.

性能指数の最悪評価

詳細つり合い条件の破れが漸近分散に与える効果をみていこう. ここでは, 詳細つり合い条件をみたす場合と破る場合で漸近分散と分散の比 (performance index [45], 性能指数) の最悪評価を与え, それぞれを比較する. 詳細な計算は付録 C に記す.

まず, 詳細つり合い条件をみたす 1次元ランダムウォーク (着目系) における性能指数の最大値 Γ は $0 < \alpha \leq 1$, $\Omega \in \mathbb{N}$ に対して

$$\Gamma = \frac{1}{\alpha \sin^2 \frac{\pi}{\Omega}} - 1 \quad (4.20)$$

で与えられる. 特に状態数 Ω が十分大きいとき, $\Gamma = O(\Omega^2)$ と振る舞う.

一方 lifting を用いて詳細つり合い条件を破った 1次元ランダムウォーク (拡張系) における漸近分散では, lifting parameter ε の値によらない物理量 f' についてのみ考察する. このとき拡張系の漸近分散 $v(f', \tilde{\boldsymbol{\pi}}, \tilde{\mathbf{T}})$ は

$$v(f', \tilde{\boldsymbol{\pi}}, \tilde{\mathbf{T}}) = \sum_{k=2}^{[\Omega/2]+1} \left[\sum_{\varepsilon = \pm} \frac{1}{1 - (A_k^\varepsilon)^2} \frac{1 + \tilde{\lambda}_k^\varepsilon}{1 - \tilde{\lambda}_k^\varepsilon} \right] \left[\sum_{\sigma=1}^{m_k} (\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^\top) (\mathbf{u}_{k,\sigma} \mathbf{f}^\top) \right] \quad (4.21)$$

で与えられることが示せる. これより性能指数の最大値 $\tilde{\Gamma}$ は $0 < \alpha \leq 1$, $\Omega \in \mathbb{N}$, $|\delta| \leq 1$, $|\delta'| \leq 1$, $0 \leq \gamma \leq 1 - (2 + \delta - \delta')\alpha/4$ に対して

$$\tilde{\Gamma} = \frac{4}{\alpha} \frac{2 + \delta - \delta'}{(2 + \delta - \delta')^2 \sin^2 \frac{\pi}{\Omega} + \phi(\delta + \delta')^2 \cos^2 \frac{\pi}{\Omega}} - 1 \quad (4.22)$$

で与えられる。ここで

$$\phi \equiv \frac{\sin^2 \frac{\pi}{\Omega}}{\sin^2 \frac{\pi}{\Omega} + \frac{2\gamma}{(2+\delta-\delta')\alpha}} \quad (4.23)$$

である。以上から、不等式 $\tilde{\Gamma} \leq \Gamma$ が成り立つのは

$$\gamma \leq \frac{\alpha}{2} \left(1 + \frac{\delta - \delta'}{2} \right) \left[\frac{(\delta + \delta')^2}{4 - (\delta + \delta')^2} \cos^2 \frac{\pi}{\Omega} - \sin^2 \frac{\pi}{\Omega} \right] \quad (4.24)$$

のときであることが導かれる。この結果を $\alpha = 1/2$ の下で、設定 (a) $\delta = \delta'$ および設定 (b) $\delta' = 1$ のそれぞれの場合で検討する。

設定 (a) $\delta = \delta'$ のとき, (4.24) 式は

$$\gamma \leq \frac{1}{4} \left(\delta^2 \cos^2 \frac{\pi}{\Omega} - \sin^2 \frac{\pi}{\Omega} \right) \quad (4.25)$$

となる。これを図示して図 4.6-(a) を得る。また設定 (b) $\delta' = 1$ のとき, (4.24) 式は

$$\gamma \leq \frac{1 + \delta}{8} \left(\frac{1 + \delta}{3 - \delta} \cos^2 \frac{\pi}{\Omega} - \sin^2 \frac{\pi}{\Omega} \right) \quad (4.26)$$

となる。これを図示して図 4.6-(b) を得る。

性能指数の最悪評価について設定 (a) と (b) で共通する特徴として

- パラメタ δ がある値より小さいと, どんな遷移確率 γ をとっても性能指数の最悪評価が改善しない
- パラメタ δ の値が大きいほど, また遷移確率 γ の値が小さいほど, 改善の度合いが大きい
- 改善の度合いが最大となるのは, $(|\delta|, \gamma) = (1, 0)$ である*1.
- SH₁, SH₂ および SH₃ 型の遷移確率 γ では性能指数の最悪評価を改善できない

ことが挙げられる。

4.3 本章のまとめ

本章では MCMC 法の性能評価の具体例として, 1次元ランダムウォークの遷移確率行列を解析した。詳細つり合い条件をみたす場合と, ねじれ詳細つり合い条件を用いた lifting によって詳細つり合い条件を破った場合のそれぞれについて1次元ランダムウォークの遷移確率行列を構成し, 厳密に対角化した。得られた固有値, 固有ベクトルか

*1 $\gamma = 0$ は TCV 型に対応し, 拡張系全体のエルゴード条件をみたさない。しかし目的のサンプリングをする分には問題ない。

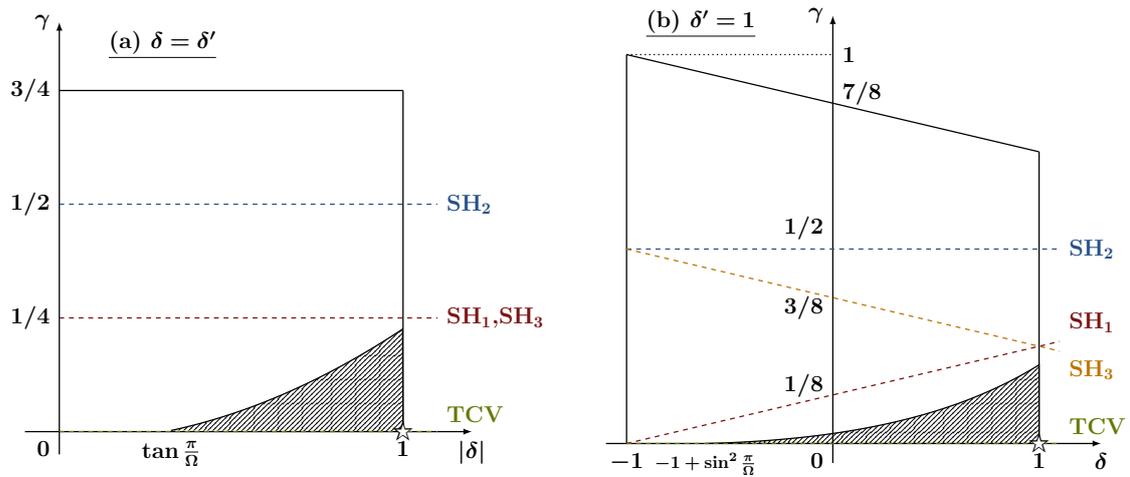


図 4.6 詳細つり合い条件を破ることで性能指数の最悪評価が改善するパラメタ領域 (δ, γ) を斜線で示した. 星印 (\star) は改善の度合いが最も大きいパラメタ (δ, γ) の位置を示す.

ら緩和時間, 漸近分散を評価し, 詳細つり合い条件の破れが 1 次元ランダムウォークに対する MCMC 法の性能に及ぼす効果を詳細に調べた. 今回の解析では, 詳細つり合い条件の破れの度合いはいくつかのパラメタによって特徴づけられた. また 2 通りの設定に限定して解析を行った. その結果, 詳細つり合い条件を破ることで 1 次元ランダムウォークの拡散的な緩和ダイナミクスが質的に変化し, 緩和時間の状態数 Ω 依存性を $O(\Omega^2)$ から $O(\Omega)$ まで改善できることが示された. この結果は先行研究で異なる解析手法から知られていた結果と整合している. さらに漸近分散の解析から, 性能指数の最悪評価もまた詳細つり合い条件の破れによって改善しうることが示した. 特に,

- パラメタ δ の値が大きいほど性能改善の傾向が大きい. すなわち, 詳細つり合い条件を破るときは「大きく」破ったほうがその効果が大きく得られる
- 着目系である可逆なランダムウォークと比較して, 緩和時間も性能指数の最悪評価もともに改善できるパラメタ領域が存在する. その一方で, 一般に機械的に構成できる遷移確率 (SH₁ 型, SH₂ 型, SH₃ 型, TCV 型) では性能改善に至らない
- 緩和時間が最も改善するパラメタの値と, 性能指数の最悪評価が最も改善するパラメタの値がそれぞれ異なる

ことがわかった. これらの結果は 1 次元ランダムウォークという特定のモデルに対する結果であり, どこまで一般的に成り立つのかは別途検討する必要がある. それでも, lifting による詳細つり合い条件の破れの効果を理解する上で解析的に得られた結果として, 今後の議論の足がかりになると期待する.

第 5 章

Irreversible simulated tempering アルゴリズム

シミュレーテッドテンパリング (simulated tempering, ST) は、レプリカ交換法やマルチカノニカル法に代表される拡張アンサンブル法のひとつである。ふたつのグループがそれぞれ独立に提案し [10, 11], その後スピングラスやタンパク質の折り畳み問題など様々な分野で応用されてきた [60, 61, 62, 63, 64, 65, 66]. また, シミュレーテッドテンパリング自体の性能改善も長らく研究されてきた [67, 68, 69, 70, 71]. その多くは詳細つり合い条件の枠組みの中で改善を試みたものであるが, 文献 [71] では Suwa-Todo アルゴリズムと組み合わせることで詳細つり合い条件を破り, 緩和時間が定数倍程度改善された例が報告されている。しかしながら, 逆温度の緩和が拡散的なダイナミクスから定性的に改善された例はこれまで報告されていなかった。

本博士論文では前章までに, lifting による詳細つり合い条件の破れが 1 次元ランダムウォークの緩和ダイナミクスを質的に改善すること, および性能指数の最悪評価を改善することを解析的に示した。シミュレーテッドテンパリングにおける逆温度空間は 1 次元ランダムウォークの状態空間と類似しており, lifting との相性がよいと考えられる。このことから, 本章ではシミュレーテッドテンパリングの逆温度更新プロセスに lifting の手法を応用し, 詳細つり合い条件を破る新たなアルゴリズムを提案する。本章で提案するアルゴリズムを irreversible simulated tempering (IST) アルゴリズムと呼ぶ。また IST アルゴリズムを用いて重み因子を推定する手法もあわせて提案する。

本章の内容は文献 [72] および [73] として出版されている。

5.1 シミュレーテッドテンパリング

本節では記号の定義も含めて、詳細つり合い条件をみたすシミュレーテッドテンパリングの概要を説明する。

5.1.1 問題設定

着目する系の状態変数を \mathbf{X} 、エネルギー関数を $E(\mathbf{X})$ とし、目的分布を逆温度 β の Gibbs–Boltzmann 分布で与える:

$$P_\beta(\mathbf{X}) = \frac{1}{Z(\beta)} \exp[-\beta E(\mathbf{X})]. \quad (5.1)$$

ここで

$$Z(\beta) \equiv \sum_{\mathbf{X}} \exp[-\beta E(\mathbf{X})] \quad (5.2)$$

は系の分配関数である。状態変数 \mathbf{X} は離散変数でも連続変数でもよく、 \mathbf{X} が連続変数の場合は (5.2) 式に現れる和を積分に置き換えればよい。

シミュレーテッドテンパリングでは系の状態変数 \mathbf{X} だけでなく、逆温度 β もまた確率変数として扱う。確率変数としての逆温度 β は、あらかじめ用意した R 個の値 $\beta_1, \beta_2, \dots, \beta_R$ を確率的にとるものとする。これにより逆温度も含めた系の状態は、着目系の状態変数 \mathbf{X} と逆温度 β の組 (\mathbf{X}, β) で記述される。さらにこれにともなって、目的分布を以下のとおり拡張する。すなわち、系が状態 (\mathbf{X}, β_r) をとる確率 $P_{\text{ST}}(\mathbf{X}, \beta_r)$ を

$$P_{\text{ST}}(\mathbf{X}, \beta_r) = \frac{1}{Z_{\text{ST}}} \exp[-\beta_r E(\mathbf{X}) + g_r] \quad (5.3)$$

と設定する。ここで実数 g_r は逆温度 β_r にのみ依存する重み因子であり、

$$Z_{\text{ST}} \equiv \sum_{r=1}^R \sum_{\mathbf{X}} \exp[-\beta_r E(\mathbf{X}) + g_r] = \sum_{r=1}^R Z(\beta_r) e^{g_r} \quad (5.4)$$

は規格化定数である。

与えられた逆温度 β_r に対して、着目系の状態が \mathbf{X} である条件つき確率 $P_{\text{ST}}(\mathbf{X}|\beta_r)$ を計算すると、

$$\begin{aligned} P_{\text{ST}}(\mathbf{X}|\beta_r) &\equiv \frac{P_{\text{ST}}(\mathbf{X}, \beta_r)}{\sum_{\mathbf{X}} P_{\text{ST}}(\mathbf{X}, \beta_r)} \\ &= \frac{1}{Z(\beta_r)} \exp[-\beta_r E(\mathbf{X})] = P_{\beta_r}(\mathbf{X}) \end{aligned} \quad (5.5)$$

となる。つまり (\mathbf{X}, β) 空間全体でサンプリングした中から逆温度が $\beta = \beta_r$ であるサンプルだけを抽出したとき、状態変数 \mathbf{X} は重み因子 g_r の値によらず、逆温度 β_r の Gibbs–Boltzmann 分布にしたがう。この意味で、重み因子 g_r の値の選び方はサンプリングの原理的な正しさには影響を与えない。一方、逆温度の周辺分布 $P_{\text{ST}}(\beta)$ は

$$P_{\text{ST}}(\beta_r) \equiv \sum_{\mathbf{X}} P_{\text{ST}}(\mathbf{X}, \beta_r) = \frac{Z(\beta_r)e^{g_r}}{Z_{\text{ST}}} \quad (5.6)$$

となり、重み因子 g_r の値に依存する。特に重み因子を $g_r = -\ln Z(\beta_r)$ と選んだとき、逆温度の周辺分布は一様になり、様々な逆温度をくまなくめぐることができると考えられる。一般に分配関数 $Z(\beta_r)$ を厳密に求めることは困難であるが、事前に近似値を見積もっておきシミュレーテッドテンパリングの重み因子に用いることで、逆温度の周辺分布をほぼ一様にできるだろう。このような重み因子の選び方はレプリカ交換法の考え方と同じように、様々な逆温度を十分にめぐることによって低温での遅い緩和の問題を改善するという発想に基づいている。重み因子 g_r の推定手法は5.3節で議論する。

5.1.2 詳細つり合い条件をみたすシミュレーテッドテンパリング アルゴリズム

シミュレーテッドテンパリングの状態更新プロセスは、

- (1) 逆温度 β_r を固定した下での、系の状態変数 \mathbf{X} の更新
- (2) 系の状態変数 \mathbf{X} を固定した下での、逆温度 β_r の更新

の2種類からなる。すなわち、系の状態変数 \mathbf{X} と逆温度 β_r を1ステップで同時に更新するプロセスは考えないということである。

更新プロセス (1) では、Metropolis–Hastings アルゴリズムやクラスターアルゴリズムなど、これまで用いられてきた通常の MCMC 法によって状態更新を行えばよい^{*1}。

更新プロセス (2) では、詳細つり合い条件をみたす Metropolis–Hastings アルゴリズムを用いることにする。状態 (\mathbf{X}, β_r) から状態 (\mathbf{X}, β_l) への遷移確率を $T(\mathbf{X}, \beta_l | \mathbf{X}, \beta_r)$ と書く。また遷移確率 T を提案確率 q と採択確率 W の積に分解し、

$$T(\mathbf{X}, \beta_l | \mathbf{X}, \beta_r) = q(\mathbf{X}, \beta_l | \mathbf{X}, \beta_r)W(\mathbf{X}, \beta_l | \mathbf{X}, \beta_r) \quad (5.7)$$

と表す。このとき、Metropolis–Hastings 型の採択確率 W は

$$W(\mathbf{X}, \beta_l | \mathbf{X}, \beta_r) = \min \left[1, \frac{q(\mathbf{X}, \beta_r | \mathbf{X}, \beta_l) P_{\text{ST}}(\mathbf{X}, \beta_l)}{q(\mathbf{X}, \beta_l | \mathbf{X}, \beta_r) P_{\text{ST}}(\mathbf{X}, \beta_r)} \right] \quad (5.8)$$

^{*1} この部分に Suwa–Todo アルゴリズムや irreversible Metropolis–Hastings アルゴリズムを用いることもできる。本博士論文では議論しない。

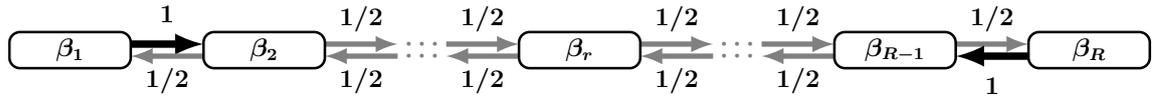


図 5.1 シミュレーテッドテンパリングにおける逆温度空間での提案確率 $q_{r,l}$ のグラフ表現. 各矢印は 1 ステップで許される提案を表し, その上 (下) の数字が提案確率を表す.

アルゴリズム 5.1: シミュレーテッドテンパリングアルゴリズム

(0) 適当に初期条件 $(\mathbf{X}^{(0)}, \beta^{(0)})$ を与える.

(1) [系の状態 \mathbf{X} の更新]

逆温度 β を固定した下で, 系の状態 \mathbf{X} を更新する. Metropolis–Hastings アルゴリズムや熱浴法, クラスタアルゴリズムなど, 従来の MCMC 法を用いる.

(2) [逆温度 β の更新]

系の状態 \mathbf{X} を固定した下で, 逆温度 β を (詳細つり合い条件をみたく) Metropolis–Hastings アルゴリズムで更新する:

- (a) 現在の状態が (\mathbf{X}, β_r) であるとする. このとき, 確率 $q_{r,l}$ で次の逆温度の候補 β_l を提案する.
- (b) 確率 $W(\mathbf{X}, \beta_l | \mathbf{X}, \beta_r)$ で状態 (\mathbf{X}, β_l) を採択する. もし棄却されたとき, 現在の状態 (\mathbf{X}, β_r) を次の状態として採択する.

(3) 手順 (1), (2) を適当な回数ずつ交互にそれぞれ繰り返す.

で与えられる. 簡単のため, 逆温度としてとりうる値は $\beta_1 < \beta_2 < \dots < \beta_R$ と並べられているとし, 提案確率 q は系の状態変数 \mathbf{X} によらず $q_{r,l} \equiv q(\mathbf{X}, \beta_l | \mathbf{X}, \beta_r)$ と書けるものとする. 特に提案確率 $q_{r,l}$ として, 最近接の逆温度に等確率で提案するものを選ぶことにする (図 5.1). 以上をふまえて, 詳細つり合い条件をみたくシミュレーテッドテンパリングアルゴリズムはアルゴリズム 5.1 のとおり与えられる.

シミュレーテッドテンパリングアルゴリズムを実行することで, マルコフ連鎖 $\{(\mathbf{X}^{(n)}, \beta^{(n)})\}_{n=0,1,2,\dots}$ が得られる. これを用いて, 逆温度 β_r における物理量 f の期待値 $\langle f \rangle_{\beta_r}$ は

$$\langle f \rangle_{\beta_r} \equiv \sum_{\mathbf{X}} f(\mathbf{X}) P_{\beta_r}(\mathbf{X}) \simeq \frac{1}{N_r} \sum_{m=1}^n f(\mathbf{X}^{(m)}) \delta(\beta^{(m)}, \beta_r) \quad (5.9)$$

と見積もることができる。ここで n は期待値評価に用いたサンプル数, $\delta(\cdot, \cdot)$ はクロネッカーのデルタであり,

$$N_r \equiv \sum_{m=1}^n \delta(\beta^{(m)}, \beta_r) \quad (5.10)$$

は逆温度 β_r の訪問回数である。

5.2 Irreversible simulated tempering アルゴリズム

前節で説明した従来のシミュレーテッドテンパリングでは, 着目系の状態変数 \mathbf{X} を固定した下での逆温度空間の状態遷移図が1次元ランダムウォークの状態遷移図と類似している。つまり逆温度 β は1次元空間上をランダムウォークする。第4章でみたとおり, 詳細つり合い条件をみたすランダムウォークは拡散的 (diffusive) であることが知られている。すなわち, 緩和時間が状態数 Ω に対して漸近的に $O(\Omega^2)$ と振る舞う。さらに第4章では, lifting による詳細つり合い条件の破れが1次元ランダムウォークにおける緩和時間の状態数依存性を $O(\Omega)$ まで改善し, 弾道的 (ballistic) な緩和ダイナミクスへと質的に変化させることを解析的に示した。このことから, シミュレーテッドテンパリングにおける逆温度の緩和ダイナミクスにも詳細つり合い条件の破れが同様の効果をもたらすのではないかと期待される。以下ではねじれ詳細つり合い条件を用いた lifting によってシミュレーテッドテンパリングの逆温度更新プロセスに詳細つり合い条件の破れを導入し, irreversible simulated tempering アルゴリズムを構成する。

5.2.1 問題設定の拡張

まず lifting の手法にそって系の状態空間を拡張する。従来のシミュレーテッドテンパリングを記述する確率変数 \mathbf{X} , β に加えて, 系に lifting parameter $\varepsilon \in \{+, -\}$ を導入する。これにより系の状態空間は2倍に拡張される。拡張系の状態は確率変数の組 $(\mathbf{X}, \beta_r, \varepsilon)$ で記述される。これにともなって, シミュレーテッドテンパリングの目的分布 $P_{\text{ST}}(\mathbf{X}, \beta)$ も拡張される。状態 $(\mathbf{X}, \beta_r, \varepsilon)$ が現れる確率 $P_{\text{IST}}(\mathbf{X}, \beta_r, \varepsilon)$ を

$$P_{\text{IST}}(\mathbf{X}, \beta_r, \varepsilon) = \frac{1}{2Z_{\text{ST}}} \exp[-\beta_r E(\mathbf{X}) + g_r] \quad (5.11)$$

と設定する。これは lifting parameter ε についての周辺分布が一様であり, lifting parameter ε を周辺化した (\mathbf{X}, β_r) の分布が拡張前の目的分布 $P_{\text{ST}}(\mathbf{X}, \beta_r)$ となるように設定したものである。

ここでは特に, ねじれ詳細つり合い条件を用いた lifting をシミュレーテッドテンパリングの逆温度更新プロセスにのみ導入する。言い換えれば, 系の状態変数 \mathbf{X} の更

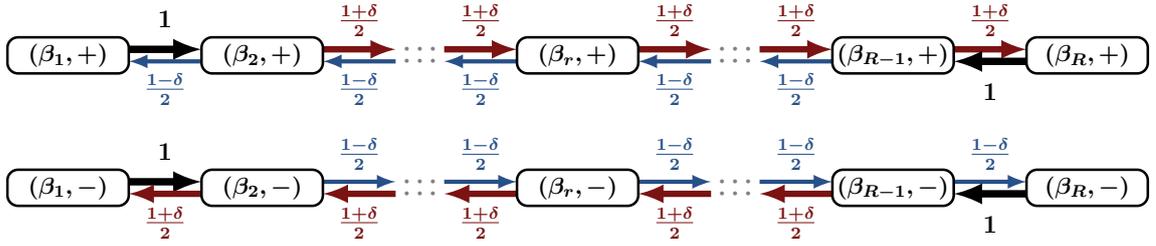


図 5.2 irreversible simulated tempering アルゴリズムにおける (β, ε) 空間での提案確率 $q_{r,l}^{(\varepsilon)}$ のグラフ表現. 各矢印は 1 ステップで許される提案を表し, その上 (下) の数字が提案確率を表す.

新は従来の手法と同様に行い, 逆温度 β と lifting parameter ε の更新に irreversible Metropolis–Hastings アルゴリズムを用いるということである.

逆温度更新におけるねじれ詳細つり合い条件は

$$T(\mathbf{X}, \beta_l, + | \mathbf{X}, \beta_r, +) P_{\text{IST}}(\mathbf{X}, \beta_r, +) = T(\mathbf{X}, \beta_r, - | \mathbf{X}, \beta_l, -) P_{\text{IST}}(\mathbf{X}, \beta_l, -) \quad (5.12)$$

で与えられる. ここで $T(\mathbf{X}, \beta_l, \varepsilon | \mathbf{X}, \beta_r, \varepsilon)$ は状態 $(\mathbf{X}, \beta_r, \varepsilon)$ から状態 $(\mathbf{X}, \beta_l, \varepsilon)$ への遷移確率を表す. これまでと同様に, 遷移確率 T を提案確率 q と採択確率 W の積に分解する:

$$T(\mathbf{X}, \beta_l, \varepsilon | \mathbf{X}, \beta_r, \varepsilon) = q(\mathbf{X}, \beta_l, \varepsilon | \mathbf{X}, \beta_r, \varepsilon) W(\mathbf{X}, \beta_l, \varepsilon | \mathbf{X}, \beta_r, \varepsilon). \quad (5.13)$$

このとき, ねじれ詳細つり合い条件をみたす採択確率 W を Metropolis–Hastings 型で与えると

$$W(\mathbf{X}, \beta_l, \varepsilon | \mathbf{X}, \beta_r, \varepsilon) = \min \left[1, \frac{q(\mathbf{X}, \beta_r, -\varepsilon | \mathbf{X}, \beta_l, -\varepsilon) P_{\text{ST}}(\mathbf{X}, \beta_l)}{q(\mathbf{X}, \beta_l, \varepsilon | \mathbf{X}, \beta_r, \varepsilon) P_{\text{ST}}(\mathbf{X}, \beta_r)} \right] \quad (5.14)$$

を得る. 以下では系の状態変数 \mathbf{X} に依存しない提案確率のみを考えることにし, $q_{r,l}^{(\varepsilon)} \equiv q(\mathbf{X}, \beta_l, \varepsilon | \mathbf{X}, \beta_r, \varepsilon)$ と略記する. $q_{r,l}^{(\varepsilon)}$ は具体的に

$$q_{1,2}^{(\varepsilon)} = q_{R,R-1}^{(\varepsilon)} = 1, \quad (5.15)$$

$r = 2, 3, \dots, R-1$ に対して

$$q_{r,r\pm 1}^{(\varepsilon)} = \frac{1 \pm \delta \varepsilon}{2}, \quad (5.16)$$

それ以外はゼロで与えられる (図 5.2). ここで δ は詳細つり合い条件の破れを特徴づけるパラメタであり, 不等式 $|\delta| \leq 1$ をみたす. 絶対値 $|\delta|$ が大きいほど詳細つり合い条件の破れの度合いが大きく, 特に $\delta = 0$ のとき詳細つり合い条件が回復する. つまり $\delta = 0$ として lifting parameter ε を無視すると, 前節で記述した詳細つり合い条件をみたすシミュレーテッドテンパリングと本質的に同じものとなる. $\delta \neq 0$ では詳細つり合い条件が破れ, たとえば $\delta > 0$ のとき図 5.2 において時計回りに確率の流れができると期待される.

アルゴリズム 5.2: irreversible simulated tempering アルゴリズム

(0) 適当に初期状態 $(\mathbf{X}^{(0)}, \beta^{(0)}, \varepsilon^{(0)})$ を与える.

(1) [系の状態 \mathbf{X} の更新]

逆温度 β と lifting parameter ε を固定した下で, 系の状態 \mathbf{X} を更新する. Metropolis–Hastings アルゴリズムや熱浴法, クラスタールゴリズムなど, 従来の MCMC 法を用いる.

(2) [逆温度 β と lifting parameter ε の更新]

系の状態 \mathbf{X} を固定した下で, 逆温度 β と lifting parameter ε を irreversible Metropolis–Hastings アルゴリズムで更新する:

(a) 現在の状態が $(\mathbf{X}, \beta_r, \varepsilon)$ であるとする. このとき, 確率 $q_{r,l}^{(\varepsilon)}$ で次の逆温度の候補 β_l を提案する.

(b) 確率 $W_{r,l}^{(\varepsilon)}$ で状態 $(\mathbf{X}, \beta_l, \varepsilon)$ を次の状態として採択する.

(c) 手順 (b) で状態更新が棄却されたとき, 確率 $\lambda_r^{(\varepsilon)}$ で状態 $(\mathbf{X}, \beta_r, -\varepsilon)$ を次の状態として採択する. ここで

$$\lambda_r^{(\varepsilon)} \equiv \frac{\Lambda_r^{(r)}}{1 - \sum_{l \neq r} q_{r,l}^{(\varepsilon)} W_{r,l}^{(\varepsilon)}}, \quad (5.17)$$

$$\Lambda_r^{(\varepsilon)} \equiv \max \left[0, \varepsilon \sum_{\varepsilon' = \pm} \sum_{l \neq r} \varepsilon' q_{r,l}^{(-\varepsilon')} W_{r,l}^{(-\varepsilon')} \right] \quad (5.18)$$

である. ここでも棄却されたとき, 現在の状態 $(\mathbf{X}, \beta_r, \varepsilon)$ を次の状態として採択する.

(3) 手順 (1), (2) を適当な回数ずつ交互にそれぞれ繰り返す.

5.2.2 Irreversible simulated tempering アルゴリズム

以下では採択確率 $W(\mathbf{X}, \beta_l, \varepsilon | \mathbf{X}, \beta_r, \varepsilon)$ を $W_{r,l}^{(\varepsilon)}$ と略記する. このとき irreversible simulated tempering (IST) アルゴリズムはアルゴリズム 5.2 のとおり与えられる.

IST アルゴリズムを実行することでマルコフ連鎖 $\{(\mathbf{X}^{(n)}, \beta^{(n)}, \varepsilon^{(n)})\}_{n=0,1,2,\dots}$ が得られる. lifting parameter ε はこちらで補助的に導入したものであり, 特に目的分布

$P_{\text{IST}}(\mathbf{X}, \beta_r, \varepsilon)$ の ε を周辺化した分布が従来のシミュレーテッドテンパリングにおける目的分布 $P_{\text{ST}}(\mathbf{X}, \beta_r)$ に一致する. このことから, 逆温度 β_r における物理量 f の期待値評価では lifting parameter ε の値を気にせず (5.9) 式を用いればよい.

5.3 重み因子の推定手法

従来のシミュレーテッドテンパリングや IST アルゴリズムを実行するにあたって, 重み因子 g_r の値を事前に与える必要がある. 重み因子 g_r の選び方はアルゴリズムの性能に大きな影響を及ぼす. 特に $g_r = -\ln Z(\beta_r)$ と選んだとき, 逆温度の周辺分布が一様となり, 効率的なサンプリングが期待できる. 一方で分配関数 $Z(\beta_r)$ の正確な値を事前を知ることは一般に難しい. そのため分配関数の近似値をなんらかの手法で評価し, 重み因子として用いることが多い [74, 75, 76, 77]. 本節ではそのような手法のひとつとして, 重み因子の逐次更新 (on-the-fly) による推定手法を考察する. この手法自体は文献 [77] で詳細つり合い条件をみたす従来のシミュレーテッドテンパリングに対し提案されている. 本博士論文ではこの手法を IST アルゴリズムと組み合わせたものを議論する. 以下で議論する重み因子推定手法を on-the-fly weight estimation と呼ぶ.

Gibbs–Boltzmann 分布 $P_\beta(\mathbf{X})$ による期待値を $\langle \dots \rangle_\beta$ と表し, 熱容量を $C(\beta) \equiv \langle E^2 \rangle_\beta - \langle E \rangle_\beta^2$ と書く. このとき統計力学でよく知られた関係式

$$\langle E \rangle_\beta = -\frac{\partial}{\partial \beta} \ln Z(\beta), \quad (5.19)$$

$$C(\beta) = \frac{\partial^2}{\partial \beta^2} \ln Z(\beta) \quad (5.20)$$

から, 逆温度 β と $\beta' \equiv \beta + d\beta$ に対して

$$-\ln Z(\beta') = -\ln Z(\beta) + \langle E \rangle_\beta d\beta - \frac{1}{2} C(\beta) (d\beta)^2 + o((d\beta)^2) \quad (5.21)$$

および

$$-\ln Z(\beta) = -\ln Z(\beta') - \langle E \rangle_{\beta'} d\beta - \frac{1}{2} C(\beta') (d\beta)^2 + o((d\beta)^2) \quad (5.22)$$

が成り立つ. 2 式を組み合わせると,

$$-\ln Z(\beta') = -\ln Z(\beta) + \frac{\langle E \rangle_\beta + \langle E \rangle_{\beta'}}{2} d\beta + \frac{C(\beta') - C(\beta)}{4} (d\beta)^2 + o((d\beta)^2) \quad (5.23)$$

を得る. ここで $(d\beta)^2$ の係数について, 熱容量 $C(\beta)$ が C^1 級ならば $C(\beta') - C(\beta) = O(d\beta)$ であるので, 結局

$$-\ln Z(\beta') = -\ln Z(\beta) + \frac{\langle E \rangle_\beta + \langle E \rangle_{\beta'}}{2} d\beta + o((d\beta)^2) \quad (5.24)$$

を得る. この式に現れるエネルギー期待値 $\langle E \rangle_{\beta}$ を IST アルゴリズムで評価しつつ重み因子を逐次更新するのが on-the-fly weight estimation のアイデアである.

on-the-fly weight estimation における n ステップ後の状態を $(\mathbf{X}^{(n)}, \beta^{(n)}, \varepsilon^{(n)})$ とする. また n ステップまでに逆温度 β_r が現れた回数 $N_r^{(n)}$ を

$$N_r^{(n)} \equiv \sum_{m=1}^n \delta(\beta^{(m)}, \beta_r) \quad (5.25)$$

とし, 逆温度 β_r のヒストグラムと呼ぶ. ここで $\delta(\cdot, \cdot)$ はクロネッカーのデルタである. このとき, 逆温度 β_r でのエネルギー期待値 $\langle E \rangle_{\beta_r}$ を

$$\langle E \rangle_{\beta_r} \simeq \bar{E}_r^{(n)} \equiv \frac{1}{N_r^{(n)}} \sum_{m=1}^n E(\mathbf{X}^{(m)}) \delta(\beta^{(m)}, \beta_r) \quad (5.26)$$

と推定し, n ステップ後の重み因子 $g_r^{(n)}$ を

$$g_r^{(n)} = g_{r+1}^{(n-1)} - \frac{\bar{E}_r^{(n)} + \bar{E}_{r+1}^{(n)}}{2} (\beta_{r+1} - \beta_r) \quad (5.27)$$

とする. 以下では簡単のため, 逆温度を $\beta_1 < \beta_2 < \dots < \beta_R$ と並べ替えて説明する. on-the-fly weight estimation による具体的な重み因子推定の手続きはアルゴリズム 5.3 のとおりである.

重み因子 g_r はその差のみが重要であるため, その値には定数分の一様な不定性がある. 次章で on-the-fly weight estimation を実行する際には, $g_1 = 0$ に固定して残りの重み因子 g_2, g_3, \dots, g_R の値を定めている. 系の自由エネルギーの値自体を決定したい場合は, 高温極限など自明に分配関数の値を決定できる逆温度を含んでシミュレーションを行えばよい.

重み因子推定を終了する基準をいくつか挙げる. たとえば逆温度空間を k 往復して得られる重み因子の推定値を g_r^k としたとき, 1 往復での修正誤差

$$\Delta^{(k)} \equiv \sum_{r=1}^R |g_r^{k+1} - g_r^k| \quad (5.28)$$

が許容する誤差 ϵ より小さくなるまで更新を行うことが考えられる. 誤差 ϵ はユーザーが適当に与える. また十分更新を行った後に逆温度のヒストグラムが一様になることが期待されるので, Wang-Landau 法 [78, 79] の停止条件と同じように, 逆温度のヒストグラム $N_r^{(n)}$ の最小値が平均の $x\%$ となるまで更新を行うことが考えられる. ただし逆温度のヒストグラム $N_r^{(n)}$ は $n \rightarrow \infty$ の極限で $Z(\beta_r)e^{g_r}$ に比例するため, 扱う系やそのサイズを考慮して x の値を適切に設定する必要がある [79].

アルゴリズム 5.3: on-the-fly weight estimation

- (1) すべての $r = 1, 2, \dots, R$ について $g_r = 0$ とする.
- (2) 初期逆温度 $\beta^{(0)} = \beta_R$ と, 適当な初期状態 $\mathbf{X}^{(0)}, \varepsilon^{(0)}$ から IST アルゴリズムを実行する.
- (3) 各ステップでエネルギー期待値の推定値 $\bar{E}_R^{(n)}$ を計算し, 重み因子 $g_{R-1}^{(n)}$ を $g_{R-1}^{(n)} = -(\beta_R - \beta_{R-1})\bar{E}_R^{(n)}/2$ で与える.
- (4) IST アルゴリズムを実行し続け, β_R から β_{R-1} へはじめて遷移したときエネルギー期待値の推定値 $\bar{E}_{R-1}^{(n)}$ を計算する. 重み因子 $g_{R-1}^{(n)}, g_{R-2}^{(n)}$ を (5.27) 式で与える.
- (5) さらに IST アルゴリズムを実行し, 各ステップごとにエネルギー期待値の推定値 $\bar{E}_R^{(n)}, \bar{E}_{R-1}^{(n)}$ を計算する. 重み因子 $g_{R-1}^{(n)}, g_{R-2}^{(n)}$ を (5.27) 式で与える.
- (6) 逆温度 β_{R-2} へはじめて遷移したとき, エネルギー期待値の推定値 $\bar{E}_{R-2}^{(n)}$ を計算する. 重み因子 $g_{R-2}^{(n)}, g_{R-3}^{(n)}$ を (5.27) 式で与える.
- (7) 同様にして, IST アルゴリズムを実行し各ステップでそのときの逆温度 β_r におけるエネルギー期待値の推定値 $\bar{E}_r^{(n)}$ を計算する. 重み因子 $g_r^{(n)}, g_{r-1}^{(n)}$ を (5.27) 式で与える.
- (8) すべての重み因子 $\{g_r\}$ が十分更新されるまで手順 (7) を繰り返す.

