

確率的セルラオートマタによる最適解の探索

北海道大学大学院 理学院 数学専攻
鎌倉 雄洋 (Katsuhiko Kamakura)

概要

組み合わせ最適化問題を直接的に解くことは一般的に難しい。これを解く手法の 1 つとして、マルコフ連鎖モンテカルロ法を使ったギブスサンプリングがある。ここでは確率的セルラオートマタと呼ばれる手法を使ったギブスサンプリングを考え、それが従来ギブスサンプラーとして用いられてきた手法であるグラウバー力学と比較してどのように優れているかを述べる。

1 はじめに

巡回セールスマン問題などを代表とする組み合わせ最適化問題は、直接的な計算をすると膨大な時間がかかる (文献 [3])。これを回避する方法として、上記のような問題をイジング模型に埋め込み、考えるコスト関数の最小を与える状態とイジング模型のエネルギー関数 $H(\sigma)$ の最小を与える状態 (基底状態) を対応させることで解く手法がある (文献 [8])。よく用いられている手法としてはマルコフ連鎖モンテカルロ法 (MCMC) を用いてギブス分布 $\pi_\beta^G(\sigma) \propto e^{-\beta H(\sigma)}$ の標本抽出を行うものである (ここで $\beta \geq 0$ は逆温度)。ギブス分布をサンプリングできる MCMC はいくつかあるが、もっともよく用いられているものとしてグラウバー力学というものがある。これは 1 ステップで 1 つずつスピンを更新するものであるが、この手法は効率が悪い可能性がある。ここでは全スピン同時更新を行えるアルゴリズムとして確率的セルラオートマタ (SCA) を導入し、グラウバー力学に比べていくつかの意味で優れていることを示す結果を紹介する。

2 マルコフ連鎖モンテカルロ法

まずマルコフ連鎖について必要となる事項を述べておく。

定義 1 (マルコフ連鎖). 有限標本空間 Ω 上の確率過程 $\{X_t\}$ が次を満たすとき $\{X_t\}$ を (時間一様) マルコフ連鎖という。

$$\begin{aligned} & \forall t \geq 0, \forall x_{t+1}, \dots, x_0 \in \Omega \\ & P(X_{t+1} = x_{t+1} | X_t = x_t, X_{t-1} = x_{t-1}, \dots, X_0 = x_0) \\ & = P(X_{t+1} = x_{t+1} | X_t = x_t) = P(X_1 = x_{t+1} | X_0 = x_t). \end{aligned}$$

特に上のマルコフ連鎖において、 $P(i, j) = P(X_1 = j | X_0 = i)$ を推移確率、 $P(i, j)$ を (i, j) 成分に持つ行列 P を推移確率行列と呼ぶ。また、確率分布 π が $\pi P = \pi$ を満たすとき、 π を P の定常分布という。

定義 2 (全変動距離). Ω 上の確率分布 μ, ν に対して, μ と ν の全変動距離を次で定義する.

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

命題 1 (マルコフ連鎖の収束定理). 推移確率行列 P を持つマルコフ連鎖 $\{X_t\}$ が既約かつ非周期であれば, 定常分布 π は一意で次を満たす.

$$\lim_{t \rightarrow \infty} d(t) = 0.$$

ただし, $d(t) = \max_{x \in \Omega} \|P^t(x, \cdot) - \pi\|_{TV}$ とおいた.

定義 3 (カップリング). X_t, Y_t を Ω 上の推移確率行列 P に従うマルコフ連鎖とする. 次を満たすとき (X_t, Y_t) をカップリングと呼ぶ.

$(X_t, Y_t)_{t \geq 0}$ は $\Omega \times \Omega$ 上のマルコフ連鎖で,

$$\begin{aligned} \forall x, y, x', y' \in \Omega, \forall t \geq 0, \\ P(X_{t+1} = x' | X_t = x, Y_t = y) &= P(x, x') \\ P(Y_{t+1} = y' | X_t = x, Y_t = y) &= P(y, y'). \end{aligned}$$

補題 1. (X_t, Y_t) をカップリングとして, 各 X_t, Y_t は Ω 上の推移確率行列 P に従い,

$X_t = Y_t \Rightarrow \forall s \geq t, X_s = Y_s$ を満たすとする.

