

イジングパーセプトロンの有限レプリカ解析と Complexity

小淵 智之^{1,A}, 樺島 祥介^B

^A東京工業大学大学院理工学研究科, ^B東京工業大学大学院総合理工学研究科

1 はじめに

レプリカ法は、モーメント計算のための強力な手法であることは間違いなく、情報統計力学における主要な解析技法として広く用いられている [1]。このレプリカ法を使う立場としては大きく分けて二つあると考えられる。一つはレプリカ法を用いて、解かれていない新たな問題、系を解析することであり、もう一つは未だ残されているレプリカ法に関するいくつかの謎を考察するために、レプリカ法をそれら問題群に適用することである。この二つはモチベーションが異なるだけで、実際にやることが大きく異なるわけではないが、レプリカ法の処方箋から有用な情報を、出来る限り引き出そうとする際には、後者の立場で問題を眺める必要がある。本稿では、この後者の立場に立ち、よく知られた系であるイジングパーセプトロンを、レプリカ法によって詳細に調べ直した結果を報告する [2]。

具体的な結果に入る前に、本稿の主役となる三つの関数を導入しておく。

まずレプリカ法で計算される母関数 $\phi(n)$

$$\phi(n) = \frac{1}{N} \log[Z^n] \quad (n \in \mathbb{R}), \quad (1)$$

ここで $[\dots]$ は配位平均である。

二つ目はレート関数 $R(f)$ である。一般にランダム系では自由エネルギーの値 $f = -(1/N) \log Z$ が揺らぐためその確率分布 $P(f)$ を考えることには意味がある。大偏差理論によるとこの分布関数は次のように指数関数形でスケールされる

$$P(f) \sim \exp\{NR(f)\}, \quad (2)$$

この指数 $R(f) (\leq 0)$ がレート関数である。レート関数の値が小さいほどそのサンプルの出現確率は小さい。母関数 $\phi(n)$ とレート関数 $R(f)$ は次のような関係にある

$$e^{N\phi(n)} = [Z^n] = \int df P(f) e^{-Nnf} \sim \int df e^{N(-nf + R(f))} \Rightarrow \phi(n) = \max_f \{-nf + R(f)\}. \quad (3)$$

レート関数が上に凸な関数であるとき、式 (3) は完全なルジャンドル変換となり母関数 $\phi(n)$ とレート関数 $R(f)$ はレプリカ数 n を媒介変数として 1 対 1 に結びつく

$$f(n) = -\frac{\partial \phi(n)}{\partial n}, \quad R(f(n)) = \phi(n) - n \frac{\partial \phi(n)}{\partial n}. \quad (4)$$

これがレプリカ法とレート関数の関係である。最も典型的な自由エネルギー値 $f = f^*$ はレート関数が零となる点 $R(f^*) = 0 (P(f^*) = 1)$ 、即ち $n = 0$ で与えられる。

最後に Complexity $\Sigma(s)$ である。スピングラスの平均場理論によれば、スピングラスを含むある種のランダム系では熱力学極限で系の取りうる状態（いわゆる Pure state）が非常にたくさん（システムサイズ N の指数関数に比例）現れることが分かっている。それぞれの Pure state はそれぞれの自由エネルギー値

¹E-mail: obuchi@stat.phys.titech.ac.jp

を持つが、その自由エネルギー値が f であるような Pure state の数 $\mathcal{N}(f)$ が次のようにスケールされると仮定する

$$\mathcal{N}(f) \sim \exp \{N\Sigma(f)\}. \quad (5)$$

この指数 $\Sigma(f) (\geq 0)$ が Complexity である。たくさんの安定な状態が現れるというスピングラスの直観的な描像から考えれば、この Complexityこそスピングラスの性質を端的に表した量と言える。Complexity $\Sigma(f)$ と母関数 $\phi(n)$ の関係はやや複雑であるので、これは第2節にて別に述べる。

本稿の目的はこれら $R(f)$ と $\Sigma(f)$ 、及び $\phi(n)$ の関係をレプリカ法の実際の処方箋の中で議論することである。この目的のためにイジングパーセプトロンを取り扱うには主に二つの理由がある。一つはまず $R(f)$ と $\Sigma(f)$ の関係が非自明であるということである。SK モデルやランダムエネルギーモデル [3, 4] などを含むいくつかの場合において、 $R(f)$ と $\Sigma(f)$ がお互いに解析接続の関係にあることが示されているが [5, 6]、イジングパーセプトロンは、このような関係が成り立たない系と考えられている。もう一つの理由は、Pure state の微視的な描像の解釈がしやすいという点にある。実際、パーセプトロンを始めとしたいくつかの全結合モデルの Pure state は、絶対零度において、シングルスピントリップに対して安定なクラスターと同一視し得ることが示唆されている [7, 8, 9]。これによりレプリカ解析の結果を数値的に検証することが可能となるため、ここでの目的に都合がよい。

本稿は以下のように構成される。次の節ではモデルの定義と具体的な定式化について説明する。RSB と Complexity との関係に関する説明もこの節で与えられる。続く第3節で解析的な結果と数値計算による結果の両方を与え比較する。双方は良い一致を示すが、一方で正しい Complexity 及びレート関数の関数形は、母関数を経由した方法（解析、数値計算の両方で用いられる）では求められないことも示唆される。これは通常仮定される、Complexity とレート関数の凸性がイジングパーセプトロンにおいて成り立たないためと考えられる。最後の節はまとめて充てられる。

2 モデルと定式化

2.1 モデル

単純パーセプトロンとは単一のニューロンの最もシンプルなモデルであり、次のような実数を符号に移す写像機能を持つ

$$y = \begin{cases} +1, & S \cdot \mathbf{x} / \sqrt{N} > 0, \\ -1, & S \cdot \mathbf{x} / \sqrt{N} < 0, \end{cases} \quad (6)$$