第 6 章

Irreversible simulated tempering アルゴリズムの性能評価

本章では IST アルゴリズムの性能を数値的に評価する。いくつかの古典スピン模型に対し、on-the-fly weight estimation を実行して重み因子を評価し、その推定効率を検証する。また推定した重み因子を用いて IST アルゴリズムを実行し、逆温度の緩和やスピン変数の自己相関時間などを評価することで、詳細つり合い条件の破れがシミュレーテッドテンパリングにもたらす効果を検証する。

6.1 設定

6.1.1 模型の設定

MCMC 法の性能は扱う模型の性質に大きく左右される。特に統計力学で現れる模型の多くは相転移を示し、相転移の種類によっては MCMC 法の緩和が著しく遅くなる現象がみられる。この現象は critical slowing down と呼ばれている。MCMC 法の新しいアルゴリズムを提案しその性能を評価する上で、critical slowing down などへの有効性を検証することは不可欠である。本博士論文ではそれぞれ性質の異なる相転移を示す模型の例として、2次元正方格子上に定義された磁場なし強磁性古典スピン模型である

- イジング模型
- 8 状態ポッツ模型
- XY 模型

の 3 種類の模型を扱う。それぞれの模型において格子の一辺の長さを L とするとき、スピンの総数は $N = L^2$ で与えられる。サイト i 上のスピン変数を S_i と書き、スピン変数の集合を $\mathbf{S} = \{S_i\}_{i=1}^N$ と記す。スピン変数 S_i がとりうる値は模型によって異なる。ま

たすべての模型について、スピンには最近接相互作用のみがはたらくとする。交換相互作用の強さを表す定数 J はすべての模型について $J = 1$ とする。それぞれの模型のエネルギー関数に現れる和 $\sum_{\langle i,j \rangle}$ は周期境界条件の下ですべての最近接スピン対にわたってとられる。

イジング模型

強磁性イジング模型のエネルギー関数は

$$E(\mathbf{S}) = - \sum_{\langle i,j \rangle} S_i S_j \quad (6.1)$$

で与えられる。ここで $S_i = \pm 1$ はイジングスピンである。

2次元正方格子上に定義された強磁性イジング模型は2次転移を示すことが知られている。その転移逆温度 β_c は双対性を用いて

$$\beta_c = \frac{1}{2} \ln(1 + \sqrt{2}) \simeq 0.4407 \quad (6.2)$$

であることが示される [80]。また Onsager によって無限系の自由エネルギーの解析解が与えられている [81]。有限系に対しても転送行列のすべての固有値が評価できるため、自由エネルギーを $O(N)$ の計算量で数値的に得ることができる [82]。

8 状態ポッツ模型

整数 q に対して、 q 状態強磁性ポッツ模型のエネルギー関数は

$$E(\mathbf{S}) = - \sum_{\langle i,j \rangle} \delta(S_i, S_j) \quad (6.3)$$

で与えられる。ここで $S_i \in \{0, 1, \dots, q-1\}$ は q 通りの状態をとりうるポッツスピンであり、 $\delta(\cdot, \cdot)$ はクロネッカーのデルタである。

この模型は $q \leq 4$ で2次転移、 $q > 4$ で1次転移を示すことが知られている。特に $q = 2$ のときは、ポッツスピン $S_i \in \{0, 1\}$ に対して

$$\sigma_i \equiv 2S_i - 1 \quad (6.4)$$

なる変換で与えられるイジングスピン $\sigma_i \in \{-1, 1\}$ を考えると、

$$\delta(S_i, S_j) = \frac{1 + \sigma_i \sigma_j}{2} \quad (6.5)$$

が成り立つ。これより、2状態強磁性ポッツ模型はイジング模型と等価であることがわかる。なおこのときの転移逆温度は (6.2) 式の2倍で与えられる。また $q > 4$ での転移逆温度 β_t は

$$\beta_t = \ln(1 + \sqrt{q}) \quad (6.6)$$

表 6.1 数値計算で用いた逆温度の値.

	Ising	Potts	XY
β_1	0.2	1.0	0.8
β_R	0.5	1.4	1.2

で与えられる [83]. 本博士論文では 1 次転移を示す $q = 8$ のときを扱う. このとき転移逆温度 β_t は $\beta_t = \ln(1 + \sqrt{8}) \simeq 1.3425$ である.

XY 模型

XY 模型のエネルギー関数は

$$E(\mathbf{S}) = - \sum_{\langle i,j \rangle} S_i \cdot S_j \quad (6.7)$$

で与えられる. ここで $S_i = (\cos \theta_i, \sin \theta_i)$, $\theta_i \in [0, 2\pi)$ は XY スピンであり, ドットはその内積を表す. エネルギー関数は $\boldsymbol{\theta} \equiv \{\theta_i\}_{i=1}^N$ の関数として

$$E(\boldsymbol{\theta}) = - \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j) \quad (6.8)$$

と書き換えることができる.

この模型は Kosterlitz–Thouless (KT) 転移を示すことが知られており^{*1}, その転移逆温度 β_{KT} は

$$\beta_{\text{KT}} \simeq 1.121 \quad (6.9)$$

である [84].

6.1.2 数値計算の詳細

逆温度の選び方

IST アルゴリズムを実行するにあたって, 確率変数としての逆温度 β がとりうる値 $\{\beta_r\}_{r=1}^R$ を事前に設定する必要がある. 以下では $\beta_1 < \beta_2 < \dots < \beta_R$ とする. 本博士論文で扱う 3 種類の模型に対して, 逆温度の両端の値 β_1 および β_R を表 6.1 のとおり選び, 固定する. 残りの値 $\beta_2, \beta_3, \dots, \beta_{R-1}$ は隣接する逆温度の間隔が一定となるように選び, 逆温度個数 R に応じて定まる. いずれの模型においても, 区間 $[\beta_1, \beta_R]$ 内に転移逆温度が含まれている. ただし逆温度 β がとりうる値の集合 $\{\beta_r\}_{r=1}^R$ に転移逆温度が含まれているわけではない.

^{*1} Berezinskii–Kosterlitz–Thouless (BKT) 転移ともいう.

スピン状態の更新方法

IST アルゴリズムでは逆温度更新プロセスに詳細つり合い条件の破れを導入した一方で, 系の状態更新には従来と同様の MCMC 法を用いている. ここでは本博士論文で用いたスピン状態の更新方法をそれぞれの模型についてまとめる.

2次元イジング模型では, シングルスピントリップに基づく Metropolis–Hastings アルゴリズムを用いてスピン状態を更新した. 2次元 8 状態ポッツ模型では, 熱浴法を用いた. 2次元 XY 模型では, Metropolis–Hastings アルゴリズムを用いた. 1 回のシングルスピントリップで用いた角度 θ_i の変化 $\Delta\theta$ の値は, 閉区間 $[-\pi/4, \pi/4]$ 上の一様分布から毎回独立にサンプリングした.

以下で行うすべての数値計算では, N 回のスピン状態更新と 1 回の逆温度更新の手続きをあわせて 1 モンテカルロステップ (MCS) とした.

6.2 重み因子の推定

まず前章で述べた on-the-fly weight estimation を用いて, それぞれの模型に対して重み因子 g_r を推定し, その効率を評価する.

6.2.1 First round-trip time

重み因子推定手法のひとつである on-the-fly weight estimation では, 初期逆温度として最大逆温度 β_R からスタートする. 数値計算を始めてからしばらくは重み因子が正しく推定できていないために, 逆温度空間での遷移が起こりにくい. ここでは詳細つり合い条件の破れが重み因子推定における初期のダイナミクスにどのような影響を及ぼすのか検証する. それぞれの模型に対して on-the-fly weight estimation を実行し, 最初に逆温度空間を 1 往復するのに要した MCS 数を図 6.1 から図 6.3 に示す. 各プロットの点および不確かさは, 独立な 16 回の試行から評価した. これらの図から, いずれの模型に対しても同様の傾向が示唆されている. 以下詳しくみていこう.

まず逆温度個数 R が小さい領域に注目する. システムサイズ L が小さいとき, この領域では逆温度個数 R に対して 1 往復に要する MCS 数が単調かつべき的に増加している. 一方システムサイズ L が大きくなると単調性が破れ, 逆温度個数 R が小さいほど要した MCS 数が大きくなっている. 本博士論文では逆温度空間の両端の値 β_1 および β_R をそれぞれ固定して数値計算を行っているため, 逆温度個数 R が小さいとき隣り合う逆温度の間隔は大きくなっている. そのため逆温度空間での遷移確率が小さくなり, より多くの MCS 数を要したと考えられる. またこの領域では詳細つり合い条件の破れによる差はみ

られない。

逆温度個数 R が大きい領域に注目しよう。いずれのシステムサイズ L でも、詳細つり合い条件をみたす場合 ($\delta = 0$) と比較して破る場合 ($\delta \neq 0$) のほうがより少ない MCS 数で逆温度空間を 1 往復していることがわかる。特に逆温度個数 R が十分大きなときの漸近形は、 $\delta = 0$ のときおおむね $O(R^2)$ であるのに対して、 $\delta \neq 0$ のとき $O(R)$ まで質的に変化している。つまり on-the-fly weight estimation における詳細つり合い条件の破れは、逆温度の拡散的なダイナミクスを質的に改善しうることを示唆している。

これらふたつの領域の間では、それぞれの領域における効果のクロスオーバーがみえており、全体として要した MCS 数が最も小さくなっている。また詳細つり合い条件の破れによる差はみられない。

以上のことから on-the-fly weight estimation で詳細つり合い条件を破ると、逆温度の間隔が十分に小さく、逆温度の個数が多いときに初期の逆温度緩和ダイナミクスが質的に改善することが数値的に示された。

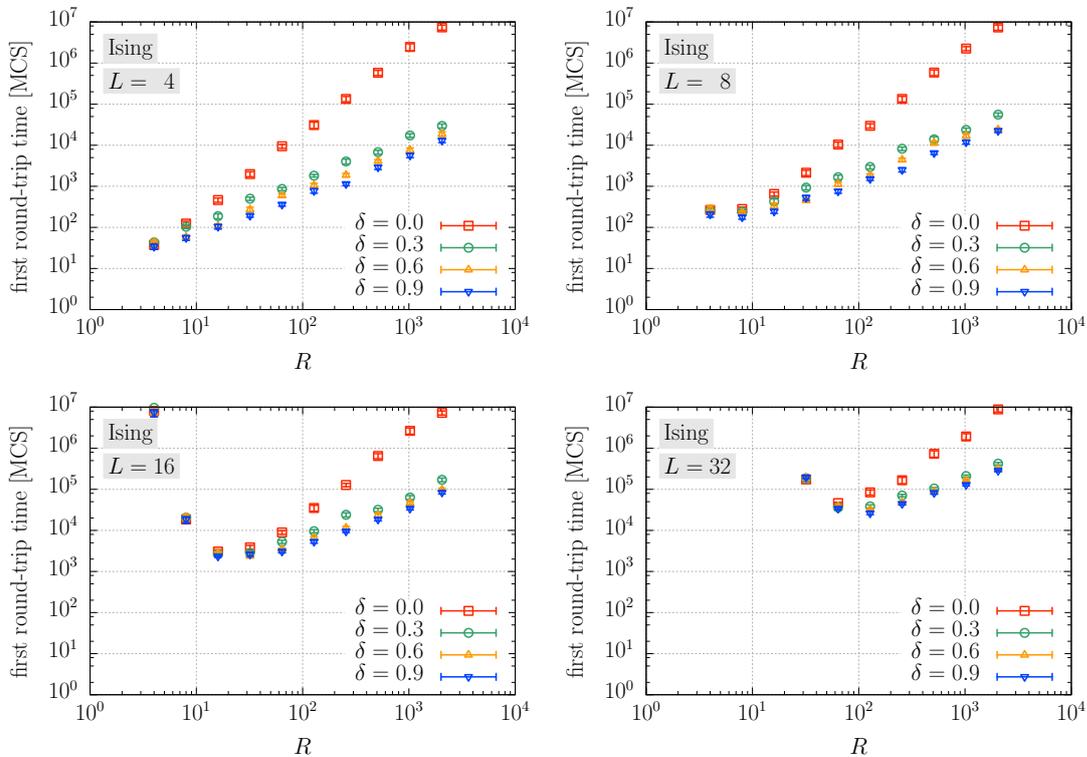


図 6.1 2次元イジング模型 (Ising) に対して on-the-fly weight estimation を実行し、逆温度空間を最初に 1 往復するのに要した MCS 数の逆温度個数 R 依存性。

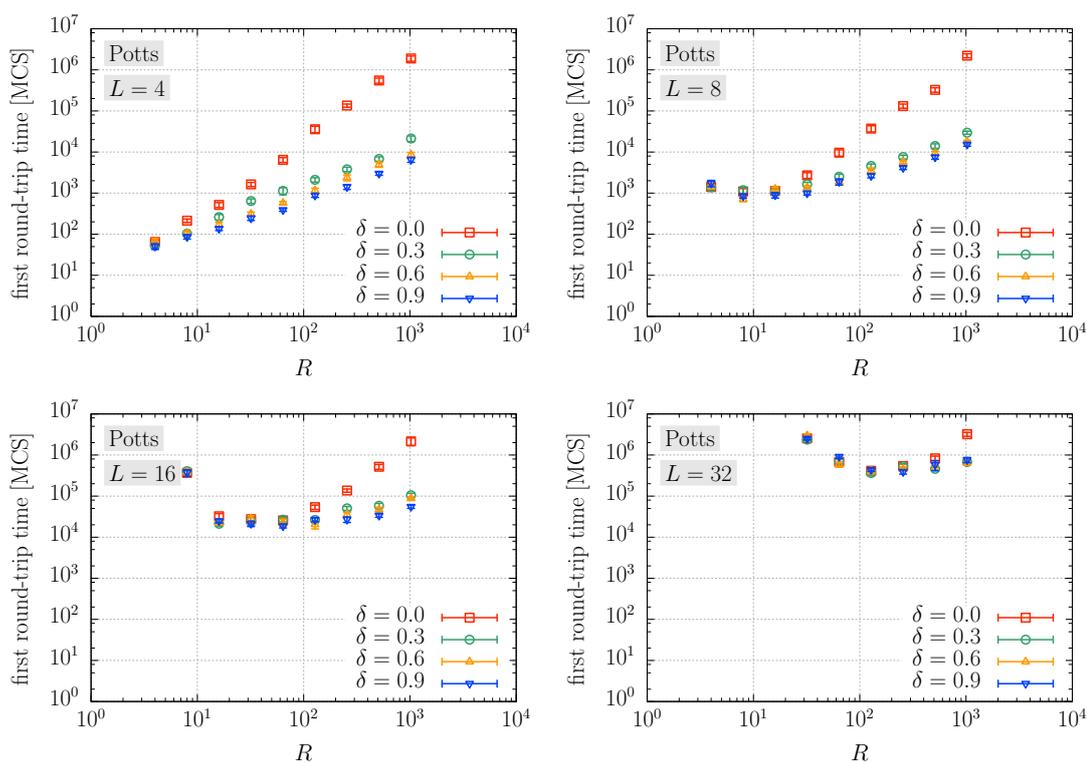


図 6.2 2次元 8 状態ポッツ模型 (Potts) に対して on-the-fly weight estimation を実行し、逆温度空間を最初に 1 往復するのに要した MCS 数の逆温度個数 R 依存性.

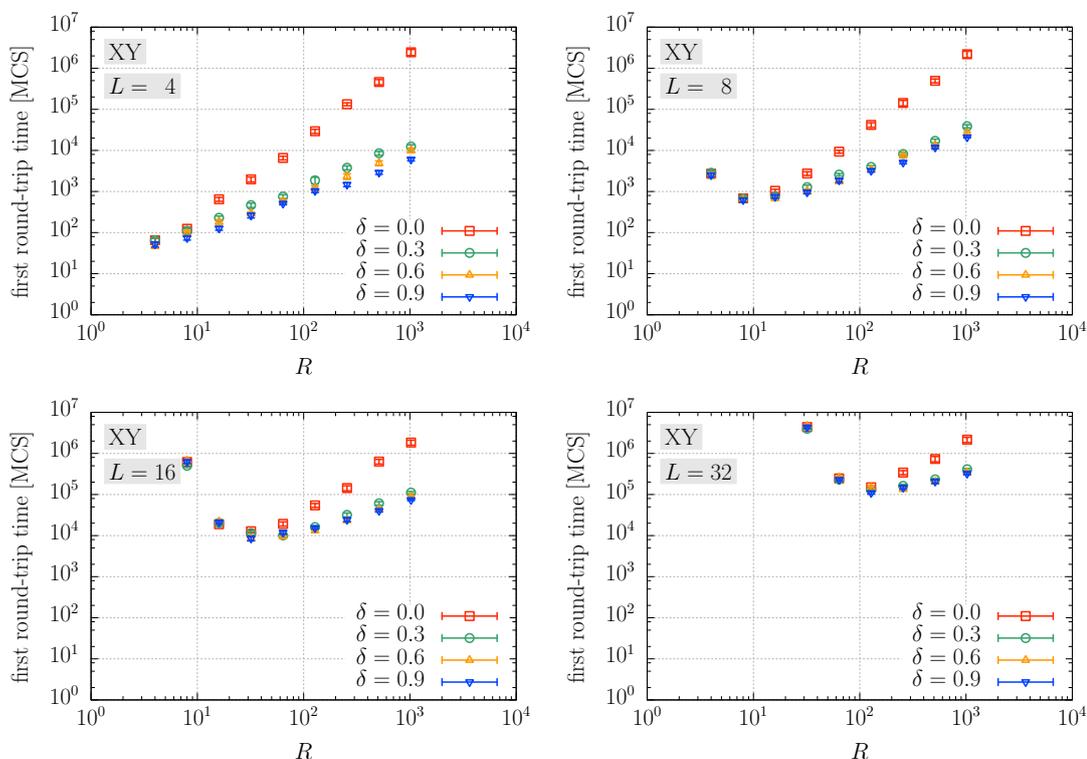


図 6.3 2次元 XY 模型 (XY) に対して on-the-fly weight estimation を実行し、逆温度空間を最初に 1 往復するのに要した MCS 数の逆温度個数 R 依存性.

6.2.2 重み因子の推定精度

次に, on-the-fly weight estimation によって得られる重み因子 g_r の推定精度を評価する [73]. 2次元イジング模型では熱力学極限での自由エネルギーが解析的に評価されているのみならず, 有限系に対しても転送行列の全固有値が評価できており, これを元にして自由エネルギーの値を数値的に得ることができる [82]. つまり任意のシステムサイズ L で厳密な重み因子の値 g_r^* が計算できる例になっている. そのためここでは特に 2次元イジング模型についてのみ考えることにする.

図 6.4 に on-the-fly weight estimation を 10^6 MCS, 10^7 MCS だけ実行して得られる重み因子 g_r の推定精度をそれぞれ示した. 不確かさは独立な 16 回の試行の結果から評価した. 重み因子の値には定数分の不定性があり, その差にのみ意味がある. そのためここでは重み因子 g_r の推定精度として, 隣接する値との差 $\Delta g_r \equiv g_{r+1} - g_r$ の推定精度を議論する.

図 6.1 から, ここで用いているパラメタ ($L = 32$, $R = 512$) では, 詳細つり合い条件をみたす場合 ($\delta = 0$) と破る場合 ($\delta = 0.9$) で逆温度空間を往復する回数に 10 倍程度の違いがある. すなわち, 10^6 MCS では $\delta = 0$ のとき逆温度空間をおおむね 1 往復するのに対して, $\delta = 0.9$ ではおおむね 10 往復する. しかし図 6.4 では重み因子の相対誤差に大きな違いはみられず, 0.1% 程度の精度で推定できていることがわかる. MCS 数を 10 倍に増やすと推定精度は改善するものの, 依然として詳細つり合い条件を破る効果はみられない. 以上のことから, on-the-fly weight estimation における詳細つり合い条件の破れは, 逆温度空間の遷移を促進しつつも重み因子の推定精度には影響を及ぼさず, その推定精度は MCS 数にのみ依存することが示唆される.

次節以降の数値計算では, それぞれの模型に対して $\delta = 0.9$ とした on-the-fly weight estimation を実行し, 逆温度空間を 100 往復して得られた重み因子を用いる (図 6.5).

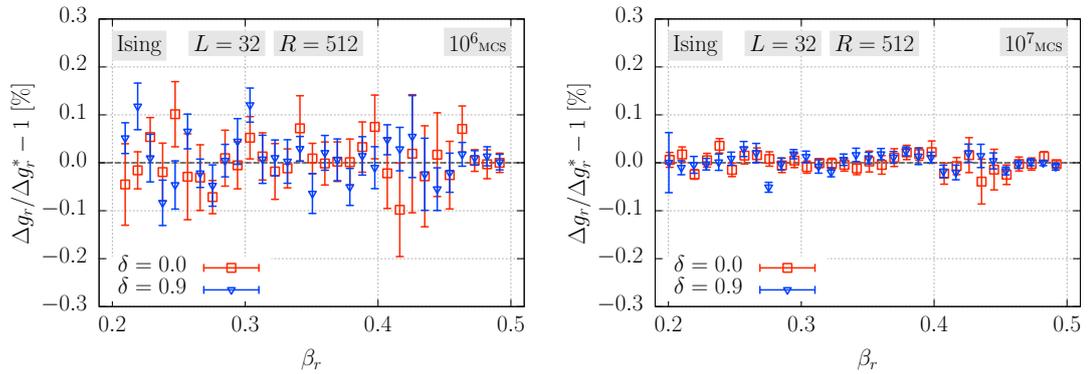


図 6.4 2次元イジング模型 (Ising) に on-the-fly weight estimation を 10^6 MCS (左) および 10^7 MCS (右) だけ実行して得られる重み因子 g_r の推定精度. 推定値 g_r と文献 [82] を元に行われる解析解 g_r^* のそれぞれにおいて, 隣接する逆温度での値の差 $\Delta g_r \equiv g_{r+1} - g_r$ および $\Delta g_r^* \equiv g_{r+1}^* - g_r^*$ を計算し, 横軸 β_r に対してそれらの相対誤差 $\Delta g_r / \Delta g_r^* - 1$ を示した. 見やすさの観点から $\delta = 0$ および $\delta = 0.9$ の場合のみ表示している. また逆温度個数 $R = 512$ での結果に対して, それぞれの δ で 512 点あるデータを等間隔に 32 点のみ表示している.

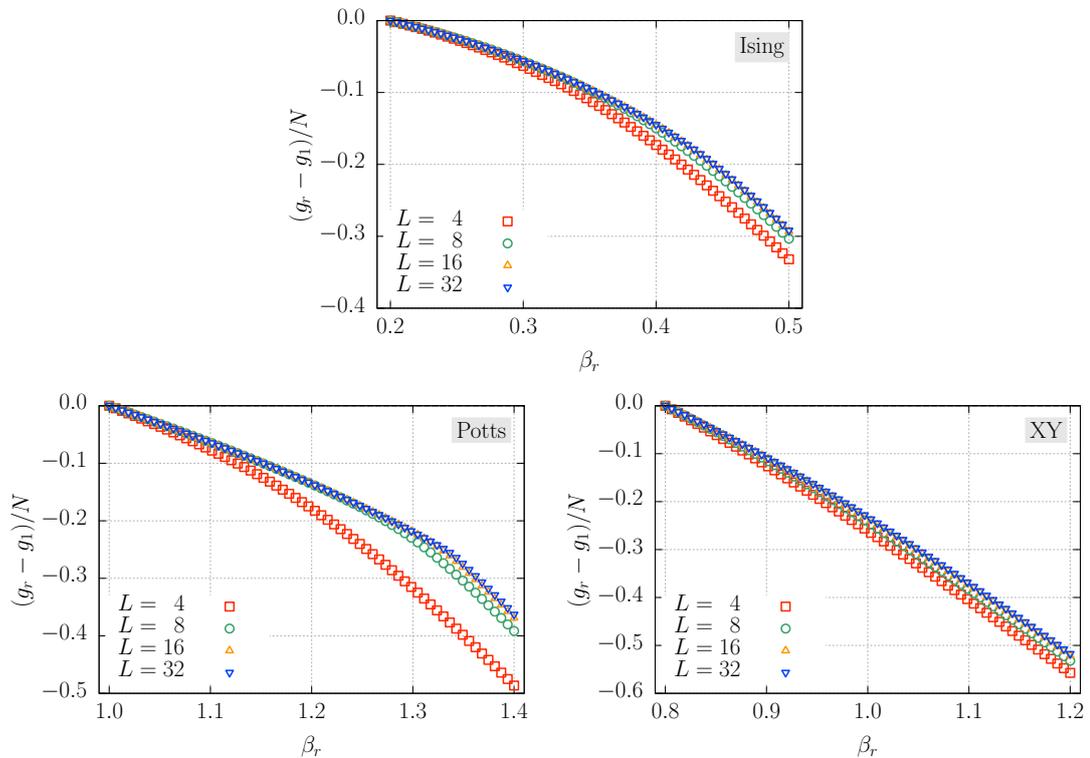


図 6.5 2次元イジング模型 (Ising), 2次元8状態ポッツ模型 (Potts), 2次元XY模型 (XY) のそれぞれに対して, on-the-fly weight estimation で推定した重み因子 g_r . それぞれのシステムサイズ L に対して逆温度個数 $R = 64$ の結果を示した.

6.3 逆温度の緩和ダイナミクス

前節では、重み因子 g_r を推定しながら IST アルゴリズムを実行したときの逆温度空間におけるダイナミクスをみた。ここでは on-the-fly weight estimation で推定した重み因子 g_r を固定して IST アルゴリズムを実行したときに、詳細つり合い条件の破れが逆温度空間でのダイナミクスに与える効果を検討する。

6.3.1 Round-trip time

まず比較的測定しやすい量として、逆温度空間での round-trip time を測定する。ここでの round-trip time τ_{rt} は、IST アルゴリズムにおける逆温度空間 1 往復あたりの平均 MCS 数と定義する。すなわち、

$$\tau_{rt} \equiv \inf \{n > 0; \beta^{(0)} = \beta_1, \beta^{(n)} = \beta_1, 0 < \exists t < n \text{ s.t. } \beta^{(t)} = \beta_R\} \quad (6.10)$$

である。本博士論文では、一様ランダムな初期状態 $\mathbf{S}^{(0)}$, $\varepsilon^{(0)}$ と初期逆温度 $\beta^{(0)} = \beta_1$ の下で IST アルゴリズムを実行し、逆温度空間を K 往復 ($K = 100$) するまでに要した MCS 数を K で割った数を τ_{rt} とした。すなわち k 往復目に要した MCS 数を n_k ($k = 1, 2, \dots, K$) として、round-trip time τ_{rt} を

$$\tau_{rt} = \frac{1}{K} \sum_{k=1}^K n_k \quad (6.11)$$

と評価した。また round-trip time τ_{rt} の不確かさ σ は、

$$\sigma = \sqrt{\frac{1}{K-1} \left[\frac{1}{K} \sum_{k=1}^K n_k^2 - \left(\frac{1}{K} \sum_{k=1}^K n_k \right)^2 \right]} \quad (6.12)$$

と評価した。

それぞれの模型に対して round-trip time τ_{rt} を測定した結果を図 6.6 から図 6.8 に示す。各プロットの点および不確かさは、独立な 16 回の試行から評価した。いずれの模型に対しても、定量的には違いがみられるものの、定性的には同様の振る舞いがみえている。また前節における重み因子推定の最初の 1 往復に要した MCS 数 (図 6.1, 6.2, 6.3) と同様の特徴がみられる。すなわち、逆温度個数 R が大きい領域で詳細つり合い条件の破れの効果が現れ、round-trip time τ_{rt} の逆温度個数 R 依存性が質的に改善している。

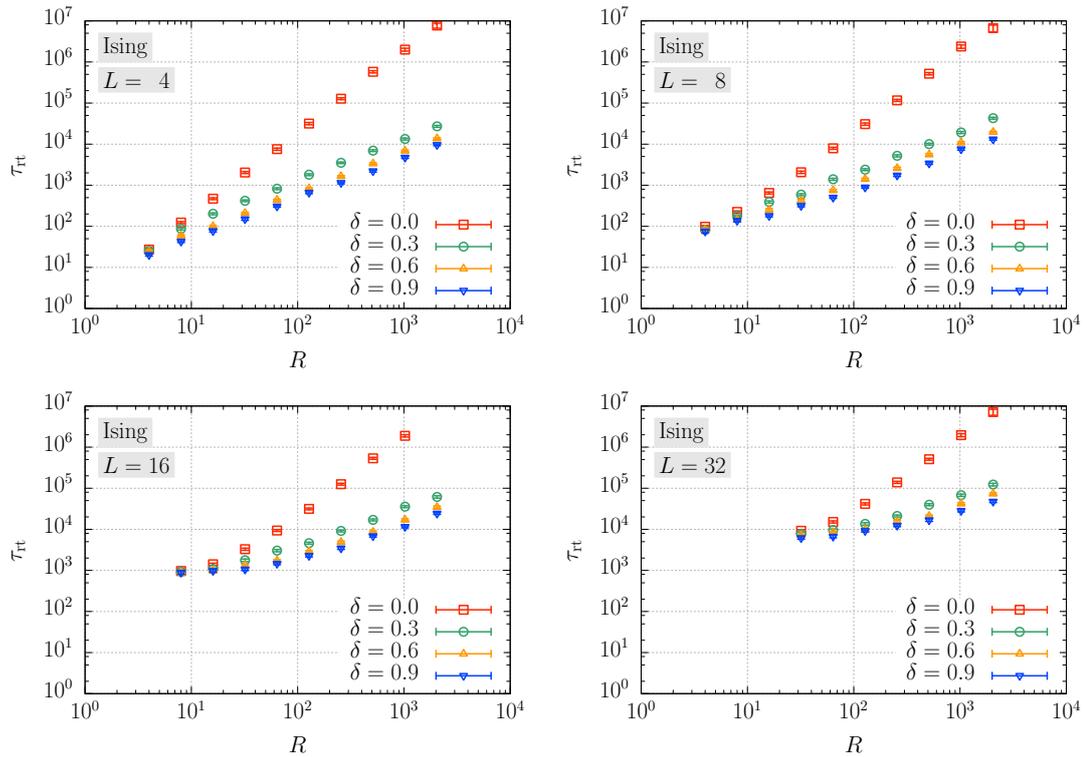


図 6.6 2次元イジング模型 (Ising) に対して IST アルゴリズムを実行したときの逆温度空間における round-trip time τ_{rt} の逆温度個数 R 依存性.

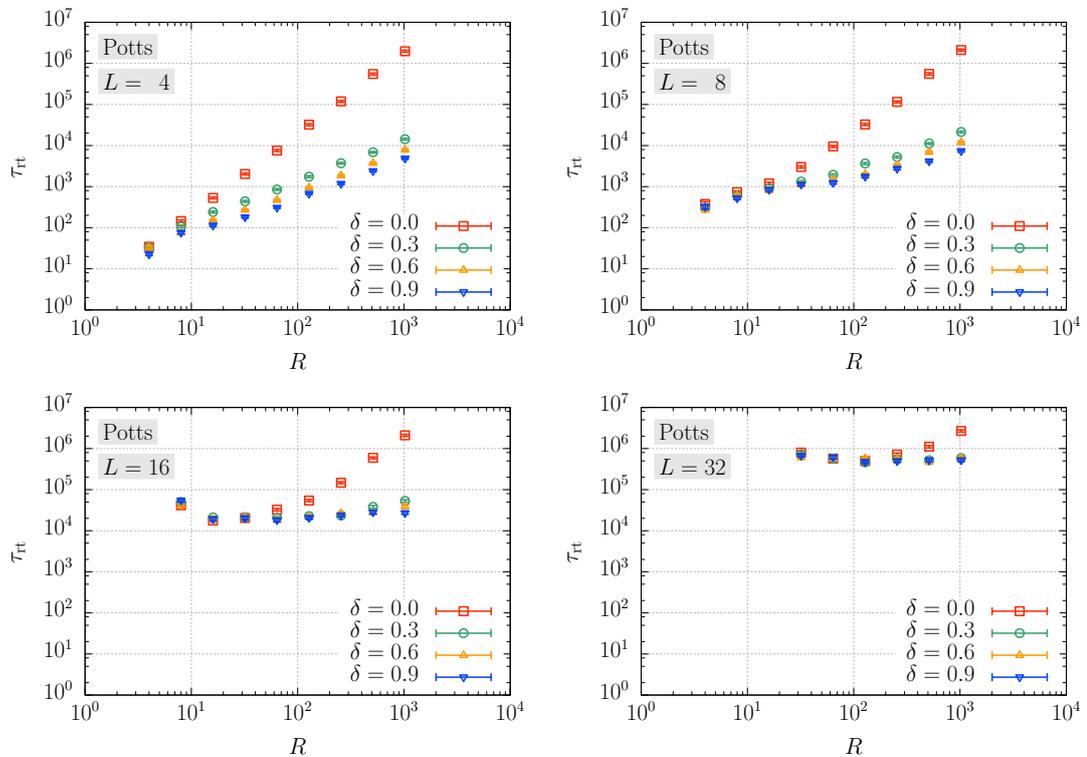


図 6.7 2次元8状態ポッツ模型 (Potts) に対して IST アルゴリズムを実行したときの逆温度空間における round-trip time τ_{rt} の逆温度個数 R 依存性.

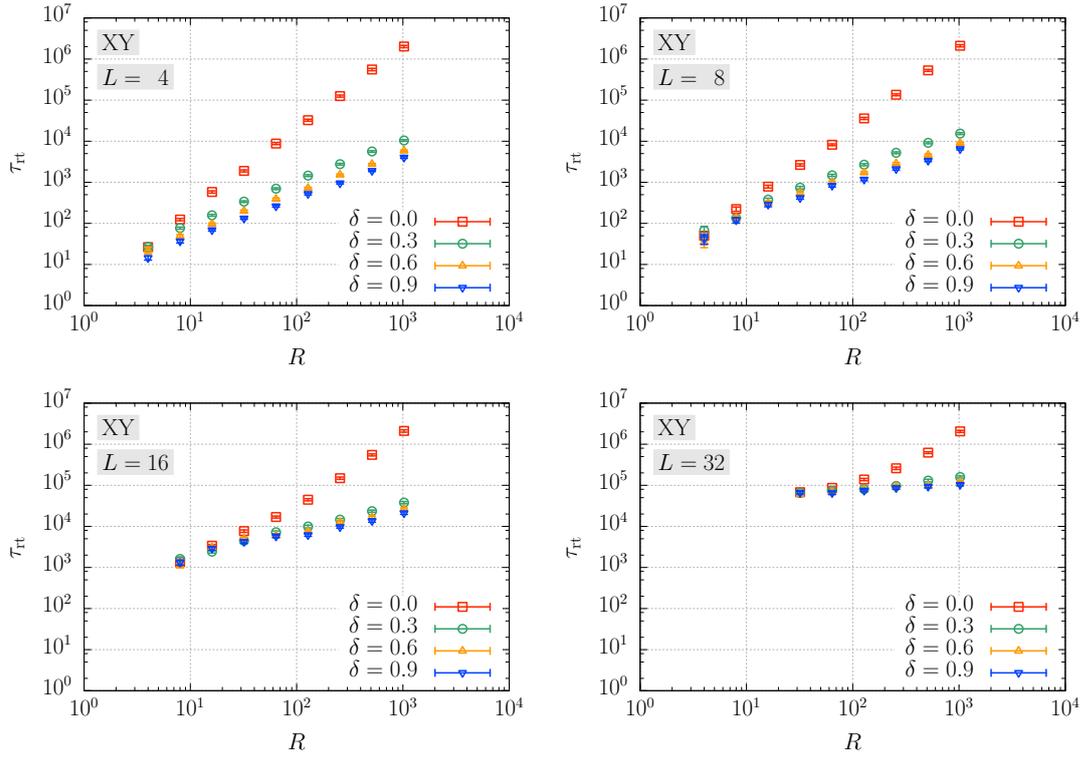


図 6.8 2 次元 XY 模型 (XY) に対して IST アルゴリズムを実行したときの逆温度空間における round-trip time τ_{rt} の逆温度個数 R 依存性.