$T = \inf\{t \geq 0 : X_t = Y_t\}$ とおくと, 次が成り立つ.

$$d(t) \leq \max_{x, y \in \Omega} P_{x, y}(T > t).$$

以上のことを前提としてマルコフ連鎖モンテカルロ法 (MCMC) について述べる.

MCMC とは標本空間上のある確率分布に従う標本を抽出するための手法であり, 以下のように構成される.

(1) 対象とする確率分布を定常分布にもつマルコフ連鎖を構成する.

(2) このマルコフ連鎖を走らせることで定常分布を近似する.

この操作によって確率的に近似された標本抽出が行える. 以下では MCMC の収束の速さやその精度に着目して 2 つのアルゴリズムの比較を行った結果について述べる.

3 アルゴリズムの定義

有限グラフを $G = (V, E)$ とし, E は自己ループや重複した辺を持たないものとする. 各辺の相互作用係数を $\{J_{x, y}\}_{\{x, y\} \in E}$, 各点における局所磁場を $\{h_x\}_{x \in V}$ で与え, スピン配置 $\sigma = \{\sigma_x\}_{x \in V} \in \{\pm 1\}^V$ に対してハミルトニアンを以下で定義する.

$$H(\sigma) = - \sum_{\{x, y\} \in E} J_{x, y} \sigma_x \sigma_y - \sum_{x \in V} h_x \sigma_x.$$

逆温度を $\beta \geq 0$ としてボルツマンウェイト w_β^G とギブス分布 π_β^G を次で定める.

$$w_\beta^G(\sigma) = e^{-\beta H(\sigma)}, \quad \pi_\beta^G(\sigma) = \frac{w_\beta^G(\sigma)}{\sum_{\sigma} w_\beta^G(\sigma)}.$$

$\sigma \in \{\pm 1\}^V$, $I \subset V$ に対して

$$(\sigma^I)_y = \begin{cases} \sigma_y & [y \notin I], \\ -\sigma_y & [y \in I], \end{cases} \quad \sigma^x = \sigma^{\{x\}}$$

と定める. グラウバー力学の推移確率を次のように定める.

$$P_\beta^G(\sigma, \tau) = \begin{cases} \frac{1}{|V|} \frac{w_\beta^G(\sigma^x)}{w_\beta^G(\sigma) + w_\beta^G(\sigma^x)} & [\tau = \sigma^x], \\ 1 - \sum_{x \in V} P_\beta^G(\sigma, \sigma^x) & [\tau = \sigma], \\ 0 & [\text{otherwise}]. \end{cases}$$

このとき P_β^G は π_β^G を定常分布に持つ. $\tilde{h}_x(\sigma)$ を

$$\tilde{h}_x(\sigma) = \sum_{y \in V} J_{x,y} \sigma_y + h_x,$$

とおくと,

$$P_\beta^G(\sigma, \sigma^x) = \frac{1}{|V|} \frac{e^{-\beta \tilde{h}_x(\sigma) \sigma_x}}{2 \cosh(\beta \tilde{h}_x(\sigma))}.$$

と書き換えられる. 次に SCA の定義をする. まず

$$\tilde{H}(\sigma, \tau) = -\frac{1}{2} \sum_{x,y \in V} J_{x,y} \sigma_x \tau_y - \frac{1}{2} \sum_{x \in V} h_x (\sigma_x + \tau_x)$$

として, グラウバーの時と同様に

$$w_{\beta,q}^{\text{SCA}}(\sigma) = \sum_{\tau} \exp \left(-\beta \tilde{H}(\sigma, \tau) + q \sum_{x \in V} \sigma_x \tau_x \right),$$

$$\pi_{\beta,q}^{\text{SCA}}(\sigma) = \frac{w_{\beta,q}^{\text{SCA}}(\sigma)}{\sum_{\sigma} w_{\beta,q}^{\text{SCA}}(\sigma)}$$

とおき, SCA の推移確率 $\pi_{\beta,q}^{\text{SCA}}$ を

$$P_{\beta,q}^{\text{SCA}}(\sigma, \tau) = \frac{1}{w_{\beta,q}^{\text{SCA}}(\sigma)} \exp \left(-\beta \tilde{H}(\sigma, \tau) + q \sum_{x \in V} \sigma_x \tau_x \right)$$