ここで $\mathbf{x} \in \mathbb{R}^N$ は問題とするパーセプトロンへの入力であり、 $y \in \{+1, -1\}$ はその出力、 S はシナプス結合である。本稿で扱うイジングパーセプトロンは、このシナプス結合のそれぞれが ± 1 を取る、特別な場合の単純パーセプトロンである。パーセプトロンを用いた情報処理は、大抵の場合、シナプス結合 S を調整することによっていくつかの入出力のパターンを再現（記憶）させることが目的となる。以下では M 個の入出力パターンを次の記号

$$D^M = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_M, y_M)\}, \quad (7)$$

で表現することにしよう。

この M 個の入出力パターンの取りうる組み合わせは様々なものが考えられるが、ここではそれぞれの入出力が独立同分布から得られる完全にランダムなパターンを記憶させることを考える。入力、出力の確率分布はそれぞれ次の式で与えられる

$$P(\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi}} \right)^N \exp \left(-\frac{\mathbf{x}^2}{2} \right), \quad (8)$$

$$P(y) = \frac{1}{2} (\delta(y-1) + \delta(y+1)). \quad (9)$$

パーセプトロンに記憶させる問題を増やしていったとき、取りうるシナプス結合の数は減っていき、最後には全くなってしまおうと考えられる。特に、埋め込むパターン数 M とシナプス結合数 N の比 $\alpha = M/N$ を有限に保って熱力学極限 $N \rightarrow \infty$ をとることで、記憶し得る問題数の典型的限界 α_s (以下 Capacity と呼ぶ) を求めることができる [10, 11]。初期のパーセプトロンの取り扱いでは、この Capacity を求めることに力点が置かれ、それがレプリカ法によって計算されていた。しかしレプリカ法から得られる情報は本来それだけではない。本稿ではこの問題を再度取り扱い、レプリカ法の最近の解釈を用いて、シナプス結合空間のマクロな特徴づけ (Complexity) 及び記憶可能なパターンが現れる確率 (レート関数) が自然に得られることを示す。

2.2 レプリカ計算

まず分配関数を定義しよう。入出力パターン D^M が与えられたとき、それを正しく再現できるシナプス結合配位の数 $Z(D^M)$ を分配関数と定義すると²、次のようになる

$$Z(D^M) = \sum_S \prod_{\mu=1}^M \Theta \left(-y_\mu \frac{\mathbf{S} \cdot \mathbf{x}_\mu}{\sqrt{N}} \right), \quad (10)$$

ここで $\Theta(u)$ はステップ関数であり $u > 0$ に対して $\Theta(u) = 1$ 、さもなければ $\Theta(u) = 0$ を取る。この分配関数に対して、母関数 $\phi(n) = (1/N) \log [Z^n(D^M)]_{D^M}$ をレプリカ法の枠組みで計算する。ここで、記号 $[\dots]_{D^M}$ は入出力パターンに関する平均である。結果だけ示すと、自然数の n に対して

$$\phi(n) = \text{Extr}_{q^{ab}, \hat{q}^{ab}} \left\{ - \sum_{a < b} \hat{q}^{ab} q^{ab} + \log \left(\sum_{S^1, S^2, \dots, S^n} e^{\sum_{a < b} \hat{q}^{ab} S^a S^b} \right) + \alpha \log \left[\prod_{a=1}^n \Theta(u^a) \right]_{\mathbf{u}} \right\}, \quad (11)$$

となる。ここで $[\dots]_{\mathbf{u}}$ は n 次元ガウス変数 u^1, u^2, \dots, u^n に対する平均であり、平均と分散は $[u^a]_{\mathbf{u}} = 0$ 、 $[u^a u^b]_{\mathbf{u}} = \delta_{ab} + (1 - \delta_{ab}) q^{ab}$ ($a, b = 1, 2, \dots, n$) のように与えられる。式 (11) はこのままでは解けないため、レプリカ対称性 (RS) やレプリカ対称性の破れ (RSB) などの仮定が必要となる。以下では RS と 1 ステップ RSB (1RSB) の範囲で母関数 $\phi(n)$ を調べる。

まず、RS 解は $q^{ab} = q, \hat{q}^{ab} = \hat{q}$ と置くことによって与えられる。この仮定の下で q, \hat{q} に関する極値条件を解くと次の二つの解が得られる:

RS1: $0 < q < 1, \hat{q} < +\infty$

$$\begin{aligned} \phi_{\text{RS1}}(n) = & -\frac{n(n-1)}{2} q \hat{q} - \frac{1}{2} n \hat{q} + \log \left(\int Dz \left(2 \cosh \left(\sqrt{\hat{q}} z \right) \right)^n \right) \\ & + \alpha \log \left(\int Dz E^n \left(\sqrt{\frac{q}{1-q}} z \right) \right), \end{aligned} \quad (12)$$

ここで $Dz = \exp(-z^2/2) / \sqrt{2\pi}$ 、 $E(u) = \int_u^{+\infty} Dz$ である。

RS2: $q = 1, \hat{q} = +\infty$.

$$\phi_{\text{RS2}}(n) = (1 - \alpha) \log 2. \quad (13)$$

次に 1RSB 解を調べる。よく知られているように 1RSB 解は n 個のレプリカをサイズ m の n/m 個のグループに分けることで得られ、このとき秩序変数はレプリカインデックスが両方とも同じグループ内にあるとき $(q^{ab}, \hat{q}^{ab}) = (q_1, \hat{q}_1)$ 、グループ間にまたがるとき $(q^{ab}, \hat{q}^{ab}) = (q_0, \hat{q}_0)$ を取ると仮定する。この条件