6.3.2 逆温度の緩和時間

逆温度空間での緩和ダイナミクスをさらに詳しく調べるため、逆温度の緩和関数の時間発展を測定する。逆温度の緩和関数 $\phi_\beta^{(n)}$ は

$$\phi_\beta^{(n)} \equiv \frac{\langle \beta^{(n)} \rangle - \langle \beta \rangle_{\text{eq}}}{\langle \beta^{(0)} \rangle - \langle \beta \rangle_{\text{eq}}} \quad (6.13)$$

と定義される。ここで $\langle \beta^{(n)} \rangle$ は n MCS 後の逆温度の値 $\beta^{(n)}$ のサンプル平均を表し、ブラケット $\langle \dots \rangle_{\text{eq}}$ は目的分布 $P_{\text{IST}}(\mathbf{S}, \beta_r, \varepsilon)$ での期待値を表す。また緩和の速さを緩和関数から定量的に評価するため、緩和時間 $\tau_{\text{relax}, \beta}$ を

$$\tau_{\text{relax}, \beta} \equiv \inf \left\{ n > 0; |\phi_\beta^{(n)}| < 0.2 \right\} \quad (6.14)$$

と定義する。

数値計算では初期逆温度を $\beta^{(0)} = \beta_R$ に固定し、初期状態 $\mathbf{S}^{(0)}$ は逆温度 β_R の平衡状態とした。具体的には、逆温度 β_R での Metropolis–Hastings アルゴリズムまたは熱浴法を 10^4 MCS 実行して得られた配置を $\mathbf{S}^{(0)}$ とした。lifting parameter の初期値 $\varepsilon^{(0)}$ は

一様ランダムである。それぞれの模型における逆温度の緩和関数 $\phi_\beta^{(n)}$ の時間発展および緩和時間 $\tau_{\text{relax},\beta}$ の逆温度個数 R 依存性を図 6.9 から図 6.14 に示す。各プロットの点および不確かさは、独立な 128 回の試行から評価した。緩和関数 $\phi_\beta^{(n)}$ の時間発展 (図 6.9, 6.11, 6.13) における $\phi_\beta^{(n)} = 0.8$ 付近の破線はそれぞれの模型の転移逆温度の値に対応する。

まず 2 次元イジング模型における逆温度の緩和ダイナミクスに注目する (図 6.9, 6.10)。図 6.9 は、システムサイズを $L = 32$ に固定して逆温度個数 R を変えると緩和関数 $\phi_\beta^{(n)}$ の時間発展がどのように変化するかを示している。詳細つり合い条件をみたす場合 ($\delta = 0$) と破る場合 ($\delta \neq 0$) の緩和ダイナミクスは逆温度個数 R を増やしていくにつれてどちらも遅くなっているものの、その遅くなり具合は $\delta = 0$ のほうが激しい。特に逆温度個数が $R = 1024$ のときは、 $\delta = 0$ よりも $\delta \neq 0$ のほうが 100 倍程度速く緩和している。つまり詳細つり合い条件を破ると緩和が加速している。

いくつかのシステムサイズ L に対して緩和時間 $\tau_{\text{relax},\beta}$ の逆温度個数依存性が図 6.10 に示されている。いずれのシステムサイズ L の値に対しても、round-trip time τ_{rt} の逆温度個数 R 依存性 (図 6.6) と定性的に同じ振る舞いを示している。すなわち、緩和時間 $\tau_{\text{relax},\beta}$ は詳細つり合い条件をみたす場合に漸近的に $\tau_{\text{relax},\beta} \simeq O(R^2)$ と振る舞うのに対して、詳細つり合い条件を破ることで $\tau_{\text{relax},\beta} \simeq O(R)$ と振る舞う。このことは、逆温度の緩和が拡散的なダイナミクスから弾道的なダイナミクスへと質的に変化していることを示している。この変化は第 4 章で議論した 1 次元ランダムウォークに対する詳細つり合い条件の破れの効果と同じものであると考えられる。

2 次元 8 状態ポッツ模型 (図 6.11, 6.12) と 2 次元 XY 模型 (図 6.13, 6.14) についても緩和時間 $\tau_{\text{relax},\beta}$ は 2 次元イジング模型と定性的に同様の振る舞いを示している。しかし緩和関数 $\phi_\beta^{(n)}$ はそれぞれの模型で異なる特徴がみられる。2 次元イジング模型と 2 次元 8 状態ポッツ模型では転移逆温度に対応する値で緩和関数にプラトーがみえているのに対して、2 次元 XY 模型ではプラトーは見当たらない。さらに 2 次元イジング模型よりも 2 次元 8 状態ポッツ模型のほうがプラトーがより鮮明にみえている。これらのことは、それぞれの模型が示す相転移の性質と関係があると考えられる。2 次元イジング模型が示す 2 次転移では比熱が発散し、2 次元 8 状態ポッツ模型が示す 1 次転移ではエネルギーが不連続に変化する。一方で 2 次元 XY 模型が示す KT 転移は比熱の発散を伴わない。このように、相転移におけるエネルギーや比熱の異常性が遷移確率を通して逆温度の緩和に影響を及ぼしていると考えられる。またプラトーは詳細つり合い条件をみたすとき ($\delta = 0$) よりも破れているとき ($\delta \neq 0$) に顕著に現れている。このことは、詳細つり合い条件の破れが緩和モードの中で初期緩和に関するモードを選択的に大きく改善していることを示唆している。

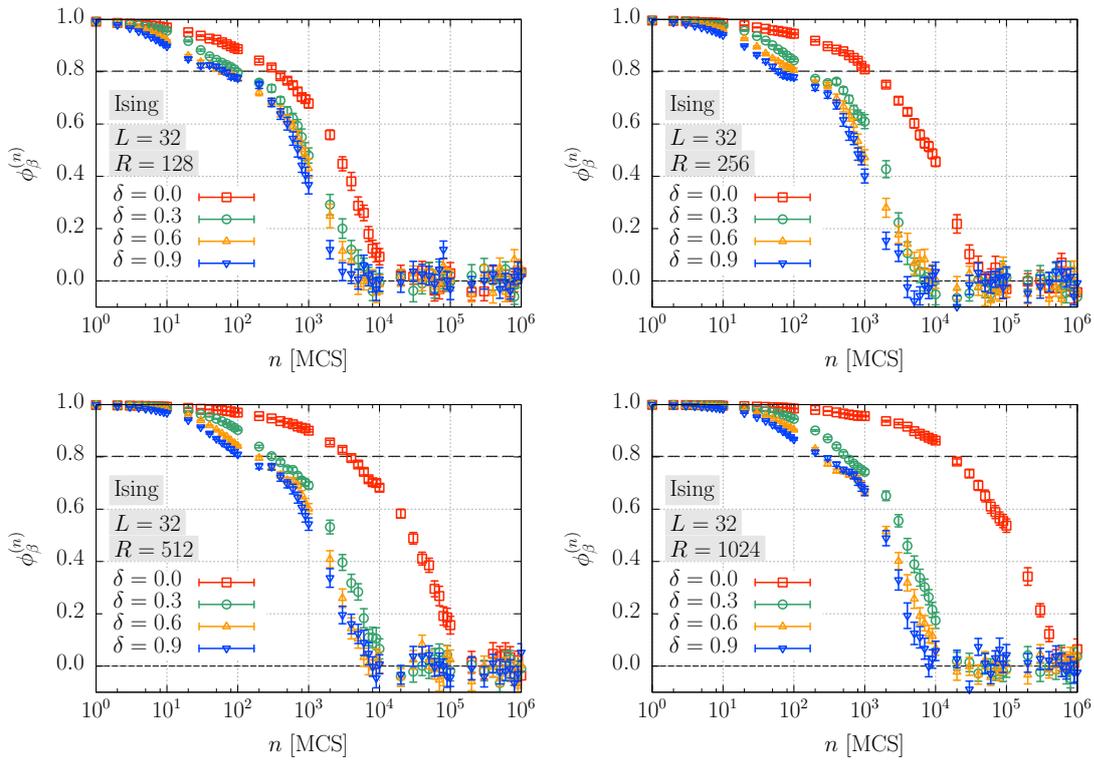


図 6.9 2次元イジング模型 (Ising) に IST アルゴリズムを実行したときの逆温度の緩和関数 $\phi_\beta^{(n)}$ の時間発展.

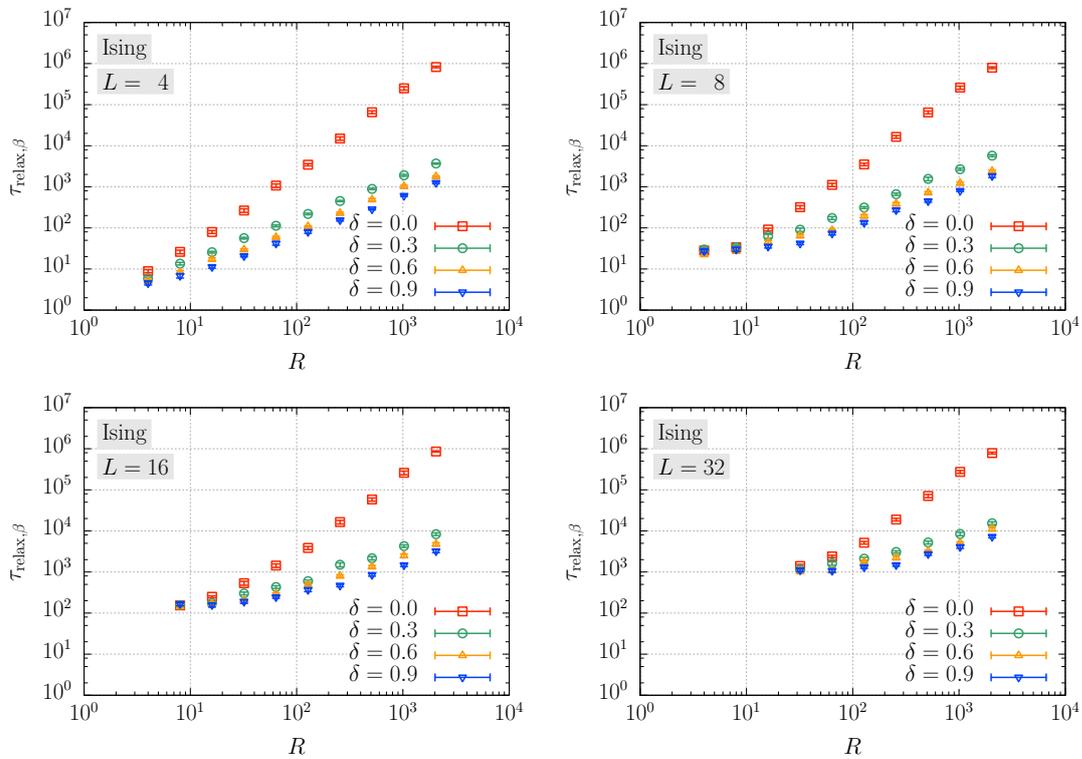


図 6.10 2次元イジング模型 (Ising) における逆温度の緩和時間 $\tau_{relax,\beta}$ の逆温度個数 R 依存性.

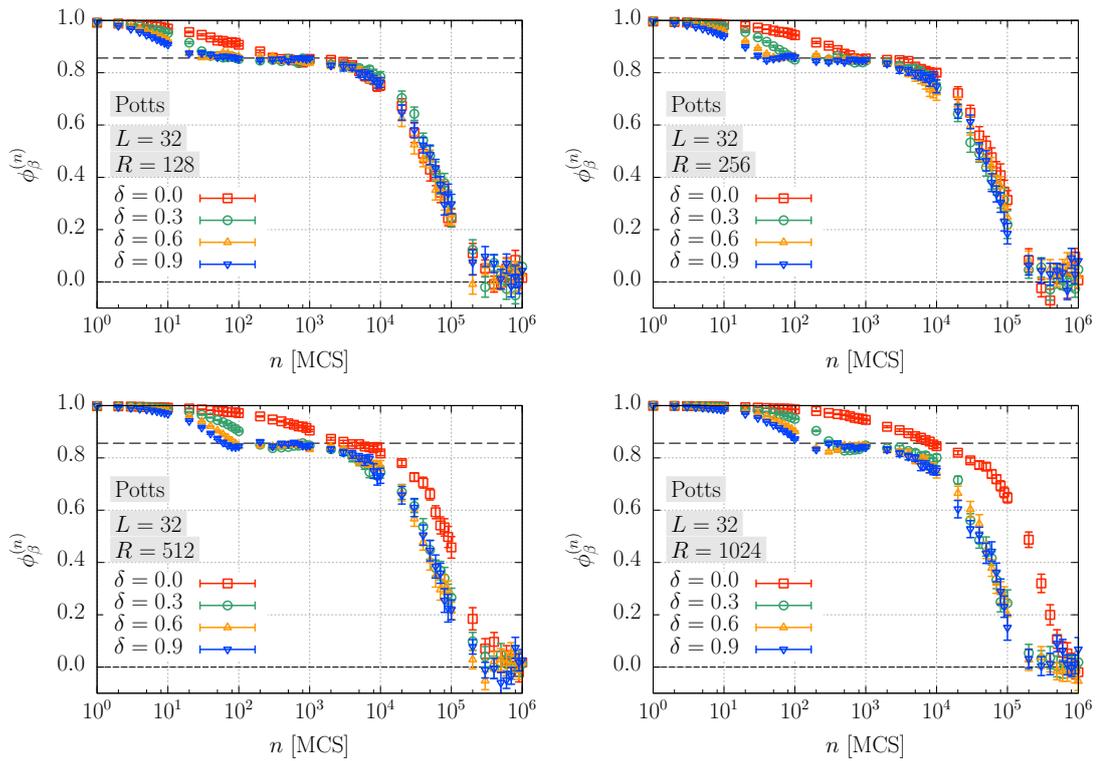


図 6.11 2次元 8 状態ポッツ模型 (Potts) に IST アルゴリズムを実行したときの逆温度の緩和関数 $\phi_\beta^{(n)}$ の時間発展.

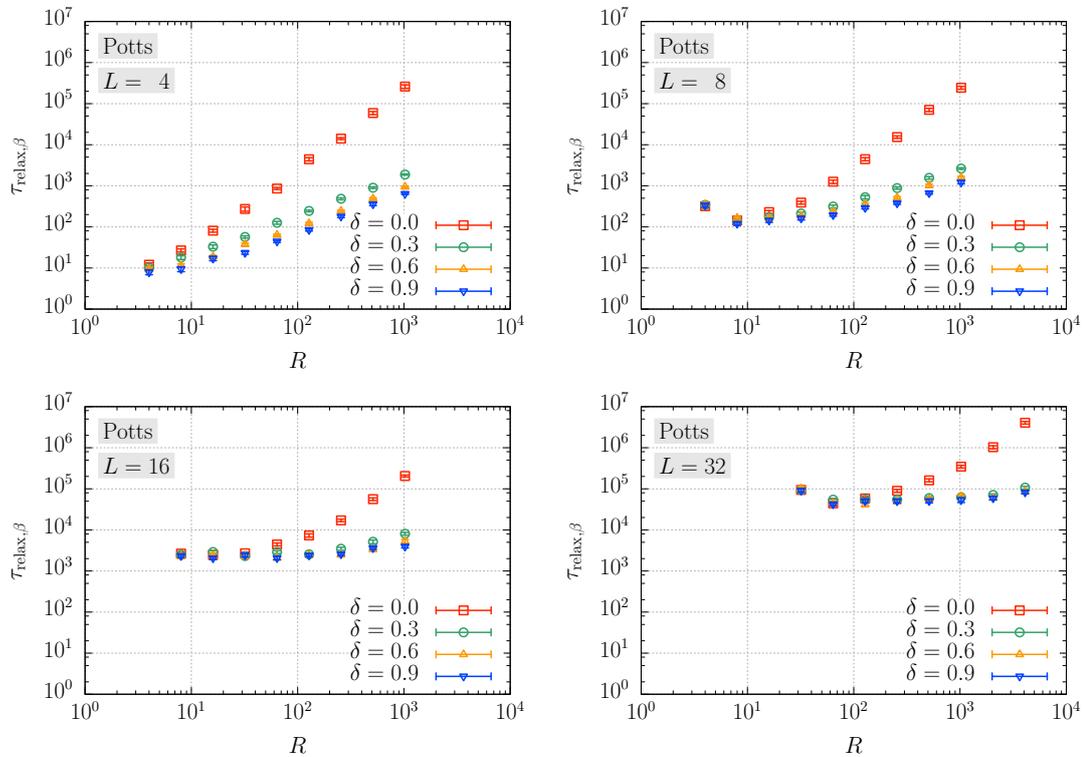


図 6.12 2次元 8 状態ポッツ模型 (Potts) における逆温度の緩和時間 $\tau_{relax,\beta}$ の逆温度個数 R 依存性.

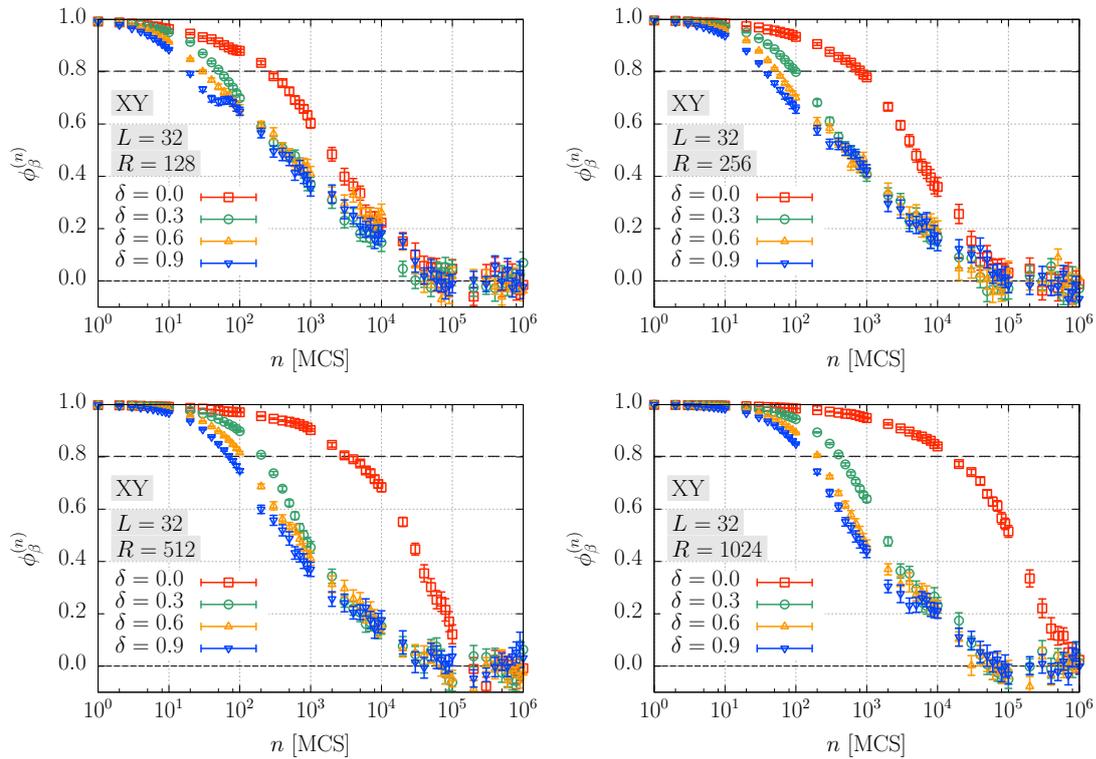


図 6.13 2次元 XY 模型 (XY) に IST アルゴリズムを実行したときの逆温度の緩和関数 $\phi_\beta^{(n)}$ の時間発展.

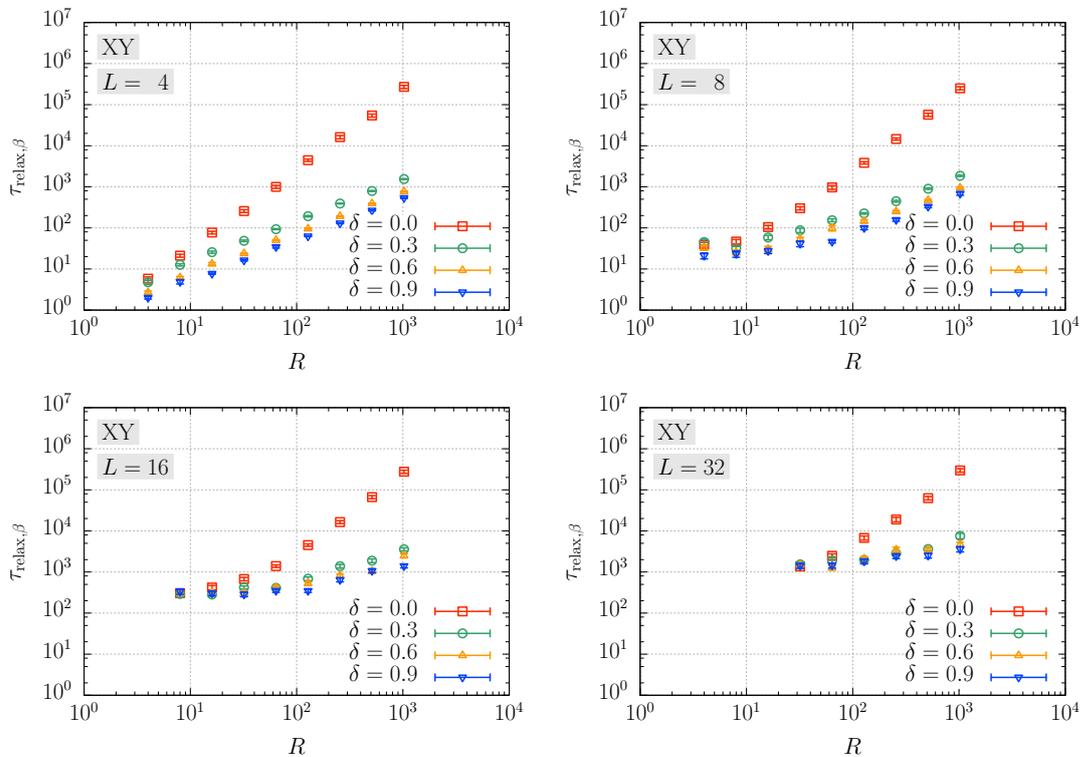


図 6.14 2次元 XY 模型 (XY) における逆温度の緩和時間 $\tau_{relax,\beta}$ の逆温度個数 R 依存性.

6.3.3 逆温度の経験遷移確率

これまでいくつかの量から IST アルゴリズムにおける逆温度空間での緩和ダイナミクスを考察してきた。これらの結果をより理解するために、逆温度空間での経験遷移確率 (empirical transition probability) を測定する。スピン数 N のイジングスピン系では状態数が 2^N 、8 状態ポッツスピン系では 8^N と、スピンの状態空間は一般に非常に大きい。また XY スピン系の状態空間は連続空間である。そのためスピンの状態空間での経験遷移確率を数値的に測定することは膨大な計算量が必要となり現実的でない。一方で逆温度 β と lifting parameter ε の組 (β, ε) がとりうる状態数は $2R$ であり、この空間での経験遷移確率は測定が容易である。そのためここでは、スピン状態 \mathbf{S} を考慮せずに (β, ε) 空間で測定した経験遷移確率を考える。

IST アルゴリズムを実行して得られるマルコフ連鎖 $\{(\mathbf{S}^{(n)}, \beta^{(n)}, \varepsilon^{(n)})\}_{n=1,2,\dots}$ に対して、状態 (β_r, ε) から状態 (β_l, ε') への経験遷移確率 $\tilde{T}(\beta_l, \varepsilon' | \beta_r, \varepsilon)$ を

$$\begin{aligned} \tilde{T}(\beta_l, \varepsilon' | \beta_r, \varepsilon) &\equiv \frac{\text{状態 } (\beta_r, \varepsilon) \text{ から状態 } (\beta_l, \varepsilon') \text{ へ直接遷移した回数}}{\text{状態 } (\beta_r, \varepsilon) \text{ の訪問回数}} \\ &= \frac{\sum_n \delta(\beta^{(n)}, \beta_r) \delta(\varepsilon^{(n)}, \varepsilon) \delta(\beta^{(n+1)}, \beta_l) \delta(\varepsilon^{(n+1)}, \varepsilon')}{\sum_n \delta(\beta^{(n)}, \beta_r) \delta(\varepsilon^{(n)}, \varepsilon)} \end{aligned} \quad (6.15)$$

と定義する。また経験遷移確率 $\tilde{T}(\beta_l, \varepsilon' | \beta_r, \varepsilon)$ を成分にもつ $2R$ 次正方行列

$$\tilde{\mathbf{T}} = (\tilde{T}(\beta_l, \varepsilon' | \beta_r, \varepsilon))_{1 \leq r, l \leq R; \varepsilon, \varepsilon' = \pm} \in \mathbb{R}^{2R \times 2R} \quad (6.16)$$

を経験遷移確率行列 (empirical transition probability matrix) と呼ぶ。本博士論文では IST アルゴリズムを 10^8 MCS 実行した後に経験遷移確率の測定を開始し、各 R に対して $2 \times 10^5 \times R$ MCS だけ測定を行った。特に逆温度空間での提案確率や ε -flip の入れ方から、経験遷移確率行列のほとんどの成分がゼロである。 $4R^2$ 個ある成分のうちゼロでない値をとりうるのは対角成分 $\tilde{T}(\beta_r, \varepsilon | \beta_r, \varepsilon)$ および $\tilde{T}(\beta_{r\pm 1}, \varepsilon | \beta_r, \varepsilon)$ と $\tilde{T}(\beta_r, -\varepsilon | \beta_r, \varepsilon)$ の計 $(8R - 4)$ 個のみである。

それぞれの模型について経験遷移確率 $\tilde{T}(\beta_{r+1}, + | \beta_r, +)$ および $\tilde{T}(\beta_r, - | \beta_r, +)$ を測定した結果を図 6.15 から図 6.20 に示す。なお、それぞれの図の両端 $\beta = \beta_1$ および $\beta = \beta_R$ 付近に特異な振る舞いがみられるが、これらは IST アルゴリズムでの逆温度の提案確率 (図 5.2) において両端を反射壁にしたことに由来するものと考えられる。

前小節での議論と対応して、2次元イジング模型と2次元8状態ポッツ模型における経験遷移確率 $\tilde{T}(\beta_{r+1}, + | \beta_r, +)$ および $\tilde{T}(\beta_r, - | \beta_r, +)$ の転移逆温度付近にそれぞれディップおよびカスパがみられる。逆温度個数 R を増やしていくと経験遷移確率 $\tilde{T}(\beta_{r+1}, + | \beta_r, +)$ のディップは浅くなり、逆温度空間全体で一様になっていく。つまり、

逆温度間隔を小さくすることで逆温度空間での遷移確率が大きくなり、比熱の異常からくる遷移確率の低下が抑えられている。この傾向は詳細つり合い条件の有無によらずみえている。一方、 ε -flip に対応する経験遷移確率 $\tilde{T}(\beta_r, -|\beta_r, +)$ にみられるカスは対数目盛でみると逆温度個数 R によらず残り、同じ形状を保っている。しかし値自体は逆温度個数 R に対しておおむね $O(R^{-1})$ で小さくなっている。系内の遷移確率がほぼ一様であること、および ε -flip の遷移確率が $O(R^{-1})$ 程度であるというふたつの特徴は、第 4 章の lifting によって詳細つり合い条件を破った 1 次元ランダムウォークで緩和時間が質的に改善する条件と整合している。このことから IST アルゴリズムにおいて逆温度の緩和ダイナミクスが質的に変化したのは、詳細つり合い条件の破れの効果であることが理論的にも説明できる。

最後に、経験遷移確率行列 \tilde{T} から逆温度の緩和速度を評価してみよう。 $\lambda_k \in \mathbb{C}$ ($k = 1, 2, \dots, 2R$) を経験遷移確率行列 $\tilde{T} \in \mathbb{R}^{2R \times 2R}$ の重複も許した固有値とする。一般性を失うことなく $1 = \lambda_1 > |\lambda_2| \geq \dots \geq |\lambda_{2R}|$ とできる。このとき逆温度空間での緩和時間 $\tilde{\tau}_{\text{relax}}$ を

$$\tilde{\tau}_{\text{relax}} \equiv -\frac{1}{\ln |\lambda_2|} \quad (6.17)$$

と定義する。それぞれの模型に対して緩和時間 $\tilde{\tau}_{\text{relax}}$ の逆温度個数 R 依存性を図 6.21 から図 6.23 に示した。ここでは数値解析ソフトウェアライブラリ LAPACK を用いて経験遷移確率行列を数値対角化し、得られた固有値から緩和時間を評価した。いずれの模型においても、緩和時間 $\tilde{\tau}_{\text{relax}}$ はこれまでみてきた round-trip time τ_{rt} や緩和関数から定まる緩和時間 $\tau_{\text{relax},\beta}$ と定性的に同じ振る舞いを示している。以上から、IST アルゴリズムにおける逆温度の緩和ダイナミクスは詳細つり合い条件の破れによって質的に変化し、その緩和が加速していることが示された。

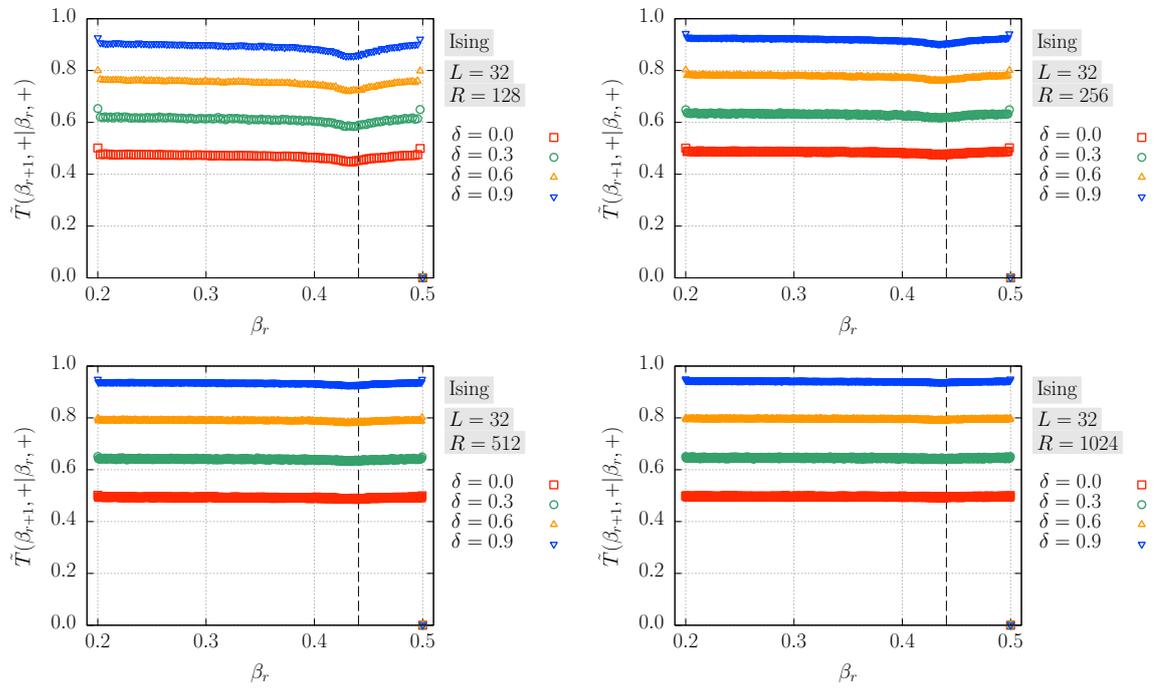


図 6.15 2次元イジング模型 (Ising) に対して IST アルゴリズムを実行したときの経験遷移確率 $\tilde{T}(\beta_{r+1}, + | \beta_r, +)$.

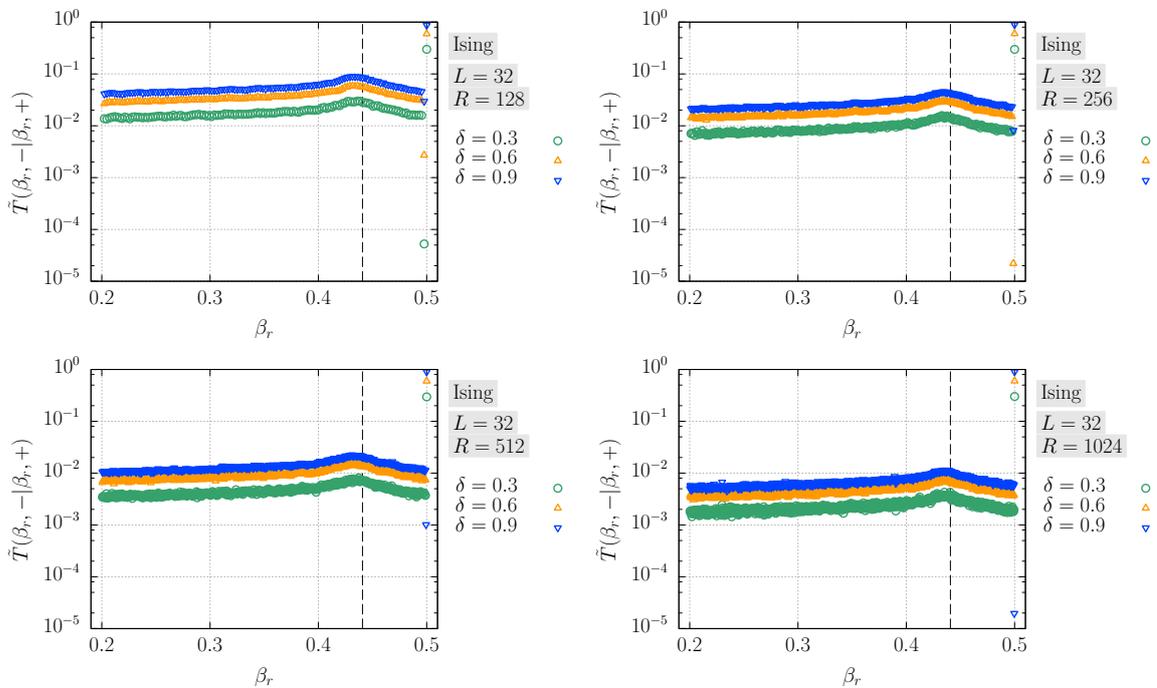


図 6.16 2次元イジング模型 (Ising) に対して IST アルゴリズムを実行したときの経験遷移確率 $\tilde{T}(\beta_r, - | \beta_r, +)$.

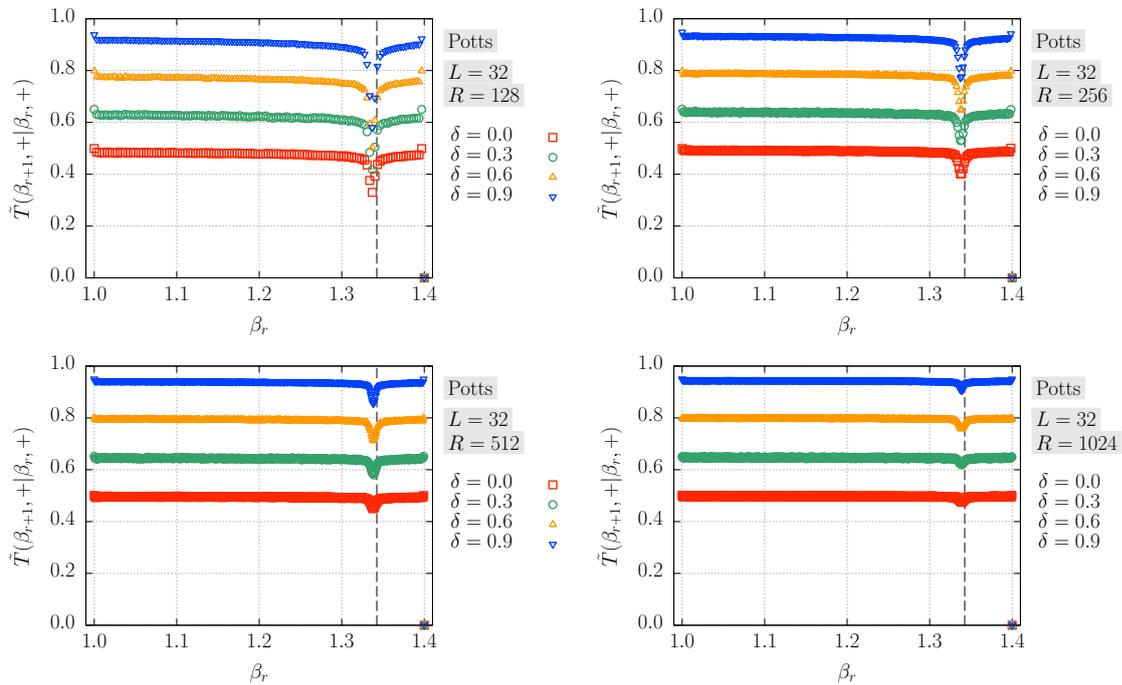


図 6.17 2次元 8 状態ポッツ模型 (Potts) に対して IST アルゴリズムを実行したときの経験遷移確率 $\tilde{T}(\beta_{r+1}, + | \beta_r, +)$.

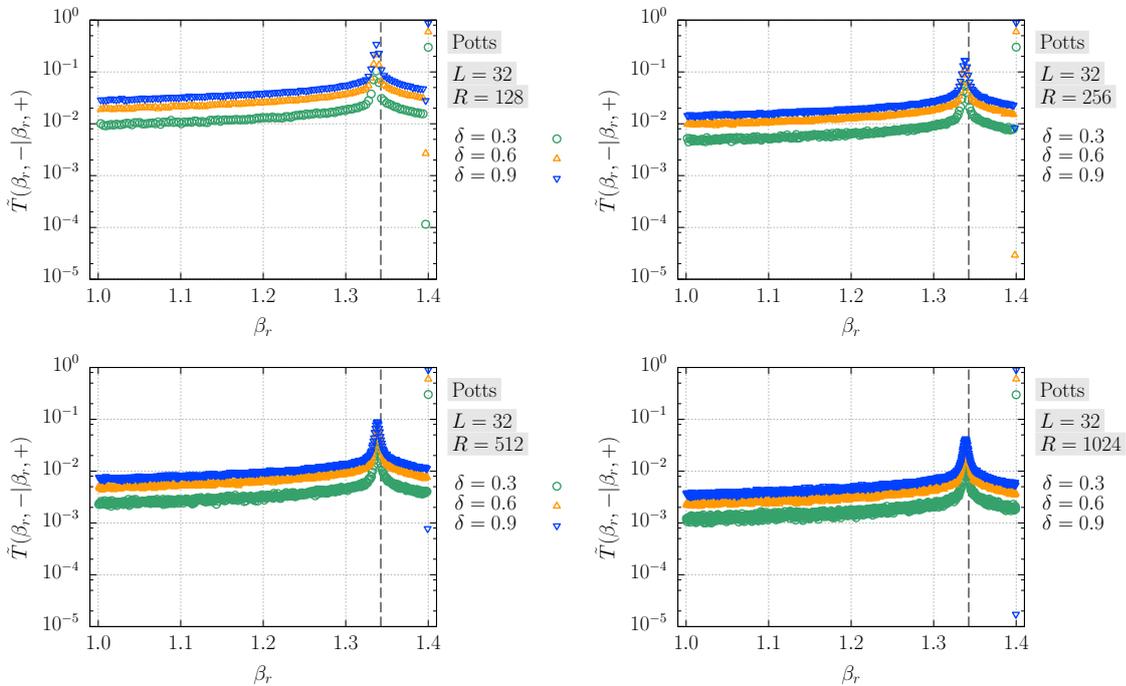


図 6.18 2次元 8 状態ポッツ模型 (Potts) に対して IST アルゴリズムを実行したときの経験遷移確率 $\tilde{T}(\beta_r, - | \beta_r, +)$.