と定める. このとき $\pi_{\beta,q}^{\text{SCA}}$ は $P_{\beta,q}^{\text{SCA}}$ の定常分布となる. $\tilde{h}_x(\sigma)$ を用いれば

$$P_{\beta,q}^{\text{SCA}}(\sigma, \tau) = \prod_{x \in V} \frac{e^{(\frac{1}{2}\beta \tilde{h}_x(\sigma) + q\sigma_x)\tau_x}}{2 \cosh(\frac{1}{2}\beta \tilde{h}_x(\sigma) + q\sigma_x)}$$

と書き換えられる. このことから SCA の各スピンの更新はほかのスピンの更新と独立に起こっていることが分かる.

また $\|\pi_{\beta,q}^{\text{SCA}} - \pi_\beta^G\|_{\text{TV}}$ は $q \uparrow \infty$ で 0 に収束することが知られている (文献 [2]). したがって SCA はギブス分布の標本抽出を近似的に行うことができる.

4 更新スピン数

まずグラウバーと SCA の 1 ステップ当たりの更新スピン数を比較する. 2 つのスピン配置 $\sigma, \tau \in \{\pm 1\}^V$ に対して $D_{\sigma, \tau}$ を σ と τ が異なる値を取っている点の集合, すなわち

$$D_{\sigma, \tau} = \{x \in V : \sigma_x \neq \tau_x\}.$$

とおく. このとき $E^*[|D_{\sigma, X_{\sigma}}|] = \sum_{\tau} |D_{\sigma, \tau}| P^*(\sigma, \tau)$ とおくと (ただし, $P^*(\sigma, \cdot)$ はグラウバー ($P^* = P_{\beta}^G$) または SCA ($P^* = P_{\beta, q}^{\text{SCA}}$) を表し, X_{σ} は $\{\pm 1\}^V$ -値の $P^*(\sigma, \cdot)$ に従う確率変数とする.), 以下の定理が成り立つ.

定理 1.

$$\bar{K} = \max \{|\tilde{h}_x(\sigma)| : x \in V, \sigma \in \{\pm 1\}^V\}.$$

とおく.

$$\begin{aligned} 2q &\leq \log |V| - \beta \bar{K} \\ \Rightarrow \forall \sigma, \quad E_{\beta}^G[|D_{\sigma, X_{\sigma}}|] &\leq E_{\beta, q}^{\text{SCA}}[|D_{\sigma, X_{\sigma}}|] \end{aligned}$$

Proof. グラウバーについては高々 1 スピンしか更新しないので,

$$\begin{aligned} E_{\beta}^G[|D_{\sigma, X_{\sigma}}|] &= \sum_{x \in V} P_{\beta}^G(\sigma, \sigma^x) = \frac{1}{|V|} \sum_{x \in V} \frac{e^{-\beta \tilde{h}_x(\sigma) \sigma_x}}{2 \cosh(\beta \tilde{h}_x(\sigma))} \\ &= \frac{1}{|V|} \sum_{x \in V} (e^{2\beta \tilde{h}_x(\sigma) \sigma_x} + 1)^{-1} \end{aligned}$$

となる. また, $|D_{\sigma, \tau}| = \sum_{x \in V} 1_{\{\sigma_x \neq \tau_x\}}$ であるので, SCA については

$$\begin{aligned} E_{\beta, q}^{\text{SCA}}[|D_{\sigma, X_{\sigma}}|] &= \sum_{x \in V} \sum_{\tau: \tau_x \neq \sigma_x} P_{\beta, q}^{\text{SCA}}(\sigma, \tau) = \sum_{x \in V} \frac{e^{-\frac{1}{2}\beta \tilde{h}_x(\sigma) \sigma_x - q}}{2 \cosh(\frac{1}{2}\beta \tilde{h}_x(\sigma) \sigma_x + q)} \\ &= \sum_{x \in V} (e^{\beta \tilde{h}_x(\sigma) \sigma_x + 2q} + 1)^{-1} \end{aligned}$$

が得られる. したがって $E_{\beta}^G[|D_{\sigma, X_{\sigma}}|] \leq E_{\beta, q}^{\text{SCA}}[|D_{\sigma, X_{\sigma}}|]$ が成り立つためには,

$$|V| (e^{2\beta \tilde{h}_x(\sigma) \sigma_x} + 1) \geq e^{\beta \tilde{h}_x(\sigma) \sigma_x + 2q} + 1$$

であればよく, $|V| - 1 \geq 0$ と $\bar{K}$ の定義から定理が示される. □

5 混合時間

次に混合時間について比較をする.

