

# ベータ自由エネルギーと Loopy belief Propagation に現れる グラフのゼータ関数について

渡辺 有祐\*  
Yusuke Watanabe

福水 健次\*  
Kenji Fukumizu

**Abstract:** In this paper, we show a new formula which connects the Hessian of the Bethe free energy and the multivariable graph zeta function. Utilizing this formula, we show new methods for analyses of Loopy Belief Propagation (LBP) algorithm. We mainly prove and discuss the formula in the case of binary pairwise model. First, we give a sufficient condition that the Hessian of the Bethe free energy is positive definite, which shows non-convexity for graphs with multiple cycles. The formula clarifies the relation between the local stability of a fixed point of LBP and local minima of the Bethe free energy. We also propose a new approach to the uniqueness of LBP fixed point, and show various conditions of uniqueness. Finally, we discuss the extension to more general class of graphical models including multinomial models and Gaussian models.

**Keywords:** Loopy Belief Propagation, Bethe free energy, graph zeta function

## 1 導入

Pearl によるビリーブプロパゲーション (Belief Propagation, BP) アルゴリズム [1] は、木構造を持つ確率モデルに対して効率的に分配関数や周辺確率分布を厳密計算することが知られていた。さらにサイクルをもつようなグラフ上の確率モデルに対してもこのアルゴリズムをそのまま適用すると、各種の応用において非常に高い近似性能を与えることが経験的に知られている [2]。このように拡張されたアルゴリズムは通常ルーピービリーブプロパゲーション (Loopy Belief Propagation, LBP) と呼ばれている。LBP は BP とは違って、複雑なふるまいをするアルゴリズムであり、それについて理解を深めることは重要な問題である。

この LBP アルゴリズムはベータ自由エネルギーと密接に関係している [3] という大変興味深い性質を持っている。すなわち、LBP アルゴリズムの固定点がベータ自由エネルギーの勾配がゼロの点と一対一に対応する。

この論文ではまず最初に、バイナリかつペアワイズなモデルに対して「ベータ自由エネルギーの行列式は、正の因子を除いて、グラフの多変数ゼータ関数の逆数に等しい」という公式を証明する。その後でこの公式を三つ

の問題に応用する。

まず第一に、この公式をベータ自由エネルギーのヘッセ行列の正定値性に関する議論に応用する。ベータ自由エネルギーは一般には凸でないことが知られており、これが LBP の振動現象や固定点が多数出現することの一つの原因になっている。したがって、ヘッセ行列が正定値である範囲を知ることは興味のある問題である。この論文ではヘッセ行列が正定値である簡単な十分条件を与える。そしてさらにベータ自由エネルギーが凸である必要十分条件はグラフがツリー、又はただ一つのサイクルを持つグラフであることを証明する。

第二の応用として、LBP の固定点の安定性とその点でのベータ自由エネルギーの局所的な構造の関係を議論する。LBP はベータ自由エネルギーを単純に小さくするようなアルゴリズムではないため、その関係は非自明である。それについて Heskes [4] は、局所安定な LBP の固定点はベータ自由エネルギーの極小であることを示している。本論文では、グラフのゼータ関数に現れる行列のスペクトルについて考え、Heskes の結果を拡張する。

第三に、LBP の固定点の一意性について議論する。まず、「LBP の固定点の指数の和は 1」という定理を示し、それとゼータ関数とヘッセ行列の公式を組み合わせるというアプローチで一意性を議論する。特にこの方法で「二つの一次独立なサイクルを持つグラフ上で、attractive でない相互作用の場合」を考えると、LBP の

\*総合研究大学院大学 複合科学研究科 統計科学専攻、統計数理研究所, 106-8569 東京都港区南麻布 4-6-7, tel.03-3446-1501, e-mail watay@ism.ac.jp, fukumizu@ism.ac.jp  
Institute of Statistical Mathematics, 4-6-7, Minami-Azabu, Minato-Ku, Tokyo, 106-8569

固定点は一意的であるということを示す。

ここまではバイナリペアワイズのモデルに対する話であった。本論文の一番最後では、ゼータ関数とベータ自由エネルギーの公式が一般の有限状態のグラフィカルモデルや、ガウス型のモデルに拡張できることを述べる。

## 2 LBP とベータ自由エネルギー

この論文を通じ、 $G = (V, E)$  は頂点集合  $V$  と無向辺集合  $E$  をもつ連結グラフとする。頂点数は  $N$ 、辺の本数は  $M$  とする。以下ではバイナリ  $x_i \in \{\pm 1\}$  の場合を考える。変数  $\mathbf{x} = (x_i)_{i \in V}$  上の確率分布が次のような形で与えられる場合（ペアワイズ型）を考えよう：

$$p(\mathbf{x}) = \frac{1}{Z} \prod_{ij \in E} \psi_{ij}(x_i, x_j) \prod_{i \in V} \psi_i(x_i), \quad (1)$$

ただしここで  $Z$  は規格化定数であり  $\psi_{ij}$  と  $\psi_i$  は正值の関数である。一般性を失うことなく  $\psi_i(x_i) = \exp(h_i x_i)$ 、 $\psi_{ij}(x_i, x_j) = \exp(J_{ij} x_i x_j)$  としよ。

各種の応用において、周辺分布  $p_i(x_i) := \sum_{\mathbf{x} \setminus \{x_i\}} p(\mathbf{x})$  と  $p_{ij}(x_i, x_j) := \sum_{\mathbf{x} \setminus \{x_i, x_j\}} p(\mathbf{x})$  を計算することが必要になる。グラフがツリーであればこれらは Pearl のベリフプロパゲーション（BP）アルゴリズムで効率的かつ厳密に計算できる [1]。グラフがサイクルを持つ場合でも BP をそのまま適用すると良い近似を与えることが経験的に知られており、この方法はルーピーベリフプロパゲーション（LBP）と呼ばれている。

LBP はメッセージを伝搬していくアルゴリズムである。すなわち、各有向辺に対して、まず最初に適当に初期化されたメッセージ  $\mu_{i \rightarrow j}(x_j)$  が割り当てられ、次のようにして更新される。

$$\mu_{i \rightarrow j}^{new}(x_j) \propto \sum_{x_i} \psi_{ji}(x_j, x_i) \psi_i(x_i) \prod_{k \in N_i \setminus j} \mu_{k \rightarrow i}(x_i), \quad (2)$$