²これはハミルトニアン $H(S|D^M)$ を、結合配位 S において再現できていないパターン数 $H(S|D^M) = \sum_{\mu=1}^M \Theta \left(-y_\mu \frac{\mathbf{S} \cdot \mathbf{x}_\mu}{\sqrt{N}} \right)$ として定義した場合の分配関数 $Z(D^M) = \sum_S e^{-\beta H(S|D^M)}$ の零温度極限 $\beta \rightarrow \infty$ に等しい。

の下で極値条件と取ると次の3つの解を見つけることができる:

1RSB1: $(q_1, q_0) = (1, q)$, $(\hat{q}_1, \hat{q}_0) = (+\infty, \hat{q})$. ここで q 及び \hat{q} は RS1 解のそれと同じ値を取る。

$$\phi_{1RSB1}(n, m) = \phi_{RS1}\left(\frac{n}{m}\right). \quad (14)$$

1RSB2: $(q_1, q_0) = (q, q)$, $(\hat{q}_1, \hat{q}_0) = (\hat{q}, \hat{q})$. ここでも q 及び \hat{q} は RS1 のそれと同じ。

$$\phi_{1RSB2}(n, m) = \phi_{RS1}(n). \quad (15)$$

1RSB3: $(q_1, q_0) = (1, 1)$ 及び $(\hat{q}_1, \hat{q}_0) = (+\infty, +\infty)$.

$$\phi_{1RSB3}(n, m) = \phi_{RS2}(n) = (1 - \alpha) \log 2. \quad (16)$$

ここで二つの重要な指摘をしておく。一つ目は、通常 Parisi の 1RSB パラメータである m は極値条件が取られるが、今回の解析ではこれを取らないこと、二つ目は、通常 RS 解と同一視され 1RSB 解として取り扱われない 1RSB2 及び 1RSB3 解を、今回の解析では 1RSB 解としてきちんと取り扱うことである。次節以降で解るように、これらは $\phi(n)$ 、 $R(f)$ 及び $\Sigma(f)$ の関係を考察するうえで重要な意味を持つてくる。

2.3 1RSB 解と Complexity、レート関数の関係

ここでは、Monasson によって議論された 1RSB 解と Complexity の関係性についての議論を、有限レプリカ数に拡張した形で簡単に紹介する [12, 13]。

さて、入出力パターン D^M の一つのサンプルが与えられたとき、Pure state がどのようになるかについて考えよう。まず、Pure state のインジケータ $\delta_\gamma(S)$ を、結合配位 S が Pure state γ に属するとき 1、さもなければ 0 を返すものとして定義しよう。このとき Pure state γ の分配関数を次のように定義できる

$$Z_\gamma = \sum_S \prod_{\mu=1}^M \Theta\left(-y_\mu \frac{S \cdot x_\mu}{\sqrt{N}}\right) \delta_\gamma(S). \quad (17)$$

この分配関数 Z_γ が $Z_\gamma \sim \exp(Ns_\gamma)$ のようにスケールされるとする。今の場合、分配関数がパターン D^M を完全に分離できるような結合配位の数であったことを考えると、 $s_\gamma \sim O(1)$ は物理的にはエントロピーに対応することが分かる。さて、エントロピーの値が s であるような Pure state の数 $\mathcal{N}(s)$ が指数関数形でスケールされると仮定する $\mathcal{N}(s) \sim \exp(N\Sigma(s))$ 。これによりエントロピーに対する Complexity $\Sigma(s) (\geq 0)$ が自然に定義される。ここで Complexity に対する母関数 $g(x|D^M) \equiv (1/N) \log[\sum_\gamma Z_\gamma^x]_{D^M}$ を導入する。これを陽に書き下し鞍点法を用いると

$$g(x|D^M) = \frac{1}{N} \log\left(\sum_\gamma Z_\gamma^x\right) \sim \frac{1}{N} \log\left(\int ds e^{N(xs + \Sigma(s))}\right) \Rightarrow \max_s \{xs + \Sigma(s)\}, \quad (18)$$

という関係が得られる。Complexity が上に凸な関数であれば、 $g(x|D^M)$ と $\Sigma(s)$ は x を媒介変数として次のような一対一の関係を持つ

$$s(x) = \frac{\partial g(x|D^M)}{\partial x}, \quad \Sigma(s(x)) = g(x|D^M) - x \frac{\partial g(x|D^M)}{\partial x}. \quad (19)$$

ところで $g(x|D^M)$ は自己平均量だと思われる。したがって、熱力学極限では $g(x|D^M)$ はその平均値 $g(x) = [g(x|D^M)]_{D^M}$ に収束するだろう。この $g(x)$ を計算するためにレプリカ法を用いると

$$g(x) = \frac{1}{N} \left[\log\left(\sum_\gamma Z_\gamma^x\right) \right]_{D^M} = \lim_{y \rightarrow 0} \frac{\partial}{\partial y} \left(\frac{1}{N} \log \left[\left(\sum_\gamma Z_\gamma^x \right)^y \right]_{D^M} \right). \quad (20)$$

実はこの右辺のファクター $\left[\left(\sum_{\gamma} Z_{\gamma}^x \right)^y \right]_{D^M}$ が 1RSB 解と以下のように関係しているのである

$$\left[\left(\sum_{\gamma} Z_{\gamma}^x \right)^y \right]_{D^M} = e^{N\phi_{1RSB}(xy,x)}. \quad (21)$$