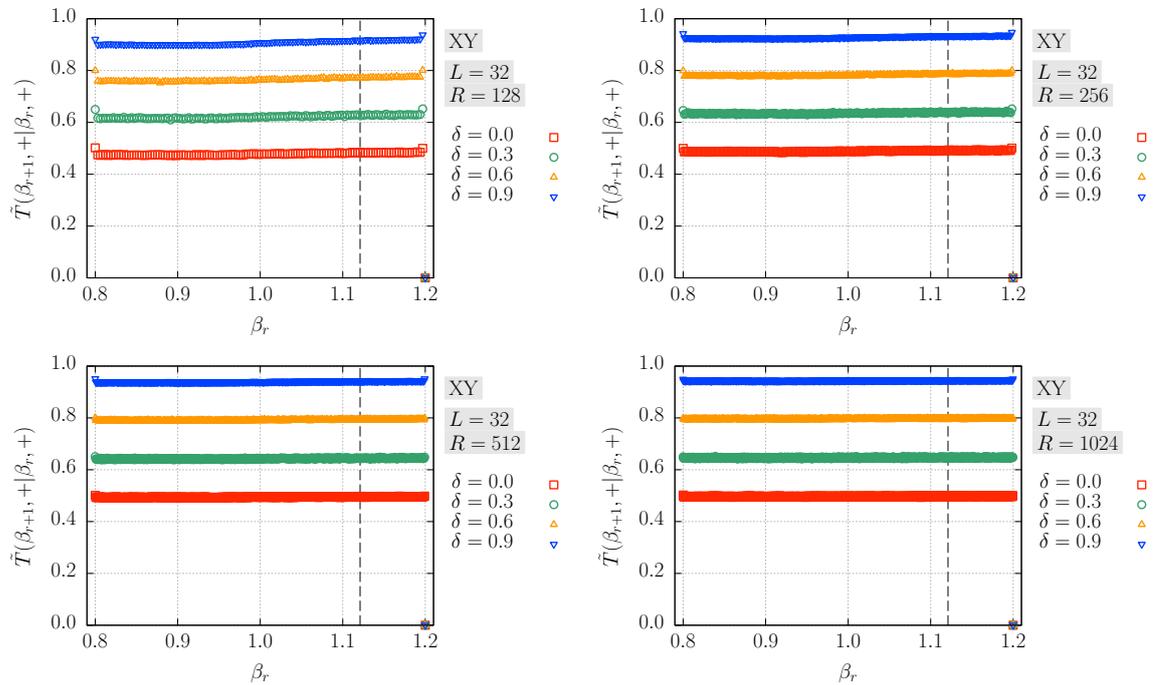


図 6.19 2次元 XY 模型 (XY) に対して IST アルゴリズムを実行したときの経験遷移確率 $\tilde{T}(\beta_{r+1}, + | \beta_r, +)$.

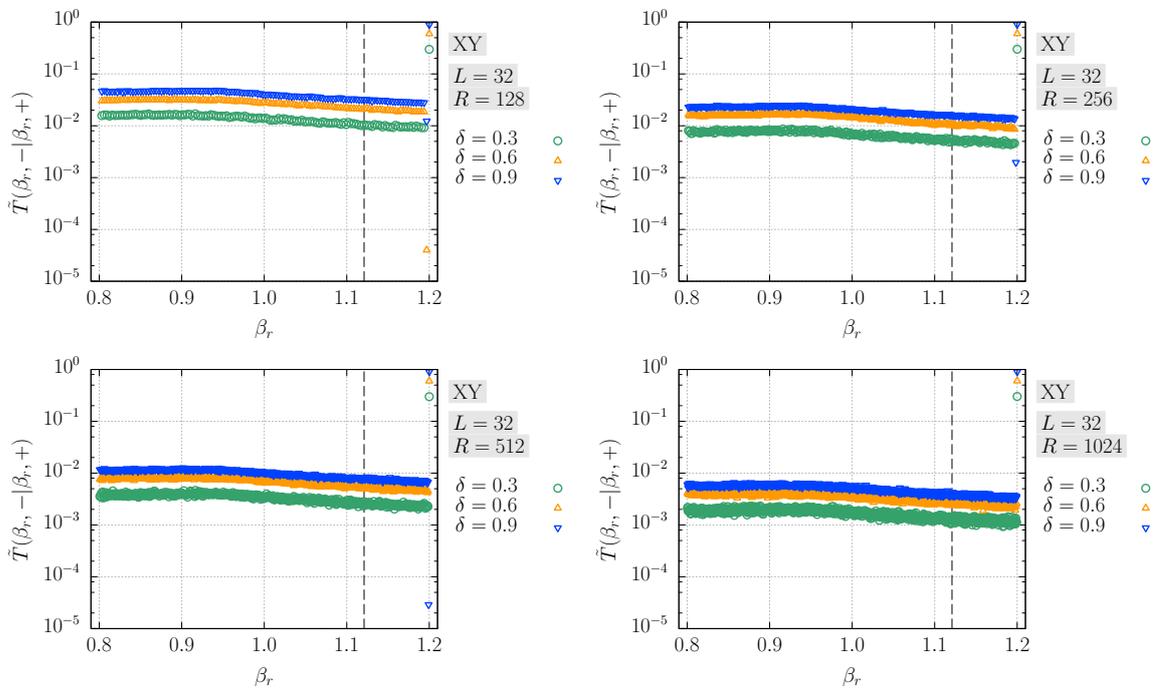


図 6.20 2次元 XY 模型 (XY) に対して IST アルゴリズムを実行したときの経験遷移確率 $\tilde{T}(\beta_r, - | \beta_r, +)$.

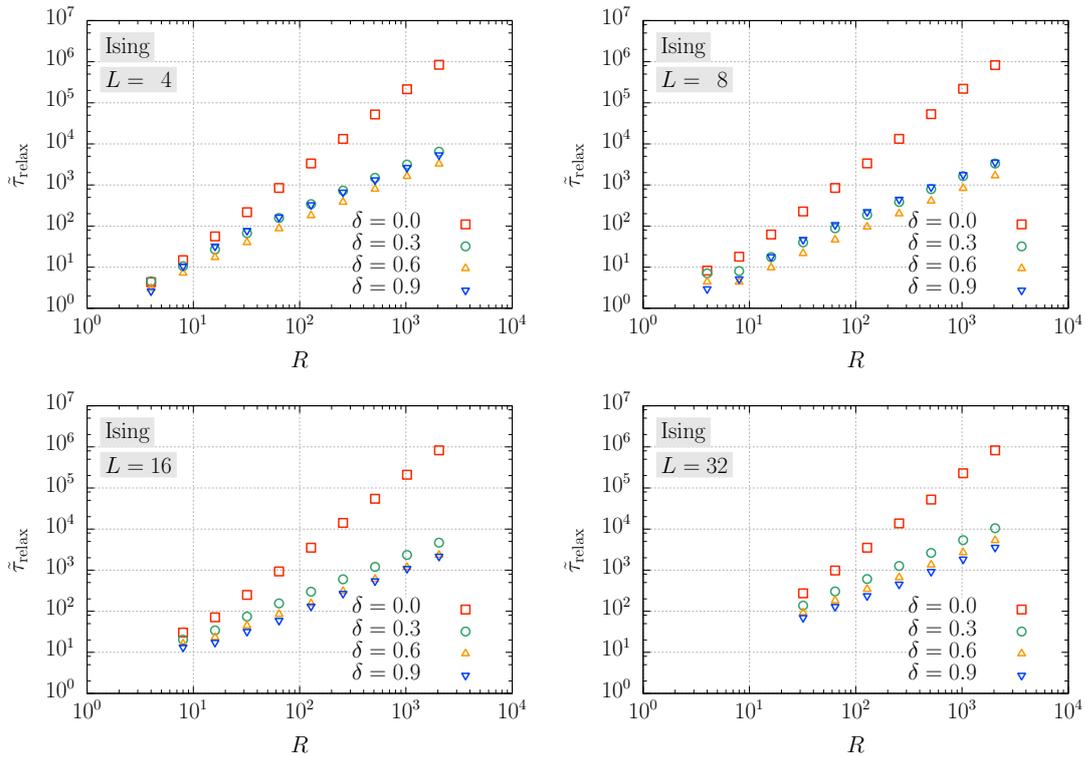


図 6.21 2次元イジング模型 (Ising) に IST アルゴリズムを適用して測定した経験遷移確率行列から定まる逆温度の緩和時間 $\tilde{\tau}_{\text{relax}}$ の逆温度個数 R 依存性.

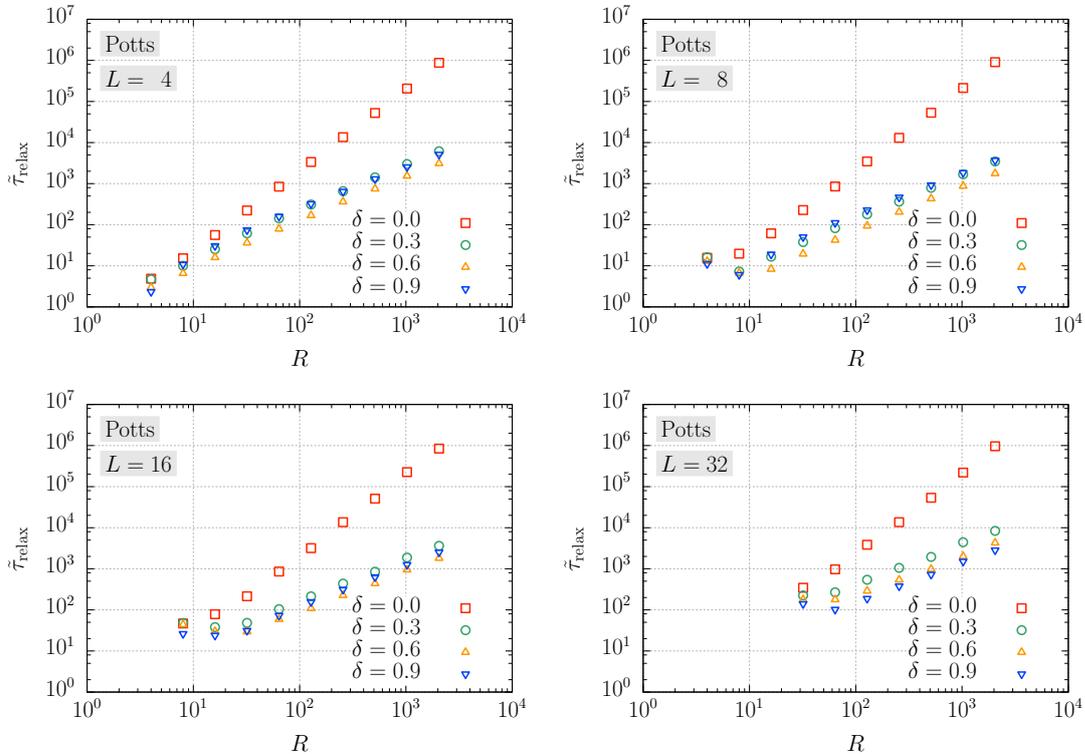


図 6.22 2次元8状態ポッツ模型 (Potts) に IST アルゴリズムを適用して測定した経験遷移確率行列から定まる逆温度の緩和時間 $\tilde{\tau}_{\text{relax}}$ の逆温度個数 R 依存性.

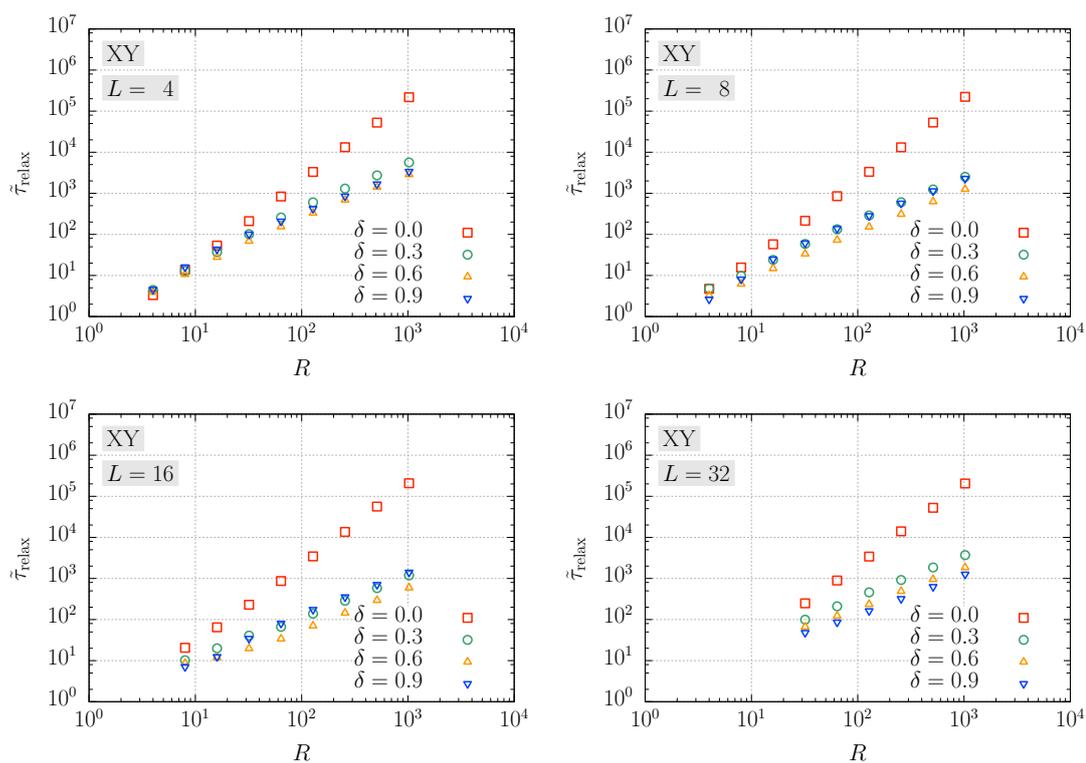


図 6.23 2次元 XY 模型 (XY) に IST アルゴリズムを適用して測定した経験遷移確率行列から定まる逆温度の緩和時間 $\tilde{\tau}_{\text{relax}}$ の逆温度個数 R 依存性.

6.4 物理量の緩和と相関

IST アルゴリズムにおける本来の目的は、目的分布 $P_{\text{IST}}(\mathbf{S}, \beta_r, \varepsilon)$ からのサンプリングや物理量の期待値評価である。前節でみたとおり、詳細つり合い条件の破れによって IST アルゴリズムにおける逆温度の緩和が加速する。このことによってより速やかに低温から高温へと遷移でき、結果として物理量の緩和も促進されると期待できる。本節ではスピン自己相関や秩序変数、エネルギーの自己相関関数、自己相関時間を測定し、詳細つり合い条件の破れが物理量の緩和や相関に及ぼす影響を評価する。

以下の数値計算では初期逆温度を $\beta^{(0)} = \beta_R$ に固定し、初期スピン配置 $\mathbf{S}^{(0)}$ は逆温度 β_R の平衡状態とした。すなわち、2次元イジング模型および XY 模型では Metropolis-Hastings アルゴリズム、2次元8状態ポッツ模型では熱浴法を逆温度 β_R で 10^4 MCS 実行して得られた配置を $\mathbf{S}^{(0)}$ とした。lifting parameter の初期値 $\varepsilon^{(0)}$ は一様ランダムとした。また各プロットおよび不確かさは独立な 128 回の試行を用いて評価した。

6.4.1 スピンの自己相関関数

まず、それぞれの模型に対してスピン自己相関関数を考える。一般にスピンの自己相関が一番遅い緩和モードに対応していると考えられており、スピン自己相関関数の緩和の加速が系全体の緩和の改善を意味すると考えることができる。

スピン自己相関関数 $C^{(n)}$ の定義は模型ごとに異なる。2次元イジング模型では

$$C^{(n)} = \left\langle \frac{1}{N} \sum_{i=1}^N S_i^{(0)} S_i^{(n)} \right\rangle, \quad (6.18)$$

2次元 q 状態ポッツ模型では

$$C^{(n)} = \left\langle \frac{1}{N} \sum_{i=1}^N \frac{q\delta(S_i^{(0)}, S_i^{(n)}) - 1}{q-1} \right\rangle, \quad (6.19)$$

2次元 XY 模型では

$$C^{(n)} = \left\langle \frac{1}{N} \sum_{i=1}^N \cos(\theta_i^{(n)} - \theta_i^{(0)}) \right\rangle \quad (6.20)$$

と定義される。ここでブラケット $\langle \dots \rangle$ はサンプル平均を表す。またスピン自己相関関数 $C^{(n)}$ から定まるスピン自己相関時間 τ_{corr} を

$$\tau_{\text{corr}} \equiv \inf \left\{ n > 0; |C^{(n)}| < 0.2 \right\} \quad (6.21)$$

と定義する. それぞれの模型に対してスピン自己相関関数 $C^{(n)}$ の時間発展およびスピン自己相関時間 τ_{corr} の逆温度個数 R 依存性を図 6.24 から図 6.29 に示す.

まず 2 次元イジング模型に注目する. 図 6.25 からスピン自己相関時間 τ_{corr} は, 詳細つり合い条件をみたま場合 ($\delta = 0$) と比較して破る場合 ($\delta \neq 0$) のほうがいずれの逆温度個数 R においても不確かさの範囲で同程度か下回っている. さらにシステムサイズ L が大きいとき, 逆温度個数 R の増加とともにスピン自己相関時間 τ_{corr} が増加していく. このとき, $\delta = 0$ での増加の割合と比べて $\delta \neq 0$ での増加の割合のほうが小さくなっている. これは前節までにみた, 詳細つり合い条件の破れによる逆温度の緩和ダイナミクスの質的变化によるものと考えられる. つまり詳細つり合い条件を破ることで逆温度の緩和が加速し, 速やかに低温から高温へと遷移することでスピンの緩和も加速されているといえる.

2 次元 8 状態ポッツ模型では図 6.11 などから, 逆温度の緩和関数 $\phi_{\beta}^{(n)}$ に 2 段階緩和がみられていた. プラトーは 1 次転移逆温度 β_t に対応する値の付近に出ており, プラトーから抜け出す時間スケールは $L = 32$ でおおむね $10^4 \sim 10^5$ MCS 程度であった. 図 6.26 でもスピン自己相関関数 $C^{(n)}$ が 2 段階緩和を示しており, その時間スケールは逆温度の緩和関数 $\phi_{\beta}^{(n)}$ のそれと同程度である. すなわち逆温度が転移点付近で著しく緩和しにくくなっているのに対応して, スピンの緩和も遅くなっている. このことは詳細つり合い条件の破れの有無によらずみえている. またスピン自己相関時間 τ_{corr} の逆温度個数 R 依存性が, イジング模型と比べて定性的には同じ振る舞いを示すものの定量的に異なる. 特にシステムサイズ L が大きいとき, 逆温度個数 R が大きくなるにつれて詳細つり合い条件の破れによるスピン自己相関時間 τ_{corr} の差がみえ始めるものの, その差が現れ始める逆温度個数 R の値はイジング模型での値よりも 10 倍以上大きい. 以上のことから詳細つり合い条件の破れは, 低温から 1 次転移点を越えるときには効果がない一方, 1 次転移点を越えたあとの緩和ダイナミクスでは逆温度の緩和を促進し, それにともなってスピンの緩和も促進すると考えられる.

最後に 2 次元 XY 模型をみてみよう. 図 6.28 では 2 次元 8 状態ポッツ模型のような 2 段階緩和はみられない. またスピン自己相関関数 $C^{(n)}$ の緩和ダイナミクスにおける詳細つり合い条件をみたま場合 ($\delta = 0$) と破る場合 ($\delta \neq 0$) の差はわずかである. さらに逆温度は図 6.13 でみたように詳細つり合い条件の破れによって緩和時間の逆温度個数 R 依存性が質的に改善していた一方で, スピン自己相関関数 τ_{corr} の逆温度個数 R 依存性は非常に小さい (図 6.29). このことは, 図 6.19 でみたように逆温度空間の遷移で KT 転移点 β_{KT} 付近に経験遷移確率が小さくなるといったことは起こらず, 速やかに高温相へと遷移できること, およびそもそもスピン自己相関関数がどの逆温度でも同じように緩和することによるものと考えられる. 結果として詳細つり合い条件の破れはほとんど影響を及ぼしていないといえる.

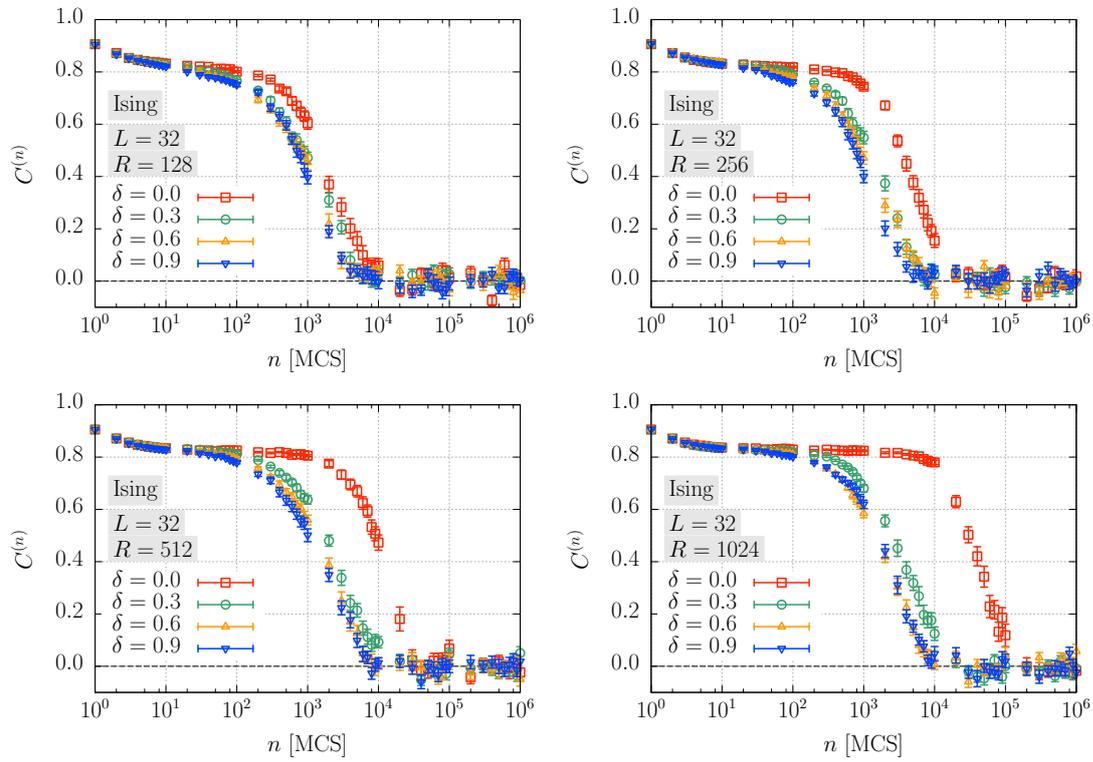


図 6.24 2次元イジング模型 (Ising) に IST アルゴリズムを実行したときのスピン自己相関関数 $C^{(n)}$ の時間発展.

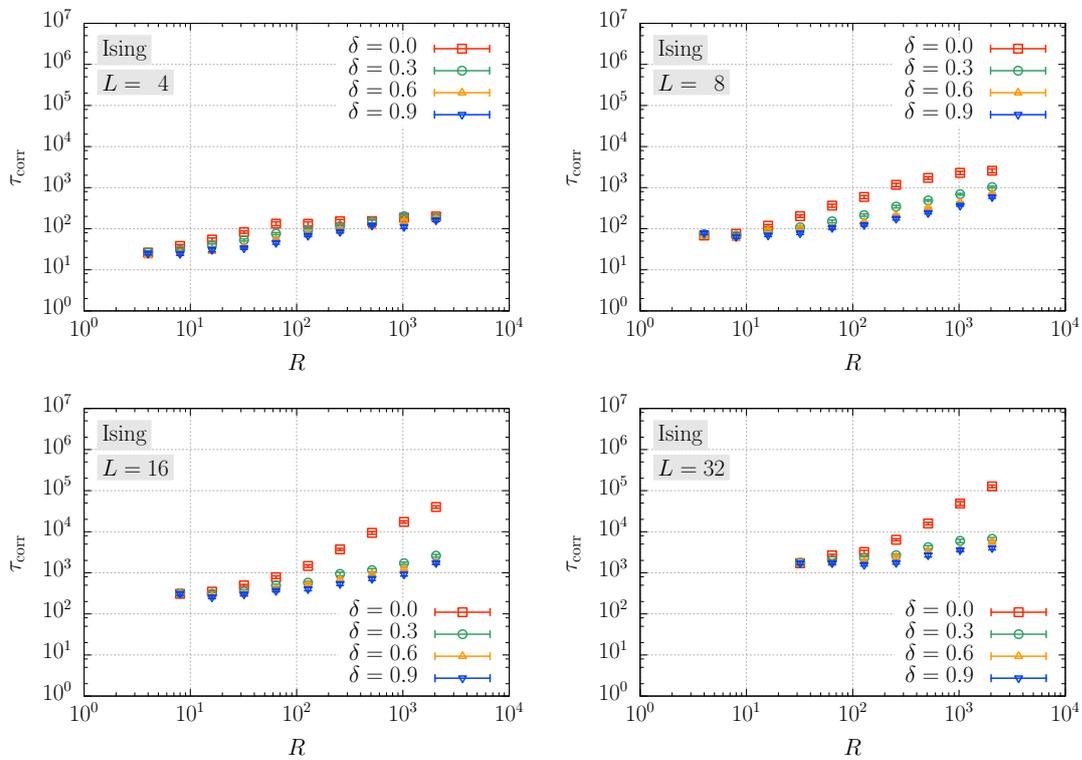


図 6.25 2次元イジング模型 (Ising) におけるスピン自己相関時間 τ_{corr} の逆温度個数 R 依存性.

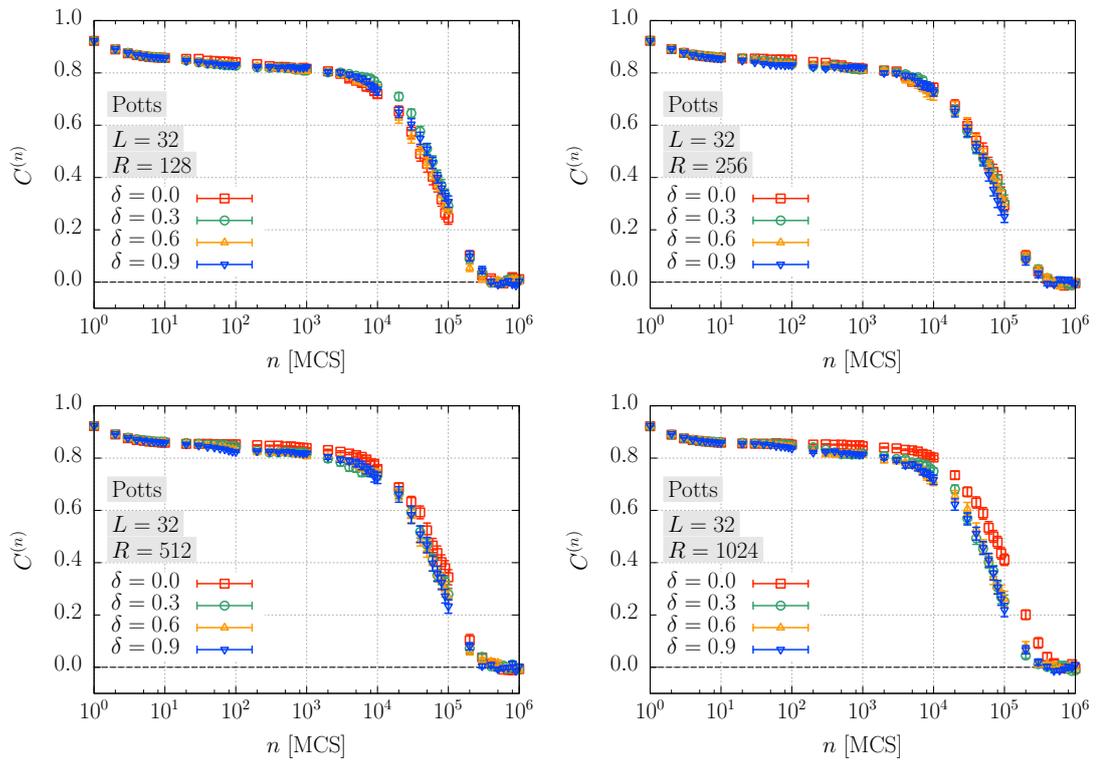


図 6.26 2次元 8 状態ポッツ模型 (Potts) に IST アルゴリズムを実行したときのスピン自己相関関数 $C^{(n)}$ の時間発展.

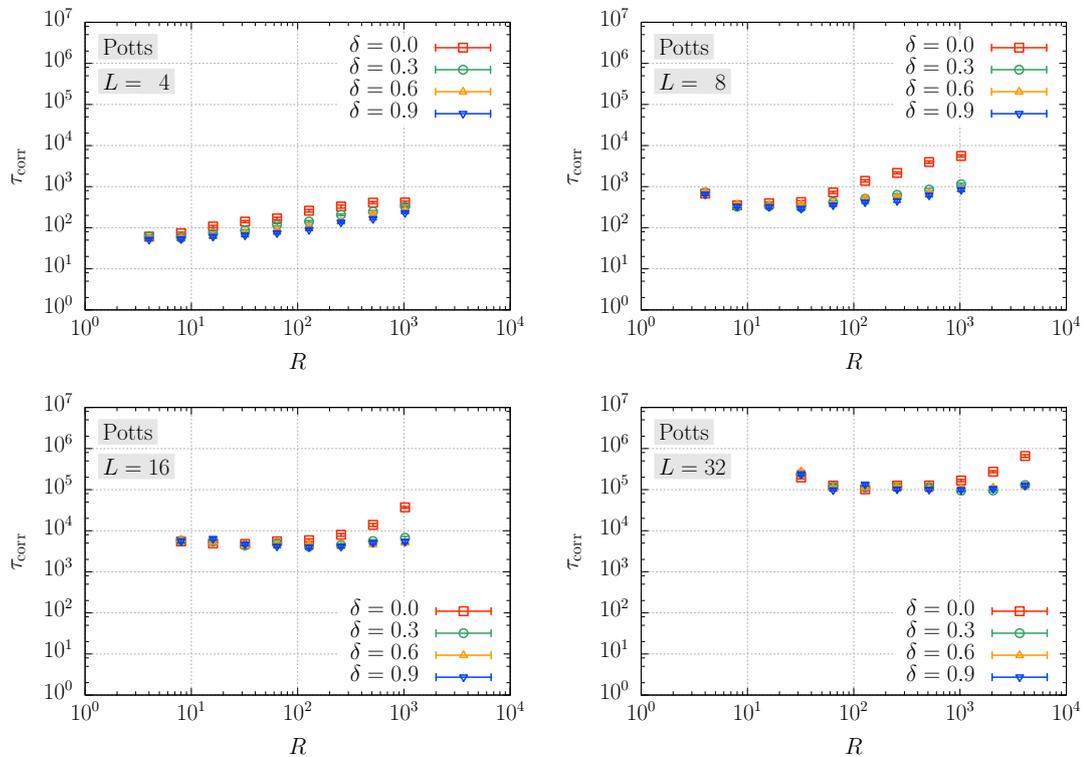


図 6.27 2次元 8 状態ポッツ模型 (Potts) におけるスピン自己相関時間 τ_{corr} の逆温度個数 R 依存性.

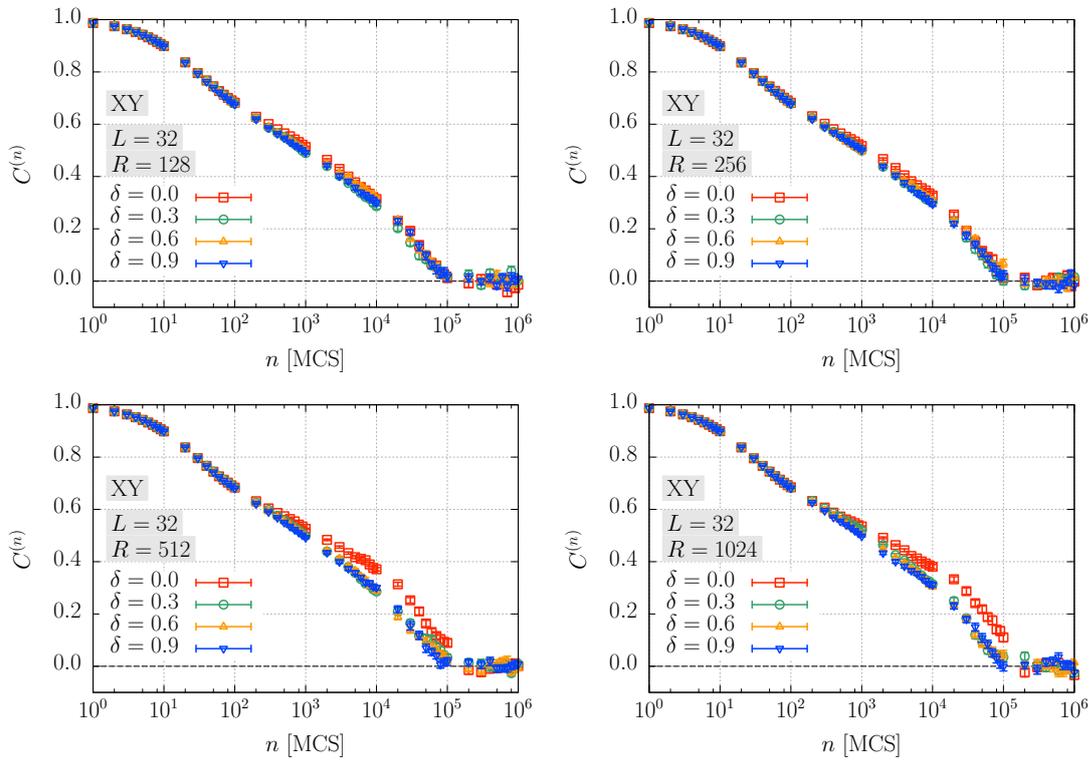


図 6.28 2次元 XY 模型 (XY) に IST アルゴリズムを実行したときのスピン自己相関関数 $C^{(n)}$ の時間発展.

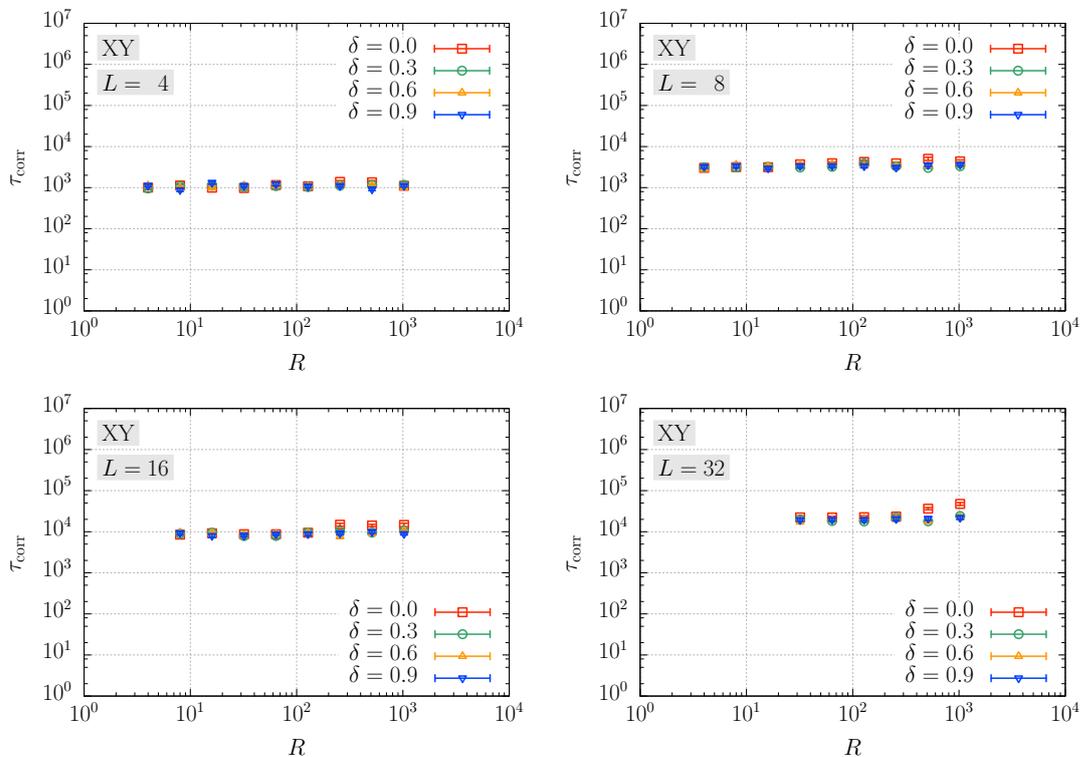


図 6.29 2次元 XY 模型 (XY) におけるスピン自己相関時間 τ_{corr} の逆温度個数 R 依存性.

