Supplementary Material for

"Graph Zeta Function in the Bethe Free Energy and Loopy Belief Propagation"

Yusuke Watanabe and Kenji Fukumizu, NIPS 2009

A   Proof of theorem 2

Theorem 2 (Multivariable version of Ihara's formula). *Let $C(V)$ be the set of functions on V. We define two linear operators on $C(V)$ by*

$$(\hat{\mathcal{D}}f)(i) := \Big(\sum_{\substack{e \in \vec{E} \\ t(e)=i}} \frac{u_e u_{\bar{e}}}{1 - u_e u_{\bar{e}}} \Big) f(i), \quad (\hat{\mathcal{A}}f)(i) := \sum_{\substack{e \in \vec{E} \\ t(e)=i}} \frac{u_e}{1 - u_e u_{\bar{e}}} f(o(e)), \quad \text{where } f \in C(V).$$

Then we have

$$\Big(\zeta_G(\boldsymbol{u})^{-1} = \Big) \det(I - \mathcal{U}\mathcal{M}) = \det(I + \hat{\mathcal{D}} - \hat{\mathcal{A}}) \prod_{[e] \in E} (1 - u_e u_{\bar{e}}).$$

Proof. First, we define three linear operators $\mathcal{O} : C(V) \to C(\vec{E})$, $\mathcal{T}^* : C(\vec{E}) \to C(V)$, and $\iota : C(\vec{E}) \to C(\vec{E})$ as follows:

$$(\mathcal{O}f)(e) := f(o(e)), \quad (\mathcal{T}^*g)(i) := \sum_{e \in \vec{E}, t(e)=i} g(e), \quad (\iota g)(e) := g(\bar{e}) \quad \text{where } f \in C(V) \text{ and } g \in C(\vec{E}).$$

We see that $\mathcal{M} = \mathcal{O}\mathcal{T}^* - \iota$, because

$$\big((\mathcal{O}\mathcal{T}^* - \iota)g\big)(e) = \sum_{e' \in \vec{E}, t(e')=o(e)} g(e') \; - g(\bar{e}) = (\mathcal{M}g)(e) \quad \text{for } g \in C(\vec{E}).$$

Then we have

$$\det(I - \mathcal{U}\mathcal{M}) = \det\Big(I - \mathcal{U}\mathcal{O}\mathcal{T}^*(I + \mathcal{U}\iota)^{-1}\Big) \det(I + \mathcal{U}\iota)$$
$$= \det\Big(I - \mathcal{T}^*(I + \mathcal{U}\iota)^{-1}\mathcal{U}\mathcal{O}\Big) \det(I + \mathcal{U}\iota).$$

In the second equality, we used $\det(I_n - AB) = \det(I_m - BA)$ for $n \times m$ and $m \times n$ matrices A and B ([S1], Lemma 8.2.4). The linear operator ι is a block diagonal matrix with standard basis. The $(e, \bar{e})$ block of $I + \mathcal{U}\iota$ is

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

Therefore, we have $\det(I + \mathcal{U}\iota) = \prod_{[e] \in E}(1 - u_e u_{\bar{e}})$.

Finally, we check that $\mathcal{T}^*(I + \mathcal{U}\iota)^{-1}\mathcal{U}\mathcal{O} = \hat{\mathcal{A}} - \hat{\mathcal{D}}$. The matrix $(I + \mathcal{U}\iota)^{-1}$ is a block diagonal matrix with $(e, \bar{e})$ block

$$\frac{1}{1 - u_e u_{\bar{e}}} \begin{bmatrix} 1 & -u_e \\ -u_{\bar{e}} & 1 \end{bmatrix}. \tag{A.1}$$

1

For $f \in C(V)$, we have

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

$\square$

B Proof of theorem 3

B.1 Explicit formula of derivatives of the Bethe free energy

In the proof of theorem 3, the graph $G = (V, E)$ is assumed to be a simple graph, i.e., there is no multiple edges and loop-edge

For the proof, we need explicit expressions of the second derivatives of the Bethe free energy. We list them below.