ただし  $N_i$  は  $i \in V$  の近傍の頂点の集合である。通常は  $\sum_{x_i = \pm 1} \mu_{i \rightarrow j}(x_j) = 1$  のように規格化される。この論文では、メッセージはすべての辺について並列に更新するものとする。

メッセージがある固定点  $\{\mu_{i \rightarrow j}^\infty(x_j)\}$  に収束したとき、 $p_i(x_i)$ 、 $p_{ij}(x_i, x_j)$  の近似は次のように計算される。

$$\begin{aligned} b_i(x_i) &\propto \psi_i(x_i) \prod_{k \in N_i} \mu_{k \rightarrow i}^\infty(x_i), \\ b_{ij}(x_i, x_j) &\propto \psi_{ij}(x_i, x_j) \psi_i(x_i) \psi_j(x_j) \\ &\quad \prod_{k \in N_i \setminus j} \mu_{k \rightarrow i}^\infty(x_i) \prod_{k \in N_j \setminus i} \mu_{k \rightarrow j}^\infty(x_j), \end{aligned} \quad (3)$$

ただし、 $\sum_{x_i, x_j} b_{ij}(x_i, x_j) = 1$ 、 $\sum_{x_i} b_i(x_i) = 1$  と規格化するものとする。式 (2) と (3) より、関係式

$\sum_{x_j} b_{ij}(x_i, x_j) = b_i(x_i)$  と  $b_{ij}(x_i, x_j) > 0$  は自動的に満たされる。

次に、ベータ自由エネルギーをギブス自由エネルギーの計算可能な近似物として導入しよう。真の確率分布 (1) は変分問題  $p(\mathbf{x}) = \operatorname{argmin}_{\hat{p}} F_{Gibbs}(\hat{p})$  によって特徴づけられる。（最小は全ての  $(x_i)_{i \in V}$  上の確率分布に関してとる。）ギブス自由エネルギー  $F_{Gibbs}(\hat{p})$  は  $F_{Gibbs}(\hat{p}) = KL(\hat{p}||p) - \log Z$  によって定義され、 $\hat{p}$  に関する凸関数である。（ただし、 $KL(\hat{p}||p) = \int \hat{p} \log(\hat{p}/p)$  は  $\hat{p}$  から  $p$  へのカルバックライブラーダイバージェンス。）

一方、ベータ近似においては制約条件  $\sum_{x_i, x_j} b_{ij}(x_i, x_j) = 1$ 、 $\sum_{x_j} b_{ij}(x_i, x_j) = b_i(x_i)$ 、 $b_{ij}(x_i, x_j) > 0$  を満たした  $\{b_i(x_i), b_{ij}(x_i, x_j)\}$  に関する最小化を考える。このような制約条件を満たす  $\{b_i(x_i), b_{ij}(x_i, x_j)\}$  のことを pseudomarginal という。次で定義されるギブス自由エネルギーの変形をベータ自由エネルギーという。（ $d_i := |N_i|$  は頂点の次数。）

$$\begin{aligned} F(b) &:= - \sum_{ij \in E} \sum_{x_i x_j} b_{ij}(x_i, x_j) \log \psi_{ij}(x_i, x_j) \\ &\quad - \sum_{i \in V} \sum_{x_i} b_i(x_i) \log \psi_i(x_i) \\ &\quad + \sum_{ij \in E} \sum_{x_i x_j} b_{ij}(x_i, x_j) \log b_{ij}(x_i, x_j) \\ &\quad + \sum_{i \in V} (1 - d_i) \sum_{x_i} b_i(x_i) \log b_i(x_i). \end{aligned} \quad (4)$$

ベータ自由エネルギー  $F$  の定義域は pseudomarginal の全体であり、極小はただ一つとは限らない。この変分問題で得られる  $\{b_i(x_i), b_{ij}(x_i, x_j)\}$  は LBP によって得られるそれと同じである [3]。正確には、ベータ自由エネルギーの勾配ゼロの点の全体と LBP の固定点の全体が一対一に対応する。

以下の議論の都合上、冗長性のない変数  $m_i = E_{b_i}[x_i]$ 、 $\chi_{ij} = E_{b_{ij}}[x_i x_j]$  を使おう。このとき pseudomarginals は以下のように表示される。：

$$\begin{aligned} b_{ij}(x_i, x_j) &= \frac{1}{4}(1 + m_i x_i + m_j x_j + \chi_{ij} x_i x_j), \\ b_i(x_i) &= \frac{1}{2}(1 + m_i). \end{aligned} \quad (5)$$

同様に  $F$  の定義域は次のようになる：

$$L(G) := \left\{ \{m_i, \chi_{ij}\} \in \mathbb{R}^{N+M}; 1 + m_i x_i + m_j x_j + \chi_{ij} x_i x_j > 0 \right. \\ \left. ij \in E \quad x_i, x_j = \pm 1 \right\}.$$

$F$  のヘッセ行列  $\nabla^2 F$  とは  $\{m_i, \chi_{ij}\}$  に関する二階微分からなる、大きさ  $N + M$  の正方行列である。これは  $L(G)$  上の行列値関数とみることができる。ちなみに、 $\nabla^2 F$  は  $J_{ij}$  と  $h_i$  に依存しないことに注意しよう。

### 3 ゼータ関数とベーテ自由エネルギーのヘッセ行列

#### 3.1 ゼータ関数と伊原の公式

まずグラフに関する用語を準備しよう。グラフ  $G$  の各無向辺を逆向きの無向辺のペアとみて、有向辺集合  $\vec{E}$  をつくる。よって  $|\vec{E}| = 2M$  である。各有向辺  $e \in \vec{E}$  に対して、 $o(e) \in V$  を  $e$  の始点とし、 $t(e) \in V$  を  $e$  の終点とする。さらに、各  $e \in \vec{E}$  に対し、その逆を  $\bar{e}$ 、対応する無向辺を  $[e] = [\bar{e}] \in E$  と書くことにする。

$G$  の閉測地線とは有向辺の列  $(e_1, \dots, e_k)$  であって  $t(e_i) = o(e_{i+1}), e_i \neq \bar{e}_{i+1}$  ( $i = 1, \dots, k-1$ ),  $t(e_k) = o(e_1), e_k \neq \bar{e}_1$  を満たすものである。二つの閉測地線は一方が他方の巡回置換で与えられる時 同値という。閉測地線  $c = (e_1, \dots, e_k)$  の  $m$  回繰り返し、 $c^m = (e_1, \dots, e_k, e_1, \dots, e_k, \dots, e_1, \dots, e_k)$  を  $c$  の multiple という。さらに、閉測地線がほかの閉測地線の multiple になっていないとき、その同値類を素サイクルという。