この式の導出はここでは省略する [2]。式 (21) を式 (20) に代入すると、 $g(x) = x(\partial/\partial n)\phi_{1RSB}(n,x)|_{n=0}$ という関係式が得られる。ここから Complexity $\Sigma(s)$ と $\phi_{1RSB}(n,m)$ を直接結びつけることができる:

$$s(x) = \frac{\partial}{\partial x} \left(x \frac{\partial \phi_{1RSB}(n,x)}{\partial n} \Big|_{n=0} \right), \quad \Sigma(s(x)) = -x^2 \frac{\partial^2 \phi_{1RSB}(n,x)}{\partial x \partial n} \Big|_{n=0}. \quad (22)$$

また一方で、 $x=1$ と置いたときには $g(x)$ の定義から $g(x)$ は分配関数と一致しなくてはいけないことが分かる。すなわち $\phi(n) = \phi_{1RSB}(n,x)|_{x=1}$ ³ が成り立つ。ゆえに式 (4) から、レート関数 $R(s)$ と $\phi_{1RSB}(n,m)$ について次の関係が成立する:

$$s_{\text{tot}}(n) = \frac{\partial \phi_{1RSB}(n,x)}{\partial n} \Big|_{x=1}, \quad R(s_{\text{tot}}(n)) = -n^2 \frac{\partial}{\partial n} \left(\frac{\phi_{1RSB}(n,x)}{n} \Big|_{x=1} \right), \quad (23)$$

ここで s_{tot} は、通常のエントロピー $(1/N) \log Z$ でありパターン D^M を完全に再現できるシナプス結合の総数と結びついているが、Pure state のエントロピーと区別するためこう書いた。以下では s_{tot} を総エントロピーと呼ぶ。Complexity や Pure state のエントロピーとは関係 $s_{\text{tot}} = \max_s \{s + \Sigma(s)\}$ で結びついている。式 (22) においてパラメータ x は不等式 $s(x) \geq 0$ and $\Sigma(s(x)) \geq 0$ が両方とも満たされる領域のみでしか値を取れないことに注意する。これはエントロピー、Complexity が共に状態数の \log を取るためである。同様に式 (23) において、不等式 $s_{\text{tot}}(n) \geq 0$ と $R(s_{\text{tot}}(n)) \leq 0$ が n の取りうる値を制限する。また、実際に $\Sigma(s)$ や $R(s)$ を計算する際には、適切な解をいくつかの 1RSB 解 (今の場合 1RSB1, 1RSB2, 1RSB3) から選び出さなければいけないことに注意する。

3 結果

3.1 解析解

さて、以上の結果を用いて、イジングパーセプトロンの Complexity とレート関数を計算する。

まずは典型領域 ($n=0$) での Complexity の振る舞いから調べよう。レプリカ法とは自然数の n から実数の n に対する $\phi(n)$ の振る舞いを知る方法であるから、典型的な系の振る舞い ($n \rightarrow 0$) が知りたいとき、最も重要なのは $n=1$ 近傍で主要な解である。比 α が $\alpha_s \approx 0.833$ より小さいとき、そのような解は $\phi_{RS1}(n)$ で与えられる。このとき 1RSB 解では、 $m=1$ と置くことによって、 $\phi_{RS1}(n)$ と一致する解が 1RSB1 と 1RSB2 とで二つ得られる。これらは $\phi(n)$ の計算の上では区別がつかないが、Complexity の母関数 $g(x)$ を考えると異なる寄与を与える

$$g_{1RSB1}(x) = x \frac{\partial \phi_{1RSB1}(n,x)}{\partial n} \Big|_{n=0} = \phi'_{RS1}(0), \quad g_{1RSB2}(x) = x \frac{\partial \phi_{1RSB2}(n,x)}{\partial n} \Big|_{n=0} = x \phi'_{RS1}(0). \quad (24)$$

この二つの解から推定される正しい $g(x)$ は

$$g(x) = \begin{cases} \phi'_{RS1}(0), & x \leq 1, \\ x \phi'_{RS1}(0), & x > 1, \end{cases} \quad (25)$$

であろう。これは次のように直感的に理解できる。パラメータ x が非常に大きいとき、和 $\sum_{\gamma} Z_{\gamma}^x$ において最も大きいエントロピーを持つ単一の Pure state が支配的になると考えられる。これは $g(x) = \max_s \{xs + \Sigma(s)\}$

³左辺の $\phi(n)$ はレプリカ法によって何らかの仮設の下計算された解候補を表しているのではなく、定義通りの $\phi(n) = (1/N) \log[Z^n]_{D^M}$ を表していることに注意。

が x の線形関数になることを意味する。また x の小さい領域では、逆にエントロピーの大きさは意味を成さなくなり Complexity の大きい、即ちエントロピーが小さくとも数の多い Pure state が利くようになると考えられる。これは $g(x)$ の x 依存性が x の小さいところで緩やかになることを示唆している。さらに $g(x)$ は定義から、 x に関する下に凸な関数でなくてはならない。式 (25) は以上の条件を自然に満たしている。

式 (25) に対応する Complexity は

$$\Sigma(s) = \begin{cases} \phi'_{RS1}(0) - s, & 0 \leq s \leq \phi'_{RS1}(0), \\ -\infty, & \text{otherwise.} \end{cases} \quad (26)$$