The first derivatives of the Bethe Free Energy are

$$
\frac{\partial F}{\partial m_i} = -h_i + (1 - d_i)\frac{1}{2}\sum_{x_i=\pm 1} x_i \log b_i(x_i) + \frac{1}{4}\sum_{k \in N_i}\sum_{x_i,x_k=\pm 1} x_i \log b_{ik}(x_i, x_k), \tag{B.1}
$$

$$
\frac{\partial F}{\partial \chi_{ij}} = -J_{ij} + \frac{1}{4}\sum_{x_i,x_j=\pm 1} x_i x_j \log b_{ij}(x_i, x_j). \tag{B.2}
$$

The second derivatives of the Bethe Free Energy are

$$
\frac{\partial^2 F}{\partial m_i \partial m_j} = \begin{cases} (1 - d_i)\frac{1}{1-m_i^2} + \frac{1}{4}\sum_{k \in N_i}\sum_{x_i,x_k}\frac{1}{1+m_i x_i + m_k x_k + \chi_{ik}x_i x_k} & \text{if } i = j, \\ \frac{1}{4}\sum_{x_i,x_j}\frac{x_i x_j}{1+m_i x_i + m_j x_j + \chi_{ij}x_i x_j} & \text{if } i \text{ and } j \text{ are adjacent } (i \neq j), \\ 0 & \text{otherwise,} \end{cases} \tag{B.3}
$$

$$
\frac{\partial^2 F}{\partial m_k \partial \chi_{ij}} = \begin{cases} \frac{1}{4}\sum_{x_i,x_j}\frac{x_j}{1+m_i x_i + m_j x_j + \chi_{ij}x_i x_j} & \text{if } k = i, \\ \frac{1}{4}\sum_{x_i,x_j}\frac{x_i}{1+m_i x_i + m_j x_j + \chi_{ij}x_i x_j} & \text{if } k = j, \\ 0 & \text{otherwise,} \end{cases} \tag{B.4}
$$

$$
\frac{\partial^2 F}{\partial \chi_{ij} \partial \chi_{kl}} = \begin{cases} \frac{1}{4}\sum_{x_i x_j}\frac{1}{1+m_i x_i + m_j x_j + \chi_{ij}x_i x_j} & \text{if } ij = kl, \\ 0 & \text{otherwise.} \end{cases} \tag{B.5}
$$

We use notations

$$
r_{ij} := \frac{1}{4}\sum_{x_i,x_j}\frac{1}{1 + m_i x_i + m_j x_j + \chi_{ij}x_i x_j}, \tag{B.6}
$$

$$
s_{ij} := \frac{1}{4}\sum_{x_i,x_j}\frac{x_j}{1 + m_i x_i + m_j x_j + \chi_{ij}x_i x_j}, \tag{B.7}
$$

$$
t_{ij} := \frac{1}{4}\sum_{x_i,x_j}\frac{x_i x_j}{1 + m_i x_i + m_j x_j + \chi_{ij}x_i x_j}. \tag{B.8}
$$

Note that $r_{ij} = r_{ji}$ and $t_{ij} = t_{ji}$, but $s_{ij} \neq s_{ji}$ in general.

2

B.2 Detailed proof of theorem 3

Theorem 3 (Main Formula). *The following equality holds at any point of $L(G)$:*

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

where $b_{ij}(x_i, x_j) = \frac{1}{4}(1 + m_i x_i + m_j x_j + \chi_{ij} x_i x_j)$, $b_i(x_i) = \frac{1}{2}(1 + m_i)$, and

$$u_{i \to j} := \frac{\chi_{ij} - m_i m_j}{1 - m_j^2}. \tag{B.9}$$

Proof. First, note that the Hessian of the Bethe free energy is a square matrix of size $N + M$:

$$\nabla^2 F(\{m_i, \chi_{ij}\}) := \begin{bmatrix} \left(\frac{\partial^2 F}{\partial m_i \partial m_j}\right) & \left(\frac{\partial^2 F}{\partial m_i \partial \chi_{st}}\right) \\ \left(\frac{\partial^2 F}{\partial \chi_{uv} \partial m_j}\right) & \left(\frac{\partial^2 F}{\partial \chi_{uv} \partial \chi_{st}}\right) \end{bmatrix}.$$

Recall that N is the number of vertices and M is the number of undirected edges.