さて、 $P$  を素サイクルの全体としよう。与えられた重み  $\mathbf{u} = (u_e)_{e \in \vec{E}}$  に対して、多変数ゼータ関数[5] は以下で定義される

$$\zeta_G(\mathbf{u}) := \prod_{p \in P} (1 - g(p))^{-1}.$$

ただしここで、 $p = (e_1, \dots, e_k)$ ,  $g(p) := u_{e_1} \cdots u_{e_k}$  とし、また  $u_e \in \mathbb{C}$  は収束性のため十分に小さいとする。

例 1.  $G$  がツリーるとき、素サイクルは存在しないので、 $\zeta_G(\mathbf{u}) = 1$  である。長さ  $N$  のサイクル状のグラフ  $C_N$  では、素サイクルは  $(e_1, e_2, \dots, e_N)$  と  $(\bar{e}_N, \bar{e}_{N-1}, \dots, \bar{e}_1)$  しかないので  $\zeta_{C_N}(\mathbf{u}) = (1 - \prod_{l=1}^N u_{e_l})^{-1} (1 - \prod_{l=1}^N u_{\bar{e}_l})^{-1}$  である。ただしこれら二つの場合を除くと、素サイクルは一般に無限個あることに注意しよう。

多変数ゼータ関数は次のような行列式表示を持つ。これにより、多変数ゼータの定義域は全  $\mathbb{C}^{2M}$  上に延長される。有向边上の関数の全体を  $C(\vec{E})$  と書こう。 $C(\vec{E})$  に作用する行列  $M$  を

$$M_{e,e'} := \begin{cases} 1 & e \neq \bar{e}' \text{ かつ } o(e) = t(e') \text{ の場合,} \\ 0 & \text{それ以外の場合} \end{cases} \quad (6)$$

と定義する。

定理 1 ([5], Theorem 3).

$$\zeta_G(\mathbf{u}) = \det(I - UM)^{-1}, \quad (7)$$

但し、 $U$  は  $U_{e,e'} := u_e \delta_{e,e'}$  なる対角行列である。

次に多変数ゼータ関数のもう一つの行列式表示を証明しよう。この公式は定理 3 の証明で非常に重要な働きを果たすことになる。

定理 2 (伊原の公式の多変数版).  $V$  上の関数の全体を  $C(V)$  と書くことにする。 $C(V)$  の二つの線形作用素を以下のように定義する。

$$\begin{aligned} (\hat{D}f)(i) &:= \left( \sum_{\substack{e \in \vec{E} \\ t(e)=i}} \frac{u_e u_{\bar{e}}}{1 - u_e u_{\bar{e}}} \right) f(i), \\ (\hat{A}f)(i) &:= \sum_{\substack{e \in \vec{E} \\ t(e)=i}} \frac{u_e}{1 - u_e u_{\bar{e}}} f(o(e)). \end{aligned} \quad (8)$$

ただしここで  $f \in C(V)$  である。このとき以下の公式が成立する。

$$\det(I - UM) = \det(I + \hat{D} - \hat{A}) \prod_{[e] \in E} (1 - u_e u_{\bar{e}}). \quad (9)$$

証明. 三つの作用素  $\mathcal{O}, \mathcal{T}^*, \iota$  を以下のように定義する:

$$\begin{aligned} (\mathcal{O}f)(e) &:= f(o(e)), & (\mathcal{T}^*g)(i) &:= \sum_{e \in \vec{E}, t(e)=i} g(e), \\ (\iota g)(e) &:= g(\bar{e}), & f &\in C(V), g \in C(\vec{E}). \end{aligned}$$

すると明らかに  $M = \mathcal{O}\mathcal{T}^* - \iota$  が成り立つ。よって、 $\det(I - UM) = \det(I - \mathcal{T}^*(I + U\iota)^{-1}U\mathcal{O}) \det(I + U\iota)$  を得る。ただしここで  $n \times m, m \times n$  の行列  $A, B$  に対して  $\det(I_n - AB) = \det(I_m - BA)$  が成り立つことを用いた。

$\iota$  は自然な基底でブロック対角行列になっている。よって、 $I + U\iota$  の  $(e, \bar{e})$  ブロックは

$$\begin{bmatrix} 1 & u_e \\ u_{\bar{e}} & 1 \end{bmatrix}$$

である。したがって  $\det(I + U\iota) = \prod_{[e] \in E} (1 - u_e u_{\bar{e}})$  が成立する。

最後に  $\mathcal{T}^*(I + U\iota)^{-1}U\mathcal{O} = \hat{A} - \hat{D}$  を確認しよう。 $f \in C(V)$  に対して、

$$\begin{aligned} & \left( \mathcal{T}^*(I + U\iota)^{-1}U\mathcal{O}f \right)(i) \\ &= \sum_{e \in \vec{E}, t(e)=i} \left( (I + U\iota)^{-1}U\mathcal{O}f \right)(e) \\ &= \sum_{e \in \vec{E}, t(e)=i} \frac{1}{1 - u_e u_{\bar{e}}} \left( (U\mathcal{O}f)(e) - u_e (U\mathcal{O}f)(\bar{e}) \right) \\ &= \sum_{e \in \vec{E}, t(e)=i} \frac{1}{1 - u_e u_{\bar{e}}} \left( u_e f(o(e)) - u_e u_{\bar{e}} f(o(\bar{e})) \right) \\ &= (\hat{A}f)(i) - (\hat{D}f)(i) \end{aligned}$$

となっていることが確認できる。  $\square$

全ての  $e \in \vec{E}$  に対して  $u_e = u$  として多変数ゼータ関数を一変数化したものは伊原ゼータ関数 [6] と呼ばれている。ここでは  $\zeta_G(u)$  と書くことにする。この場合、定理 2 は

$$\zeta_G(u)^{-1} = (1-u^2)^M \det\left(I + \frac{u^2}{1-u^2} \mathcal{D} - \frac{u}{1-u^2} \mathcal{A}\right) \quad (10)$$