推移確率行列 P に従い, 定常分布 π を持つマルコフ連鎖に対して, 混合時間を

$$t_{\text{mix}}(\varepsilon) = \inf \{t \geq 0 : d(t) \leq \varepsilon\}$$

と定める. 特にグラウバーと SCA に関する混合時間をそれぞれ $t_{mix}^G(\varepsilon)$, $t_{mix}^{SCA}(\varepsilon)$ とする. また

$$J = \max_{\{x,y\} \in E} |J_{x,y}|, \quad \Delta = \max_{x \in V} \deg(x)$$

とおく. このとき次の定理が成り立つ.

定理 2. $c_\beta = 1 - \Delta \tanh(\frac{1}{2}\beta J)(1 + \tanh(\frac{1}{2}\beta J))^{\Delta-1}$ として, $c_\beta < 1$ を満たすならば次の評価が成り立つ.

$$t_{mix}^{SCA}(\varepsilon) \leq \left\lceil \frac{\log |V| - \log \varepsilon}{c_\beta} \right\rceil.$$

一般にグラウバーの混合時間についてはすべての β で, $t_{mix}^G \geq (\frac{|V|}{2} - 1) \log(2\varepsilon)^{-1}$ となることが知られている (文献 [7]). これと定理 2 を合わせると, 高温状態では SCA のほうがグラウバーより収束が速いことが言える.

Proof. 証明には補題 1 を用いる. $\{\pm 1\}^V$ 上のカップリング (X_t, Y_t) を次のように定める.

$$p(\sigma, x) = \frac{e^{\frac{\beta}{2}\tilde{h}_x(\sigma) + q\sigma_x}}{2 \cosh(\frac{\beta}{2}\tilde{h}_x(\sigma) + q\sigma_x)} = \frac{1 + \tanh(\frac{\beta}{2}\tilde{h}_x(\sigma) + q\sigma_x)}{2}$$

と定め, $X_0 = \sigma, Y_0 = \tau$ とし, $[0, 1]$ 上の一様分布に従うそれぞれ独立な確率変数列 $\{U_x\}_{x \in V}$ に対して.

$$X_t(x) = \begin{cases} +1 & [U_x \leq p(X_{t-1}, x)] \\ -1 & [U_x > p(X_{t-1}, x)] \end{cases} \quad Y_t(x) = \begin{cases} +1 & [U_x \leq p(Y_{t-1}, x)] \\ -1 & [U_x > p(Y_{t-1}, x)] \end{cases}$$

と更新するものとする. 補題 1 から $\rho(\sigma, \tau) = \sum_{x \in V} 1_{\{\sigma_x \neq \tau_x\}}$ とおくと,

$$\begin{aligned} P_{\sigma, \tau}(T > t) &= P_{\sigma, \tau}(\rho(X_t, Y_t) \geq 1) \\ &\leq E_{\sigma, \tau}[\rho(X_t, Y_t)] \end{aligned}$$

であるから, $E_{\sigma, \tau}[\rho(X_t, Y_t)]$ を計算すればよい. まず $\rho(\sigma, \tau) = 1$ の時を考える. 特に $\sigma_v = +1, \tau_v = -1$ と置いても一般性を失わない. カップリングの定義からスピンの X_1 と Y_1 で異なる値に更新されるのは $N(v) = \{x \in V : \{x, v\} \in E\}$ 内の点に限るので,

$$\begin{aligned} E_{\sigma, \tau}[\rho(X_1, Y_1)] &= \sum_{j=0}^{\deg(v)} j P(\rho(X_1, Y_1) = j) \\ &= \sum_{j=0}^{\deg(v)} j \sum_{\substack{S \subseteq N(v) \\ |S|=j}} \prod_{w \in N(v) \setminus S} (1 - |p(\sigma, w) - p(\tau, w)|) \prod_{w \in S} |p(\sigma, w) - p(\tau, w)| \end{aligned}$$