Step1: Computation of Y
From (B.5), the (E,E)-block of the Hessian is a diagonal matrix given by

$$\frac{\partial^2 F}{\partial \chi_{ij} \partial \chi_{kl}} = \delta_{ij,kl} r_{ij}.$$

Using this diagonal block, we erase (V,E)-block and (E,V)-block of the Hessian. Thus, we obtain a square matrix X such that $\det X = 1$ and

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

Applying an identity

$$\begin{pmatrix} 1 & 0 & \frac{-s_{ij}}{r_{ij}} \\ 0 & 1 & \frac{-s_{ji}}{r_{ij}} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} w_i & t_{ij} & s_{ij} \\ t_{ij} & w_j & s_{ji} \\ s_{ij} & s_{ji} & r_{ij} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{-s_{ij}}{r_{ij}} & \frac{-s_{ji}}{r_{ij}} & 1 \end{pmatrix} = \begin{pmatrix} w_i - \frac{s_{ij}^2}{r_{ij}} & t_{ij} - \frac{s_{ij}s_{ji}}{r_{ij}} & 0 \\ t_{ij} - \frac{s_{ij}s_{ji}}{r_{ij}} & w_j - \frac{s_{ji}^2}{r_{ij}} & 0 \\ 0 & 0 & r_{ij} \end{pmatrix}$$

for each edge, we have

$$(Y)_{i,j} = \begin{cases} (1 - d_i)\frac{1}{1-m_i^2} + \sum_{k \in N_i}\left(r_{ik} - \frac{s_{ik}^2}{r_{ik}}\right) & \text{if } i = j, \\ t_{ij} - \frac{s_{ij}s_{ji}}{r_{ij}} & \text{if } i \text{ and } j \text{ are adjacent}, \\ 0 & \text{otherwise.} \end{cases}$$

The elements of Y are represented in terms of $\{m_i, \chi_{ij}\}$ as follows:

$$(Y)_{i,i} = \frac{1}{1 - m_i^2} + \sum_{k \in N_i}\left(r_{ik} - \frac{s_{ik}^2}{r_{ik}} - \frac{1}{1 - m_i^2}\right)$$

$$= \frac{1}{1 - m_i^2} + \sum_{k \in N_i} \frac{(\chi_{ik} - m_i m_k)^2}{(1 - m_i^2)(1 - m_i^2 - m_k^2 + 2m_i m_k \chi_{ik} - \chi_{ik}^2)} \quad \text{and,}$$

$$(Y)_{i,j} = t_{ij} - \frac{s_{ij}s_{ji}}{r_{ij}}$$

$$= \frac{-(\chi_{ij} - m_i m_j)}{(1 - m_i^2 - m_j^2 + 2m_i m_j \chi_{ij} - \chi_{ij}^2)} \quad \text{for adjacent } i \text{ and } j.$$

Step2: Computation of $I_N + \hat{\mathcal{D}} - \hat{\mathcal{A}}$
From the definition (B.9) of $u_{j \to i}$, we see that

$$\frac{u_{i \to j} u_{j \to i}}{1 - u_{i \to j} u_{j \to i}} = \frac{(\chi_{ij} - m_i m_j)^2}{(1 - m_i^2 - m_j^2 + 2m_i m_j \chi_{ij} - \chi_{ij}^2)},$$

$$\frac{u_{i \to j}}{1 - u_{i \to j} u_{j \to i}} = \frac{(1 - m_i^2)(\chi_{ij} - m_i m_j)}{(1 - m_i^2 - m_j^2 + 2m_i m_j \chi_{ij} - \chi_{ij}^2)}.$$

3

Therefore, the diagonal element is

$$(I_N + \hat{\mathcal{D}} - \hat{\mathcal{A}})_{i,i} = (I_N + \hat{\mathcal{D}})_{i,i} = 1 + \sum_{k \in N_i} \frac{u_{i \to k} u_{k \to i}}{1 - u_{i \to k} u_{k \to i}}$$

$$= 1 + \sum_{k \in N_i} \frac{(\chi_{ik} - m_i m_k)^2}{(1 - m_i^2 - m_k^2 + 2 m_i m_k \chi_{ik} - \chi_{ik}^2)},$$

and for adjacent i and j,

$$(I_N + \hat{\mathcal{D}} - \hat{\mathcal{A}})_{i,j} = -(\hat{\mathcal{A}})_{i,j} = \frac{-u_{j \to i}}{1 - u_{i \to j} u_{j \to i}}$$

$$= \frac{-(1 - m_j^2)(\chi_{ij} - m_i m_j)}{(1 - m_i^2 - m_j^2 + 2 m_i m_j \chi_{ij} - \chi_{ij}^2)}.$$