となる。この式 (10) は伊原の公式などと呼ばれている。ただし  $\mathcal{D}$  は次数行列、 $\mathcal{A}$  は隣接行列と呼ばれ以下で定義される

$$(\mathcal{D}f)(i) := d_i f(i), \quad (\mathcal{A}f)(i) := \sum_{e \in \vec{E}, t(e)=i} f(o(e)), \quad f \in C(V).$$

### 3.2 主公式

定理 3 (主公式). 以下の公式が  $L(G)$  の任意の点で成り立つ:

$$\det(I - \mathcal{U}\mathcal{M}) = \det(\nabla^2 F) \prod_{ij \in E} \prod_{x_i, x_j = \pm 1} b_{ij}(x_i, x_j) \times \prod_{i \in V} \prod_{x_i = \pm 1} b_i(x_i)^{1-d_i} 2^{2N+4M} \quad (11)$$

ここで  $b_{ij}$ ,  $b_i$  は式 (5) によって与えられる。また、

$$u_{i \rightarrow j} := \frac{\chi_{ij} - m_i m_j}{1 - m_j^2} \quad (12)$$

と定めている。

略証. 定義から容易にヘッセ行列の (E,E)-ブロックは対角行列であることが分かる。この対角成分を使って (V,E)-ブロックと (E,V)-ブロックの成分を消す。言い換えると、正方行列  $X$  を  $\det X = 1$  でかつ

$$X^T (\nabla^2 F) X = \begin{bmatrix} Y & 0 \\ 0 & \left( \frac{\partial^2 F}{\partial \chi_{ij} \partial \chi_{kl}} \right) \end{bmatrix}$$

となるようにとる。長い計算を実行すると、

$$(Y)_{i,j} = \begin{cases} \frac{1}{1-m_i^2} + \sum_{k \in N_i} \frac{(\chi_{ik} - m_i m_k)^2}{(1-m_i^2)(1-m_i^2 - m_k^2 + 2m_i m_k \chi_{ik} - \chi_{ik}^2)} & : i = j \text{ の場合,} \\ -\mathcal{A}_{i,j} \frac{\chi_{ik} - m_i m_k}{1-m_i^2 - m_j^2 + 2m_i m_j \chi_{ij} - \chi_{ij}^2} & : \text{そのほかの場合} \end{cases}$$

を得る。式  $u_{j \rightarrow i} = \frac{\chi_{ij} - m_i m_j}{1 - m_i^2}$  を使えば  $I_N + \hat{\mathcal{D}} - \hat{\mathcal{A}} = YW$  となることが分かる。ただしここで  $\hat{\mathcal{A}}$ ,  $\hat{\mathcal{D}}$  は式 (8) で定義されたものであり、 $W$  は  $W_{i,j} := \delta_{i,j}(1 - m_i^2)$  で定義される対角行列である。以上より

$$\det(I - \mathcal{U}\mathcal{M}) = \det(Y) \prod_{i \in V} (1 - m_i^2) \prod_{[e] \in E} (1 - u_e u_{\bar{e}}) = (11) \text{ の右辺}$$

が成立する。第一の等式には定理 2 が使われていることに注意しておく。□

定理 3 はベータ自由エネルギーのヘッセ行列の行列式は本質的に  $\det(I - \mathcal{U}\mathcal{M})$  (すなわち多変数ゼータの逆数) に等しいことを言っている。節 5 で示すように  $\mathcal{U}\mathcal{M}$  は LBP の固定点においては、更新則の線形化に他ならないので、この公式は LBP に関する種々の性質を導く。

## 4 正定値性条件への応用

ベータ自由エネルギーの凸性は LBP の解の一意性を保証することもあり興味を持たれてきた。Pakzad ら [7] と Heskes [8] は凸性の十分条件を示し、ツリー又はただ一つのサイクルを持つグラフではベータ自由エネルギーが凸であることを証明した。この節では主公式の一つの応用として、そのような大域的な構造の代わりに局所的な構造について議論する。

以下、正方行列  $X$  に対して、 $\text{Spec}(X) \subset \mathbb{C}$  はその固有値全体を表す。また、 $X$  のスペクトル半径 (固有値の絶対値の最大値) を  $\rho(X)$  と書く。

定理 4.  $\mathcal{M}$  は式 (6) で与えられたものとする。さらに  $\{m_i, \chi_{ij}\} \in L(G)$  に対して  $\mathcal{U}$  は式 (12) で与えられるものとする。このとき、 $\text{Spec}(\mathcal{U}\mathcal{M}) \subset \mathbb{C} \setminus \mathbb{R}_{\geq 1}$  ならば  $\nabla^2 F$  が点  $\{m_i, \chi_{ij}\}$  で正定値。

証明. 点  $t \in [0, 1]$  に対して  $m_i(t) := m_i$ ,  $\chi_{ij}(t) := t\chi_{ij} + (1-t)m_i m_j$  と定義する。すると  $\{m_i(t), \chi_{ij}(t)\} \in L(G)$ ,  $\{m_i(1), \chi_{ij}(1)\} = \{m_i, \chi_{ij}\}$  が成り立つ。この区間では、 $\mathcal{U}(t)$  と  $\nabla^2 F(t)$  が  $\{m_i(t), \chi_{ij}(t)\}$  によって同様に定まる。明らかに  $\mathcal{U}(t) = t\mathcal{U}$  である。よって  $\text{Spec}(\mathcal{U}\mathcal{M}) \subset \mathbb{C} \setminus \mathbb{R}_{\geq 1}$  より、 $\det(I - t\mathcal{U}\mathcal{M}) \neq 0 \quad \forall t \in [0, 1]$  が成立する。よって定理 3 より  $\det(\nabla^2 F(t)) \neq 0$  がこの区間上で成立する。一方  $\nabla^2 F(0)$  が正定値行列であることは容易に確認できる。対称行列  $\nabla^2 F(t)$  の固有値は実数で、なおかつ  $t$  に関して連続なので  $\nabla^2 F(1)$  の固有値はすべて正の実数であることが分かる。□

$u_{i \rightarrow j}$  と  $u_{j \rightarrow i}$  の対称化を

$$\beta_{i \rightarrow j} = \beta_{j \rightarrow i} := \frac{\chi_{ij} - m_i m_j}{\{(1 - m_i^2)(1 - m_j^2)\}^{1/2}} = \frac{\text{Cov}_{b_{ij}}[x_i, x_j]}{\{\text{Var}_{b_i}[x_i] \text{Var}_{b_j}[x_j]\}^{1/2}} \quad (13)$$