ここで $w \in N(v)$ に対して $s = \frac{\beta}{2}(\tilde{h}_w(\sigma) - J_{v,w}) + q\sigma_w = \frac{\beta}{2}(\tilde{h}_w(\tau) - J_{v,w}) + q\tau_w$ とおけば,

$$|p(\sigma, w) - p(\tau, w)| = \frac{1}{2} \left| \tanh\left(s + \frac{\beta}{2}J\right) - \tanh\left(s - \frac{\beta}{2}J\right) \right|$$

であり, $a \in \mathbb{R}$ に対して $|\tanh(x+a) - \tanh(x-a)|$ は $x=0$ で最大値 $2|\tanh(a)|$ を取るから,

$$\begin{aligned} |p(\boldsymbol{\sigma}, w) - p(\boldsymbol{\tau}, w)| &\leq \left| \tanh\left(\frac{\beta}{2} J_{v,w}\right) \right| \\ &\leq \tanh\left(\frac{\beta}{2} J\right) \end{aligned}$$

となる. これと $1 - |p(\boldsymbol{\sigma}, z) - p(\boldsymbol{\tau}, z)| \leq 1$ を合わせると,

$$\begin{aligned} E_{\boldsymbol{\sigma}, \boldsymbol{\tau}}[\rho(X_1, Y_1)] &\leq \sum_{j=0}^{\deg(v)} j \sum_{\substack{S \subset N(v) \\ |S|=j}} \prod_{w \in S} \tanh\left(\frac{\beta}{2} J\right) \\ &= \sum_{j=0}^{\deg(v)} j \binom{\deg(v)}{j} \tanh^{\deg(v)}\left(\frac{\beta}{2} J\right) \\ &= \deg(v) \tanh\left(\frac{\beta}{2} J\right) \left(1 + \tanh\left(\frac{\beta}{2} J\right)\right)^{\deg(v)-1} \\ &\leq \Delta \tanh\left(\frac{\beta}{2} J\right) \left(1 + \tanh\left(\frac{\beta}{2} J\right)\right)^{\Delta-1} \\ &= 1 - c_\beta \leq e^{-c_\beta} \end{aligned}$$

が得られる. 次に $\rho(\boldsymbol{\sigma}, \boldsymbol{\tau}) = r$ とする. $\boldsymbol{\eta}^{(0)} = \boldsymbol{\sigma}, \boldsymbol{\eta}^{(1)}, \dots, \boldsymbol{\eta}^{(r)} = \boldsymbol{\tau}$ をすべての $i = 1, 2, \dots, r$ で $\rho(\boldsymbol{\eta}^{(i-1)}, \boldsymbol{\eta}^{(i)}) = 1$ となるようにとる. このとき三角不等式から

$$\begin{aligned} E_{\boldsymbol{\sigma}, \boldsymbol{\tau}}[\rho(X_1, Y_1)] &\leq \sum_{i=1}^r E_{\boldsymbol{\eta}^{(i-1)}, \boldsymbol{\eta}^{(i)}}[\rho(X_1, Y_1)] \\ &\leq r e^{-c_\beta} = \rho(\boldsymbol{\sigma}, \boldsymbol{\tau}) e^{-c_\beta} \end{aligned}$$

が得られる. さらに条件付き期待値の性質とマルコフ性から

$$E_{\boldsymbol{\sigma}, \boldsymbol{\tau}}[\rho(X_t, Y_t)] = E_{\boldsymbol{\sigma}, \boldsymbol{\tau}}[E_{X_{t-1}, Y_{t-1}}[\rho(X_t, Y_t)]] \leq E_{\boldsymbol{\sigma}, \boldsymbol{\tau}}[\rho(X_{t-1}, Y_{t-1}) e^{-c_\beta}] \leq |V| e^{-c_\beta t}$$

となる. したがって, $d(t) \leq \varepsilon$ の十分条件は $t \geq \frac{\log |V| - \log \varepsilon}{c_\beta}$ であり, 混合時間の定義から定理が示される. $\square$

6 ギブス分布との近さ

今までは分布同士の近さを全変動距離で測っていたが, 基底状態のみを求めるには次のものを考えれば十分である.