6.4.2 秩序変数の自己相関関数

次に秩序変数の緩和ダイナミクスをみるために、秩序変数の自己相関関数を評価する。強磁性秩序に対応する物理量 m を、2次元イジング模型に対して

$$m = \frac{1}{N} \sum_{i=1}^N S_i, \quad (6.22)$$

2次元 q 状態ポッツ模型に対して

$$m = \frac{q-1}{q} \|\mathbf{M}\|_2^2, \quad (6.23)$$

2次元 XY 模型に対して

$$m = \frac{1}{N} \sum_{i=1}^N \cos \theta_i \quad (6.24)$$

と定義する。ここで $\|\cdot\|_2$ は L_2 -ノルムを表し、

$$\mathbf{M} = (M_0, M_1, \dots, M_{q-1}), \quad (6.25)$$

$$M_\alpha = \frac{1}{N} \sum_{i=1}^N \frac{q\delta(S_i, \alpha) - 1}{q-1}, \quad \alpha = 0, 1, \dots, q-1 \quad (6.26)$$

である。これらに対して、秩序変数の自己相関関数 $C_m^{(n)}$ を

$$C_m^{(n)} \equiv \frac{\langle m^{(0)} m^{(n)} \rangle - \langle m^{(0)} \rangle \langle m \rangle_{\text{eq}}}{\langle (m^{(0)})^2 \rangle - \langle m^{(0)} \rangle \langle m \rangle_{\text{eq}}} \quad (6.27)$$

と定義する。ここでブラケット $\langle \dots \rangle$ はサンプル平均を表し、ブラケット $\langle \dots \rangle_{\text{eq}}$ は目的分布 $P_{\text{IST}}(\mathbf{S}, \beta_r, \varepsilon)$ による期待値を表す。さらに秩序変数の自己相関時間 $\tau_{\text{corr},m}$ を

$$\tau_{\text{corr},m} \equiv \inf \left\{ n > 0; |C_m^{(n)}| < 0.2 \right\} \quad (6.28)$$

と定義する。

それぞれの模型に対して秩序変数の自己相関関数 $C_m^{(n)}$ の時間発展および秩序変数の自己相関時間 $\tau_{\text{corr},m}$ の逆温度個数 R 依存性を図 6.30 から図 6.35 に示す。いずれの模型に対しても、磁化密度の自己相関関数 $C_m^{(n)}$ やその自己相関時間 $\tau_{\text{corr},m}$ はスピン自己相関関数 $C^{(n)}$ とスピン自己相関時間 τ_{corr} とそれぞれ同様の振る舞いを示している。すなわち、2次元イジング模型と2次元8状態ポッツ模型では詳細つり合い条件の破れによる緩和の加速がみられるのに対して、2次元 XY 模型では観測した逆温度個数 R の範囲では詳細つり合い条件の破れによる影響はみられなかった。

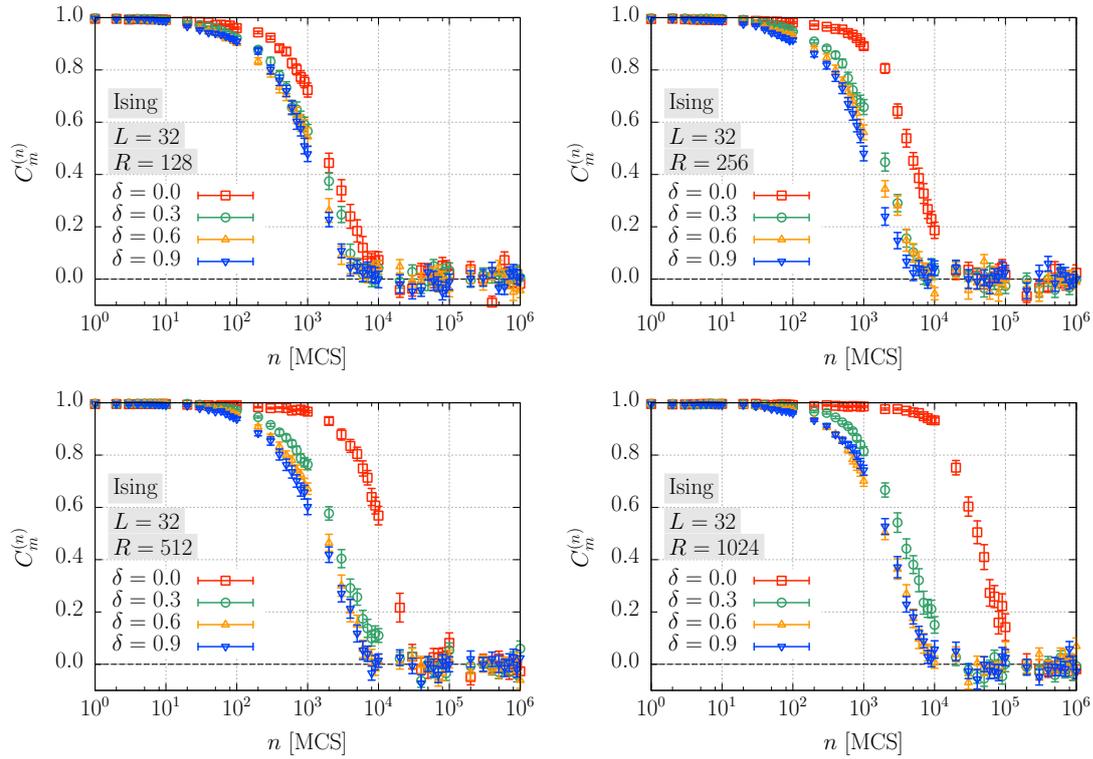


図 6.30 2次元イジング模型 (Ising) に IST アルゴリズムを実行したときの秩序変数の自己相関関数 $C_m^{(n)}$ の時間発展.

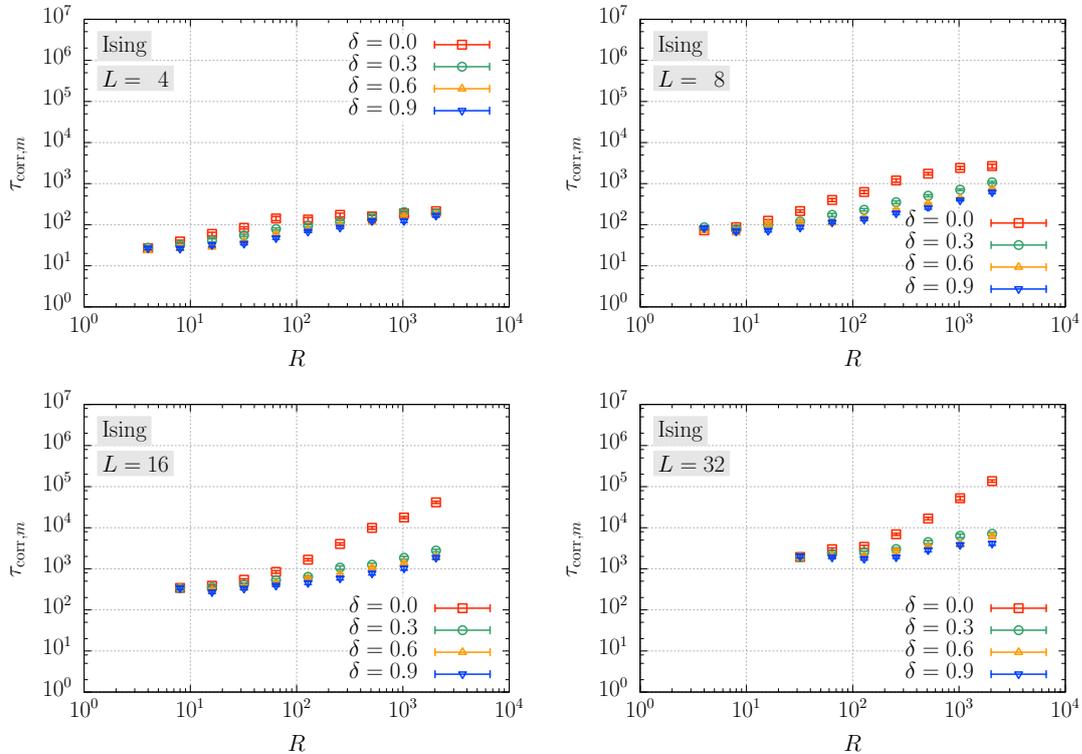


図 6.31 2次元イジング模型 (Ising) における秩序変数の自己相関時間 $\tau_{\text{corr},m}$ の逆温度個数 R 依存性.

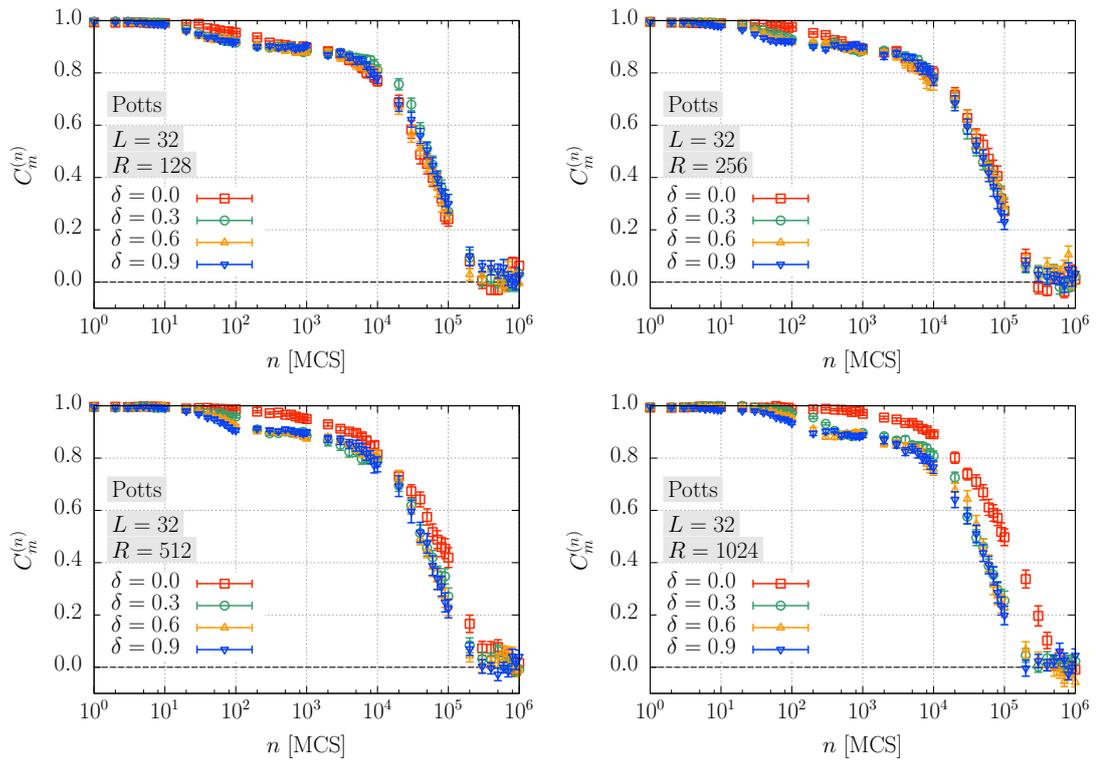


図 6.32 2次元 8 状態ポッツ模型 (Potts) に IST アルゴリズムを実行したときの秩序変数の自己相関関数 $C_m^{(n)}$ の時間発展.

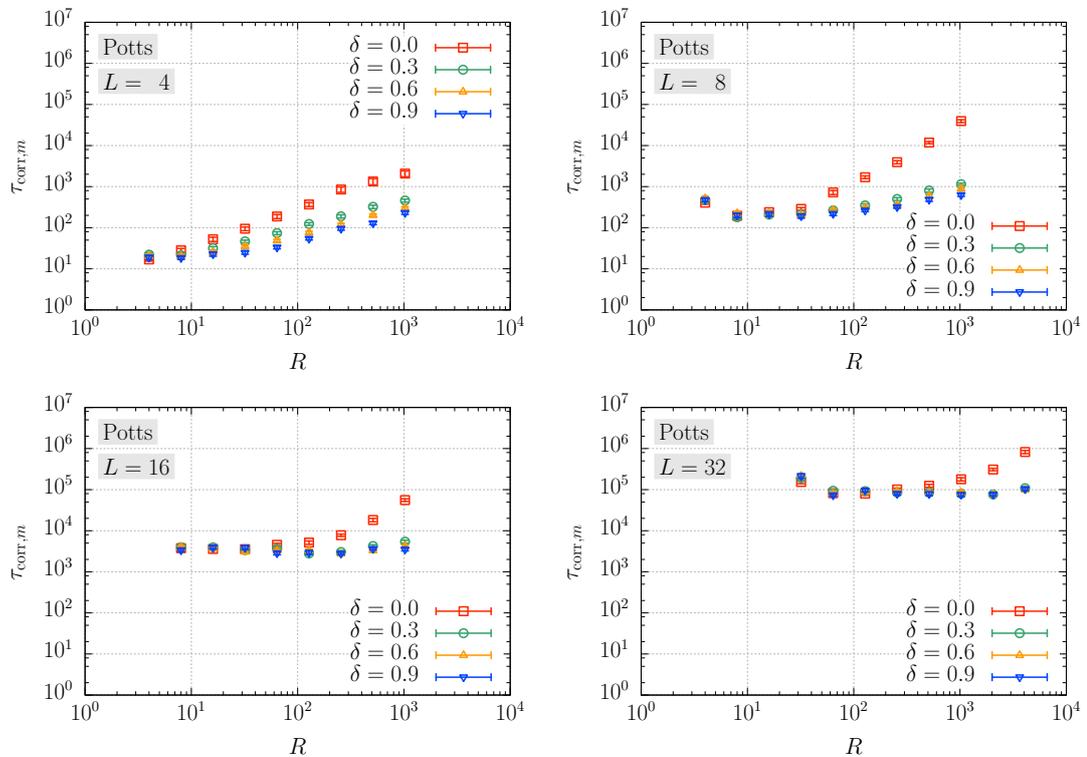


図 6.33 2次元 8 状態ポッツ模型 (Potts) における秩序変数の自己相関時間 $\tau_{\text{corr},m}$ の逆温度個数 R 依存性.

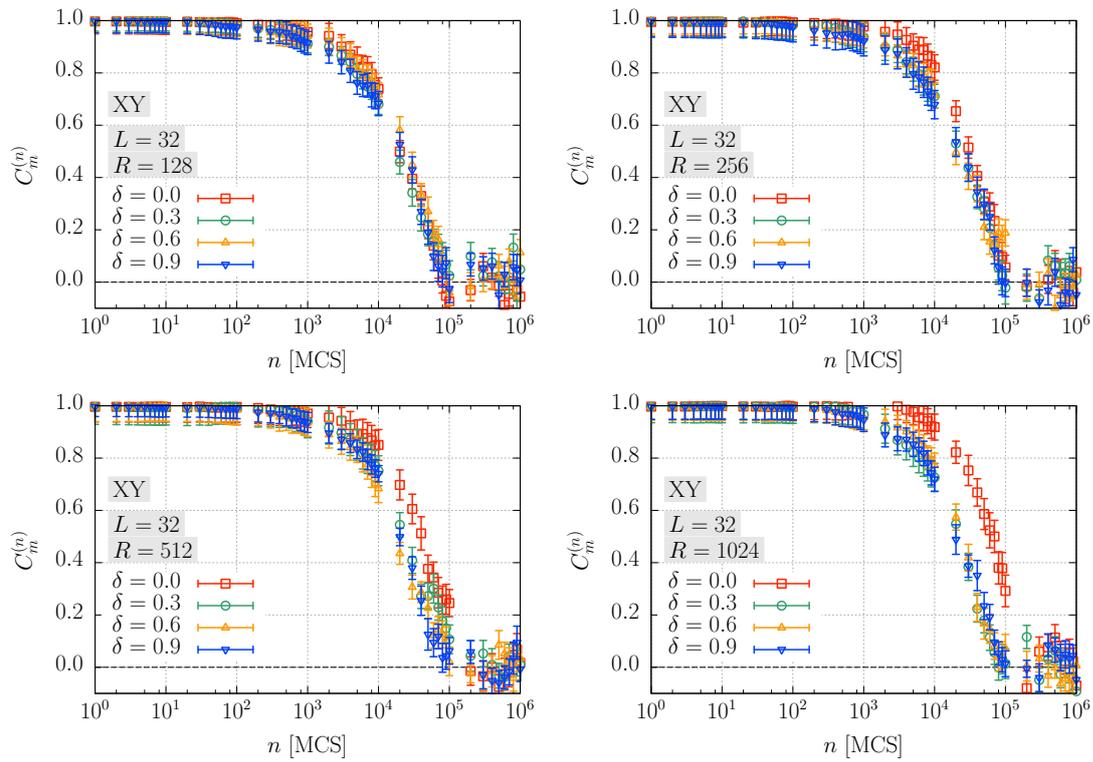


図 6.34 2次元 XY 模型 (XY) に IST アルゴリズムを実行したときの秩序変数の自己相関関数 $C_m^{(n)}$ の時間発展.

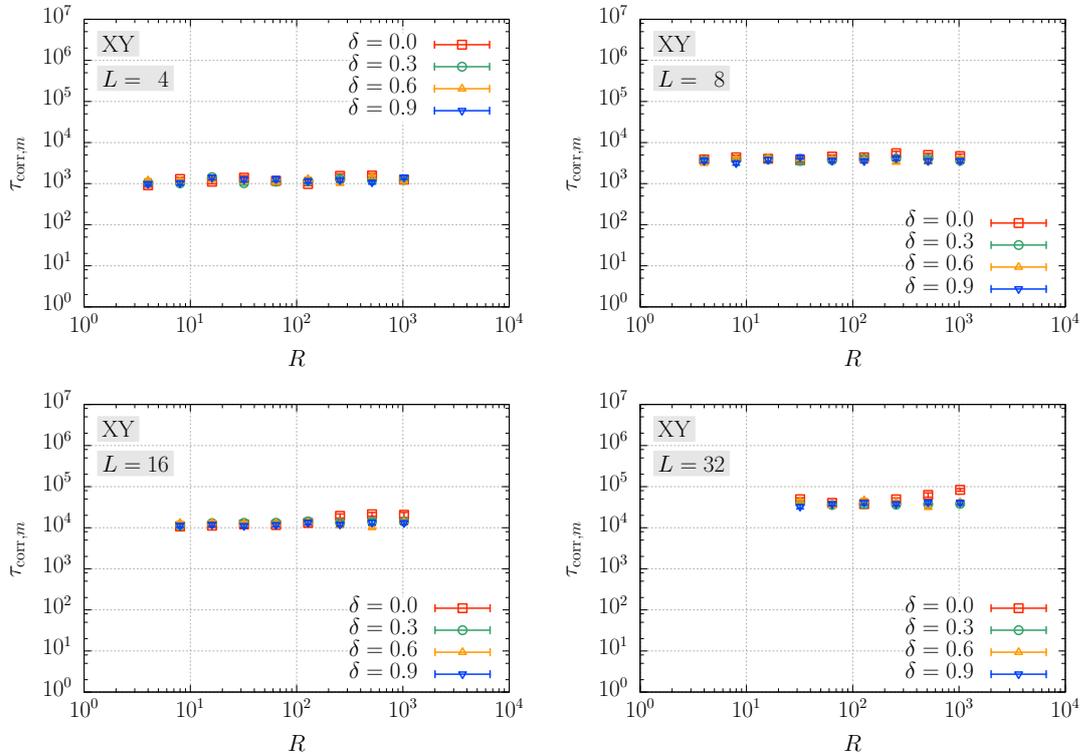


図 6.35 2次元 XY 模型 (XY) における秩序変数の自己相関時間 $\tau_{\text{corr},m}$ の逆温度個数 R 依存性.

6.4.3 エネルギーの自己相関関数

秩序変数と同様に、エネルギーについても自己相関関数と自己相関時間を考えよう。エネルギーの自己相関関数 $C_E^{(n)}$ は

$$C_E^{(n)} = \frac{\langle E^{(0)} E^{(n)} \rangle - \langle E^{(0)} \rangle \langle E \rangle_{\text{eq}}}{\langle (E^{(0)})^2 \rangle - \langle E^{(0)} \rangle \langle E \rangle_{\text{eq}}} \quad (6.29)$$

と定義される。またエネルギーの自己相関時間 $\tau_{\text{corr},E}$ は

$$\tau_{\text{corr},E} \equiv \inf \left\{ n > 0; |C_E^{(n)}| < 0.2 \right\} \quad (6.30)$$

で与えられる。

それぞれの模型に対して、エネルギーの自己相関関数 $C_E^{(n)}$ の時間発展およびエネルギーの自己相関時間 $\tau_{\text{corr},E}$ の逆温度個数 R 依存性を図 6.36 から図 6.41 に示す。2次元イジング模型や2次元8状態ポッツ模型におけるエネルギーの自己相関関数 $C_E^{(n)}$ は、スピン自己相関関数 $C^{(n)}$ や磁化密度の自己相関関数 $C_m^{(n)}$ と定性的に同様な振る舞いを示し、その自己相関時間は詳細つり合い条件の破れによって質的な変化をみせている。一方で2次元XY模型では異なる振る舞いがみえている。エネルギーの自己相関関数 $C_E^{(n)}$ やその自己相関時間 $\tau_{\text{corr},E}$ には詳細つり合い条件の破れによる変化が明確に現れている。そしてその振る舞いはシステムサイズ L が大きいとき、逆温度の緩和関数 $\phi_\beta^{(n)}$ や緩和時間 $\tau_{\text{relax},\beta}$ などと定性的に似かよっている。このことは、詳細つり合い条件の破れが逆温度の緩和を加速させることで、エネルギーの緩和もまた加速していると考えられる。

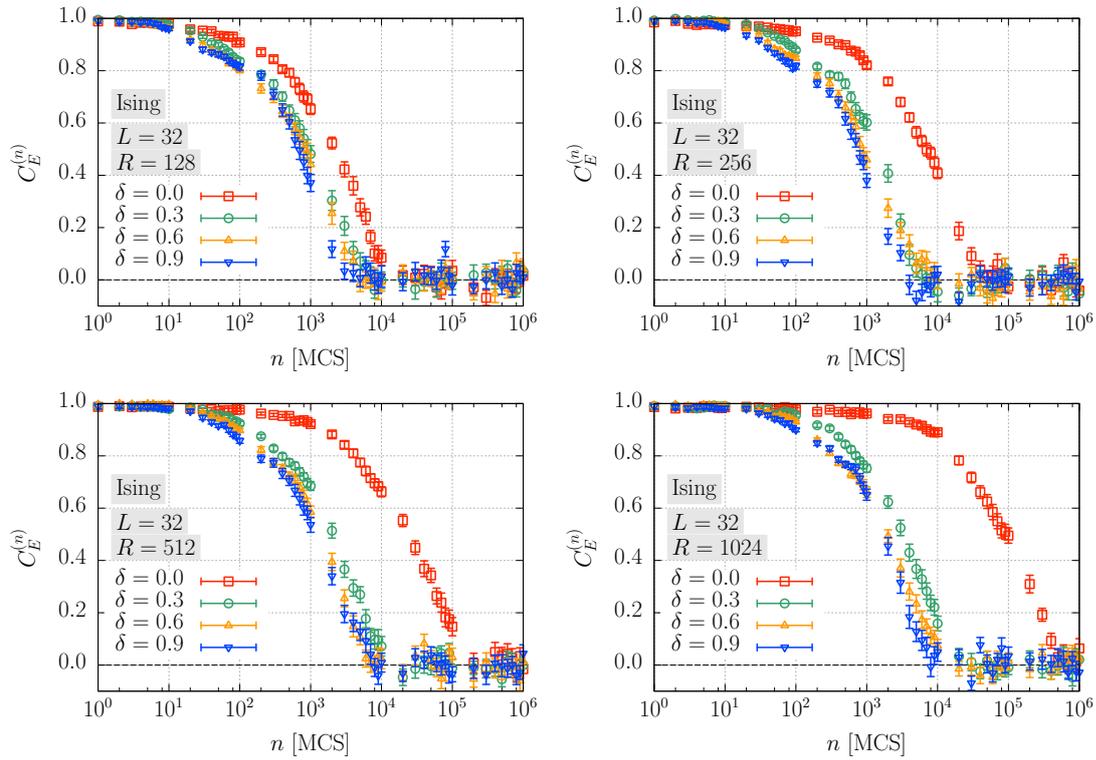


図 6.36 2次元イジング模型 (Ising) に IST アルゴリズムを実行したときのエネルギーの自己相関関数 $C_E^{(n)}$ の時間発展.

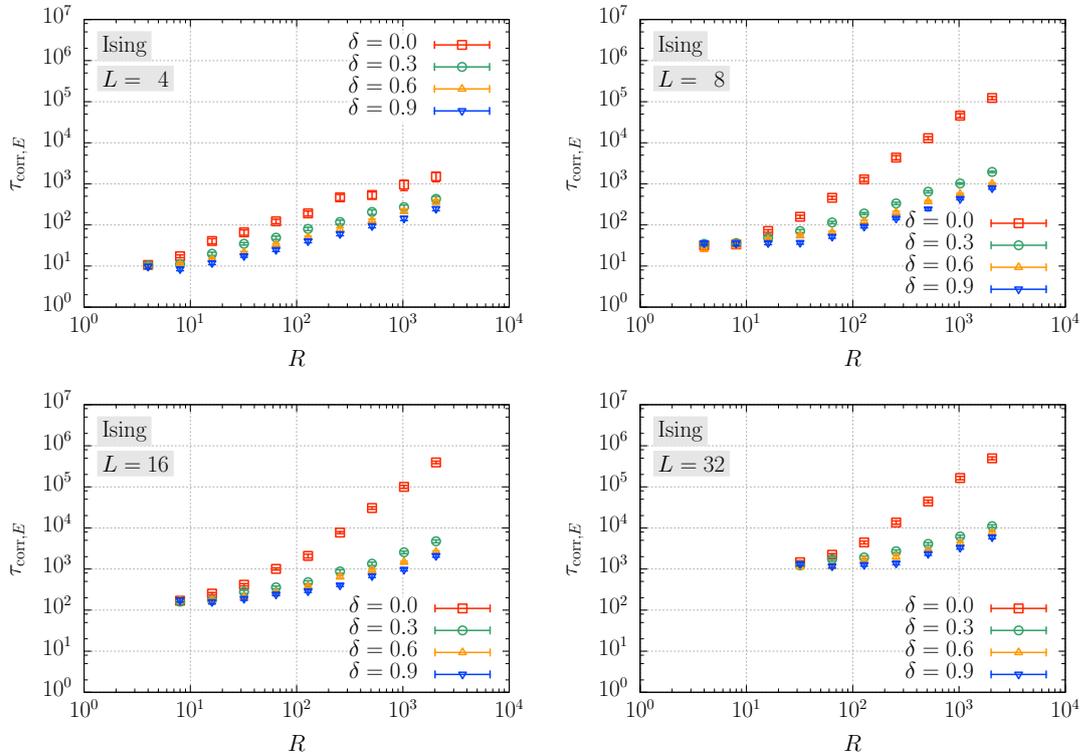


図 6.37 2次元イジング模型 (Ising) におけるエネルギーの自己相関時間 $\tau_{\text{corr},E}$ の逆温度個数 R 依存性

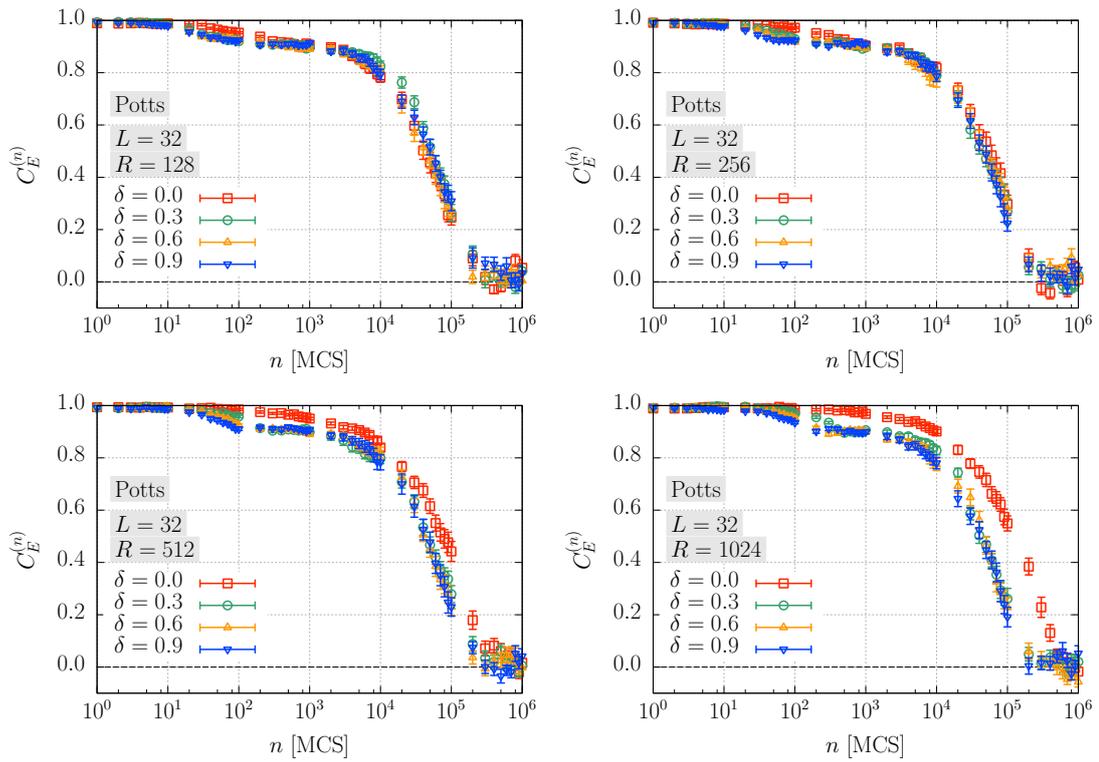


図 6.38 2次元8状態ポッツ模型 (Potts) に IST アルゴリズムを実行したときのエネルギーの自己相関関数 $C_E^{(n)}$ の時間発展.

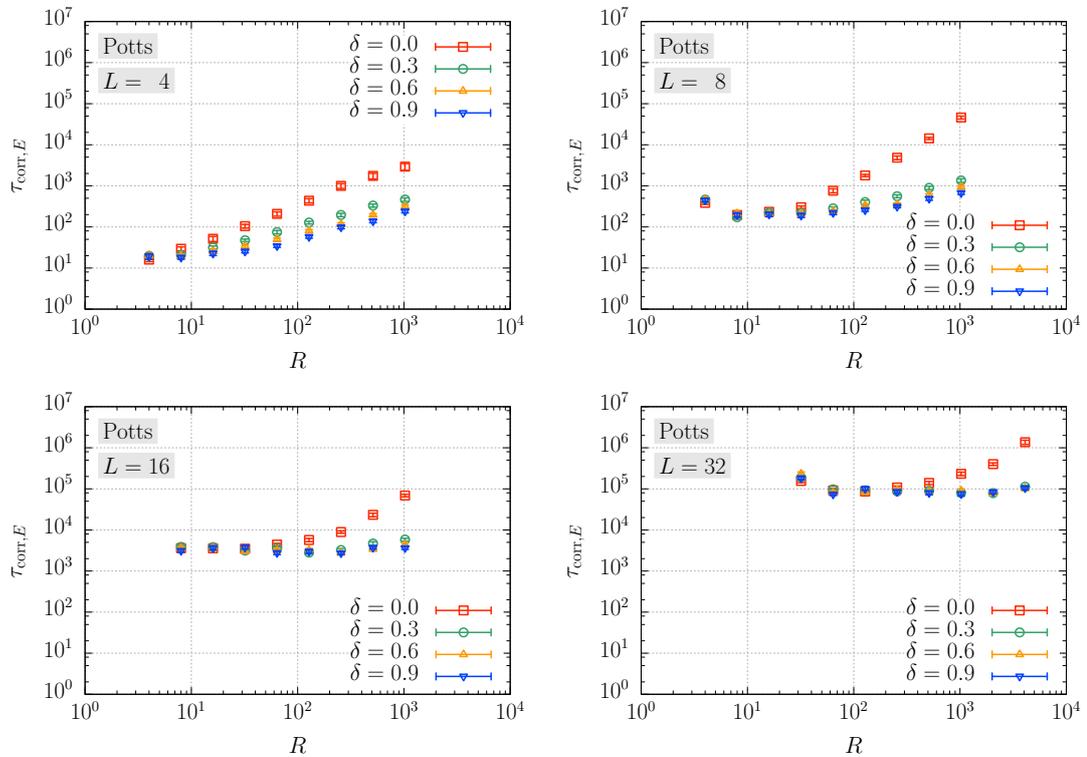


図 6.39 2次元8状態ポッツ模型 (Potts) におけるエネルギーの自己相関時間 $\tau_{\text{corr},E}$ の逆温度個数 R 依存性.

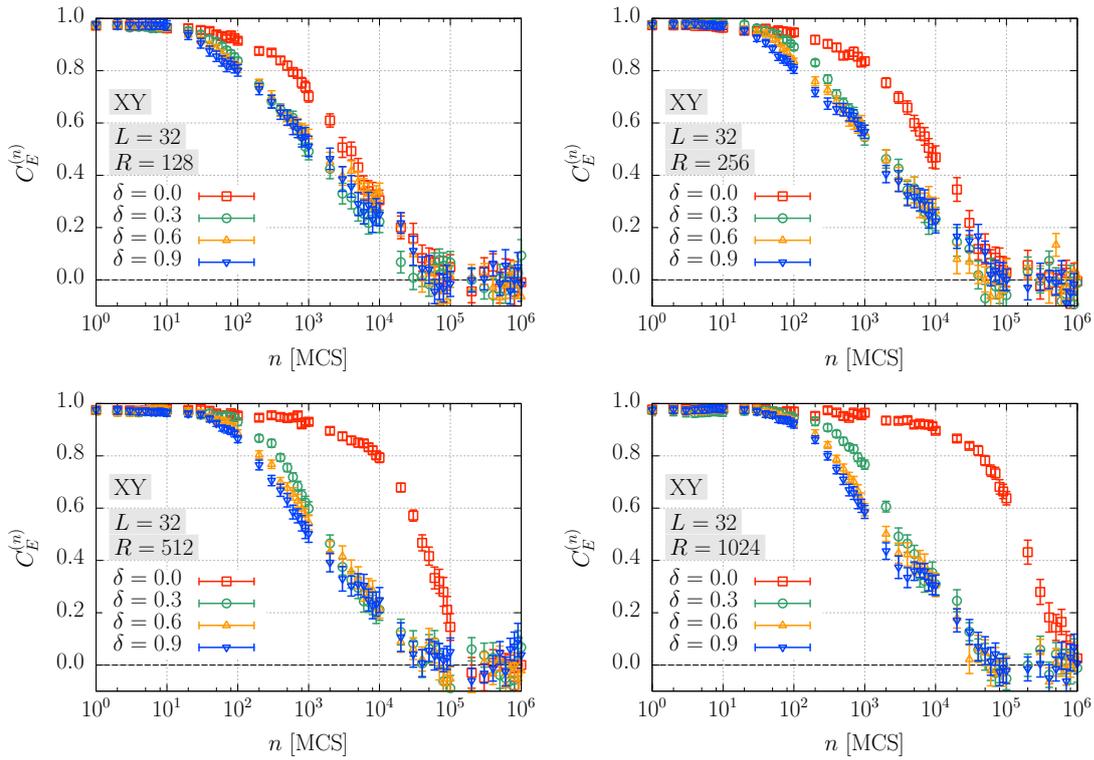


図 6.40 2次元 XY 模型 (XY) に IST アルゴリズムを実行したときのエネルギーの自己相関関数 $C_E^{(n)}$ の時間発展.

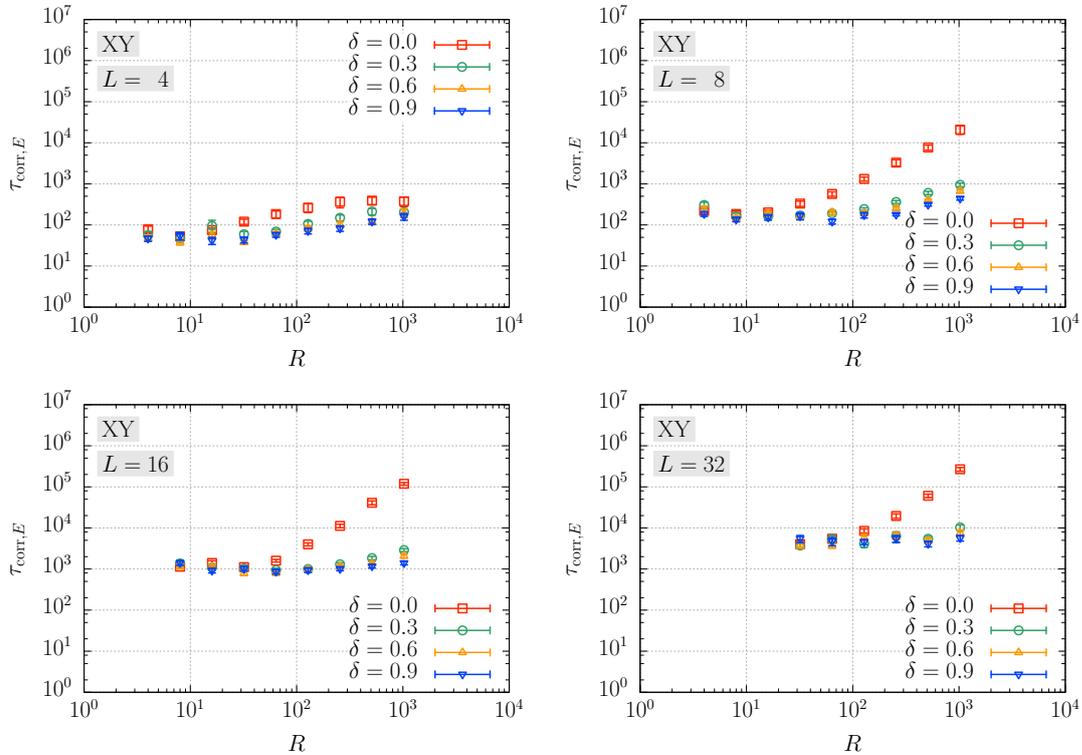


図 6.41 2次元 XY 模型 (XY) におけるエネルギーの自己相関時間 $\tau_{\text{corr},E}$ の逆温度個数 R 依存性.