のようにして定める。つまり、 $u_{i \rightarrow j} u_{j \rightarrow i} = \beta_{i \rightarrow j} \beta_{j \rightarrow i}$  が成り立つ。 $\beta_{i \rightarrow j} = \beta_{j \rightarrow i}$  なので、しばしば  $\beta_{i \rightarrow j}$  を  $\beta_{ij}$  と略記する。最後の式より、 $|\beta_{ij}| < 1$  が成立することが見て取れる。対角行列  $\mathcal{Z}, \mathcal{B}$  を  $(\mathcal{Z})_{e,e'} := \delta_{e,e'}(1 - m_{t(e)}^2)^{1/2}$ ,  $(\mathcal{B})_{e,e'} := \delta_{e,e'} \beta_e$  のように定義しよう。すると  $\mathcal{B}\mathcal{M} = \mathcal{Z}\mathcal{U}\mathcal{M}\mathcal{Z}^{-1}$  が成立する。よって  $\text{Spec}(\mathcal{U}\mathcal{M}) = \text{Spec}(\mathcal{B}\mathcal{M})$  が成り立つ。

次の系はヘッセ行列が正定値であることのより明示的な条件を pseudomarginals の相関係数の言葉で与える。

系 1.  $\alpha$  を  $M$  のペロン・フロベニウス固有値とする。 $L_{\alpha^{-1}}(G) := \{m_i, \chi_{ij} \in L(G); |\beta_e| < \alpha^{-1} \ \forall e \in \vec{E}\}$  と定義する。このとき  $\nabla^2 F$  は  $L_{\alpha^{-1}}(G)$  上で正定値である。

証明.  $|\beta_e| < \alpha^{-1}$  であるので、 $\rho(\mathcal{BM}) < \rho(\alpha^{-1}\mathcal{M}) = 1$  が成立する ([9] Theorem 8.1.18)。よって  $\text{Spec}(\mathcal{BM}) \cap \mathbb{R}_{\geq 1} = \emptyset$ 。□

$\alpha^{-1}$  は原点からそれに一番近い伊原ゼータの極への距離に他ならない。例 1 より、 $\zeta_G(u) = 1$  ( $G$  はツリー)、 $\zeta_{C_N}(u) = (1 - u^N)^{-2}$  である。よって  $\alpha^{-1}$  はそれぞれ  $\infty$ 、1 である。これらの場合  $|\beta_e| < 1$  が常に成り立つことより、 $L_{\alpha^{-1}}(G) = L(G)$  で  $F$  は  $L(G)$  上で凸である。これは [8] の結果の別証明になっている。一般に [9] の定理 8.1.22 を使うと  $\min_{i \in V} d_i - 1 \leq \alpha \leq \max_{i \in V} d_i - 1$  が言える。

非凸性に関しては主公式から次のことが分かる。

系 2.  $t < 1$  で  $\{m_i(t) := 0, \chi_{ij}(t) := t\} \in L(G)$  とおく。このとき、

$$\lim_{t \rightarrow 1} \det(\nabla^2 F(t))(1-t)^{M+N-1} = -2^{-M-N+1}(M-N)\kappa(G),$$

が成立する。ただし  $\kappa(G)$  は全域木の個数である。特にグラフが連結で二つ以上の一次独立なサイクルをもつとき (つまり  $M - N \geq 1$ )、 $F$  は  $L(G)$  上で凸ではない。

証明. 主張の式は橋本の定理 [10] から導かれる。橋本の定理は伊原ゼータ関数の  $u \rightarrow 1$  での極限を与える。(詳細略) 後半の主張は明らか。□

この節の結果をまとめると、 $F$  が  $L(G)$  が凸である必要十分条件はグラフがツリーまたは 1 - サイクルのグラフであることである。我々の知る限り、これは新しい結果である。

## 5 安定性への応用

この節では LBP の局所安定性とベータ自由エネルギーの LBP 固定点周りでの局所的な構造を議論する。Heskes [4] は (十分に緩和して) 局所安定な固定点はベータ自由エネルギーの極小であることを示した。その逆は必ずしも成り立たないことが知られている。このギャップがどれ程のものなのか以下で見ていこう。

まず最初に LBP の更新式を離散力学系としてみよう。今考えているモデルはバイナリだったので、各メッセージ  $\mu_{i \rightarrow j}(x_j)$  はスカラーパラメタ  $\eta_{i \rightarrow j}$  でパラメトライズできる。各時刻における LBP アルゴリズムの状

態は  $\eta = (\eta_e)_{e \in \vec{E}} \in C(\vec{E})$ 、よって記述され、更新式 (2) はその上の変換  $T$  とみられる。LBP の固定点は  $\{\eta^\infty \in C(\vec{E}); T(\eta^\infty) = \eta^\infty\}$  と書ける。

固定点  $\eta^\infty$  はこの点の十分近くから出発すると必ずこの点に収束するとき、局所安定という。局所安定性は固定点における  $T$  の線形化  $T'$  によって決定される。論文 [11] でも議論されているとおり、 $\eta^\infty$  が局所安定である必要十分条件は  $\text{Spec}(T'(\eta^\infty)) \subset \{\lambda \in \mathbb{C}; |\lambda| < 1\}$  である。

LBP の振動現象を抑えるため、更新式の緩和 (damp)  $T_\epsilon := (1 - \epsilon)T + \epsilon I$  はしばしば有用である。ただしここで  $0 \leq \epsilon < 1$  は緩和の強さを表すパラメタで  $I$  は単位行列である。固定点がある強さの緩和のもとで局所安定である必要十分条件は  $\text{Spec}(T'(\eta^\infty)) \subset \{\lambda \in \mathbb{C}; \text{Re} \lambda < 1\}$  である。

LBP 更新式の線形化  $T'(\eta^\infty)$  を Furtlehner らに従って良い座標系で表示しよう。

定理 5 ([12], Proposition 4.5).  $u_{i \rightarrow j}$  は LBP 固定点  $\eta^\infty$  において式 (3), (5), (12) で与えられたものとする。線形化  $T'(\eta^\infty)$  は  $UM$  に相似である。つまりある正則行列が存在して  $PUM = PT'(\eta^\infty)P^{-1}$  が成立する。

定理 3 において  $\det(I - T'(\eta^\infty)) = \det(I - UM)$  となるので、定理 3 は線形化行列とベータ自由エネルギーの局所的な構造の直接的な関係を表していると言える。さらに定理 4 より、LBP の固定点は  $\text{Spec}(T'(\eta^\infty)) \subset \mathbb{C} \setminus \mathbb{R}_{\geq 1}$  のときベータ自由エネルギーの極小であることがわかる。