となる。この区分的に線形な Complexity の振る舞いは通常見られない異常な振る舞いである。この性質は Complexity が上に凸な関数で無いことを示唆している。Complexity $\Sigma(s)$ が上に凸な場合、 $\Sigma(s)$ の情報は $g(x)$ に全て書きこまれ $g(x)$ と $\Sigma(s)$ が一対一に対応するが、そうでない場合は、本来の Complexity における上に凸で無い部分の情報は全て失われ、その凸包部のみが $g(x)$ からの変換で得られるのである。現在のモデルでは Complexity が下に凸な関数になっておりその結果、凸包を成す二点 $(0, \phi'_{RS1}(0))$ 及び $(\phi'_{RS1}(0), 0)$ を直線で結んだ解が得られていると考えられる (図 1)。またこの凸包の座標と直線の傾き $-x = -1$ から、エントロピー 0 の小さな多数のクラスターと非常に大きいエントロピーを与える単一のクラスターとが等しい寄与をしていることがわかる。非常に大きいクラスターが存在することは、局所探索アルゴリズムによって与えられた入出力パターン D^M を完全に再現する結合配位を探し出しやすいことを意味する。これは先行研究の結果を支持している [14]。

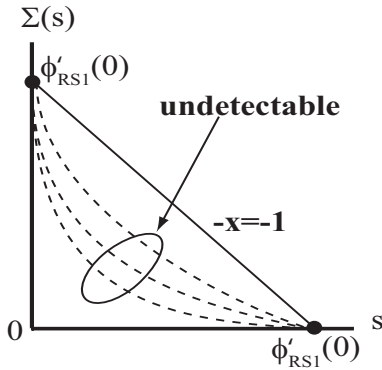


図 1: Complexity(26) の概形。本来の Complexity の概形が点線のように下に凸であったとしても、鞍点法を用いる現在の定式化では全て実線で書かれている、凸包を結ぶ直線として検出されてしまう。

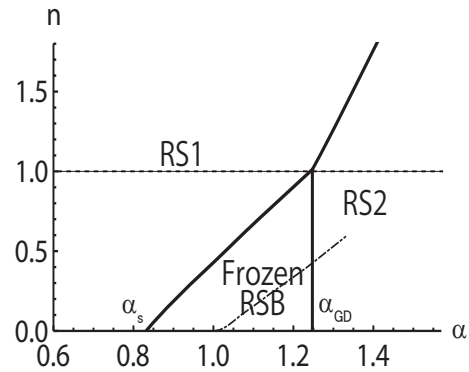


図 2: イジングパーセプトロンの n - α 平面上の相図。実線が相境界を表す。点線は $n = 1$ の線を表し、一点鎖線は AT 線を示すが、この場合相境界の決定には有意でない。この AT 線はある α を境に消失しているが、RS 1 解を与える $0 < q < 1$ の範囲の q の解が、そこを境に無くなるためである。

さて、次にレート関数の議論に移ろう。領域 $\alpha > \alpha_s = 0.833\dots$ では、Capacity を超えているため典型的にはもはや記憶可能なサンプル D^M は現れない。そのため Complexity の典型解析 ($n = 0$) ではこの領域を調べることは出来ない。しかし $n > 0$ として非典型的な領域を調べることで、Capacity を超えて系の情報を得ることが出来る。そこでこの領域のレート関数を調べてみる。

領域 $\alpha_s < \alpha < \alpha_{GD} = 1.245\dots$ においては、 $\alpha < \alpha_s$ の領域と同じように $\phi_{RS1}(n)$ が $n \approx 1$ で有意である。これは解として $\phi_{1RSB1}(n, m = 1) = \phi_{1RSB2}(n, m = 1) = \phi_{RS1}(n)$ を $R(s)$ の評価に用いるべきであること意味する。ところが、この解を単純に $n < 1$ の領域へ接続していくと、ある $0 < n_s(\alpha) < 1$ において $\phi_{RS1}(n)$ が最小値を取ることが分かる。これより小さな n へさらに接続しようとする、総エントロピー $s_{tot}(n) = (\partial/\partial n)\phi_{RS}(n)$ が負になってしまう。これに対する妥当な解として、 $n \leq n_s$ で $\phi(n) = \phi_{RS1}(n_s(\alpha))$ ($n \leq n_s$) のように $\phi(n)$ の値を定数に保ってしまうというのが考えられる。これをここでは

Frozen RSB 解と呼ぶ。この解を採用すると対応するレート関数は

$$R(s_{\text{tot}} = 0) = \phi_{\text{RS1}}(n_s(\alpha)) = \min_n \{\phi_{\text{RS1}}(n)\}, \quad (27)$$

となる。これは、ある入出力パターン D^M が与えられたとき、それが記憶可能である確率を表している。

さらに α を大きくすると、ある値 $\alpha_{\text{GD}} = 1.245\dots$ を超えたところで、 $\phi(n)$ の $n \approx 1$ での主要な解が $\phi_{\text{RS1}}(n)$ から $\phi_{\text{RS2}}(n)$ へと移る。その結果レート関数は

$$R(0) = \phi_{\text{RS2}}(n) = (1 - \alpha) \log 2. \quad (28)$$

となる。以上の $\phi(n)$ の振る舞いを図 3 にまとめておく。領域 $\alpha > \alpha_{\text{GD}}$ においては、ある $n = n_{\text{GD}}$ におい

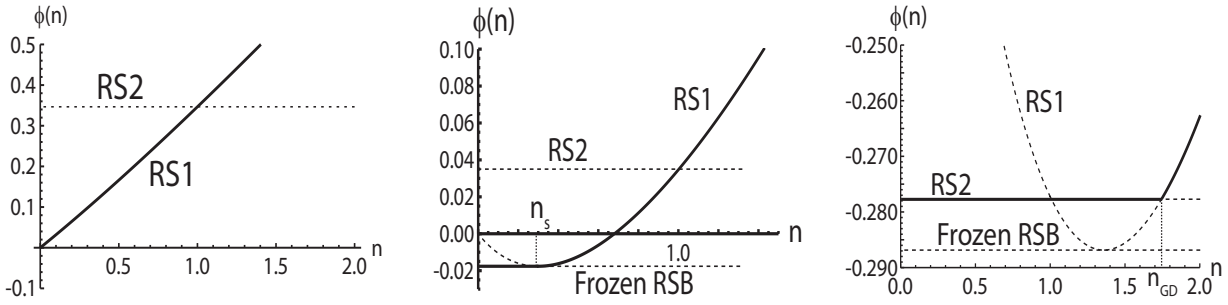


図 3: 母関数 $\phi(n)$ の各領域における振る舞い。実線が正しい解を表し、点線は RS 及び frozen RSB の解候補たちを表す。それぞれ α の値は左から順に 0.5, 0.95, 1.4 である。