Combining the results of step 1 and 2, we have

$$I_N + \hat{\mathcal{D}} - \hat{\mathcal{A}} = Y \begin{bmatrix} 1 - m_1^2 & 0 & \dots & 0 \\ 0 & 1 - m_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 - m_N^2 \end{bmatrix}.$$

Step3: Final step
We see that

$$\zeta_G(\boldsymbol{u})^{-1} = \det(I - \mathcal{U}\mathcal{M}) \tag{B.10}$$

$$= \det(I_N + \hat{\mathcal{D}} - \hat{\mathcal{A}}) \prod_{[e] \in E} (1 - u_e u_{\bar{e}}) \tag{B.11}$$

$$= \det(Y) \prod_{i \in V} (1 - m_i^2) \prod_{[e] \in E} (1 - u_e u_{\bar{e}})$$

$$= \det(\nabla^2 F) \prod_{i \in V} (1 - m_i^2) \prod_{ij \in E} \frac{1 - u_{i \to j} u_{j \to i}}{r_{ij}}$$

$$= \det(\nabla^2 F) \prod_{i \in V} (1 - m_i^2)^{1 - d_i} \prod_{ij \in E} \frac{(1 - u_{i \to j} u_{j \to i})(1 - m_i^2)(1 - m_j^2)}{r_{ij}}. \tag{B.12}$$

From (B.10) to (B.11), we used the edge zeta version of Ihara's formula (theorem 3).

Furthermore, with a straightforward computation we see that

$$\frac{(1 - u_{i \to j} u_{j \to i})(1 - m_i^2)(1 - m_j^2)}{r_{ij}} = 4^4 \prod_{x_i, x_j = \pm 1} b_{ij}(x_i, x_j),$$

$$(1 - m_i^2)^{1 - d_i} = 2^{2 - 2d_i} \prod_{x_i = \pm 1} b_i(x_i)^{1 - d_i},$$

where $b_{ij}(x_i, x_j) = \frac{1}{4}(1 + m_i x_i + m_j x_j + \chi_{ij} x_i x_j)$ and $b_i(x_i) = \frac{1}{2}(1 + m_i)$.

Therefore,

$$(B.12) = 2^{\sum_{i \in V}(2 - 2d_i)} 4^{4M} \det(\nabla^2 F) \prod_{i \in V} \prod_{x_i = \pm 1} b_i(x_i)^{1 - d_i} \prod_{ij \in E} \prod_{x_i, x_j = \pm 1} b_{ij}(x_i, x_j)$$

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

$$\square$$

4

C Proof of corollary 2

Here, we prove the limit formula in corollary 2.

Corollary 2. *Let* $\{m_i(t) := 0, \chi_{ij}(t) := t\} \in L(G)$ *for* $t < 1$. *Then we have*

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

where $\kappa(G)$ *is the number of spanning trees in* G.

Proof. We can easily check that $u_{i \to j}(t) = t$,

$$\prod_{ij \in E} \prod_{x_i, x_j = \pm 1} b_{ij}(x_i, x_j) = 4^{-4M}(1-t)^{2M}(1+t)^{2M}, \text{ and}$$

$$\prod_{i \in V} \prod_{x_i = \pm 1} b_i(x_i)^{1-d_i} = 2^{-2N+4M}$$

on this interval. Therefore

$$
\begin{aligned}
\lim_{t \to 1} \det(\nabla^2 F(t))(1-t)^{M+N-1} &= \lim_{t \to 1} \zeta_G(\boldsymbol{u}(t))^{-1}(1-t)^{M+N-1} \\
&\qquad \left(4^{-4M}(1-t)^{2M}(1+t)^{2M}2^{-2N+4M}2^{2N+4M} \right)^{-1} \\
&= \lim_{t \to 1} \zeta_G(t)^{-1}(1-t)^{-M+N-1}2^{-2M} \\
&= -(M-N)\kappa(G)2^{-M-N+1}.
\end{aligned}
$$

On the final equality, we used Hashimoto's formula:

$$\lim_{u \to 1} \zeta_G(u)^{-1}(1-u)^{-M+N-1} = -2^{M-N+1}(M-N)\kappa(G).$$

We refer to [S2–4] for this formula. $\qquad\square$

D Transformation of messages and proof of theorem 5

D.1 Transformation of messages

First, we make an easy observation on the LBP update.