以上をまとめると以下のことが分かる。 $\{\lambda \in \mathbb{C}; \text{Re} \lambda < 1\}$  は  $\mathbb{C} \setminus \mathbb{R}_{\geq 1}$  に含まれるので、(十分に緩和して) 局所安定な固定点はベータ自由エネルギーの極小である。これは Heskes [4] によって示されている。さらに、 $\text{Spec}(T'(\eta^\infty))$  が  $\mathbb{C} \setminus \mathbb{R}_{\geq 1}$  に含まれるが、 $\{\lambda \in \mathbb{C}; \text{Re} \lambda < 1\}$  に含まれないときその固定点はベータ自由エネルギーの極小であるのにどう緩和しても局所安定ではないことがいえる。

どのような条件のもとではベータ自由エネルギーの極小は (緩和された) LBP の局所安定な固定点になっているのかは興味ある問題であろう。現時点ではこの問題に完全には答えられないが、attractive なモデル ( $J_{ij} \geq 0$ ) に関しては次の結果が得られる。この定理は要するに、相互作用と外場の強さを連続的に動かしていくとき、安定な固定点が不安定化する点はベータ自由エネルギーの極小が鞍点化する点に等しいことを言っている。

定理 6. 連続的にパラメトライズされた attractive なモデル  $\{\psi_{ij}(t), \psi_i(t)\}$  を考える。(例えば  $t$  は温度:  $\psi_{ij}(t) = \exp(t^{-1} J_{ij} x_i x_j)$ ,  $\psi_i(t) = \exp(t^{-1} h_i x_i)$ ) と与えられた  $t$  に対して、LBP を走らせ、安定な固定点に到達したとす

る。ここで、 $t$  を動かし、 $t = t_0$  でこの固定点が不安定化したとすると、そのベータ自由エネルギーの極小が  $t = t_0$  で鞍点に変化する。

略証. まず、attractive であることから  $u_{i \rightarrow j} \geq 0$  が常に成り立つ。あとはペロン・フロベニウスの定理と定理 3 より明らか。□

定理 6 は [11] の定理 2 の拡張になっている。(彼らは  $h_i = 0$  で  $m_i = 0$  の場合のみを考えている。)

## 6 LBP 固定点の一意性問題への応用

LBP 固定点の一意性は、LBP が収束するときにはその固定点がベータ自由エネルギーの最小であることを保証するので、過去たくさん議論されてきた。その方法としては minmax 問題に変形する [8]、縮小写像になる条件を調べる [13, 14]、Gibbs 測度の一意性を考える [15] などが考えられてきた。ここではそれらとは異なり、微分トポロジ的な手法を使う。我々の手法においては定理 3 と合わせて、以下の定理が重要である。

定理 7. 任意の  $q \in (\nabla F)^{-1}(0)$  に対して  $\det \nabla^2 F(q) \neq 0$  と仮定する。

$$\sum_{q: \nabla F(q)=0} \operatorname{sgn}(\det \nabla^2 F(q)) = 1,$$

ここで  $x > 0$  のとき  $\operatorname{sgn}(x) = 1$ 、 $x < 0$  のとき  $\operatorname{sgn}(x) = -1$  である。この和の各項 (+1 または -1) を  $q$  における  $F$  の指数という。

ベータ自由エネルギーの勾配がゼロの点  $(\nabla F)^{-1}(0)$  は LBP の固定点に他ならないことに注意しよう。この定理は LBP の固定点全てに関して指数を足し上げると 1 になることを言っている。

この定理の証明は省略する。この定理を証明するにあたり、ベータ自由エネルギーの重要な性質は  $q_n \rightarrow \partial L(G)$  に対し  $\|\nabla F(q_n)\| \rightarrow \infty$  が成り立つことである。ただしここで  $\partial L(G)$  は  $L(G) \subset \mathbb{R}^{N+M}$  の境界集合である。

一般に、関数  $F$  の定義域が一次元の場合は勾配ゼロの点における指数とは二回微分の符号に他ならない。この場合に適切な境界条件のもとで定理 7 が成り立つことは図 1 から直感的に納得されるであろう。

もし何らかの方法で、前もって LBP の固定点の指数が必ず 1 であることを保証できれば LBP の固定点はただ一つであることが証明できる。LBP を実行する前からわかる知識としては以下がある。

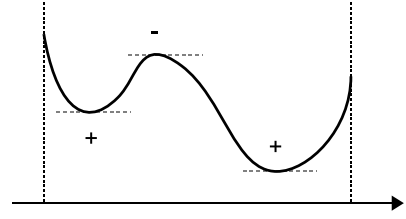


図 1: 勾配 0 の点の指数を足すと 1

補題 1. 任意の LBP 固定点において  $\beta_{ij}$  は式 (13) によって与えられたものとする。このとき  $|\beta_{ij}| \leq \tanh(|J_{ij}|)$ 、 $\operatorname{sgn}(\beta_{ij}) = \operatorname{sgn}(J_{ij})$  が成立する。

証明. 式 (3) より、ある  $\theta_i, \theta_j$  が存在して  $b_{ij}(x_i, x_j) \propto \exp(J_{ij}x_i x_j + \theta_i x_i + \theta_j x_j)$  が成立する。等号は  $\theta_i = 0$  かつ  $\theta_j = 0$  のとき成立する。□

定理 7 と補題 3 より論文 [14] で示されている一意性の条件が直ちに出てくる。(ただし [14] ではこの条件のもとでもっと強く、ただ一つの解への収束性まで示している。)

系 3 ([14]).  $\rho(\mathcal{J}\mathcal{M}) < 1$  ならば LBP の固定点はただ一つである。ただしここで  $\mathcal{J}$  は  $\mathcal{J}_{e,e'} = \tanh(|J_e|)\delta_{e,e'}$  は定まる対角行列である。

証明.  $|\beta_{ij}| \leq \tanh(|J_{ij}|)$  であるので  $\rho(\mathcal{B}\mathcal{M}) \leq \rho(\mathcal{J}\mathcal{M}) < 1$  が成立する。よって  $\det(I - \mathcal{B}\mathcal{M}) = \det(I - \mathcal{U}\mathcal{M}) > 0$  であり、LBP の固定点の指数は必ず 1 である。□