てレート関数の飛びを伴う転移が起こる。これは、Complexity の場合と同じようにレート関数においても、あるエントロピーの幅 $0 < s < (\partial/\partial n)\phi_{\text{RS1}}(n_{\text{GD}}(\alpha))$ の部分で凸性が破れていること示唆している。

以上の議論から n - α 平面上の相図を書くことができる。結果は図 2 である。RS1 解の Almeida-Thouless (AT) 条件 [15] も評価してあるが、AT 不安定性が起こる以前に他の物理的理由から異なる解に転移しているため、相境界の決定にはなにも寄与していない。この場合 AT 条件が意味があるとすれば、RS2 解は $n > 0$ の領域で AT 安定であるが、 $n = 0$ 直上でのみ AT 条件を満たさない [10]。この不安定性については立ち入った議論が必要な可能性があるが、 $n \rightarrow 0$ での極限の振る舞いのみが重要である、というレプリカ法の精神に従えば有意でない可能性が高く、本稿では無視することにする。

3.2 数値解

第 1 節でも述べたが、イジングパーセプトロンの場合 Pure state がシングルスピントリップに対して安定なクラスターと同一視できるということが、これまで示唆されている [7, 8, 9]。ここではそれを認め、数値的に母関数 $g(x)$ や Complexity $\Sigma(s)$ 、レート関数 $R(s)$ を計算することを試みる。絶対零度を取っているので零エネルギークラスター、すなわち入出力パターン D^M を完全に再現するシナプス結合配位が成すクラスターにのみ以下では着目する。

クラスターのサイズ (含まれているシナプス結合配位数) を Q 、サンプル D^M におけるサイズ Q のクラスターの数 $C(Q|D^M)$ と書こう。エントロピー s は $s = (1/N) \log Q$ 、Complexity $\Sigma(s|D^M)$ は $(1/N) \log C(Q|D^M)$ にそれぞれ対応すると考えられる。クラスターのサイズや数は数値的に計算できるのでここから母関数 $\phi_{\text{1RSB}}(n = xy, m = x)$ を構成することが出来る

$$\phi_{\text{1RSBnum}}(n = xy, m = x) = \frac{1}{N} \log \left[\left(\sum_{\gamma} Q_{\gamma}^x \right)^y \right], \quad (29)$$

ここで [...] はサンプル平均を表す。典型極限 $y \rightarrow 0$ をとると母関数 $g(x)$ が得られる

$$g_{\text{num}}(x) = \lim_{y \rightarrow 0} \frac{\partial}{\partial y} \phi_{\text{1RSBnum}}(xy, x) = \frac{1}{N} \frac{\left[\Theta \left(\sum_{\gamma} Q_{\gamma}^x \right) \log \left(\sum_{\gamma} Q_{\gamma}^x \right) \right]}{\left[\Theta \left(\sum_{\gamma} Q_{\gamma}^x \right) \right]}, \quad (30)$$

このステップ関数 $\Theta(x)$ は $\log \left[\left(\sum_{\gamma} Q_{\gamma}^x \right)^y \right]$ の y に関する微分から現れ、クラスターが全く現れないようなサンプルが出たとき $\Theta \left(\sum_{\gamma} Q_{\gamma}^x \right) \log \left(\sum_{\gamma} Q_{\gamma}^x \right)$ の寄与が 0 となるように働く。

以上の量は数値的に全て評価できる量であるので、ここからレプリカ解析の結果を検証することが出来る。実際に $g_{\text{num}}(x)$ を評価し解析解 (25) と比較した物を図 4 に与えた。実際の数値計算はサンプル D^M を分布 (9) に従って生成し、シナプス配位を全件探索してクラスターに分類し、そこから $g_{\text{num}}(x)$ をいくつかの x (実際は $x = 0$ から 2.0 まで 41 点取った) で評価し、それらを十分な回数繰り返すという力づくの計算である。システムサイズ N が大きくなるにつれ $x \leq 1$ の領域のデータは平坦になっていき、解析解

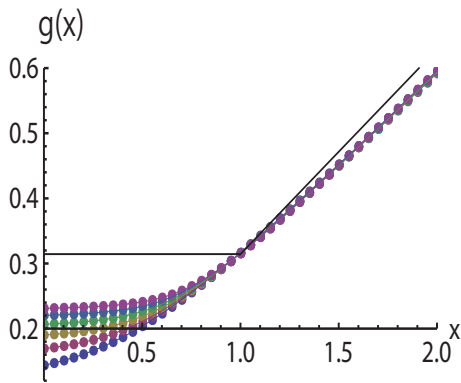


図 4: 数値的に構成した母関数 $g_{\text{num}}(x)$ の $\alpha = 0.5$ における振る舞い。下から順にシステムサイズは $N = 14, 16 \dots 24$ である。実線は解析解 (25) である。サンプル数はそれぞれの N に対して 32,000 取った。統計誤差はマーカーのサイズより小さい。システムサイズが大きくなるにつれ $g_{\text{num}}(x)$ が解析解に近づいている様子が見える。