Proposition D.1. *Let* $\{\pi_{i \to j}\}$ *be any set of messages. We define a transformation from messages* $\{\mu_{i \to j}^t\}$ *to messages* $\{\tilde{\mu}_{i \to j}^t\}$ *by*

$$\tilde{\mu}_{i \to j}^t(x_j) \propto \frac{\mu_{i \to j}^t(x_j)}{\pi_{i \to j}(x_j)}. \tag{D.1}$$

We also define transformation from functions $\{\psi_{ij}, \psi_i\}$ *to functions* $\{\tilde{\psi}_{ij}, \tilde{\psi}_i\}$ *by*

$$\tilde{\psi}_{ij}(x_i, x_j) \propto \frac{\psi_{ij}(x_i, x_j)}{\pi_{i \to j}(x_j)\pi_{j \to i}(x_i)}, \tag{D.2}$$

$$\tilde{\psi}_i(x_i) \propto \psi_i(x_i) \prod_{k \in N_i} \pi_{k \to i}(x_i). \tag{D.3}$$

Then the update

$$\mu_{i \to j}^{t+1}(x_j) \propto \sum_{x_i} \psi_{ji}(x_j, x_i)\psi_i(x_i) \prod_{k \in N_i \setminus j} \mu_{k \to i}^t(x_i), \tag{D.4}$$

is equivalent to

$$\tilde{\mu}_{i \to j}^{t+1}(x_j) \propto \sum_{x_i} \tilde{\psi}_{ji}(x_j, x_i)\tilde{\psi}_i(x_i) \prod_{k \in N_i \setminus j} \tilde{\mu}_{k \to i}^t(x_i). \tag{D.5}$$

Proof. The equivalence of (D.4) and (D.5) is easily checked by (D.1), (D.2), and (D.3). $\qquad\square$

5

Symbolically, proposition D.1 implies that

$$\Pi \circ T \circ \Pi^{-1} = \tilde{T}, \tag{D.6}$$

where Π is the transformation of the messages by $\pi_{i \to j}$, T is the LBP update with $\{\psi_{ij}, \psi_i\}$, and $\tilde{T}$ is the LBP update with $\{\tilde{\psi}_{ij}, \tilde{\psi}_i\}$. Differentiation of this relation gives the transformation of the linearization matrix.

If we choose $\{\pi_{i \to j}\}$ as $\pi_{i \to j}(x_j) = \mu_{i \to j}^{\infty}(x_j)$, then (D.1), (D.2) and (D.3) becomes

$$\tilde{\mu}_{i \to j}^{t}(x_j) \propto \frac{\mu_{i \to j}^{t}(x_j)}{\mu_{i \to j}^{\infty}(x_j)} \tag{D.7}$$

$$\tilde{\psi}_{ij}(x_i, x_j) \propto \frac{b_{ij}(x_i, x_j)}{b_i(x_i) b_j(x_j)} \tag{D.8}$$

$$\tilde{\psi}_i(x_i) \propto b_i(x_i). \tag{D.9}$$

This is the transformation used in the paper.

D.2   Proof of theorem 5

Theorem 5 ([S5], Proposition 4.5). *At a LBP fixed point $\boldsymbol{\eta}^{\infty}$, the linearization $T'(\boldsymbol{\eta}^{\infty})$ is similar to $\mathcal{UM}$, i.e. $\mathcal{UM} = PT'(\boldsymbol{\eta}^{\infty})P^{-1}$ with an invertible matrix P.*

Proof. Let $\{\mu_{i \to j}^{\infty}(x_j)\}$ be the set of messages at the fixed point and let Π be the transformation of messages defined by the fixed point messages. We parameterize the messages by $\eta_{i \to j}^{t} = \mu_{i \to j}^{t}(+)/\mu_{i \to j}^{t}(-)$. It is enough to prove the assertion after the transformation and in this parameterization, because these operations cause similar linearization matrices.

After the transformation, the LBP update is given in terms of $\tilde{\eta}$ as follows:

$$\tilde{\eta}_{i \to j}^{t+1} = \frac{\sum_{x_i} \tilde{\psi}_{ji}(+, x_i) \tilde{\psi}_i(x_i) \prod_{k \in N_i \backslash j} \tilde{\mu}_{k \to i}(x_i)}{\sum_{x_i} \tilde{\psi}_{ji}(-, x_i) \tilde{\psi}_i(x_i) \prod_{k \in N_i \backslash j} \tilde{\mu}_{k \to i}(x_i)}$$

$$= \frac{\frac{b_{ji}(+,+)}{b_j(+)} \prod_{k \in N_i \backslash j} \tilde{\eta}_{k \to i}(x_i) + \frac{b_{ji}(+,-)}{b_j(+)}}{\frac{b_{ji}(-,+)}{b_j(-)} \prod_{k \in N_i \backslash j} \tilde{\eta}_{k \to i}(x_i) + \frac{b_{ji}(-,-)}{b_j(-)}}.$$