系 3 の証明では、絶対値の評価のみを用いた。以下の系 4 では符号の情報も使う。この系を述べるために用語を準備しよう。相互作用  $\{J_{ij}, h_i\}$  と  $\{J'_{ij}, h'_i\}$  は  $(s_i) \in \{\pm 1\}^V$  が存在して  $J'_{ij} = J_{ij}s_i s_j$ 、 $h'_i = h_i s_i$  を満たすとき同値であるという。同値なモデルはゲージの変換  $x_i \rightarrow x_i s_i$  で得られるので、LBP 固定点の一意性は等しいことに注意しよう。

系 4. 連結グラフ  $G$  の一次独立なサイクルの個数が 2 であるとする。(つまり  $M - N + 1 = 2$ 。) このとき、attractive なモデルに同値でない限り、LBP 固定点はただ一つ存在する。

完全な証明を与える代わりに一つの例で示そう。

例 2.  $V := \{1, 2, 3, 4\}$ ,  
 $E := \{12, 13, 14, 23, 34\}$  なるグラフを考える。相互作用は任意の  $\{h_i\}$  と  $\{-J_{12}, J_{13}, J_{14}, J_{23}, J_{34}\}$  ( $J_{ij} \geq 0$ ) で与えられるものとする。(図 2 参照。) 任意の  $0 \leq \beta_{13}, \beta_{23}, \beta_{14}, \beta_{34} < 1$  と  $-1 < \beta_{12} \leq 0$  に対して  $\det(I -$

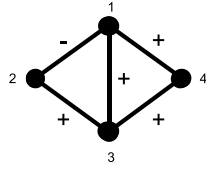


図 2: 例 2 のグラフ

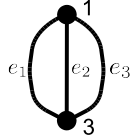


図 3: グラフ  $\hat{G}$

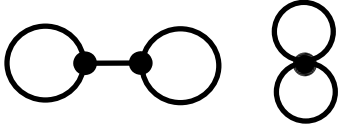


図 4: 他の二種類

$\mathcal{BM}) > 0$  を示せば十分。  $G$  の素サイクルの集合は  $\hat{G}$  (図 3) のそれと一対一に対応する。したがって、  $\det(I - \mathcal{BM}) = \det(I - \hat{\mathcal{B}}\hat{\mathcal{M}})$  が成立する。ただしここで  $\hat{\beta}_{e_1} = \beta_{12}\beta_{23}$ ,  $\hat{\beta}_{e_2} = \beta_{13}$ ,  $\hat{\beta}_{e_3} = \beta_{34}$ 。  $-1 < \hat{\beta}_{e_1} \leq 0$  と  $0 \leq \hat{\beta}_{e_2}, \hat{\beta}_{e_3} < 1$  より、  $\det(I - \hat{\mathcal{B}}\hat{\mathcal{M}}) = (1 - \hat{\beta}_{e_1}\hat{\beta}_{e_2} - \hat{\beta}_{e_1}\hat{\beta}_{e_3} - \hat{\beta}_{e_2}\hat{\beta}_{e_3} - 2\hat{\beta}_{e_1}\hat{\beta}_{e_2}\hat{\beta}_{e_3})(1 - \hat{\beta}_{e_1}\hat{\beta}_{e_2} - \hat{\beta}_{e_1}\hat{\beta}_{e_3} - \hat{\beta}_{e_2}\hat{\beta}_{e_3} + 2\hat{\beta}_{e_1}\hat{\beta}_{e_2}\hat{\beta}_{e_3}) > 0$  が言える。ほかの場合は  $\hat{G}$  または図 4 のグラフに同様に帰着できる。

一方 attractive なモデルに関しては LBP の固定点は必ずしも一つではない。 [11, 14]

過去の LBP 固定点の一意性に関する結果はすべて、  $|J_{ij}|$  を何らかの意味で上から抑える条件のもとで得られている。系 4 はそれとはまったく異なるタイプの条件のもとで一意性を示していることに注意しておく。

## 7 一般的なモデルに対する拡張

ここまでのところ、バイナリ、ペアワイズのモデルのみを考えてきた。しかし、定理 3 はもっと一般のモデルにまで拡張できる。この節では、この拡張について手短かに述べる。詳細は今後の論文に譲りたい。

$G = (V \cup F, \vec{E})$  をファクターグラフとしよう。  $e \in \vec{E}$  はファクター  $o(e)$  を始点とし、変数頂点  $t(e)$  を終点とする有向辺であるとする。  $G$  の閉測地線とは  $(e_1, e_2, \dots, e_k)$  なる有向辺の列であって  $i \in \mathbb{Z}/k\mathbb{Z}$  で  $t(e_i) \in o(e_{i+1})$ ,  $t(e_i) \neq t(e_{i+1})$  を満たすものである。  $G$  の素サイクルも普通のグラフのときと同様に定義される。与えられた正の整数の集合  $\{r_i\}_{i \in V}$ ,  $r_j \times r_i$  行列の集合  $\{u_{i \rightarrow j}^\alpha\}_{i, j \in \alpha}$  に対し、ゼータ関数の拡張は

$$\zeta_G(\mathbf{u}) := \prod_{[C] \in \mathcal{P}} \frac{1}{\det(I - u_{t(e_{k-1}) \rightarrow t(e_k)}^{o(e_k)} \cdots u_{t(e_k) \rightarrow t(e_1)}^{o(e_1)})}$$

によって定義される。ただしここで  $C = (e_1, e_2, \dots, e_k)$  とした。この場合も定理 1 と同様な行列式表示が存在する。

各ファクター  $\alpha \in F$ 、変数頂点  $i \in V$  に対して、  $x_\alpha = (x_i)_{i \in \alpha}$  上の指数型分布族  $\mathcal{E}_\alpha$  と  $x_i$  上の指数型分布族  $\mathcal{E}_i$  が与えられているとする。これらの十分統計量を  $\phi_\alpha(x_\alpha)$ ,  $\phi_i(x_i)$  と書くことにし、次元を  $r_\alpha, r_i$  とする。さらに  $\phi_\alpha$  が  $i \in \alpha$  で  $\phi_i$  を含むとする。すなわち

$$\phi_\alpha(x_\alpha) = (\phi_{i_1}(x_{i_1}), \dots, \phi_{i_k}(x_{i_k}), \hat{\phi}_\alpha(x_\alpha))$$