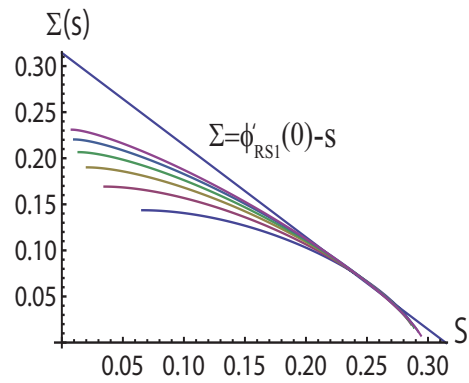


図 5: 母関数 $g_{\text{num}}(x)$ から式 (22) によって構成した Complexity の、 $\alpha = 0.5$ における振る舞い。システムサイズ N は下から上に大きくなる。実線は解析解であり、 N が大きくなるにつれ、これに近づいている様子が見える。

$g(x) = \phi'_{\text{RS1}}(0)$ に近づいていることが見て取れる。一方で $x > 1$ の領域では、解析解より傾きが若干緩やかに見えるが、システムサイズが増えるにつれ上昇する傾向はデータから見えている。以上の結果から、数値計算の結果は解析解をよく再現していると言える。母関数が構成出来たので Complexity もまた評価することが出来る。結果は図 5 に与えた。この結果もまた、サイズが増えるにつれレプリカ解析の結果に近づいている。実際の Complexity の評価は、 $g_{\text{num}}(x)$ の得られたデータを多項式補完しその多項式を用いて、式 (22) から Complexity $\Sigma(s)$ を構成した。多項式の次数は赤池情報量基準 [16] を用いて決定し、最終的に 27 次多項式を用いたが、補完式の振る舞いは次数にあまり影響されないことを確認している。

以上では、レプリカ解析と数値解析の結果が整合することを確認した。しかし前述したように、この母関数の方法で構成した Complexity は本来の Complexity の分布とは異なっている可能性がある。これを検証するために、直接クラスターの分布を見ることを行った。まず、クラスターと Pure state の対応を直観的に利用し、以下のような Raw Complexity $\Sigma_r(s|D^M)$ を定義する

$$\Sigma_r(s = (1/N) \log Q|D^M) = \frac{1}{N} \Theta \left(C(Q|D^M) \right) \log \left(C(Q|D^M) \right). \quad (31)$$

サンプル平均を取ることで、これの典型的な概形 $\Sigma_r(s) = [\Sigma_r(s|D^M)]$ を得ることができる。その結果を $\alpha = 0.5$ の場合について図 6 に与えた。本来の Raw Complexity は離散的な関数だが、滑らかな概形を得る

ために各点を結んでいる。エントロピーが0のクラスターに対応する Raw Complexity $\Sigma_r(0)$ はサイズが大

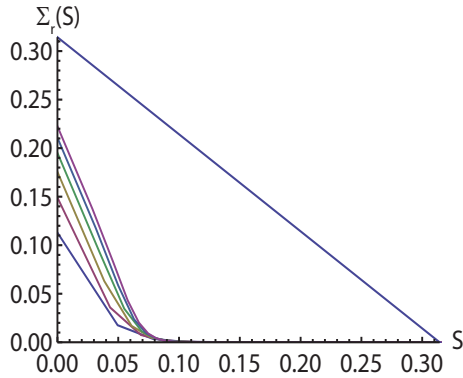


図 6: Raw complexity $\Sigma_r(s)$ の $\alpha = 0.5$ におけるプロット。下から上に行くにつれシステムサイズは大きくなり、それぞれ $N = 14, 16, \dots, 24$ に対応する。実線は解析解 $\Sigma(s)$ の解析解。サンプル数はそれぞれのサイズに対して 32,000 ずつ取っている。

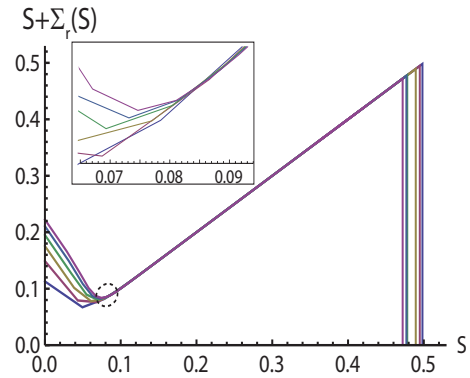


図 7: エントロピーと Raw Complexity の和 $s + \Sigma_r(s)$ のプロット。システムサイズが増えるにつれ V の字型の関数形に収束していくように見える。これは $\Sigma_r(s)$ が下に凸であり、 $g(x=1)$ への寄与が V の字の頂点、Complexity の凸包から来ていることを示唆する。挿入図は点線で囲んだ部分の拡大図。

きくなるにつれ $\phi'_{RS1}(0)|_{\alpha=0.5} = 0.314\dots$ に下から近づいているように見える。しかし、 $x \geq 0.1$ における $\Sigma_r(x)$ はサイズが多くなってほぼ恒等的に 0 に張り付いている。このことは $\Sigma_r(x)$ が下に凸な関数に収束することを示唆している。このことをもうすこし見やすくするため、Raw Complexity を用いて $s + \Sigma_r(s)$ を計算しそれを図 7 に与えた。この図は二つのピークと一つの谷を持つ V の字型の関数形となっている。このことは $\Sigma_r(s)$ が下に凸な関数であり、 $g(x=1)$ への主要な寄与が二つのピークから来ていることを示唆している。右側のピークはサイズが増えるにつれ左に寄ってきている。収束する点に関してはサイズが小さいため確たることは言えないが、解析解 $s = 0.314\dots$ の点に近づいているといっても矛盾はない。一方で谷の位置は、図 7 の挿入図で解るように、 $x = 0.084$ 近傍の点で頭打ちになるように見える。以上から、イジングパーセプトロンの Complexity は下に凸な関数形になっており、母関数から構成される $\Sigma(s)$ と Raw Complexity $\Sigma_r(s)$ 一致しないということが示された。