定義 4. $\varepsilon > 0$ として,

$$R_H = \max_{\boldsymbol{\sigma}, \boldsymbol{\tau} \in \{\pm 1\}^V} |H(\boldsymbol{\sigma}) - H(\boldsymbol{\tau})|.$$

とおき, 分布 μ が次を満たすとき, μ は *order-preservation* の意味でギブス分布に近いと定義する.

$$\mu(\boldsymbol{\sigma}) \geq \mu(\boldsymbol{\tau}) \quad \Rightarrow \quad H(\boldsymbol{\sigma}) \leq H(\boldsymbol{\tau}) + \varepsilon R_H.$$

この定義を用いると次の定理が成り立つ.

定理 3.

$$v = \sum_{\{x,y\} \in E} J_{x,y}^2 + \sum_{x \in V} h_x^2,$$

$\bar{K}$ を定理 1 で定義したものとする. このとき,

$$2q \geq \log |V| + \beta \bar{K} - \log \frac{\varepsilon \sqrt{v}}{2\bar{K}}.$$

を満たせば, $\pi_{\beta,q}^{\text{SCA}}$ は π_{β}^{G} に *order-preservation* の意味で近くなる.

Proof. まず $w_{\beta,q}^{\text{SCA}}(\sigma)/w_{\beta}^{\text{G}}(\sigma)$ は

$$\begin{aligned} \frac{w_{\beta,q}^{\text{SCA}}(\sigma)}{w_{\beta}^{\text{G}}(\sigma)} &= \sum_{\tau} \exp \left(\beta (\tilde{H}(\sigma, \sigma) - \tilde{H}(\sigma, \tau)) + q \sum_{x \in V} \sigma_x \tau_x \right) \\ &= e^{|V|q} \sum_{I \subset V} \sum_{\tau = \sigma^I} \exp \left(\frac{\beta}{2} \sum_{x \in V} \tilde{h}_x(\sigma) (\tau_x - \sigma_x) + q \sum_{x \in V} (\sigma_x \tau_x - 1) \right) \\ &= e^{|V|q} \sum_{I \subset V} \prod_{x \in I} e^{-\beta \tilde{h}_x(\sigma) \sigma_x - 2q} = e^{|V|q} \prod_{x \in V} (1 + \delta \phi_x(\sigma)) \end{aligned}$$

と書き換えられる. ただし,

$$\delta = e^{-2q}, \quad \phi_x(\sigma) = e^{-\beta \tilde{h}_x(\sigma) \sigma_x}$$

とおいた. したがって

$$\frac{\pi_{\beta,q}^{\text{SCA}}(\sigma)}{\pi_{\beta,q}^{\text{SCA}}(\tau)} = \frac{w_{\beta,q}^{\text{SCA}}(\sigma)}{w_{\beta,q}^{\text{SCA}}(\tau)} = \frac{w_{\beta}^{\text{G}}(\sigma)}{w_{\beta}^{\text{G}}(\tau)} \prod_{x \in V} \frac{1 + \delta \phi_x(\sigma)}{1 + \delta \phi_x(\tau)}$$

となる. ここで $\pi_{\beta,q}^{\text{SCA}}(\sigma) \geq \pi_{\beta,q}^{\text{SCA}}(\tau)$ とおくと,

$$\frac{w_{\beta}^{\text{G}}(\sigma)}{w_{\beta}^{\text{G}}(\tau)} \geq \prod_{x \in V} \frac{1 + \delta \phi_x(\tau)}{1 + \delta \phi_x(\sigma)}.$$

であるから, 定理を示すには

$$\prod_{x \in V} \frac{1 + \delta \phi_x(\tau)}{1 + \delta \phi_x(\sigma)} \geq e^{-\beta \varepsilon R_H},$$