ここで、  $\alpha = \{i_1, \dots, i_k\}$  とした。さらに「与えられた指数型分布族は周辺化でとじている」ということを仮定する。すなわち、もし  $p(x_\alpha) \in \mathcal{E}_\alpha$  ならば  $\sum_{x_\alpha \setminus i} p(x_\alpha) \in \mathcal{E}_i$  である。たとえば、有限状態で任意のファクターグラフで書かれるモデルやガウス型モデルはこれらの状況に合う。LBP のアルゴリズムは自然にこのクラスのモデルに対して定義される。

LBP の固定点では期待値パラメタ  $\eta_i = E_{b_i}[\phi_i]$  と  $\eta_\alpha = E_{b_\alpha}[\phi_\alpha]$  は局所的な整合性条件  $(\eta_\alpha)_i = \eta_i$  を各  $\alpha \ni i$  で満たす。ベータ自由エネルギーはこのような条件を満たす期待値パラメタの関数になっている。また、LBP の固定点はベータ自由エネルギーの勾配ゼロの点に一致する。

定理 8 (定理 3 の拡張). 整合性条件を満たすような期待値パラメタ  $\{\eta_i, \eta_\alpha\}$  に対して以下の公式が成立。

$$\zeta_G(\mathbf{u})^{-1} = \det(\nabla^2 F) \prod_{\alpha \in F} \det(\text{Var}_{b_\alpha}[\phi_\alpha]) \prod_{i \in V} \det(\text{Var}_{b_i}[\phi_i])^{1-d_i},$$

ただしここで  $u_{i \rightarrow j}^\alpha := \text{Var}_{b_j}[\phi_j]^{-1} \text{Cov}_{b_\alpha}[\phi_j, \phi_i]$  は  $r_j \times r_i$  の行列である。

最後に付け加えると fractional belief propagation [16] 型のベータ自由エネルギーの変形に対しても定理 8 は拡張できることが分かっている。この場合はバクトラックを許す Bartholdi 型のゼータ関数が現れる。 [17]

## 8 おわりに

最後に今後の課題について述べる。

まず定理 8 を Generalized Belief Propagation (GBP) や Expectation Propagation の場合に対して、拡張ができるのかは興味のある問題である。特に、GBP に拡張するのに必要になるゼータ関数の拡張は数学的にも興味深い対象である可能性がある。

一方 LBP の重要な応用の一つである LDPC 符号についての研究では、Koetter [18] らは pseudo-codewords と多変数のゼータ関数との関係を示している。ガウス型のモデルでは、Johnson [19] らがゼータ的な分配関数の展開公式を与えている。これらの仕事と我々の仕事は基本的には大きく違っているが、その背後に何らかの関係性があるかどうかは今後の興味深い研究課題である。

本論文ではベータ自由エネルギーに関する議論に終始し、真の自由エネルギー（分配関数）についてはほとんど触れなかった。池田らの論文 [20, 21] では、情報幾何的な LBP の解釈を導入し、ベータ近似と厳密解の間の展開を導出している。我々の方法の中においても、ベータ近似と真の値を関係付けるような何らかの明確な定理を見出すことは重要な課題である。

## Acknowledgements

本研究は学術振興会特別研究員奨励費 (20-993) と基盤研究 C (19500249) の助成を受けたものである。

## References

- [1] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, 1988.
- [2] K. Murphy, Y. Weiss, and M.I. Jordan. Loopy belief propagation for approximate inference: An empirical study. *UAI*, 15:467–475, 1999.
- [3] J.S. Yedidia, W.T. Freeman, and Y. Weiss. Generalized belief propagation. *NIPS*, 13:689–95, 2001.
- [4] T. Heskes. Stable fixed points of loopy belief propagation are minima of the Bethe free energy. *NIPS*, pages 343–350, 2002.
- [5] H.M. Stark and A.A. Terras. Zeta functions of finite graphs and coverings. *Adv. in Math.*, 121(1):124–165, 1996.
- [6] Y. Ihara. On discrete subgroups of the two by two projective linear group over p-adic fields. *Journal of the Mathematical Society of Japan*, 18(3):219–235, 1966.
- [7] P. Pakzad and V. Anantharam. Belief propagation and statistical physics. *CISS*, 2002.
- [8] T. Heskes. On the uniqueness of loopy belief propagation fixed points. *Neural Comput.*, 16(11):2379–2413, 2004.
- [9] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, 1990.
- [10] K. Hashimoto. On zeta and L-functions of finite graphs. *Internat. J. Math*, 1(4):381–396, 1990.
- [11] JM Mooij and HJ Kappen. On the properties of the Bethe approximation and loopy belief propagation on binary networks. *J. Stat. Mech: Theor. Exp.*, (11):P11012, 2005.
- [12] C. Furtlehner, J.M. Lasgouttes, and A. De La Fortelle. Belief propagation and Bethe approximation for traffic prediction. *INRIA RR-6144*, 2007.
- [13] A.T. Ihler, JW Fisher, and A.S. Willsky. Loopy belief propagation: Convergence and effects of message errors. *JMLR*, 6(1):905–936, 2006.
- [14] J. M. Mooij and H. J. Kappen. Sufficient Conditions for Convergence of the Sum-Product Algorithm. *IEEE Trans. on Inf. Th.*, 53(12):4422–4437, 2007.
- [15] S. Tatikonda and M.I. Jordan. Loopy belief propagation and Gibbs measures. *UAI*, 18:493–500, 2002.
- [16] W. Wiegerinck and T. Heskes. Fractional belief propagation. *Advances in Neural Information Processing Systems*, pages 455–462, 2003.
- [17] L. Bartholdi. Counting Paths in Graphs. *Enseign. Math.*, 45:83–131, 1999.
- [18] R. Koetter, W.C.W. Li, PO Vontobel, and JL Walker. Pseudo-codewords of cycle codes via zeta functions. *IEEE Information Theory Workshop*, 2004.
- [19] J.K. Johnson, V.Y. Chernyak, and M. Chertkov. Orbit-Product Representation and Correction of Gaussian Belief Propagation. *ICML*, 2009.
- [20] S. Ikeda, T. Tanaka, and S. Amari. Stochastic reasoning, free energy, and information geometry. *Neural Computation*, 16(9):1779–1810, 2004.
- [21] S. Ikeda, T. Tanaka, and S. Amari. Information geometry of turbo and low-density parity-check codes. *IEEE Transactions on Information Theory*, 50(6):1097–1114, 2004.