最後に、 $\alpha > \alpha_s$ の領域においてレート関数を評価する。この領域では典型的には、イジングパーセプトロンによって記憶可能な入出力パターンは現れず、また現れても小さなクラスター、即ち $s = 0$ のクラスターが支配的だと考えられる。そこで数値的に記憶可能なパターンが現れる確率 P_{sep} を評価することで、零エントロピーのレート関数 $R(0)$ を $R(0) = (1/N) \log P_{sep}$ のように評価することが出来る。結果を図 8 及び 9 に与えた。それぞれ $\alpha = 1.0, 1.5$ に対応する。図 8 及び 9 の実線は数値データの線形フィットである。これらの結果はレプリカ解析の結果とよく合っており、レプリカ解析が正しいことを示唆している。統計誤差は小さいため、解析解と数値解の小さいが有意な差異は、有限サイズ効果による系統誤差と思われる。

4 終わりに

本稿では、レプリカ法の近年の成果を用いて、イジングパーセプトロンを詳細に検討し直した。その結果、イジングパーセプトロンはその相空間構造がやや異常であり Complexity が上に凸でない関数形を持つことを示した。また、有限レプリカ数の領域を積極的に活用することで典型解析では調べることができない領域を系統的に調べることが可能にし、非典型に現れる記憶可能なサンプルの出現確率を解析的に評価することに成功した。今回の成果は、RSB の意味が曖昧でその運用法を手探りで探さなければいけなかった時代では、得ることが難しかったであろう内容である。レプリカ法の理解が深まることで、多くのことが分かるようになる問題はパーセプトロン以外にも恐らくたくさんあるであろう。また一方で現在のレプリカ

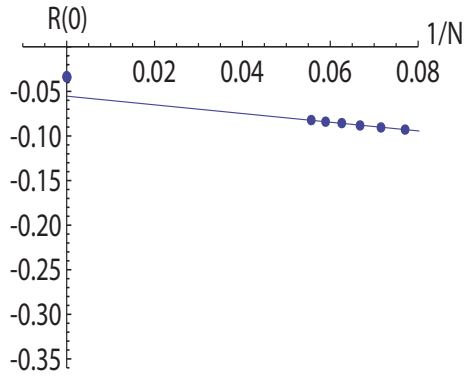


図 8: レート関数の $\alpha = 1.0$ におけるサイズ依存性。横座標 $1/N = 0$ における点は frozen RSB から計算される値である。システムサイズは $N = 12, 13, \dots, 18$ である。サンプル数はそれぞれの N に対して 320,000 である。統計誤差はマーカーより小さい。

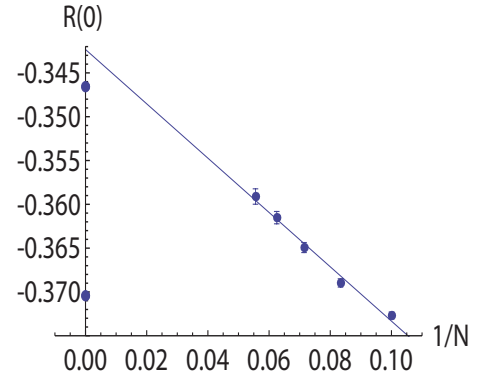


図 9: レート関数の $\alpha = 1.5$ におけるサイズ依存性。横座標 $1/N = 0$ における二つの点は上が RS2 解、下が frozen RSB 解に対応する。システムサイズは $N = 10, 12, \dots, 18$ 。サンプル数はそれぞれの N に対し 25, 600, 000。

法の処方箋では対処しきれないような問題もまだまだ眠っていると思われる。それらの系へのレプリカ法を用いた積極的な解析は、問題の理解のみならずレプリカ法の理解にとっても有益であり今後の重要な課題である。

参考文献

- [1] H. Nishimori: *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford University Press, Oxford, 2001).
- [2] T. Obuchi and Y. Kabashima: arXiv:0910.2281
- [3] B. Derrida: Phys. Rev. B **24** (1981) 2613
- [4] E. Gardner and B. Derrida: J. Phys. A **22** (1989) 1975
- [5] T. Nakajima and K. Hukushima: J. Phys. Soc. Jpn. **77** (2008) 074718
- [6] T. Obuchi, Y. Kabashima and H. Nishimori: J. Phys. A **42** (2009) 075004
- [7] S. Cocco, R. Monasson and R. Zecchina: Phys. Rev. E **54** (1996) 717
- [8] G. Biroli and R. Monasson: Europhys. Lett. **50** (2000) 155
- [9] J. Ardelius and L. Zdeborova: Phys. Rev. E **78** (2008) 040101
- [10] E. Gardner and B. Derrida: J. Phys. A **21** (1988) 271
- [11] W. Krauth and M. Mézard: J. Phys. France **50** (1989) 3057
- [12] R. Monasson and D. O’Kane: Europhys. Lett. **27** (1994) 85
- [13] R. Monasson: Phys. Rev. Lett. **75** (1995) 2847
- [14] A. Braunstein and R. Zecchina: Phys. Rev. Lett. **96** (2006) 030201
- [15] J. R. L. de Almeida and D. J. Thouless: J. Phys. A **11** (1978) 983
- [16] H. Akaike: IEEE Trans. on Automatic Control **19** (1974) 716