6.5 本章のまとめ

本章では、前章で構成した irreversible simulated tempering (IST) アルゴリズムを 3 種類の古典スピンモデルに適用し、その性能を評価した。

まず重み因子 g_r を on-the-fly weight estimation により推定し、あわせてその精度を検証した。その結果、詳細つり合い条件を破ることで推定における逆温度空間での緩和ダイナミクスは加速され、より少ない MCS 数で逆温度空間を往復できることがわかった。しかしその場合でも逆温度空間の往復数ではなく推定に要した MCS 数によってのみ推定精度が決まり、詳細つり合い条件の破れが必ずしも重み因子の推定精度を向上させるわけではないことも明らかとなった。

次に推定した重み因子を固定した下で IST アルゴリズムを実行したときの、逆温度の緩和ダイナミクスを評価した。逆温度空間での round-trip time や緩和関数、経路遷移確率などの測定を通して、詳細つり合い条件の破れは逆温度の拡散的な緩和ダイナミクスを弾道的なダイナミクスへと質的に変化させ、逆温度の緩和を加速させることが明らかとなった。IST アルゴリズムにおける逆温度空間は、第 4 章での詳細つり合い条件の破れた 1 次元ランダムウォークと類似の構造をもっており、詳細つり合い条件の破れの効果もそこで得られた結果から理論的に説明できる。

逆温度の緩和とあわせて、スピン自己相関や磁化密度、エネルギーの自己相関を評価した。モデルによって改善の度合いは異なるものの、詳細つり合い条件の破れによってこれらの量は質的に改善しうるということが数値計算から示唆された。以上のことから、シミュレーテッドテンパリングアルゴリズムにねじれ詳細つり合い条件を用いた lifting によって詳細つり合い条件の破れを導入すると、その性能を改善できることが数値的に示された。

第 7 章

結論と今後の展望

本博士論文で議論し明らかにしたことを章ごとにまとめ、今後の展望を述べる。

7.1 本博士論文で明らかにしたこと

本博士論文は本章も含めて全 7 章から構成されている。第 1 章で導入を行った後、第 2 章および第 3 章では議論の土台として先行研究の結果をまとめている。本博士論文の主要な結果は第 4 章、第 5 章、および第 6 章で述べられている。ここでは第 4 章、第 5 章、および第 6 章の結果をいま一度まとめておく。

第 4 章 詳細つり合い条件を破る 1 次元ランダムウォーク

詳細つり合い条件の破れが MCMC 法の性能に与える影響を具体的に考察するために、1 次元リング上のランダムウォーク問題を扱った。ねじれ詳細つり合い条件を用いた lifting によって、1 次元ランダムウォークの遷移確率行列に詳細つり合い条件の破れを導入した。その結果、詳細つり合い条件の破れはいくつかのパラメタ δ , δ' , γ によって特徴づけられた。本博士論文では類似の先行研究における解析手法と異なり、対応する遷移確率行列を解析的に対角化しその固有値、固有ベクトルをすべて導出することによって緩和時間や漸近分散を評価した。特にパラメタ δ , δ' の値を制限した 2 通りの設定で解析を行った結果、そのいずれにおいても詳細つり合い条件をみだす場合と比較して、緩和時間や性能指数の最悪評価を改善できることが示された。さらにその改善は定量的なもの (定数倍程度の改善) にとどまらず、緩和時間の状態数依存性が変化するという質的な改善をもたらすことも明らかとなった。一方で詳細つり合い条件の破れが性能を悪化させるパラメタ領域の存在も明らかとなり、一般に詳細つり合い条件の破れが MCMC 法の性能改善をもたらすわけではないことも示唆された。

第5章 Irreversible simulated tempering アルゴリズム

シミュレーテッドテンパリングの逆温度更新プロセスに irreversible Metropolis–Hastings アルゴリズムを用いて詳細つり合い条件の破れを導入し, irreversible simulated tempering (IST) アルゴリズムを構成した. 特に逆温度空間での提案確率を lifting parameter に依存させ, 詳細つり合い条件の破れをパラメタ δ で表現した. あわせてシミュレーテッドテンパリングを実行する上で必要となる重み因子について, その推定手法である on-the-fly weight estimation に IST アルゴリズムを応用した.

第6章 Irreversible simulated tempering アルゴリズムの性能評価

前章で提案した IST アルゴリズムにおいて詳細つり合い条件の破れがもたらす影響を考察するため, 2次元イジング模型, 2次元8状態ポッツ模型, 2次元XY模型のそれぞれに対して IST アルゴリズムを実行し, その性能を数値的に評価した. on-the-fly weight estimation による重み因子の推定では詳細つり合い条件を破ることで逆温度空間での緩和ダイナミクスが加速した. その一方で重み因子の推定精度は推定に要した MCS 数に大きく依存し, 詳細つり合い条件の破れによる影響は見いだせなかった. 次に IST アルゴリズムにおける逆温度の緩和ダイナミクスを評価するため, 逆温度空間での round-trip time, 緩和関数, 経験遷移確率を測定した. これらの量から逆温度の緩和時間を見積もった. その結果, 詳細つり合い条件の破れは逆温度の緩和を加速させることがわかった. 特に緩和時間の逆温度個数 R 依存性が $\delta = 0$ では漸近的に $O(R^2)$ である一方, $\delta \neq 0$ では漸近的に $O(R)$ となることを見出した. この結果は詳細つり合い条件の破れが逆温度の緩和を質的に改善することを示唆している. これにともなって, スピンの自己相関時間や秩序変数, エネルギーの自己相関時間にも改善がみられた. つまり IST アルゴリズムは従来のシミュレーテッドテンパリングよりも性能が向上していることが数値的に示された.

7.2 今後の展望

本博士論文での議論をふまえて, 今後の展望を述べる. また本博士論文で扱いきれなかったいくつかの事項を指摘し, 論点を明らかにした上で今後の課題としたい.

詳細つり合い条件の破れと緩和時間・性能指数の最悪評価

詳細つり合い条件の破れた MCMC 法は, その additive reversibilization によって特徴づけられる可逆な MCMC 法と比較して緩和時間 (第2固有値) や漸近分散が改善す

ることが解析的に示されていることを第3章で議論した。第4章では、ねじれ詳細つり合い条件を用いた lifting によって詳細つり合い条件を破っても性能が改善しない例があることを見出した。この例は、「ある可逆な MCMC 法にねじれ詳細つり合い条件を用いた lifting をほどこして irreversible Metropolis–Hastings (IMH) アルゴリズムを構成したとき、IMH アルゴリズムの性能は元となった可逆な MCMC 法の性能と比較して改善することもあれば悪化することもある」ことを示している。つまり、additive reversibilization での議論のように一般的なことは言えないことを示唆している。しかしこれはねじれ詳細つり合い条件を用いた lifting での結果であり、Suwa–Todo アルゴリズムに代表される他の非可逆な MCMC 法にまで成り立つかどうかは明らかではない。つまり今回の結果は、詳細つり合い条件の破り方を特定した枠組みで緩和時間や性能指数の最悪評価について一般的な主張が成り立つことを妨げるものではない。今後はねじれ詳細つり合い条件を用いた lifting 以外の詳細つり合い条件の破り方についても、(additive reversibilization ではなく) 比較すべき可逆な MCMC 法との性能比較を行う必要がある。また理論的にも詳細つり合い条件の破り方をある程度制限した上で何が成り立つのかを模索していく必要がある。

詳細つり合い条件を「どれくらい」破るとよいのかについても議論すべきだろう。第2章では、詳細つり合い条件をみたく遷移確率行列は実対称行列と相似であり、すべての固有値が実固有値であることをみた。詳細つり合い条件を破ることは、遷移確率行列の定常分布を変えないまま反対称行列を加えることに相当している*1。一般に反対称行列の各成分の絶対値は、「遷移確率行列の成分が非負であり、エルゴード条件を破らない範囲」で大きくとることができる。詳細つり合い条件の破れが漸近分散を改善することを示した文献 [24] では、詳細つり合い条件を破れば破るほど改善の度合いが大きくなることが示されている。すなわち、遷移確率行列のいずれかの成分がゼロとなるまで反対称成分を大きくとることが、漸近分散の観点では一般に最良である。しかし一方で、緩和時間の観点では詳細つり合い条件を「ほどよく」破ったほうが最良となることがありうる。たとえば第4章のランダムウォークの解析では、lifting parameter ε を flip する遷移確率 γ を固定して詳細つり合い条件の破れを特徴づけるパラメタ δ, δ' を変えたとき、設定 (a) $\delta = \delta'$ において緩和時間が最小となるのは、詳細つり合い条件を最も大きく破る $|\delta| = 1$ のときではなく、 $\gamma = (1/4)|\delta| \sin(2\pi/\Omega)$ をみたくパラメタ δ の値を選んだときであった。このパラメタの値は第2固有値が縮退しちょうど複素固有値となり始める値でもある。このことは詳細つり合い条件の破れが第2固有値を改善することを議論した文献 [44] でも示唆されていたことでもあった。この値以上にパラメタ δ の絶対値を大きくすると第2固有値が複素固有値となり、第2固有値の絶対値から定まる緩和時間はむしろ大きくなってしま

*1 この反対称行列は第3章で言及した vorticity matrix [25] に対応する。

う。このように、最適な詳細つり合い条件の破れ具合は緩和時間と漸近分散で一般に一致しない。しかしこの結果は緩和時間・漸近分散のどちらの観点でも最適である詳細つり合い条件の破り方が存在することを否定するものではない。今後はどのように詳細つり合い条件を破れば緩和時間と漸近分散をともに改善できるのか、これまでに提案されてきた破り方も含めて数値的・理論的に検討する必要がある。

IST アルゴリズムにおける逆温度セットの選び方

第 6 章では、IST アルゴリズムの性能を数値的に評価した。その際、逆温度がとりうる値を簡単のため等間隔に設定していた。逆温度がとりうる値の個数が少ないとき、等間隔な選び方が最適ではないことは経験遷移確率の測定から明らかである。特に 1 次転移や 2 次転移を示す逆温度の近傍では逆温度の遷移確率が著しく小さくなり (図 6.15, 6.17), 逆温度空間での緩和ダイナミクスを悪化させている。逆温度の遷移確率は逆温度間隔に対して指数関数的に小さくなるため、逆温度間隔を転移点付近で小さく、転移点から離れた領域で大きくとることで、逆温度の遷移確率をより一様にできると考えられる。従来のシミュレーテッドテンパリングでも、逆温度間隔を調節することで逆温度の遷移を促している例もある [61, 85]。

遷移確率が一律な 1 次元ランダムウォークでは詳細つり合い条件の破れによって緩和が加速しうることを第 4 章でみた。IST アルゴリズムにおいても、逆温度の遷移確率が一律となるように逆温度間隔を調節できれば、より逆温度の緩和が促進されると期待される。しかしその際、逆温度間隔を調節するための計算コストが問題になる。たとえば文献 [61] の方法で逆温度間隔を調節するには、一般に

- はじめに用意した逆温度セットで重み因子を推定する
- 推定した重み因子で IST アルゴリズムを実行し、round-trip time を測定する
- 測定した round-trip time を用いて逆温度間隔を調節する

という手続きが必要となる。ここでは

[方策 1] 逆温度の個数が少ないまま、逆温度間隔を調節してシミュレーションする

[方策 2] 逆温度を十分な個数使い、等間隔に並べてシミュレーションする

というふたつの方策を考えることにする。このとき [方策 1] に要する計算コストと [方策 2] で逆温度の個数を増やしたことによる緩和時間の増加を比較すると、[方策 2] のほうが全体として少ない計算コストで済む可能性がある。本博士論文の結果によれば、IST アルゴリズムにおける詳細つり合い条件の破れは緩和時間・自己相関時間の漸近的な逆温度個数 R 依存性を $O(R^2)$ から $O(R)$ へと質的に改善する。このことから、従来のシミュ

レーテッドテンパリングでは [方策 2] を採ると緩和時間が $O(R^2)$ で増加してしまう一方, IST アルゴリズムで詳細つり合い条件を破ることで, 緩和時間の増加を $O(R)$ に抑えられる. よって IST アルゴリズムでは [方策 2] が最善となる場合が多くなると考えられる.

IST アルゴリズムの他のモデルへの応用

もともとシミュレーテッドテンパリングは適用範囲が広く, カノニカル分布で記述されるモデルであって従来の MCMC 法を実行できる系ならば適用可能である. IST アルゴリズムはシミュレーテッドテンパリングの逆温度更新プロセスに手を加えただけであるので, シミュレーテッドテンパリングが備えている汎用性をそのまま引き継いでいる. たとえばスピングラスやタンパク質の折り畳み問題などより興味のある系へ IST アルゴリズムを応用することで, シミュレーテッドテンパリングにおける詳細つり合い条件の破れが系の緩和を加速させ, これらの系における平衡状態の理解に貢献できると考えている.

本博士論文ではねじれ詳細つり合い条件という特定の詳細つり合い条件の破り方に注目し, 1次元ランダムウォークという特定のモデルに対する解析から, 詳細つり合い条件の破れがもたらす効果について一定の知見を得た. またその知見をシミュレーテッドテンパリングの性能改善に活かし, 従来よりも効率的なアルゴリズムの提案に成功した. しかしより一般のモデルに対して, より一般に詳細つり合い条件を破った MCMC 法がどのような性能を発揮するのかについては, まだ十分に理解できたとはいえない. 本博士論文を議論の土台として, 詳細つり合い条件の破れがもたらす効果を理論的により深く理解するための研究が今後より一層進展することを願って筆をおく.

付録 A

行列のスペクトル分解

本付録では, 第 4 章に現れる行列 $J_\Omega(a, b)$ および行列 $\tilde{J}_{2\Omega}(a, b; c)$ のすべての固有値および正規直交条件と完全性条件をみたす固有ベクトルの組を与え, そのスペクトル分解を書き下す. 特に行列 $\tilde{J}_{2\Omega}(a, b; c)$ ではパラメタ a, b, c の値に応じて対角化可能でない場合がありうる. そのときでも Jordan 標準形を考慮することで, スペクトル分解に対応する式を導けることをみる.

本付録の内容は文献 [50] の APPENDIX B および APPENDIX C として出版されている.

A.1 行列 $J_\Omega(a, b)$ のスペクトル分解

実数 a, b に対して, Ω 次正方行列 $J_\Omega(a, b)$ は

$$J_\Omega(a, b) \equiv \begin{pmatrix} 0 & a & & & b \\ b & 0 & a & & \\ & \ddots & \ddots & \ddots & \\ & & b & 0 & a \\ a & & & b & 0 \end{pmatrix} \in \mathbb{R}^{\Omega \times \Omega} \quad (\text{A.1})$$

と定義される.

A.1.1 $a \neq b$ のとき

$a \neq b$ のとき, 行列 $J_\Omega(a, b)$ の固有値 $\mu_k(a, b) \in \mathbb{C}$ は $k = 1, 2, \dots, \Omega$ に対して

$$\begin{aligned} \mu_k(a, b) &= ae^{i\theta_k} + be^{-i\theta_k} \\ &= (a + b) \cos \theta_k + i(a - b) \sin \theta_k \end{aligned} \quad (\text{A.2})$$

で与えられる. ここで i は虚数単位であり, $\theta_k \equiv 2\pi(k-1)/\Omega$ である. 固有値 $\mu_k(a, b)$ の重複度はすべて 1 である. 固有値 $\mu_k(a, b)$ に属する左固有ベクトル $\mathbf{u}_k \in \mathbb{C}^\Omega$ および右固有ベクトル $\mathbf{v}_k \in \mathbb{C}$ はそれぞれ

$$\mathbf{u}_k = \left(\frac{1}{\Omega} e^{-in\theta_k} \right)_{n=1}^{\Omega}, \quad \mathbf{v}_k = \left(e^{in\theta_k} \right)_{n=1}^{\Omega} \quad (\text{A.3})$$

で与えられる. これらの固有ベクトルは正規直交条件

$$\mathbf{u}_k \mathbf{v}_l^\top = \delta_{kl} \quad (\text{A.4})$$

および完全性条件

$$\sum_{k=1}^{\Omega} \mathbf{v}_k^\top \mathbf{u}_k = \mathbf{I}_\Omega \quad (\text{A.5})$$

をみtas. したがって行列 $\mathbf{J}_\Omega(a, b)$ のスペクトル分解は

$$\mathbf{J}_\Omega(a, b) = \sum_{k=1}^{\Omega} \mu_k(a, b) \mathbf{v}_k^\top \mathbf{u}_k \quad (\text{A.6})$$

で与えられる.

A.1.2 $a = b$ のとき

次に $a = b$ の場合を考える. 以下では実数 x に対して, x を超えない最大の整数を $[x]$ と表す. このとき整数 $k = 2, 3, \dots, [(\Omega + 1)/2]$ に対して関係式 $\theta_{\Omega+2-k} = 2\pi - \theta_k$ が成り立つことから, 行列 $\mathbf{J}_\Omega(a, a)$ の固有値 $\mu_k(a, a)$ と $\mu_{\Omega+2-k}(a, a)$ は縮退する. このことを考慮すると, 行列 $\mathbf{J}_\Omega(a, a)$ の固有値 $\mu_k(a, a) \in \mathbb{R}$ は $k = 1, 2, \dots, [\Omega/2] + 1$ に対して

$$\mu_k(a, a) = 2a \cos \theta_k \quad (\text{A.7})$$

で与えられる. またその重複度 m_k はそれぞれ

$$m_k = \begin{cases} 1 & k = 1 \text{ のとき,} \\ 2 & 2 \leq k \leq [(\Omega + 1)/2] \text{ のとき,} \\ 1 & \Omega \text{ が偶数かつ } k = \Omega/2 + 1 \text{ のとき} \end{cases} \quad (\text{A.8})$$

で与えられる. 行列 $\mathbf{J}_\Omega(a, a)$ は実対称行列であることから, その固有ベクトルとして実ベクトルをとることができる. 縮退を区別する添字を σ として, 行列 $\mathbf{J}_\Omega(a, a)$ の固有値 $\mu_k(a, a)$ に属する左固有ベクトル $\mathbf{u}_{k,\sigma} \in \mathbb{R}^\Omega$ および右固有ベクトル $\mathbf{v}_{k,\sigma} \in \mathbb{R}^\Omega$ はそれぞれ以下のとおり与えられる:

$k = 1$ のとき $\mathbf{1} \equiv (1, 1, \dots, 1)$ として

$$\mathbf{u}_{1,1} = \frac{1}{\Omega} \mathbf{1}, \quad \mathbf{v}_{1,1} = \mathbf{1}. \quad (\text{A.9})$$

$2 \leq k \leq [(\Omega + 1)/2]$ のとき

$$\mathbf{u}_{k,1} = \left(\frac{\sqrt{2}}{\Omega} \cos n\theta_k \right)_{n=1}^{\Omega}, \quad \mathbf{v}_{k,1} = \left(\sqrt{2} \cos n\theta_k \right)_{n=1}^{\Omega}, \quad (\text{A.10})$$

$$\mathbf{u}_{k,2} = \left(\frac{\sqrt{2}}{\Omega} \sin n\theta_k \right)_{n=1}^{\Omega}, \quad \mathbf{v}_{k,2} = \left(\sqrt{2} \sin n\theta_k \right)_{n=1}^{\Omega}. \quad (\text{A.11})$$

Ω が偶数かつ $k = \Omega/2 + 1$ のとき

$$\mathbf{u}_{\Omega/2+1,1} = \left(\frac{1}{\Omega} (-1)^n \right)_{n=1}^{\Omega}, \quad \mathbf{v}_{\Omega/2+1,1} = \left((-1)^n \right)_{n=1}^{\Omega}. \quad (\text{A.12})$$

これらの固有ベクトルは正規直交条件

$$\mathbf{u}_{k,\sigma} \mathbf{v}_{l,\rho}^{\top} = \delta_{kl} \delta_{\sigma\rho} \quad (\text{A.13})$$

および完全性条件

$$\sum_{k=1}^{[\Omega/2]+1} \sum_{\sigma=1}^{m_k} \mathbf{v}_{k,\sigma}^{\top} \mathbf{u}_{k,\sigma} = \mathbf{I}_{\Omega} \quad (\text{A.14})$$

をみます。したがって行列 $J_{\Omega}(a, a)$ のスペクトル分解は

$$J_{\Omega}(a, a) = \sum_{k=1}^{[\Omega/2]+1} \mu_k(a, a) \left(\sum_{\sigma=1}^{m_k} \mathbf{v}_{k,\sigma}^{\top} \mathbf{u}_{k,\sigma} \right) \quad (\text{A.15})$$

で与えられる。

A.2 行列 $\tilde{J}_{2\Omega}(a, b; c)$ のスペクトル分解

実数 a, b, c に対して, 2Ω 次正方行列 $\tilde{J}_{2\Omega}(a, b; c)$ は

$$\tilde{J}_{2\Omega}(a, b; c) \equiv \begin{pmatrix} J_{\Omega}(a, b) & c\mathbf{1}_{\Omega} \\ c\mathbf{1}_{\Omega} & J_{\Omega}(b, a) \end{pmatrix} \in \mathbb{R}^{2\Omega \times 2\Omega} \quad (\text{A.16})$$

で定義される。以下では行列 $\tilde{J}_{2\Omega}(a, b; c)$ のすべての固有値, 固有ベクトルを列挙し, そのスペクトル分解を与える。ただしある整数 $q \in \{2, 3, \dots, [(\Omega + 1)/2]\}$ に対して実数 a, b, c に $c = \pm(a - b) \sin \theta_q$ なる関係があるとき, 行列 $\tilde{J}_{2\Omega}(a, b; c)$ は対角化可能でない。この場合でもスペクトル分解に対応する式を構成できることをみる。

A.2.1 $c = 0$ のとき

まずはじめに $c = 0$ の場合を考える. このとき行列 $\tilde{J}_{2\Omega}(a, b; 0)$ はブロック対角行列 $J_{\Omega}(a, b) \oplus J_{\Omega}(b, a)$ となる. よって, 行列 $\tilde{J}_{2\Omega}(a, b; 0)$ の固有値 $\tilde{\mu}_k(a, b; 0) \in \mathbb{C}$ は $k = 1, 2, \dots, \Omega$ に対して

$$\tilde{\mu}_k(a, b; 0) = (a + b) \cos \theta_k + i(a - b) \sin \theta_k \quad (\text{A.17})$$

で与えられる. すべての固有値は 2 重に縮退している.

零ベクトルを $\mathbf{0} \equiv (0, 0, \dots, 0) \in \mathbb{R}^{\Omega}$ と書き, ベクトル $\mathbf{u}_k \in \mathbb{C}^{\Omega}$ および $\mathbf{v}_k \in \mathbb{C}^{\Omega}$ をそれぞれ付録 A.1 で与えた行列 $J_{\Omega}(a, b)$ の左固有ベクトルおよび右固有ベクトルとする. またベクトル \mathbf{a} に対して \mathbf{a}^* はその複素共役を表すとする. このとき, 固有値 $\tilde{\mu}_k(a, b; 0)$ に属する左固有ベクトル $\tilde{\mathbf{u}}_{k,\sigma} \in \mathbb{C}^{2\Omega}$ および右固有ベクトル $\tilde{\mathbf{v}}_{k,\sigma} \in \mathbb{C}^{2\Omega}$ はそれぞれ $k = 1, 2, \dots, \Omega$ に対して

$$\tilde{\mathbf{u}}_{k,1} = (\mathbf{u}_k, \mathbf{0}), \quad \tilde{\mathbf{v}}_{k,1} = (\mathbf{v}_k, \mathbf{0}) \quad (\text{A.18})$$

および

$$\tilde{\mathbf{u}}_{k,2} = (\mathbf{0}, \mathbf{u}_k^*), \quad \tilde{\mathbf{v}}_{k,2} = (\mathbf{0}, \mathbf{v}_k^*) \quad (\text{A.19})$$

で与えられる. これらの固有ベクトルは正規直交関係

$$\tilde{\mathbf{u}}_{k,\sigma} \tilde{\mathbf{v}}_{l,\rho}^{\top} = \delta_{kl} \delta_{\sigma\rho} \quad (\text{A.20})$$

および完全性条件

$$\sum_{k=1}^{\Omega} \sum_{\sigma=1}^2 \tilde{\mathbf{v}}_{k,\sigma}^{\top} \tilde{\mathbf{u}}_{k,\sigma} = I_{2\Omega} \quad (\text{A.21})$$

をみたす. したがって行列 $\tilde{J}(a, b; 0)$ のスペクトル分解は

$$\tilde{J}_{2\Omega}(a, b; 0) = \sum_{k=1}^{\Omega} \tilde{\mu}_k(a, b; 0) \left(\sum_{\sigma=1}^2 \tilde{\mathbf{v}}_{k,\sigma}^{\top} \tilde{\mathbf{u}}_{k,\sigma} \right) \quad (\text{A.22})$$

で与えられる.

A.2.2 $c \neq 0$ かつ $a = b$ の場合

次に $c \neq 0$ かつ $a = b$ のときを考えよう. このとき行列 $\tilde{J}_{2\Omega}(a, a; c)$ は実対称行列となるので, その固有値はすべて実数である. また固有ベクトルもすべて実ベクトルにとるこ

とができる. 行列 $\tilde{J}_{2\Omega}(a, a; c)$ の固有値 $\tilde{\mu}_k^\pm(a, a; c) \in \mathbb{R}$ は $k = 1, 2, \dots, [\Omega/2] + 1$ に対して

$$\tilde{\mu}_k^\pm(a, a; c) = 2a \cos \theta_k \pm c \quad (\text{A.23})$$

で与えられる. 固有値 $\tilde{\mu}_k^\pm(a, a; c)$ の重複度 m_k は添字 k のみに依存し,

$$m_k = \begin{cases} 1 & k = 1 \text{ のとき,} \\ 2 & 2 \leq k \leq [(\Omega + 1)/2] \text{ のとき,} \\ 1 & \Omega \text{ が偶数かつ } k = \Omega/2 + 1 \text{ のとき} \end{cases} \quad (\text{A.24})$$

で与えられる. ベクトル $\mathbf{u}_{k,\sigma} \in \mathbb{R}^\Omega$ および $\mathbf{v}_{k,\sigma} \in \mathbb{R}^\Omega$ をそれぞれ付録 A.1 で与えた行列 $J_\Omega(a, a)$ の左固有ベクトルおよび右固有ベクトルとすると, 行列 $\tilde{J}_{2\Omega}(a, a; c)$ の固有値 $\tilde{\mu}_k^\pm(a, a; c)$ に属する左固有ベクトル $\tilde{\mathbf{u}}_{k,\sigma}^\pm \in \mathbb{R}^{2\Omega}$ および右固有ベクトル $\tilde{\mathbf{v}}_{k,\sigma}^\pm \in \mathbb{R}^{2\Omega}$ はそれぞれ以下のとおり与えられる:

$k = 1$ のとき

$$\tilde{\mathbf{u}}_{1,1}^\pm = \frac{1}{2\Omega}(\mathbf{1}, \pm\mathbf{1}), \quad \tilde{\mathbf{v}}_{1,1}^\pm = (\mathbf{1}, \pm\mathbf{1}). \quad (\text{A.25})$$

$2 \leq k \leq [\Omega/2] + 1$ のとき $\sigma = 1, 2, \dots, m_k$ に対して

$$\tilde{\mathbf{u}}_{k,\sigma}^\pm = \frac{1}{\sqrt{2}}(\mathbf{u}_{k,\sigma}, \pm\mathbf{u}_{k,\sigma}), \quad \tilde{\mathbf{v}}_{k,\sigma}^\pm = \frac{1}{\sqrt{2}}(\mathbf{v}_{k,\sigma}, \pm\mathbf{v}_{k,\sigma}). \quad (\text{A.26})$$

これらの固有ベクトルは正規直交関係

$$\tilde{\mathbf{u}}_{k,\sigma}^\varepsilon \tilde{\mathbf{v}}_{l,\rho}^{\mu\top} = \delta_{kl} \delta_{\sigma\rho} \delta_{\varepsilon\mu} \quad (\text{A.27})$$

および完全性条件

$$\sum_{k=1}^{[\Omega/2]+1} \sum_{\varepsilon=\pm} \sum_{\sigma=1}^{m_k} \tilde{\mathbf{v}}_{k,\sigma}^{\varepsilon\top} \tilde{\mathbf{u}}_{k,\sigma}^\varepsilon = I_{2\Omega} \quad (\text{A.28})$$

をみだす. したがって行列 $\tilde{J}_{2\Omega}(a, a; c)$ のスペクトル分解は

$$\tilde{J}_{2\Omega}(a, b; c) = \sum_{k=1}^{[\Omega/2]+1} \left[\sum_{\varepsilon=\pm} \tilde{\mu}_k^\varepsilon(a, a; c) \left(\sum_{\sigma=1}^{m_k} \tilde{\mathbf{v}}_{k,\sigma}^{\varepsilon\top} \tilde{\mathbf{u}}_{k,\sigma}^\varepsilon \right) \right] \quad (\text{A.29})$$

で与えられる.

A.2.3 $c \neq 0, a \neq b$ かつ対角化可能な場合

残る場合分けは「 $c \neq 0$ かつ $a \neq b$ のとき」である. いま, ある整数 $q \in \{2, 3, \dots, [(\Omega + 1)/2]\}$ を用いて実数 a, b, c に $c = \pm(a - b) \sin \theta_q$ なる関係が成り立つとき, 行列 $\tilde{J}_{2\Omega}(a, b; c)$ は対角化可能でない. まずはこの関係が成り立たないとき, すなわち対角化可能な場合について述べる.

$c \neq 0, a \neq b$ かつすべての整数 $k = 2, 3, \dots, [(\Omega + 1)/2]$ について $c \neq \pm(a - b) \sin \theta_k$ のとき, 行列 $\tilde{J}_{2\Omega}(a, b; c)$ の固有値 $\tilde{\mu}_k^\pm(a, b; c) \in \mathbb{C}$ は $k = 1, 2, \dots, [\Omega/2] + 1$ に対して

$$\tilde{\mu}_k^\pm(a, b; c) = (a + b) \cos \theta_k \pm \sqrt{c^2 - (a - b)^2 \sin^2 \theta_k} \quad (\text{A.30})$$

で与えられる. 特に $|c| < |a - b| \sin \theta_k$ のとき固有値 $\tilde{\mu}_k^\pm(a, b; c)$ は複素数となる. 固有値 $\tilde{\mu}_k^\pm(a, b; c)$ の重複度 m_k は添字 k のみに依存し,

$$m_k = \begin{cases} 1 & k = 1 \text{ のとき,} \\ 2 & 2 \leq k \leq [(\Omega + 1)/2] \text{ のとき,} \\ 1 & \Omega \text{ が偶数かつ } k = \Omega/2 + 1 \text{ のとき} \end{cases} \quad (\text{A.31})$$

で与えられる. 固有値 $\tilde{\mu}_k^\pm(a, b; c)$ に属する左固有ベクトル $\tilde{\mathbf{u}}_{k,\sigma}^\pm \in \mathbb{C}^{2\Omega}$ および右固有ベクトル $\tilde{\mathbf{v}}_{k,\sigma}^\pm \in \mathbb{C}^{2\Omega}$ はそれぞれ以下のとおり与えられる:

$k = 1$ のとき

$$\tilde{\mathbf{u}}_{1,1}^\pm = \frac{1}{2\Omega} (\mathbf{1}, \pm \mathbf{1}), \quad \tilde{\mathbf{v}}_{1,1}^\pm = (\mathbf{1}, \pm \mathbf{1}). \quad (\text{A.32})$$

$2 \leq k \leq [(\Omega + 1)/2]$ のとき

$$A_k^\pm \equiv \frac{c \mp \sqrt{c^2 - (a - b)^2 \sin^2 \theta_k}}{(a - b) \sin \theta_k} \quad (\text{A.33})$$

を用いて

$$\begin{cases} \tilde{\mathbf{u}}_{k,1}^\pm = \frac{1}{\sqrt{2}[1 - (A_k^\pm)^2]} (\mathbf{u}_{k,1} + A_k^\pm \mathbf{u}_{k,2}, \mathbf{u}_{k,1} - A_k^\pm \mathbf{u}_{k,2}), \\ \tilde{\mathbf{v}}_{k,1}^\pm = \frac{1}{\sqrt{2}} (\mathbf{v}_{k,1} - A_k^\pm \mathbf{v}_{k,2}, \mathbf{v}_{k,1} + A_k^\pm \mathbf{v}_{k,2}), \end{cases} \quad (\text{A.34})$$

$$\begin{cases} \tilde{\mathbf{u}}_{k,2}^\pm = \frac{1}{\sqrt{2}[1 - (A_k^\pm)^2]} (\mathbf{u}_{k,2} - A_k^\pm \mathbf{u}_{k,1}, \mathbf{u}_{k,2} + A_k^\pm \mathbf{u}_{k,1}), \\ \tilde{\mathbf{v}}_{k,2}^\pm = \frac{1}{\sqrt{2}} (\mathbf{v}_{k,2} + A_k^\pm \mathbf{v}_{k,1}, \mathbf{v}_{k,2} - A_k^\pm \mathbf{v}_{k,1}). \end{cases} \quad (\text{A.35})$$

Ω が偶数かつ $k = \Omega/2 + 1$ のとき

$$\begin{cases} \tilde{\mathbf{u}}_{\Omega/2+1,1}^{\pm} = \frac{1}{\sqrt{2}}(\mathbf{u}_{\Omega/2+1,1}, \pm\mathbf{u}_{\Omega/2+1,1}), \\ \tilde{\mathbf{v}}_{\Omega/2+1,1}^{\pm} = \frac{1}{\sqrt{2}}(\mathbf{v}_{\Omega/2+1,1}, \pm\mathbf{v}_{\Omega/2+1,1}). \end{cases} \quad (\text{A.36})$$

これらの固有ベクトルは正規直交条件

$$\tilde{\mathbf{u}}_{k,\sigma}^{\varepsilon} \tilde{\mathbf{v}}_{l,\rho}^{\mu\top} = \delta_{kl} \delta_{\sigma\rho} \delta_{\varepsilon\mu} \quad (\text{A.37})$$

および完全性条件

$$\sum_{k=1}^{[\Omega/2]+1} \sum_{\varepsilon=\pm} \sum_{\sigma=1}^{m_k} \tilde{\mathbf{v}}_{k,\sigma}^{\varepsilon\top} \tilde{\mathbf{u}}_{k,\sigma}^{\varepsilon} = \mathbf{I}_{2\Omega} \quad (\text{A.38})$$

をみます。したがって行列 $\tilde{J}_{2\Omega}(a, b; c)$ のスペクトル分解は

$$\tilde{J}_{2\Omega}(a, b; c) = \sum_{k=1}^{[\Omega/2]+1} \left[\sum_{\varepsilon=\pm} \tilde{\mu}_k^{\varepsilon}(a, b; c) \left(\sum_{\sigma=1}^{m_k} \tilde{\mathbf{v}}_{k,\sigma}^{\varepsilon\top} \tilde{\mathbf{u}}_{k,\sigma}^{\varepsilon} \right) \right] \quad (\text{A.39})$$

で与えられる。

A.2.4 対角化可能でない場合

最後に、 $a \neq b$ かつ、ある整数 $q \in \{2, 3, \dots, [(\Omega+1)/2]\}$ に対して $c = \epsilon(a-b) \sin \theta_q \equiv c_q$ なるときを考える。ここで $\epsilon = \pm 1$ である。このとき、(A.30) 式で与えられた固有値 $\tilde{\mu}_q^+(a, b; c_q)$ と $\tilde{\mu}_q^-(a, b; c_q)$ は縮退する。すなわち、固有値 $\tilde{\mu}_q \equiv \tilde{\mu}_q^{\pm}(a, b; c_q)$ の重複度は 4 となる。しかし一方で係数 $A_q^{\pm} = \epsilon$ が複号によらなくなり、固有ベクトル $\tilde{\mathbf{u}}_{q,\sigma}^+$ と $\tilde{\mathbf{u}}_{q,\sigma}^-$ 、 $\tilde{\mathbf{v}}_{q,\sigma}^+$ と $\tilde{\mathbf{v}}_{q,\sigma}^-$ がそれぞれ同じになる。これより、4 重縮退した固有値 $\tilde{\mu}_q \in \mathbb{R}$ に属する線形独立な固有ベクトルを 2 組しかとることができないことがわかる。このことから、行列 $\tilde{J}_{2\Omega}(a, b; c_q)$ は対角化可能でない。それでも一般に Jordan 標準形へ変形することはできる。ベクトル $\mathbf{u}_{q,\sigma} \in \mathbb{C}^{\Omega}$ および $\mathbf{v}_{q,\sigma} \in \mathbb{C}^{\Omega}$ をそれぞれ付録 A.1 で与えた行列 $J_{\Omega}(a, b)$ の固有値 $\mu_q(a, b)$ に属する左固有ベクトルおよび右固有ベクトルとするとき、

$$\begin{cases} \tilde{\mathbf{u}}_q^{(i)} = B_q^{(i)}(\mathbf{u}_{q,1} - \epsilon\mathbf{u}_{q,2}, \mathbf{u}_{q,1} + \epsilon\mathbf{u}_{q,2}), \\ \tilde{\mathbf{u}}_q^{(ii)} = B_q^{(ii)}(\mathbf{u}_{q,1} + \epsilon\mathbf{u}_{q,2}, \mathbf{u}_{q,1} - \epsilon\mathbf{u}_{q,2}), \\ \tilde{\mathbf{u}}_q^{(iii)} = B_q^{(iii)}(\mathbf{u}_{q,2} + \epsilon\mathbf{u}_{q,1}, \mathbf{u}_{q,2} - \epsilon\mathbf{u}_{q,1}), \\ \tilde{\mathbf{u}}_q^{(iv)} = B_q^{(iv)}(\mathbf{u}_{q,2} - \epsilon\mathbf{u}_{q,1}, \mathbf{u}_{q,2} + \epsilon\mathbf{u}_{q,1}) \end{cases} \quad (\text{A.40})$$

および

$$\begin{cases} \tilde{\mathbf{v}}_q^{(i)} = C_q^{(i)}(\mathbf{v}_{q,1} - \epsilon\mathbf{v}_{q,2}, \mathbf{v}_{q,1} + \epsilon\mathbf{v}_{q,2}), \\ \tilde{\mathbf{v}}_q^{(ii)} = C_q^{(ii)}(\mathbf{v}_{q,1} + \epsilon\mathbf{v}_{q,2}, \mathbf{v}_{q,1} - \epsilon\mathbf{v}_{q,2}), \\ \tilde{\mathbf{v}}_q^{(iii)} = C_q^{(iii)}(\mathbf{v}_{q,2} + \epsilon\mathbf{v}_{q,1}, \mathbf{v}_{q,2} - \epsilon\mathbf{v}_{q,1}), \\ \tilde{\mathbf{v}}_q^{(iv)} = C_q^{(iv)}(\mathbf{v}_{q,2} - \epsilon\mathbf{v}_{q,1}, \mathbf{v}_{q,2} + \epsilon\mathbf{v}_{q,1}) \end{cases} \quad (\text{A.41})$$

なるベクトルを考える. ここで係数 $B_q^{(x)}, C_q^{(x)}$ ($x = i, ii, iii, iv$) は

$$B_q^{(x)}C_q^{(x)} = \frac{1}{4} \quad (\text{A.42})$$

および

$$B_q^{(ii)}C_q^{(i)} = B_q^{(iv)}C_q^{(iii)} = \frac{1}{2}C_q \quad (\text{A.43})$$

をみたすように定める. このとき, これらのベクトルは正規直交条件

$$\tilde{\mathbf{u}}_q^{(x)}\tilde{\mathbf{v}}_q^{(y)\top} = \delta_{xy} \quad (\text{A.44})$$

をみたす. いま, 行列 $\mathbf{U}, \mathbf{V} \in \mathbb{R}^{4 \times 2\Omega}$ を

$$\mathbf{U} \equiv \begin{pmatrix} \tilde{\mathbf{u}}_q^{(i)} \\ \tilde{\mathbf{u}}_q^{(ii)} \\ \tilde{\mathbf{u}}_q^{(iii)} \\ \tilde{\mathbf{u}}_q^{(iv)} \end{pmatrix}, \quad \mathbf{V} \equiv \begin{pmatrix} \tilde{\mathbf{v}}_q^{(i)} \\ \tilde{\mathbf{v}}_q^{(ii)} \\ \tilde{\mathbf{v}}_q^{(iii)} \\ \tilde{\mathbf{v}}_q^{(iv)} \end{pmatrix} \quad (\text{A.45})$$

と定義する. このとき, 正規直交条件 (A.44) は

$$\mathbf{UV}^\top = \mathbf{I}_4 \quad (\text{A.46})$$

と書き換えられる. 以上の関係式を用いると, 対角化可能な場合の行列 $\tilde{\mathbf{J}}_{2\Omega}(a, b; c)$ で成り立つスペクトル分解 (A.39) において, 固有値 $\tilde{\mu}_q$ に対応する項

$$\sum_{\epsilon=\pm} \tilde{\mu}_q^\epsilon(a, b; c) \left(\sum_{\sigma=1,2} \tilde{\mathbf{v}}_{q,\sigma}^{\epsilon\top} \tilde{\mathbf{u}}_{q,\sigma}^\epsilon \right) \quad (\text{A.47})$$

を

$$\mathbf{V}^\top \begin{pmatrix} \tilde{\mu}_q & 1 & & 0 \\ 0 & \tilde{\mu}_q & & 0 \\ & & \tilde{\mu}_q & 1 \\ 0 & & 0 & \tilde{\mu}_q \end{pmatrix} \mathbf{U} = \tilde{\mu}_q \sum_x \tilde{\mathbf{v}}_q^{(x)\top} \tilde{\mathbf{u}}_q^{(x)} + \tilde{\mathbf{v}}_q^{(i)\top} \tilde{\mathbf{u}}_q^{(ii)} + \tilde{\mathbf{v}}_q^{(iii)\top} \tilde{\mathbf{u}}_q^{(iv)} \quad (\text{A.48})$$

に置き換えた式が, 対角化可能でない場合の行列 $\tilde{\mathbf{J}}_{2\Omega}(a, b; c_q)$ で成り立つことが示される. これは行列 $\tilde{\mathbf{J}}_{2\Omega}(a, b; c_q)$ が対角化可能でないものの, 2 次の Jordan 細胞をふたつもつ Jordan 標準形に変形できることと対応している.