を示せばよく, これは

$$\beta \varepsilon R_H \geq \sum_{x \in V} \log \underbrace{\frac{1 + \delta \phi_x(\sigma)}{1 + \delta \phi_x(\tau)}}_{>0} = \sum_{x \in V} \log \left(1 + \delta \underbrace{\frac{\phi_x(\sigma) - \phi_x(\tau)}{1 + \delta \phi_x(\tau)}}_{>-1} \right).$$

と書き換えられる. 今

$$R_H \geq \sqrt{v} \tag{1}$$

が成り立つから,

$$\mathbb{E}_\mu [(H - \mathbb{E}_\mu[H])^2] \leq R_H^2$$

がすべての $\{\pm 1\}^V$ 上の確率分布 μ について成り立つ. ただし $\mathbb{E}_\mu[f] = \sum_{\sigma} f(\sigma)\mu(\sigma)$ である. したがって, μ を一様分布とすれば

$$\begin{aligned}\mathbb{E}_\mu[H] &= \frac{1}{2^{|V|}} \sum_{\sigma \in \{\pm 1\}^V} H(\sigma) = 0, \\ \mathbb{E}_\mu[H^2] &= \frac{1}{2^{|V|}} \sum_{\sigma \in \{\pm 1\}^V} H(\sigma)^2 = v\end{aligned}$$

がわかる. また

$$\log \left(1 + \delta \frac{\phi_x(\sigma) - \phi_x(\tau)}{1 + \delta \phi_x(\tau)} \right) \leq \left| \delta \frac{\phi_x(\sigma) - \phi_x(\tau)}{1 + \delta \phi_x(\tau)} \right| \leq \delta |\phi_x(\sigma) - \phi_x(\tau)|$$

が成り立つことがテイラー展開から示せるので, 実数 X, Y に対して, 不等式 $|e^X - e^Y| \leq |X - Y|(e^X \vee e^Y)$ が成り立つことを使えば,

$$\beta \delta |\tilde{h}_x(\sigma)\sigma_x - \tilde{h}_x(\tau)\tau_x|(\phi_x(\sigma) \vee \phi_x(\tau)) \leq 2\beta \delta \bar{K} e^{\beta \bar{K}}$$

となり,

$$\sum_{x \in V} \log \left(1 + \delta \frac{\phi_x(\sigma) - \phi_x(\tau)}{1 + \delta \phi_x(\tau)} \right) \leq 2|V|\beta \delta \bar{K} e^{\beta \bar{K}} \quad (2)$$

がわかる. 式 (1), (2) を合わせれば定理が得られる. $\square$

参考文献

- [1] V. Černý. Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. *J. Optimiz. Theory App.*, **45** (1985): 41–51.
- [2] P. Dai Pra, B. Scoppola and E. Scoppola. Sampling from a Gibbs measure with pair interaction by means of PCA. *J. Stat. Phys.*, **149** (2012): 722–737.
- [3] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman and Company, San Francisco (1979).
- [4] C. Giraud. *Introduction to High-Dimensional Statistics*. Chapman & Hall/CRC, Boca Raton (2015).
- [5] R.J. Glauber. Time-dependent statistics of the Ising Model. *J. Math. Phys.*, **4** (1963): 294–307.
- [6] S. Kirkpatrick, C.D. Gelatt Jr., M.P. Vecchi. Optimization by simulated annealing. *Science*, New Series, **220** (1983): 671–680.
- [7] D.A. Levin, Y. Peres and E.L. Wilmer. *Markov Chains and Mixing Times*. AMS, Providence, Rhode Island (2009).

- [8] A. Lucas. Ising formulations of many NP problems. *Front. Phys.*, **12** (2014): <https://doi.org/10.3389/fphy.2014.00005>.
- [9] M. Normann, A. Sakai, H. Toyokawa and Y. Ueda. Stability of energy landscape. In preparation.
- [10] I.N. Sanov. On the probability of large deviations of random variables. *Mat. Sbornik*, **42** (1957): 11–44. English translation in *Sel. Transl. Math. Stat. Prob.*, **1** (1961): 213–244.
- [11] B. Scoppola and A. Troiani. Gaussian mean field lattice gas. *J. Stat. Phys.*, **170** (2018): 1161–1176.