付録 B

第 2 固有値の特定

本付録では、第 4 章でねじれ詳細つり合い条件の手法を用いた lifting により拡張された 1 次元ランダムウォークの遷移確率行列 \tilde{T} における、絶対値の意味での第 2 固有値 (second largest eigenvalue modulus, SLEM) $\tilde{\eta}$ を特定する.

第 4 章の設定と同じく、簡単のため $\alpha = 1/2$ とし、 Ω は十分大きな偶数とする. さらに設定 (a) $\delta = \delta'$, および設定 (b) $\delta' = 1$ の 2 通りに限定して議論をすすめる.

本付録の内容は文献 [50] の APPENDIX D として出版されている.

B.1 設定 (a) $\delta = \delta'$ の第 2 固有値

設定 (a) $\delta = \delta'$ における遷移確率行列 \tilde{T} の固有値 $\tilde{\lambda}_k^\pm$ は, $k = 1, 2, \dots, \Omega/2 + 1$ に対して

$$\tilde{\lambda}_k^\pm = \frac{3}{4} + \frac{1}{4} \cos \left[\frac{2\pi}{\Omega} (k-1) \right] - \gamma \pm \sqrt{\gamma^2 - \delta^2 \left\{ \frac{1}{4} \sin \left[\frac{2\pi}{\Omega} (k-1) \right] \right\}^2} \quad (\text{B.1})$$

で与えられる. パラメタはそれぞれ不等式 $0 \leq \gamma \leq 3/4$, $|\delta| \leq 1$ をみます. 以下では固有値に現れる根号内の符号に基づいて場合分けし, そのそれぞれにおいて SLEM を明らかにする.

B.1.1 すべての固有値が実数のとき

$\gamma \geq |\delta|/4$ のとき, すべての固有値が実数であることが保証される. このとき SLEM の候補は

- $\tilde{\lambda}_2^+ = \frac{3}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} - \gamma + \sqrt{\gamma^2 - \delta^2 \left(\frac{1}{4} \sin \frac{2\pi}{\Omega} \right)^2}$
- $\tilde{\lambda}_1^- = 1 - 2\gamma$
- $|\tilde{\lambda}_{\Omega/2+1}^-| = \left| \frac{1}{2} - 2\gamma \right|$

である. 特にいま考えているすべてのパラメタ領域で不等式 $\tilde{\lambda}_2^+ > 0$, $\min\{\tilde{\lambda}_2^+, \tilde{\lambda}_1^-\} > \tilde{\lambda}_{\Omega/2+1}^-$ が成り立つことに注意する.

$\tilde{\lambda}_2^+ \geq \tilde{\lambda}_1^-$ となる領域

固有値 $\tilde{\lambda}_2^+$ と $\tilde{\lambda}_1^-$ の大小を比較する.

$$\begin{aligned} \tilde{\lambda}_2^+ - \tilde{\lambda}_1^- &= \left[\frac{3}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} - \gamma + \sqrt{\gamma^2 - \delta^2 \left(\frac{1}{4} \sin \frac{2\pi}{\Omega} \right)^2} \right] - (1 - 2\gamma) \\ &= \gamma - \frac{1}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} + \sqrt{\gamma^2 - \delta^2 \left(\frac{1}{4} \sin \frac{2\pi}{\Omega} \right)^2} \end{aligned} \quad (\text{B.2})$$

より,

$$\gamma \geq \frac{1}{2} \sin^2 \frac{\pi}{\Omega} \quad (\text{B.3})$$

では不等式 $\tilde{\lambda}_2^+ \geq \tilde{\lambda}_1^-$ が成り立つ. 一方, $\gamma \leq (1/2) \sin^2(\pi/\Omega)$ のとき

$$\gamma \geq \frac{1}{4} \left(\sin^2 \frac{\pi}{\Omega} + \delta^2 \cos^2 \frac{\pi}{\Omega} \right) \quad (\text{B.4})$$

で不等式 $\tilde{\lambda}_2^+ \geq \tilde{\lambda}_1^-$ が成り立つ.

$\tilde{\lambda}_2^+ \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ となる領域

つねに $\tilde{\lambda}_2^+ > \tilde{\lambda}_{\Omega/2+1}^-$ なので, $|\tilde{\lambda}_{\Omega/2+1}^-|$ については絶対値の中身が負の場合のみ比較すれば十分である. このとき

$$\begin{aligned} \tilde{\lambda}_2^+ - |\tilde{\lambda}_{\Omega/2+1}^-| &= \frac{3}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} - \gamma + \sqrt{\gamma^2 - \delta^2 \left(\frac{1}{4} \sin \frac{2\pi}{\Omega} \right)^2} + \frac{1}{2} - 2\gamma \\ &= \frac{5}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} - 3\gamma + \sqrt{\gamma^2 - \delta^2 \left(\frac{1}{4} \sin \frac{2\pi}{\Omega} \right)^2} \end{aligned} \quad (\text{B.5})$$

であるので,

$$\gamma \leq \frac{5}{12} + \frac{1}{12} \cos \frac{2\pi}{\Omega} \quad (\text{B.6})$$

のときは不等式 $\tilde{\lambda}_2^+ \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ が成り立つ. 一方, $\gamma \geq 5/12 + (1/12) \cos(2\pi/\Omega)$ のときは

$$\begin{aligned} &\left(\sqrt{\gamma^2 - \delta^2 \left(\frac{1}{4} \sin \frac{2\pi}{\Omega} \right)^2} \right)^2 - \left(\frac{5}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} - 3\gamma \right)^2 \\ &= -8\gamma^2 + 6\gamma \left(\frac{5}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} \right) - \left(\frac{5}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} \right)^2 - \frac{1}{16} \delta^2 \sin^2 \frac{2\pi}{\Omega} \geq 0 \end{aligned} \quad (\text{B.7})$$

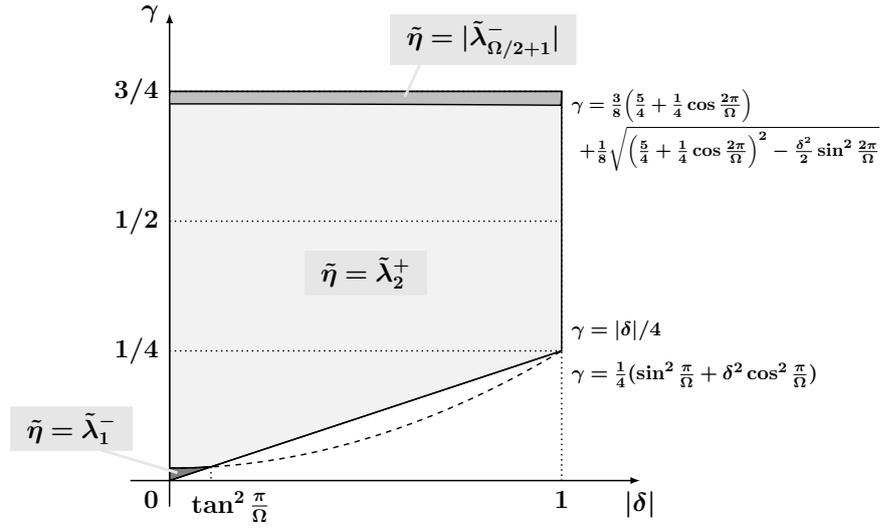


図 B.1 すべての固有値が実数のときに SLEM を与えるパラメタ領域.

すなわち

$$\begin{aligned} & \frac{3}{8} \left(\frac{5}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} \right) - \frac{1}{8} \sqrt{\left(\frac{5}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} \right)^2 - \frac{1}{2} \delta^2 \sin^2 \frac{2\pi}{\Omega}} \\ & \leq \gamma \leq \frac{3}{8} \left(\frac{5}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} \right) + \frac{1}{8} \sqrt{\left(\frac{5}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} \right)^2 - \frac{1}{2} \delta^2 \sin^2 \frac{2\pi}{\Omega}} \end{aligned} \quad (\text{B.8})$$

のとき不等式 $\tilde{\lambda}_2^+ \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ が成り立つ.

$\tilde{\lambda}_1^- \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ となる領域

つねに $\tilde{\lambda}_1^- > \tilde{\lambda}_{\Omega/2+1}^-$ であるので, $|\tilde{\lambda}_{\Omega/2+1}^-|$ については絶対値の中身が負の場合のみ比較すれば十分である. このとき

$$\tilde{\lambda}_1 - \tilde{\lambda}_{\Omega/2+1}^- = 1 - 2\gamma - \left(2\gamma - \frac{1}{2} \right) = \frac{3}{2} - 4\gamma \quad (\text{B.9})$$

であるので,

$$\gamma \leq \frac{3}{8} \quad (\text{B.10})$$

のとき不等式 $\tilde{\lambda}_1^- \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ が成り立つ. 以上をあわせて図 B.1 を得る.

B.1.2 ほぼすべての固有値が複素数のとき

$k = 1, \Omega/2 + 1$ に対応する 4 つの固有値 $\tilde{\lambda}_1^\pm, \tilde{\lambda}_{\Omega/2+1}^\pm$ はつねに実数である. それ以外の固有値は, $\gamma \leq (1/4)|\delta| \sin(2\pi/\Omega)$ のときすべて複素数となる. $2 \leq k \leq \Omega/2$ のとき固有値 $\tilde{\lambda}_k^\pm$ の絶対値 $|\tilde{\lambda}_k^\pm|$ は, $x_k \equiv \cos[2\pi(k-1)/\Omega]$ とおけば

$$\begin{aligned} |\tilde{\lambda}_k^\pm|^2 &= \left(\frac{3}{4} + \frac{1}{4}x_k - \gamma\right)^2 + \frac{\delta^2}{16}(1 - x_k^2) - \gamma^2 \\ &= \frac{1}{16}(1 - \delta^2)x_k^2 + \frac{1}{2}\left(\frac{3}{4} - \gamma\right)x_k + \frac{9}{16} - \frac{3}{2}\gamma + \frac{\delta^2}{16} \end{aligned} \quad (\text{B.11})$$

と計算できる. 特に $\gamma \leq 3/4, |\delta| \leq 1$ に注意すると, これは $x_k > 0$ で x_k について単調増加関数であることがわかる. すなわち $2 \leq k \leq \Omega/2$ における固有値の絶対値は $k = 2$ で最大値をとる. また, $\gamma \leq (1/4)|\delta| \sin(2\pi/\Omega)$ では $k = \Omega/2 + 1$ に対応する固有値は正であり, $\tilde{\lambda}_1^-$ のほうが大きい. よってこのとき SLEM の候補は

- $|\tilde{\lambda}_2^\pm| = \sqrt{\frac{1}{16}(1 - \delta^2) \cos^2 \frac{2\pi}{\Omega} + \frac{1}{2}\left(\frac{3}{4} - \gamma\right) \cos \frac{2\pi}{\Omega} + \frac{9}{16} - \frac{3}{2}\gamma + \frac{\delta^2}{16}}$
- $\tilde{\lambda}_1^- = 1 - 2\gamma > 0$

のいずれかである. これらを比較すると, $|\tilde{\lambda}_2^\pm| \geq \tilde{\lambda}_1^-$ となるのは

$$\begin{aligned} |\tilde{\lambda}_2^\pm|^2 - |\tilde{\lambda}_1^-|^2 &= \left(\frac{3}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} - \gamma\right)^2 + \delta^2 \left(\frac{1}{4} \sin \frac{2\pi}{\Omega}\right)^2 - \gamma^2 - (1 - 2\gamma)^2 \\ &= -4 \left\{ \left[\gamma - \frac{1}{16} \left(5 - \cos \frac{2\pi}{\Omega}\right) \right]^2 - \frac{1}{256} \left(5 - \cos \frac{2\pi}{\Omega}\right)^2 \right. \\ &\quad \left. - \frac{1}{64} \left(3 + \cos \frac{2\pi}{\Omega}\right)^2 + \frac{\delta^2}{64} \sin^2 \frac{2\pi}{\Omega} - \frac{1}{4} \right\} \end{aligned} \quad (\text{B.12})$$

より,

$$\begin{aligned} \frac{1}{16} \left(5 - \cos \frac{2\pi}{\Omega}\right) - \frac{1}{16} \sqrt{\left(5 \cos \frac{2\pi}{\Omega} - 1\right) \left(\cos \frac{2\pi}{\Omega} + 3\right) + 4\delta^2 \sin^2 \frac{2\pi}{\Omega}} &\leq \gamma \\ &\leq \frac{1}{16} \left(5 - \cos \frac{2\pi}{\Omega}\right) + \frac{1}{16} \sqrt{\left(5 \cos \frac{2\pi}{\Omega} - 1\right) \left(\cos \frac{2\pi}{\Omega} + 3\right) + 4\delta^2 \sin^2 \frac{2\pi}{\Omega}} \end{aligned} \quad (\text{B.13})$$

をみたとときである. ここで上の不等式における根号の中身は十分大きな整数 Ω に対してつねに正である. また特に最右辺の不等号は $\gamma \leq (1/4)|\delta| \sin(2\pi/\Omega)$ の下でつねにみたされる. 以上から図 B.2 を得る.

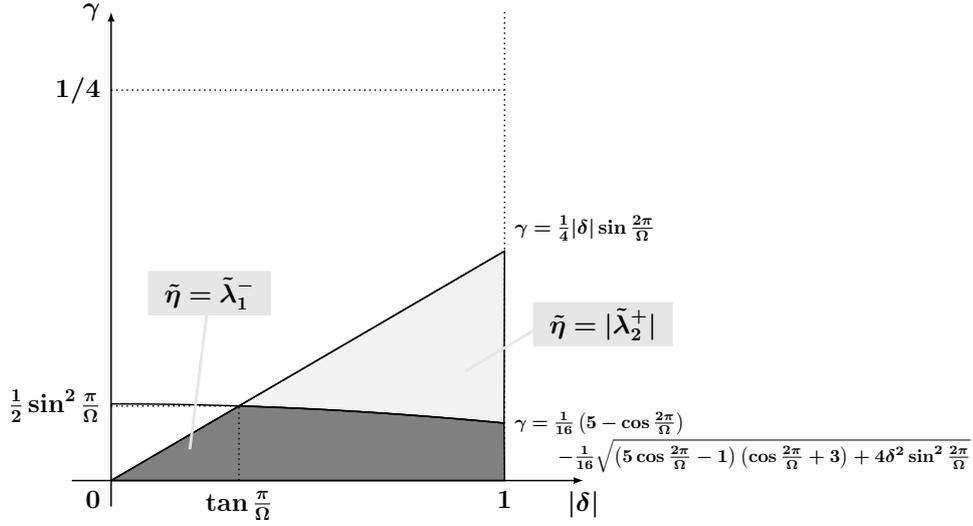


図 B.2 $\tilde{\lambda}_1^\pm, \tilde{\lambda}_{\Omega/2+1}^\pm$ 以外の固有値がすべて複素数のときに SLEM を与えるパラメタ領域.

B.1.3 それ以外の場合

$(1/4)|\delta| \sin(2\pi/\Omega) \leq \gamma \leq |\delta|/4$ のとき, 実固有値も複素固有値もともに現れる. $k = 3, 4, \dots, \Omega/2$ に対して $\tilde{\lambda}_k^\pm$ が実固有値のときは, $\tilde{\lambda}_2^+$ のほうが絶対値が大きいことがわかる. したがって B.1.1 節と同様の結論を得る. 次に $\tilde{\lambda}_k^\pm$ が複素固有値のときを考える. このとき $\tilde{\lambda}_2^+$ は実固有値であり, パラメタ Ω, δ を固定したときの $|\tilde{\lambda}_2^+|$ は $\gamma = (1/4)|\delta| \sin(2\pi/\Omega)$ のとき最小値をとるので, 他の複素固有値の絶対値がこれより小さいことを示せば, SLEM の候補は B.1.1 節と同じになる. 以下この事実を示す.

$k = 3, 4, \dots, \Omega/2$ に対して $\theta_k \equiv 2\pi(k-1)/\Omega$ とおくと, $\tilde{\lambda}_k^\pm$ が複素固有値のとき, すなわち $(1/4)|\delta| \sin(2\pi/\Omega) \leq \gamma \leq (1/4)|\delta| \sin \theta_k$ のとき

$$\begin{aligned}
& \left(\min_{\gamma} |\tilde{\lambda}_2^+| \right)^2 - |\tilde{\lambda}_k^\pm|^2 \\
&= \left(\frac{3}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} - \frac{1}{4} |\delta| \sin \frac{2\pi}{\Omega} \right)^2 - \left[\left(\frac{3}{4} + \frac{1}{4} \cos \theta_k - \gamma \right)^2 + \delta^2 \left(\frac{1}{4} \sin \theta_k \right)^2 - \gamma^2 \right] \\
&= 2\gamma \left(\frac{3}{4} + \frac{1}{4} \cos \theta_k \right) + \left(\frac{3}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} - \frac{1}{4} |\delta| \sin \frac{2\pi}{\Omega} \right)^2 \\
&\quad - \left(\frac{3}{4} + \frac{1}{4} \cos \theta_k \right)^2 - \delta^2 \left(\frac{1}{4} \sin \theta_k \right)^2 \tag{B.14}
\end{aligned}$$

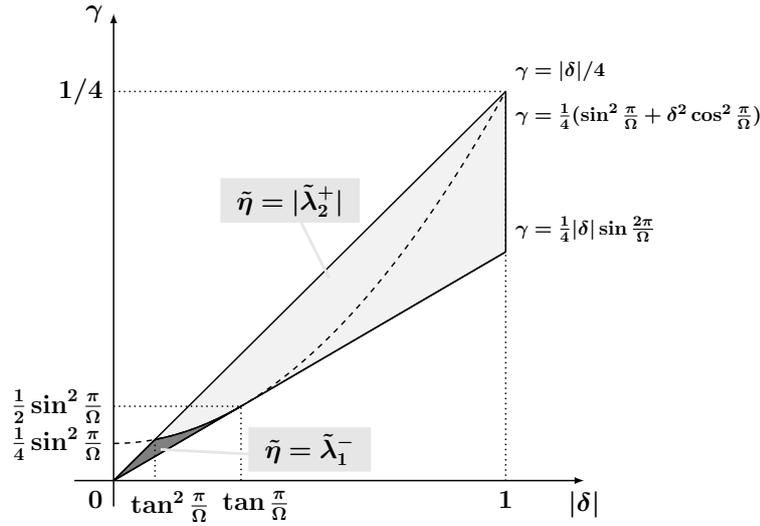
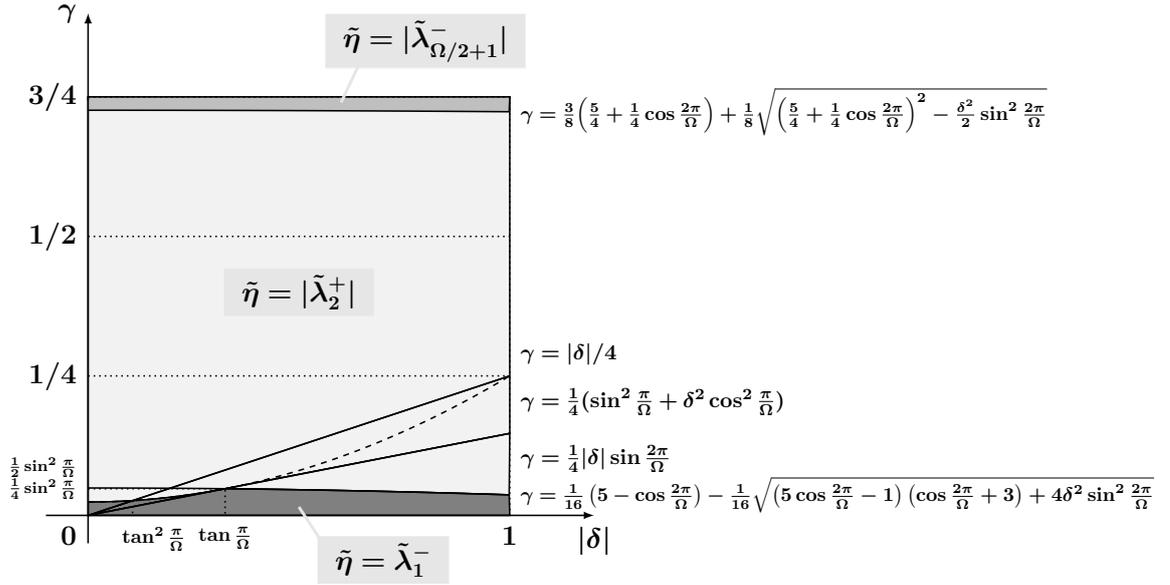


図 B.3 $(1/4)|\delta| \sin(2\pi/\Omega) \leq \gamma \leq |\delta|/4$ で SLEM を与える領域.

となる. この右辺は $\gamma = (1/4)|\delta| \sin(2\pi/\Omega)$ のとき最小となり,

$$\begin{aligned}
& \left(\min_{\gamma} |\tilde{\lambda}_2^+| \right)^2 - |\tilde{\lambda}_k^\pm|^2 \\
& \geq 2 \cdot \frac{1}{4} |\delta| \sin \frac{2\pi}{\Omega} \left(\frac{3}{4} + \frac{1}{4} \cos \theta_k \right) + \left(\frac{3}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} - \frac{1}{4} |\delta| \sin \frac{2\pi}{\Omega} \right)^2 \\
& \quad - \left(\frac{3}{4} + \frac{1}{4} \cos \theta_k \right)^2 - \delta^2 \left(\frac{1}{4} \sin \theta_k \right)^2 \\
& = \left(\frac{3}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} \right)^2 - \left(\frac{3}{4} + \frac{1}{4} \cos \theta_k \right)^2 \\
& \quad - \frac{1}{8} |\delta| \sin \frac{2\pi}{\Omega} \left(\cos \frac{2\pi}{\Omega} - \cos \theta_k \right) - \frac{1}{16} |\delta| \left(\cos^2 \frac{2\pi}{\Omega} - \cos^2 \theta_k \right) \\
& = \frac{1}{8} \left(3 - |\delta| \sin \frac{2\pi}{\Omega} \right) \left(\cos \frac{2\pi}{\Omega} - \cos \theta_k \right) + \frac{1}{16} (1 - |\delta|) \left(\cos^2 \frac{2\pi}{\Omega} - \cos^2 \theta_k \right) \quad (\text{B.15})
\end{aligned}$$

を得る. $\cos(2\pi/\Omega) > \cos \theta_k$ および $|\delta| \geq 1$ からこれは正である. 以上から図 B.3 を得る.

図 B.4 設定 (a) $\delta = \delta'$ の下で SLEM を与える領域.

B.1.4 まとめ

B.1.1 節から B.1.3 節の結果をまとめると, 設定 (a) $\delta = \delta'$ における SLEM $\tilde{\eta}$ は

$$\left\{ \begin{array}{l} \tilde{\eta} = \tilde{\lambda}_1^- \quad 0 \leq \gamma \leq \min \left[\frac{1}{4} \left(\sin^2 \frac{\pi}{\Omega} + \delta^2 \cos^2 \frac{\pi}{\Omega} \right), \right. \\ \quad \left. \frac{1}{16} \left(5 - \cos \frac{2\pi}{\Omega} \right) - \frac{1}{16} \sqrt{\left(5 \cos \frac{2\pi}{\Omega} - 1 \right) \left(\cos \frac{2\pi}{\Omega} + 3 \right) + 4\delta^2 \sin^2 \frac{2\pi}{\Omega}} \right] \\ \quad \text{のとき,} \\ \tilde{\eta} = |\tilde{\lambda}_{\Omega/2+1}^-| \quad \frac{3}{8} \left(\frac{5}{3} + \frac{1}{2} \cos \frac{2\pi}{\Omega} \right) + \frac{1}{8} \sqrt{\left(\frac{5}{4} + \frac{1}{4} \cos \frac{2\pi}{\Omega} \right)^2 - \frac{1}{2} \delta^2 \sin^2 \frac{2\pi}{\Omega}} \leq \gamma \leq \frac{3}{4} \\ \quad \text{のとき,} \\ \tilde{\eta} = |\tilde{\lambda}_2^+| \quad \text{上記以外するとき} \end{array} \right. \quad (\text{B.16})$$

である. またこれらを図示して図 B.4 を得る.

B.2 設定 (b) $\delta' = 1$ の第 2 固有値

次に設定 (b) $\delta' = 1$ のときを考察する. このとき遷移確率行列 \tilde{T} の固有値 $\tilde{\lambda}_k^\pm$ は, $k = 1, 2, \dots, \Omega/2 + 1$ に対して

$$\tilde{\lambda}_k^\pm = 1 - a + a \cos \left[\frac{\pi}{\Omega} (k-1) \right] - \gamma \pm \sqrt{\gamma^2 - a^2 \sin^2 \left[\frac{2\pi}{\Omega} (k-1) \right]} \quad (\text{B.17})$$

で与えられる. ここで $a \equiv (1 + \delta)/8$ である. パラメタはそれぞれ $0 \leq \gamma \leq (7 - \delta)/8$, $|\delta| \leq 1$ をみたす. 以下では固有値に現れる根号内の符号に基づいて場合分けし, そのそれぞれにおいて SLEM を明らかにする.

B.2.1 すべての固有値が実数のとき

$\gamma \geq (1 + \delta)/8$ のとき, すべての固有値が実数であることが保証される. このとき SLEM の候補は

- $\tilde{\lambda}_2^+ = 1 - a + a \cos \frac{2\pi}{\Omega} - \gamma + \sqrt{\gamma^2 - a^2 \sin^2 \frac{2\pi}{\Omega}}$
- $\tilde{\lambda}_1^- = 1 - 2\gamma$
- $|\tilde{\lambda}_{\Omega/2+1}^-| = |1 - 2a - 2\gamma|$

である. 特に, いま考えているすべてのパラメタ領域で $\min\{\tilde{\lambda}_2^+, \tilde{\lambda}_1^-\} > \tilde{\lambda}_{\Omega/2+1}^-$ が成り立つことから, $|\tilde{\lambda}_{\Omega/2+1}^-|$ については $1 - 2a - 2\gamma < 0$, すなわち $\gamma \geq (3 - \delta)/8$ の場合のみ考えればよい.

$\tilde{\lambda}_2^+ \geq \tilde{\lambda}_1^-$ となる領域

固有値 $\tilde{\lambda}_2^+$ と $\tilde{\lambda}_1^-$ の大小を比較する.

$$\tilde{\lambda}_2^+ - \tilde{\lambda}_1^- = \gamma - 2a \sin^2 \frac{\pi}{\Omega} + \sqrt{\gamma^2 - a^2 \sin^2 \frac{2\pi}{\Omega}} \quad (\text{B.18})$$

より,

$$\gamma \geq 2a \sin^2 \frac{\pi}{\Omega} = \frac{1 + \delta}{4} \sin^2 \frac{\pi}{\Omega} \quad (\text{B.19})$$

では不等式 $\tilde{\lambda}_2^+ \geq \tilde{\lambda}_1^-$ が成り立つ. 一方, $\gamma \leq 2a \sin^2(\pi/\Omega)$ では

$$\left(\sqrt{\gamma^2 - a^2 \sin^2 \frac{2\pi}{\Omega}} \right)^2 - \left(\gamma - 2a \sin^2 \frac{\pi}{\Omega} \right)^2 = 4a(\gamma - a) \sin^2 \frac{\pi}{\Omega} \quad (\text{B.20})$$

より,

$$\gamma \geq a = \frac{1 + \delta}{8} \quad (\text{B.21})$$

のとき不等式 $\tilde{\lambda}_2^+ \geq \tilde{\lambda}_1^-$ が成り立つ. しかし Ω が十分大きいとき $2 \sin^2(\pi/\Omega) < 1$ となるので, 結局 $\tilde{\lambda}_2^+ \geq \tilde{\lambda}_1^-$ となるのは

$$\gamma \geq 2a \sin^2 \frac{\pi}{\Omega} = \frac{1 + \delta}{4} \sin^2 \frac{\pi}{\Omega} \quad (\text{B.22})$$

のときである.

$\tilde{\lambda}_2^+ \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ となる領域

$\gamma \geq (3 - \delta)/8$ の下で

$$\begin{aligned} \tilde{\lambda}_2^+ - |\tilde{\lambda}_{\Omega/2+1}^-| &= 1 - 2a \sin^2 \frac{\pi}{\Omega} - \gamma + \sqrt{\gamma^2 - a^2 \sin^2 \frac{2\pi}{\Omega}} - (2a + 2\gamma - 1) \\ &= \sqrt{\gamma^2 - a^2 \sin^2 \frac{2\pi}{\Omega}} - \left[3\gamma - 2 + 2a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right] \end{aligned} \quad (\text{B.23})$$

であるので,

$$\gamma \leq \frac{2}{3} \left[1 - a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right] = \frac{2}{3} \left[1 - \frac{1 + \delta}{8} \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right] \quad (\text{B.24})$$

のときは不等式 $\tilde{\lambda}_2^+ \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ が成り立つ。一方, $\gamma \geq (2/3)\{1 - a[1 + \sin^2(\pi/\Omega)]\}$ のときは

$$\begin{aligned} &\left(\sqrt{\gamma^2 - a^2 \sin^2 \frac{2\pi}{\Omega}} \right)^2 - \left[3\gamma - 2 + 2a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right]^2 \\ &= -8\gamma^2 + 12\gamma \left[1 - a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right] - 4 \left[1 - a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right]^2 + a^2 \sin^2 \frac{2\pi}{\Omega} \\ &= -8 \left\{ \gamma - \frac{3}{4} \left[1 - a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right] \right\}^2 + \frac{1}{2} \left[1 - a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right]^2 - a^2 \sin^2 \frac{2\pi}{\Omega}, \end{aligned} \quad (\text{B.25})$$

すなわち

$$\begin{aligned} &\frac{3}{4} \left[1 - a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right] - \frac{1}{4} \sqrt{\left[1 - a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right]^2 - 2a^2 \sin^2 \frac{2\pi}{\Omega}} \\ &\leq \gamma \leq \frac{3}{4} \left[1 - a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right] + \frac{1}{4} \sqrt{\left[1 - a \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right]^2 - 2a^2 \sin^2 \frac{2\pi}{\Omega}} \end{aligned} \quad (\text{B.26})$$

のとき不等式 $\tilde{\lambda}_2^+ \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ が成り立つ*1。これに $a = (1 + \delta)/8$ を代入して δ と γ の関係式を得る。

$\tilde{\lambda}_1^- \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ となる領域

$\gamma \geq (3 - \delta)/8$ の下で

$$\tilde{\lambda}_1^- - |\tilde{\lambda}_{\Omega/2+1}^-| = 1 - 2\gamma - (2a + 2\gamma - 1) = -4 \left(\gamma - \frac{1 - a}{2} \right) \quad (\text{B.27})$$

*1 根号の中が負のときは, 不等式 $|\tilde{\lambda}_2^+| \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ を成り立たせる γ は存在しない。しかし Ω を十分大きくとれば, 根号の中は正になる。

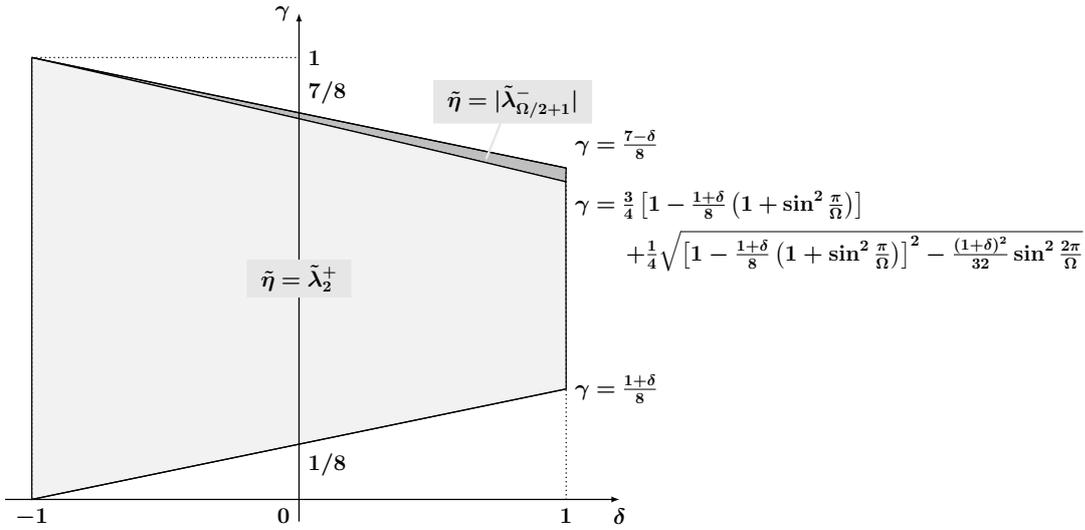


図 B.5 すべての固有値が実数のとき SLEM を与える領域.

であるので,

$$\gamma \leq \frac{1-a}{2} = \frac{7-\delta}{16} \quad (\text{B.28})$$

のとき不等式 $\tilde{\lambda}_1^- \geq |\tilde{\lambda}_{\Omega/2+1}^-|$ が成り立つ. 以上をあわせて図 B.5 を得る.

B.2.2 ほぼすべての固有値が複素数のとき

$k = 1, \Omega/2 + 1$ に対応する 4 つの固有値 $\tilde{\lambda}_1^\pm, \tilde{\lambda}_{\Omega/2+1}^\pm$ はつねに実数である. それ以外の固有値は, $\gamma \leq a \sin(2\pi/\Omega) = [(1+\delta)/8] \sin(2\pi/\Omega)$ のときすべて複素数となる. $2 \leq k \leq [(\Omega+1)/2]$ のとき固有値 $\tilde{\lambda}_k^\pm$ の絶対値 $|\tilde{\lambda}_k^\pm|$ は, $x \equiv \cos \theta_k$ とおけば

$$\begin{aligned} |\tilde{\lambda}_k^\pm|^2 &= (1-a+ax-\gamma)^2 + a^2(1-x^2) - \gamma^2 \\ &= 2a(1-a-\gamma)x + (1-a-\gamma)^2 + a^2 - \gamma^2 \end{aligned} \quad (\text{B.29})$$

と計算できる. 特に $1-a-\gamma > 0$ に注意すると, これは x について単調増加関数であることがわかる. すなわち $2 \leq k \leq [(\Omega+1)/2]$ における固有値 $\tilde{\lambda}_k^\pm$ の絶対値 $|\tilde{\lambda}_k^\pm|$ は $k=2$ で最大値をとる. また γ は十分小さいので $|\tilde{\lambda}_{\Omega/2+1}^-|$ はここでの SLEM の候補から除外してよい. よってこのとき SLEM の候補は

- $|\tilde{\lambda}_2^\pm| = \sqrt{\left(1-a+a \cos \frac{2\pi}{\Omega} - \gamma\right)^2 + a^2 \sin^2 \frac{2\pi}{\Omega} - \gamma^2}$
- $\tilde{\lambda}_1^- = 1 - 2\gamma > 0$

のいずれかである.

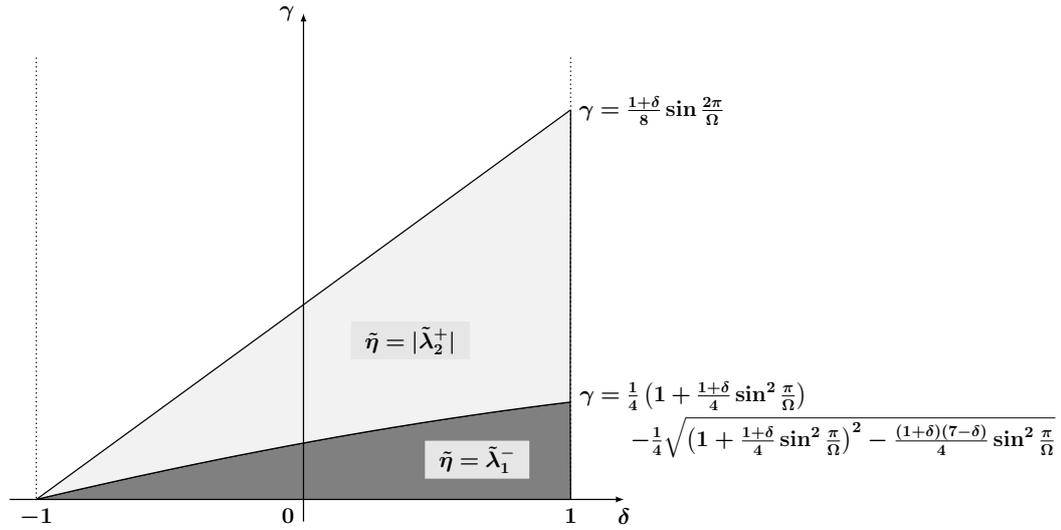


図 B.6 $\tilde{\lambda}_1^\pm, \tilde{\lambda}_{\Omega/2+1}^\pm$ 以外の固有値がすべて複素数のときに SLEM を与える領域.

$|\tilde{\lambda}_2^+| \geq |\tilde{\lambda}_1^-|$ となる領域

$$\begin{aligned}
 |\tilde{\lambda}_2^\pm|^2 - |\tilde{\lambda}_1^-|^2 &= \left(1 - a + a \cos \frac{2\pi}{\Omega} - \gamma\right)^2 + a^2 \sin^2 \frac{2\pi}{\Omega} - \gamma^2 - (1 - 2\gamma)^2 \\
 &= -4 \left[\gamma - \frac{1}{4} \left(1 + 2a \sin^2 \frac{\pi}{\Omega}\right) \right]^2 + \frac{1}{16} \left(1 + 2a \sin^2 \frac{\pi}{\Omega}\right)^2 - a(1 - a) \sin^2 \frac{\pi}{\Omega}
 \end{aligned} \tag{B.30}$$

より,

$$\begin{aligned}
 &\frac{1}{4} \left(1 + 2a \sin^2 \frac{\pi}{\Omega}\right) - \frac{1}{4} \sqrt{\left(1 + 2a \sin^2 \frac{\pi}{\Omega}\right)^2 - 16a(1 - a) \sin^2 \frac{\pi}{\Omega}} \\
 &\leq \gamma \leq \frac{1}{4} \left(1 + 2a \sin^2 \frac{\pi}{\Omega}\right) + \frac{1}{4} \sqrt{\left(1 + 2a \sin^2 \frac{\pi}{\Omega}\right)^2 - 16a(1 - a) \sin^2 \frac{\pi}{\Omega}}
 \end{aligned} \tag{B.31}$$

のとき不等式 $|\tilde{\lambda}_2^+| \geq |\tilde{\lambda}_1^-|$ が成り立つ*2. 以上から図 B.6 を得る.

B.2.3 それ以外の場合

$a \sin(2\pi/\Omega) \leq \gamma \leq a$ のとき, 実固有値も複素固有値もともに現れる. ある整数 $k = 2, 3, \dots, [(\Omega + 1)/2]$ に対して, $\tilde{\lambda}_k^\pm$ が実固有値のときは, $\tilde{\lambda}_2^+$ のほうが絶対値が大きいことがわかる. したがって B.2.1 節と同様の結論を得る. 次に $\tilde{\lambda}_k^\pm$ が複素固有値のと

*2 根号の中が負のときは不等式 $|\tilde{\lambda}_2^+| \geq |\tilde{\lambda}_1^-|$ を成り立たせる γ は存在しない. しかし Ω が十分大きければ, 根号の中は正になる.

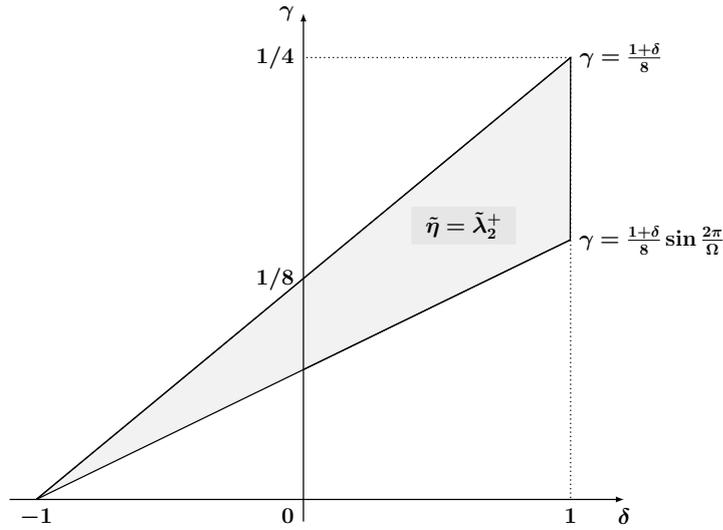


図 B.7 $a \sin(2\pi/\Omega) \leq \gamma \leq a$ のとき SLEM を与える領域.

きを考える. このとき $\tilde{\lambda}_2^+$ は実固有値であり, パラメタ Ω, a を固定したときの $|\tilde{\lambda}_2^+|$ は $\gamma = a \sin(2\pi/\Omega)$ のとき最小値をとるので, 他の複素固有値の絶対値がこれより小さいことを示せば, SLEM の候補は B.2.1 節と同じになる. 以下この事実を示す.

$k = 3, 4, \dots, [(\Omega + 1)/2]$ に対して $\tilde{\lambda}_k^\pm$ が複素固有値のとき, すなわち $a \sin(2\pi/\Omega) \leq \gamma \leq a \sin \theta_k$ のとき

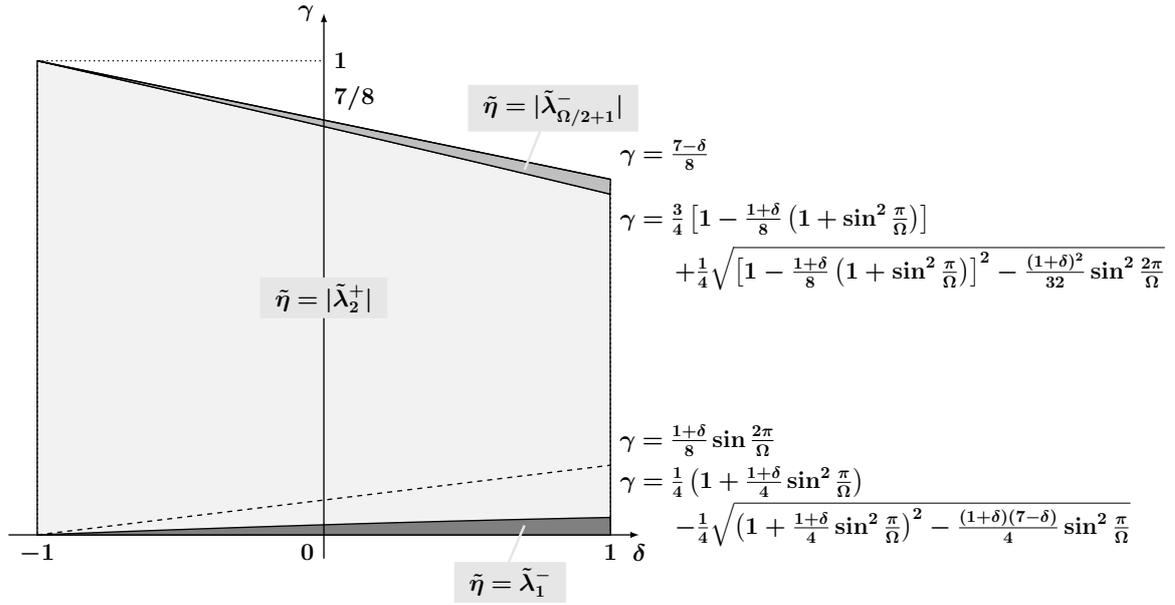
$$\begin{aligned}
 & \left(\min_{\gamma} |\tilde{\lambda}_2^+| \right)^2 - |\tilde{\lambda}_k^\pm|^2 \\
 &= \left(1 - a + a \cos \frac{2\pi}{\Omega} - a \sin \frac{2\pi}{\Omega} \right)^2 - \left[(1 - a + a \cos \theta_k - \gamma)^2 + a^2 \sin^2 \theta_k - \gamma^2 \right] \\
 &= 2(1 - a + a \cos \theta_k) \gamma + 2a(1 - a) \left(\cos \frac{2\pi}{\Omega} - \cos \theta_k \right) \\
 & \quad - 2a(1 - a) \sin \frac{2\pi}{\Omega} - 2a^2 \sin \frac{2\pi}{\Omega} \cos \frac{2\pi}{\Omega} \quad (\text{B.32})
 \end{aligned}$$

である. 右辺における γ の係数は, $a = (1 + \delta)/8$ のとき $0 \leq a \leq 1/4$ なので正である. よって右辺は $\gamma = a \sin(2\pi/\Omega)$ のとき最小となり,

$$\left(\min_{\gamma} |\tilde{\lambda}_2^+| \right)^2 - |\tilde{\lambda}_k^\pm|^2 \geq 2a(1 - a) \left(\cos \frac{2\pi}{\Omega} - \cos \theta_k \right) \left(1 - a - a \sin \frac{2\pi}{\Omega} \right) \quad (\text{B.33})$$

と評価できる. $\cos(2\pi/\Omega) > \cos \theta_k$ および $0 \leq a \leq 1/4$ からこれは正である.

以上から, $k \geq 3$ に対応する固有値が複素固有値になっても, $\tilde{\lambda}_2^+$ が実固有値である限りは絶対値は $\tilde{\lambda}_2^+$ のほうが大きいことが示された. これと B.2.1 節での結果をあわせると, 図 B.7 を得る.

図 B.8 設定 (b) $\delta' = 1$ のとき SLEM を与える領域.

B.2.4 まとめ

B.2.1 節から B.2.3 節までの結果をまとめると、設定 (b) $\delta' = 1$ における SLEM $\tilde{\eta}$ は

$$\left\{ \begin{array}{ll} \tilde{\eta} = \tilde{\lambda}_1^- & 0 \leq \gamma \leq \frac{1}{4} \left(1 + \frac{1+\delta}{4} \sin^2 \frac{\pi}{\Omega} \right) \\ & -\frac{1}{4} \sqrt{\left(1 + \frac{1+\delta}{4} \sin^2 \frac{\pi}{\Omega} \right)^2 - \frac{(1+\delta)(7-\delta)}{4} \sin^2 \frac{\pi}{\Omega}} \end{array} \right. \quad \text{のとき,}$$

$$\left\{ \begin{array}{ll} \tilde{\eta} = |\tilde{\lambda}_{\Omega/2+1}^-| & \frac{3}{4} \left[1 - \frac{1+\delta}{8} \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right] \\ & + \frac{1}{4} \sqrt{\left[1 - \frac{1+\delta}{8} \left(1 + \sin^2 \frac{\pi}{\Omega} \right) \right]^2 - \frac{(1+\delta)^2}{32} \sin^2 \frac{2\pi}{\Omega}} \leq \gamma \leq \frac{7-\delta}{8} \end{array} \right. \quad \text{のとき,}$$

$$\left\{ \begin{array}{ll} \tilde{\eta} = |\tilde{\lambda}_2^+| & \text{上記以外} \end{array} \right. \quad \text{のとき,}$$

(B.34)

である。またこれらを図示して図 B.8 を得る。

付録 C

性能指数の最悪評価

本付録では, 第 4 章で議論した詳細つり合い条件をみたす 1 次元ランダムウォークおよびねじれ詳細つり合い条件をみたす 1 次元ランダムウォークのそれぞれについて, 漸近分散と分散の比 (performance index [45], 性能指数) の最悪評価を与える.

C.1 詳細つり合い条件をみたす場合

ある目的分布 π に対して詳細つり合い条件をみたす (1 次元ランダムウォークに限らない) 一般の遷移確率行列 \mathbf{T} を考える. このとき, 物理量 f の漸近分散 $v(f, \pi, \mathbf{T})$ は遷移確率行列 \mathbf{T} の固有値 λ_k , 左固有ベクトル $\mathbf{u}_{k,\sigma}$ および右固有ベクトル $\mathbf{v}_{k,\sigma}$ と, 物理量 f に対応するベクトル \mathbf{f} を用いて

$$v(f, \pi, \mathbf{T}) = \sum_{k=2}^K \frac{1 + \lambda_k}{1 - \lambda_k} \left[\sum_{\sigma=1}^{m_k} (\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^\top) (\mathbf{u}_{k,\sigma} \mathbf{f}^\top) \right] \quad (\text{C.1})$$

で与えられる. ここで行列 $\mathbf{B} = \text{diag}(\pi_i)$ は目的分布 π から定まる対角行列である. また物理量 f の分散 $\text{var}_\pi[f]$ は

$$\text{var}_\pi[f] = \sum_{k=2}^K \left[\sum_{\sigma=1}^{m_k} (\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^\top) (\mathbf{u}_{k,\sigma} \mathbf{f}^\top) \right] \quad (\text{C.2})$$

と表せる. これより性能指数 $\Gamma(f, \pi, \mathbf{T})$ は

$$\Gamma(f, \pi, \mathbf{T}) \equiv \frac{v(f, \pi, \mathbf{T})}{\text{var}_\pi[f]} = \sum_{k=2}^K \frac{1 + \lambda_k}{1 - \lambda_k} a_k \quad (\text{C.3})$$

と表せる. ここで

$$a_k \equiv \sum_{\sigma=1}^{m_k} \frac{(\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^\top) (\mathbf{u}_{k,\sigma} \mathbf{f}^\top)}{\text{var}_\pi[f]} \quad (\text{C.4})$$

とおいた. 特に $\sum_{k=2}^K a_k = 1$ である.

遷移確率行列 \mathbf{T} が目的分布 $\boldsymbol{\pi}$ に対して詳細つり合い条件をみたすことから, 左固有ベクトル $\mathbf{u}_{k,\sigma}$ と右固有ベクトル $\mathbf{v}_{k,\sigma}$ には $\mathbf{u}_{k,\sigma} = \mathbf{v}_{k,\sigma} \mathbf{B}$ なる関係が成り立つ. これを用いると, 性能指数 $\Gamma(f, \boldsymbol{\pi}, \mathbf{T})$ の表式 (C.3) に現れる項 $(\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^\top)(\mathbf{u}_{k,\sigma} \mathbf{f}^\top)$ は

$$(\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^\top)(\mathbf{u}_{k,\sigma} \mathbf{f}^\top) = (\mathbf{u}_{k,\sigma} \mathbf{f}^\top)^2 \geq 0 \quad (\text{C.5})$$

より非負であることがわかる. したがってこのとき性能指数の最悪評価とは, 条件 $a_k \geq 0$ かつ $\sum_{k=2}^K a_k = 1$ の下で性能指数の最大値を考えることにほかならない. いま関数 $f(x) = (1+x)/(1-x)$ が $-1 \leq x < 1$ で単調増加関数であること, および a_k に対する条件から, (C.3) 式の最大値は絶対値をとらない意味での第 2 固有値を与える k について $a_k = 1$ であるときに実現する. すなわち遷移確率行列 \mathbf{T} の固有値 $\lambda_k \in \mathbb{R}$ を $1 = \lambda_1 > \lambda_2 > \dots > \lambda_K > -1$ と並べたとき, 性能指数の最大値 Γ は

$$\Gamma \equiv \max \Gamma(f, \boldsymbol{\pi}, \mathbf{T}) = \frac{1 + \lambda_2}{1 - \lambda_2} \quad (\text{C.6})$$

で与えられる. また最大を実現する物理量 f は, 対応するベクトル \mathbf{f} が第 2 固有値 λ_2 に属する右固有ベクトル $\{\mathbf{v}_{2,\sigma}\}_{\sigma=1}^{m_2}$ の線形結合で与えられるものである.

以上の議論から, 4.1 節における詳細つり合い条件をみたす 1 次元ランダムウォークでは, 性能指数の最悪評価は

$$\Gamma = \frac{1}{\alpha \sin^2 \frac{\pi}{\Omega}} - 1 \quad (\text{C.7})$$

で与えられる.

C.2 詳細つり合い条件を破る場合

目的分布 $\boldsymbol{\pi}$ に対してつり合い条件をみたすものの詳細つり合い条件を破る遷移確率行列について考える. このとき漸近分散の表式は詳細つり合い条件をみたす場合に成り立つ (C.1) 式をそのまま使えるものの, 性能指数の最悪評価は一般に難しい. 漸近分散の非負性は定義から保証されるものの, 複素固有値が出現しうることや $a_k \in \mathbb{C}$ となりうることなどから a_k の非負性を保証できず, 詳細つり合い条件をみたす場合の議論をそのまま当てはめることはできないのだ. ここでは 4.2 節で扱った, ねじれ詳細つり合い条件をみたす 1 次元ランダムウォークにおける性能指数を考察する.

以下では物理量として lifting parameter ε によらないものだけを考える. すなわち, 拡張前の空間に定義された物理量 f に対応するベクトルを $\mathbf{f} \in \mathbb{R}^\Omega$ とするとき, ここで考える物理量 f' に対応するベクトルは $\mathbf{f}' = (\mathbf{f}, \mathbf{f}) \in \mathbb{R}^{2\Omega}$ で与えられる.

C.2.1 対角化可能である場合

拡張後の遷移確率行列 $\tilde{\mathbf{T}}$ が対角化可能であるとき、漸近分散 $v(f', \tilde{\boldsymbol{\pi}}, \tilde{\mathbf{T}})$ は付録 A の結果を用いて

$$\begin{aligned} v(f', \tilde{\boldsymbol{\pi}}, \tilde{\mathbf{T}}) &= \sum_{(k,\varepsilon) \neq (1,+)} \frac{1 + \tilde{\lambda}_k^\varepsilon}{1 - \tilde{\lambda}_k^\varepsilon} \left[\sum_{\sigma=1}^{m_k} (\mathbf{f}' \tilde{\mathbf{B}} \tilde{\mathbf{v}}_{k,\sigma}^\varepsilon) (\tilde{\mathbf{u}}_{k,\sigma}^\varepsilon \mathbf{f}'^\top) \right] \\ &= \sum_{k=2}^{[\Omega/2]+1} \left[\sum_{\varepsilon=\pm} \frac{1}{1 - (A_k^\varepsilon)^2} \frac{1 + \tilde{\lambda}_k^\varepsilon}{1 - \tilde{\lambda}_k^\varepsilon} \right] \left[\sum_{\sigma=1}^{m_k} (\mathbf{f} \mathbf{B} \mathbf{v}_{k,\sigma}^\top) (\mathbf{u}_{k,\sigma} \mathbf{f}^\top) \right] \end{aligned} \quad (\text{C.8})$$

と導かれる。ここで

$$\tilde{\mathbf{B}} \equiv \frac{1}{2} \begin{pmatrix} \mathbf{B} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{pmatrix} \in \mathbb{R}^{2\Omega \times 2\Omega} \quad (\text{C.9})$$

である。また $\mathbf{O} \in \mathbb{R}^{\Omega \times \Omega}$ は零行列を表す。漸近分散 $v(f', \tilde{\boldsymbol{\pi}}, \tilde{\mathbf{T}})$ は拡張後の遷移確率行列 $\tilde{\mathbf{T}}$ の固有ベクトルで構成されていたが、それを拡張前の遷移確率行列 \mathbf{T} の固有ベクトルだけで、前節と同じように書けることがポイントである。すなわち、性能指数 $\tilde{\Gamma}(f', \tilde{\boldsymbol{\pi}}, \tilde{\mathbf{T}})$ は前節と同じ a_k を用いて

$$\tilde{\Gamma}(f', \tilde{\boldsymbol{\pi}}, \tilde{\mathbf{T}}) = \sum_{k=2}^{[\Omega/2]+1} V_k a_k \quad (\text{C.10})$$

で与えられる。ここで

$$V_k \equiv \frac{(a+b) \sin^2 \frac{\theta_k}{2} + c}{(a+b)^2 \sin^4 \frac{\theta_k}{2} + (a-b)^2 \sin^2 \frac{\theta_k}{2} \cos^2 \frac{\theta_k}{2} + (a+b)c \sin^2 \frac{\theta_k}{2}} - 1 \quad (\text{C.11})$$

および $a = (1 + \delta)\alpha/4$, $b = (1 - \delta')\alpha/4$, $c = \gamma$ である。以上から性能指数の最悪評価 $\tilde{\Gamma}$ は、 V_k の最大を与える k について $a_k = 1$ とすることで得られることがわかる。よって以下では V_k の最大値を考察する。

$x = \sin^2(\theta_k/2)$ とおくと、 $2 \leq k \leq [\Omega/2] + 1$ で $\sin^2(\pi/\Omega) \leq x \leq 1$ であり、 x は k について単調増加関数である。また

$$\begin{aligned} V_k + 1 &= \frac{(a+b)x + c}{(a+b)^2 x^2 + (a-b)^2 x(1-x) + (a+b)cx} \\ &= \frac{(a+b)x + c}{x[4abx + (a+b)c + (a-b)^2]} \end{aligned} \quad (\text{C.12})$$

である。さらにこれを部分分数分解すると

$$V_k + 1 = \left[\frac{c}{(a+b)c + (a-b)^2} \right] \frac{1}{x} + \left[\frac{(a-b)^2(a+b+c)}{(a+b)c + (a-b)^2} \right] \frac{1}{4abx + (a+b)c + (a-b)^2} \quad (\text{C.13})$$

を得る. これより, V_k は $x > 0$ で単調減少関数であることがわかる. したがって V_k は $k = 2$ のときに最大値をとる. 以上から, 性能指数の最悪評価 $\tilde{\Gamma}$ は

$$\begin{aligned} \tilde{\Gamma} &= \max \tilde{\Gamma}(f', \tilde{\pi}, \tilde{T}') \\ &= \frac{(a+b) \sin^2 \frac{\pi}{\Omega} + c}{(a+b)^2 \sin^4 \frac{\pi}{\Omega} + (a-b)^2 \sin^2 \frac{\pi}{\Omega} \cos^2 \frac{\pi}{\Omega} + (a+b)c \sin^2 \frac{\pi}{\Omega}} - 1 \end{aligned} \quad (\text{C.14})$$

が得られる. ここに $a = (1+\delta)\alpha/4$, $b = (1-\delta')\alpha/4$, $c = \gamma$ を代入して,

$$\tilde{\Gamma} = \frac{4}{\alpha} \frac{2 + \delta - \delta'}{(2 + \delta - \delta')^2 \sin^2 \frac{\pi}{\Omega} + \phi(\delta + \delta')^2 \cos^2 \frac{\pi}{\Omega}} - 1 \quad (\text{C.15})$$

および

$$\phi \equiv \frac{\sin^2 \frac{\pi}{\Omega}}{\sin^2 \frac{\pi}{\Omega} + \frac{2\gamma}{(2+\delta-\delta')\alpha}} \quad (\text{C.16})$$

を得る.

C.2.2 対角化可能でない場合

ある整数 $q \in \{2, 3, \dots, [(\Omega+1)/2]\}$ に対して $c = |a-b| \sin \theta_q$ としたとき, 遷移確率行列 \tilde{T} は対角化可能でなくなる. それでも遷移確率行列 \tilde{T} の Jordan 標準形を考えることで, スペクトル分解に修正を加えたものが成り立つことを付録 A でみた. ここでは (C.10) 式の性能指数 $\tilde{\Gamma}(f', \tilde{\pi}, \tilde{T})$ に現れる係数 V_q について, 遷移確率行列 \tilde{T} が対角化可能でない場合でも (C.11) 式がそのまま成り立つことを示す. これより前小節の結論もそのまま成り立つことがわかる.

$c = |a-b| \sin \theta_q$ のとき, 遷移確率行列 \tilde{T} の修正されたスペクトル分解における固有値 $\tilde{\lambda}_q \equiv \tilde{\lambda}_q^\pm$ に対応する項は, 付録 A から

$$\tilde{\lambda}_q \sum_{x=i, ii, iii, iv} \tilde{\mathbf{v}}_q^{(x)\top} \tilde{\mathbf{u}}_q^{(x)} + \tilde{\mathbf{v}}_q^{(i)\top} \tilde{\mathbf{u}}_q^{(ii)} + \tilde{\mathbf{v}}_q^{(iii)\top} \tilde{\mathbf{u}}_q^{(iv)} \quad (\text{C.17})$$

で与えられる. ここで

$$\tilde{\lambda}_q = 1 - 2(a+b) \sin^2 \frac{\theta_q}{2} - |a-b| \sin \theta_q \quad (\text{C.18})$$

である. 修正されたスペクトル分解から遷移確率行列 \tilde{T} の n 乗を計算することで, 性能指数 $\tilde{\Gamma}(f', \tilde{\pi}, \tilde{T})$ の固有値 $\tilde{\lambda}_q$ に対応する項における a_k の係数 V'_q が

$$V'_q = \frac{1 + \tilde{\lambda}_q}{1 - \tilde{\lambda}_q} + \frac{2|a-b| \sin \theta_q}{(1 - \tilde{\lambda}_q)^2} \quad (\text{C.19})$$

となることがわかる. これに $\tilde{\lambda}_q$ の表式を代入すると

$$V'_q = \frac{(a+b)\sin\frac{\theta_q}{2} + 2|a-b|\cos\frac{\theta_q}{2}}{\sin\frac{\theta_q}{2}\left[(a+b)\sin\frac{\theta_q}{2} + |a-b|\cos\frac{\theta_q}{2}\right]^2} - 1 \quad (\text{C.20})$$

と表せる. 一方対角化可能な場合の係数 V_k (C.11) において $k = q$, $c = |a-b|\sin\theta_q$ を代入すると

$$\begin{aligned} V_q &= \frac{(a+b)\sin\frac{\theta_q}{2} + |a-b|\sin\theta_q}{(a+b)^2\sin^4\frac{\theta_q}{2} + (a-b)^2\sin^2\frac{\theta_q}{2}\cos^2\frac{\theta_q}{2} + (a+b)|a-b|\sin^2\frac{\theta_q}{2}\sin\theta_q} - 1 \\ &= \frac{(a+b)\sin\frac{\theta_q}{2} + 2|a-b|\cos\frac{\theta_q}{2}}{\sin\frac{\theta_q}{2}\left[(a+b)\sin\frac{\theta_q}{2} + |a-b|\cos\frac{\theta_q}{2}\right]^2} - 1 \end{aligned} \quad (\text{C.21})$$

となり, V'_q に一致する. すなわち, 遷移確率行列 $\tilde{\mathbf{T}}$ が対角化可能でない場合でも (C.10) 式はそのまま成り立つことが示された.

参考文献

- [1] 藤原邦男, 基礎物理学 1 『物理学序論としての力学』 (東京大学出版会, 1984).
- [2] 夏目雄平・小川建吾, 基礎物理学シリーズ 13 『計算物理 I』 (朝倉書店, 2002).
- [3] 齊藤宣一, 大学数学の入門 9 『数値解析入門』 (東京大学出版会, 2012).
- [4] 田中和之 編著, 臨時別冊・数理科学 SGC ライブラリ-50 『確率的情報処理と統計力学 様々なアプローチとそのチュートリアル』 (サイエンス社, 2006).
- [5] S. Ulam, R. D. Richtmyer, and J. von Neumann, *Statistical methods in neutron diffusion*, Report LAMS-551 (Los Alamos Scientific Laboratory, Los Alamos, 1947).
- [6] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [7] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, 4th ed. (Cambridge University Press, Cambridge, 2014).
- [8] J. S. Liu, *Monte Carlo Strategies in Scientific Computing* (Springer-Verlag, New York, 2001).
- [9] A. Doucet, N. De Freitas, and N. Gordon (editors), *Sequential Monte Carlo Methods in Practice* (Springer-Verlag, New York, 2001).
- [10] A. P. Lyubartsev, A. A. Martinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, *J. Chem. Phys.* **96**, 1776 (1992).
- [11] E. Marinari and G. Parisi, *Europhys. Lett.* **19**, 451 (1992).
- [12] B. A. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68**, 9 (1992).
- [13] K. Hukushima and K. Nemoto, *J. Phys. Soc. Jpn.* **65**, 1604 (1996).
- [14] R. H. Swendsen and J. -S. Wang, *Phys. Rev. Lett.* **58**, 86 (1987).
- [15] U. Wolff, *Phys. Rev. Lett.* **62**, 361 (1989).
- [16] P. H. Peskun, *Biometrika* **60**, 607 (1973).
- [17] W. K. Hastings, *Biometrika* **57**, 97 (1970).
- [18] A. A. Barker, *Aust. J. Phys.* **18**, 119 (1965).
- [19] R. Ren and G. Orkoulas, *J. Chem. Phys.* **124**, 064109 (2006).

- [20] H. Suwa and S. Todo, Phys. Rev. Lett. **105**, 120603 (2010).
- [21] S. Todo and H. Suwa, J. Phys.:Conf. Ser. **473**, 012013 (2013).
- [22] H. Suwa, *Geometrically Constructed Markov Chain Monte Carlo Study of Quantum Spin-phonon Complex Systems* (Springer Japan, Tokyo, 2014).
- [23] R. M. Neal, Technical Report, No. 0406, Department of Statistics, University of Toronto (2004) [arXiv:0407281].
- [24] Y. Sun, J. Schmidhuber, and F. Gomez, in *Advances in Neural Information Processing Systems 23*, edited by J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta (Curran Associates, Inc., New York, 2010), p. 2235.
- [25] J. Bierkens, Stat. Comput. **26**, 1213 (2016).
- [26] K. S. Turitsyn, M. Chertkov, and M. Vucelja, Physica D **240**, 410 (2011).
- [27] Y. Sakai and K. Hukushima, J. Phys. Soc. Jpn. **82**, 064003 (2016).
- [28] K. Hukushima and Y. Sakai, J. Phys.: Conf. Ser. **473**, 012012 (2013).
- [29] 酒井佑士, 修士論文『ねじれ詳細つり合い条件を用いたマルコフ連鎖モンテカルロ法』(東京大学, 2014).
- [30] H. C. Fernandes and M. Weigel, Comput. Phys. Commun. **182**, 1856 (2011).
- [31] R. D. Schram and G. T. Barkema, Physica A **418**, 88 (2015).
- [32] E. P. Bernard, W. Krauth, and D. B. Wilson, Phys. Rev. E **80**, 056704 (2009).
- [33] E. A. J. F. Peters and G. de With, Phys. Rev. E **85**, 026703 (2012).
- [34] M. Michel, S. C. Kapfer, and W. Krauth, J. Chem. Phys. **140**, 054116 (2014).
- [35] M. Michel, J. Mayer, and W. Krauth, Europhys. Lett. **112**, 20003 (2015).
- [36] Y. Nishikawa, M. Michel, W. Krauth, and K. Hukushima, Phys. Rev. E **92**, 063306 (2015).
- [37] Y. Nishikawa and K. Hukushima, Phys. Rev. B **94**, 064428 (2016).
- [38] Y. Nishikawa and K. Hukushima, J. Phys.: Conf. Ser. **750**, 012014 (2016).
- [39] A. Ichiki and M. Ohzeki, Phys. Rev. E **91**, 062105 (2015).
- [40] M. Ohzeki and A. Ichiki, Phys. Rev. E **92**, 012105 (2015).
- [41] M. Ohzeki and A. Ichiki, J. Phys.: Conf. Ser. **638**, 012003 (2015).
- [42] M. Ohzeki, J. Phys.: Conf. Ser. **699**, 012019 (2016).
- [43] K. Takahashi and M. Ohzeki, Phys. Rev. E **93**, 012129 (2016).
- [44] A. Ichiki and M. Ohzeki, Phys. Rev. E **88**, 020101(R) (2013).
- [45] P. Brémaud, *Markov Chains, Gibbs Field, Monte Carlo Simulation, and Queues* (Springer, New York, 1999).
- [46] 舟木直久, 講座〈数学の考え方〉20『確率論』(朝倉書店, 2004).

-
- [47] 山本哲朗, 臨時別冊・数理科学 SGC ライブラリ-79 『行列解析の基礎 Advanced 線形代数』 (サイエンス社, 2010).
- [48] 齋藤正彦, 基礎数学 1 『線形代数入門』 (東京大学出版会, 1966).
- [49] 伏見正則, シリーズ〈金融工学の基礎〉3 『確率と確率過程』 (朝倉書店, 2004).
- [50] Y. Sakai and K. Hukushima, Phys. Rev. E **93**, 043318 (2016).
- [51] L. Tierney, Ann. Statist. **22**, 1701 (1994).
- [52] A. Frigessi, C. -R. Hwang, and L. Younes, Ann. Appl. Probab. **2**, 610 (1992).
- [53] A. Mira, Stat. Sci. **16**, 340 (2001).
- [54] L. Pollet, S. M. A. Rombouts, K. V. Houcke, and K. Heyde, Phys. Rev. E **70**, 056705 (2004).
- [55] T. -L. Chen, W. -K. Chen, C. -R. Hwang, and H. -M. Pai, SIAM J. Control Optim. **50**, 2743 (2012).
- [56] J. Bierkens and G. Roberts, to be published in Ann. Appl. Probab. [arXiv:1509.00302].
- [57] P. Diaconis, S. Holmes, and R. M. Neal, Tech. Rep. BU-13850-M, Biometrics Unit, Cornell University (1997); Ann. Appl. Probab. **10**, 726 (2000).
- [58] F. Chen, L. Lovász, and I. Pak, *Proceedings of the 31th Annual ACM Symposium on Theory of Computing* (ACM, New York, 1999), p. 275.
- [59] M. Vucelja, Am. J. Phys. **84**, 958 (2016).
- [60] E. Vicari, Phys. Lett. B **309**, 139 (1993).
- [61] W. Kerler and P. Rehberg, Phys. Rev. E **50**, 4220 (1994).
- [62] B. Coluzzi, J. Phys. A: Math. Gen. **28**, 747 (1995).
- [63] L. A. Fernández, E. Marinari, and J. J. Ruiz-Lorenzo, J. Phys. I France **5**, 1247 (1995).
- [64] M. Picco and F. Ritort, Physica A **250**, 46 (1998).
- [65] A. Irbäck and F. Potthast, J. Chem. Phys. **103**, 10298 (1995).
- [66] A. Irbäck, C. Peterson, and F. Potthast, Phys. Rev. E **55**, 860 (1997).
- [67] C. J. Geyer and E. A. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
- [68] A. Mitsutake and Y. Okamoto, Chem. Phys. Lett. **332**, 131 (2000).
- [69] A. Mitsutake and Y. Okamoto, J. Chem. Phys. **121**, 2491 (2004).
- [70] Y. Li, V. A. Protopopescu, and A. Gorin, Phys. Lett. A **328**, 274 (2004).
- [71] Y. Mori and H. Okumura, J. Comput. Chem. **36**, 2344 (2015).
- [72] Y. Sakai and K. Hukushima, J. Phys. Soc. Jpn. **85**, 104002 (2016).
- [73] Y. Sakai and K. Hukushima, J. Phys.: Conf. Ser. **750**, 012013 (2016).
- [74] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **18**, 920 (1997).

- [75] S. Park and V. S. Pande, Phys. Rev. E **76**, 016703 (2007).
- [76] C. E. Fiore and M. G. E. de Luz, J. Chem. Phys. **133**, 244102 (2010).
- [77] P. H. Nguyen, Y. Okamoto, and P. Derreumaux, J. Chem. Phys. **138**, 061102 (2013).
- [78] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
- [79] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
- [80] H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 252 (1941).
- [81] L. Onsager, Phys. Rev. **65**, 117 (1944).
- [82] B. Kastening, Phys. Rev. E **64**, 066106 (2001).
- [83] F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).
- [84] P. Olsson, Phys. Rev. B **52**, 4526 (1995).
- [85] A. Valentim, M. G. E. da Luz, and C. E. Fiore, Comput. Phys. Commun. **185**, 2046 (2014).

List of Publications

Yuji Sakai and Koji Hukushima

Eigenvalue analysis of an irreversible random walk with skew detailed balance conditions

Phys. Rev. E **93**, 043318 (2016) [arXiv:1511.08100].

Yuji Sakai and Koji Hukushima

Irreversible Simulated Tempering

J. Phys. Soc. Jpn. **85**, 104002 (2016) [arXiv:1601.04286].

Yuji Sakai and Koji Hukushima

Irreversible simulated tempering algorithm with skew detailed balance conditions: a learning method of weight factors in simulated tempering

J. Phys: Conf. Ser. **750**, 012013 (2016).

謝辞

本博士論文の執筆にあたり、指導教員である福島孝治准教授に感謝いたします。学部の卒業研究から約5年半にわたってご指導いただきました。本博士論文におけるアイデアのほとんどは福島さんとの議論によるものです。ありがとうございました。

福島研究室のスタッフ、同期、後輩に感謝します。みなさんには毎週の進捗報告会で数多くの有益なコメントをいただきました。日頃の他愛のない会話もいい気分転換になりました。特に同期の高邊賢史氏には海外出張をはじめとして多くの場面で助けられました。福島研究室の院生が彼と私のふたりだったときの進捗報告会やM1セミナーは、私にとってかけがえのない財産となっています。

駒場キャンパス所属の同期に感謝します。不定期に開かれる同期の親睦会は大変よい息抜きとなりました。また院生部屋で同室となった加藤雄介研究室・国場敦夫研究室のみなさんにはお世話になりました。

物理学会やさまざまな研究会で私の発表にコメントいただいたすべての方に感謝します。みなさんからいただいたコメントの多くは本博士論文に反映され、内容をより一層充実したものにしてくれました。

筆者は平成26年度から3年間、独立行政法人日本学術振興会の特別研究員(DC1)として採用されました。本研究は「詳細つり合い条件を用いないマルコフ連鎖モンテカルロ法の理論的研究」として特別研究員奨励費(課題番号26・7868)の助成を受けています。

その他にも多くの方に、私の研究生活を支えていただきました。感謝いたします。最後に、私をここまで育て、経済的・精神的に支えてくれた家族に感謝します。本博士論文を妻・清香に捧げます。