Let $\tilde{\boldsymbol{\eta}}^{\infty} := \Pi(\boldsymbol{\eta}^{\infty})$, then $\tilde{\eta}_e^{\infty} = 1$ for all $e \in \vec{E}$. We can compute $\tilde{T}'(\tilde{\boldsymbol{\eta}}^{\infty})$ as follows:

$$\tilde{T}'(\tilde{\boldsymbol{\eta}}^{\infty}) = \frac{\partial \tilde{\eta}_{i \to j}^{t+1}}{\partial \tilde{\eta}_{k \to l}^{t}} \Big|_{\tilde{\boldsymbol{\eta}}^t = 1}$$

$$= \left(\frac{b_{ji}(+,+)}{b_j(+)} - \frac{b_{ji}(-,+)}{b_j(-)} \right) \mathcal{M}_{i \to j, k \to l}$$

$$= \frac{\chi_{ij} - m_i m_j}{1 - m_j^2} \mathcal{M}_{i \to j, k \to l}.$$

$\square$

E   Idea and proof of theorem 7

E.1   Idea of theorem 7

In theorem 7, we show that the sum of indexes is equal to one. This is not so special. The simplest example that illustrate the idea of the theorem is sketched in figure E.1. For each stationary point, plus or minus sign is assigned depending on the sign of the second derivative. When we deform the function, the sum is still equal to one as long as the outward gradients are positive at the boundaries. (See figure E.2.)

Lemma 1, combined with lemma 2, describes the behavior of the Bethe free energy near the boundary of $L(G)$.

6
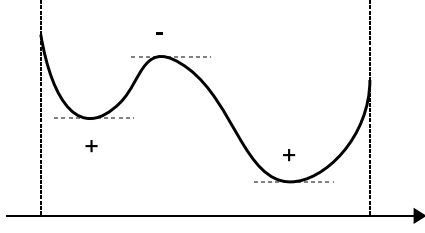
Figure E.1: The sum of indexes is one.
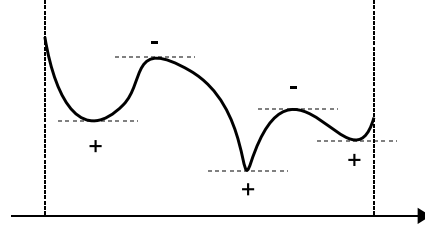
Figure E.2: The sum of indexes is still one.

E.2 Proof of lemma 1

Lemma 1. *If a sequence $\{q_n\} \subset L(G)$ converges to a point $q_* \in \partial L(G)$, then $\|\nabla F(q_n)\| \to \infty$, where $\partial L(G)$ is the boundary of $L(G) \subset \mathbb{R}^{N+M}$.*

Proof. First, note that it is enough to prove the assertion when $h_i = 0$ and $J_{ij} = 0$.

We prove by contradiction. Assume that $\|\nabla F(q_n)\| \not\to \infty$. Then, there exists $R > 0$ such that $\|\nabla F(q_n)\| \leq R$ for infinitely many n. Let $B_0(R)$ be the closed ball of radius R centered at the origin. Taking subsequences, if necessary, we can assume that

$$\nabla F(q_n) \to^{\exists} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \in B_0(R), \tag{E.1}$$

because of the compactness of $B_0(R)$. Let $b_{ij}^{(n)}(x_i, x_j)$ and $b_i^{(n)}(x_i)$ be the pseudomarginals corresponding to q_n. Since $q_n \to q_* \in \partial L(G)$, there exist $ij \in E$, x_i and x_j such that

$$b_{ij}^{(n)}(x_i, x_j) \to 0.$$

Without loss of generality, we assume that $x_i = +1$ and $x_j = +1$. From (E.1), we have

$$\nabla F(q_n)_{ij} = \frac{1}{4} \log \frac{b_{ij}^{(n)}(+,+)b_{ij}^{(n)}(-,-)}{b_{ij}^{(n)}(+,-)b_{ij}^{(n)}(-,+)} \longrightarrow \eta_{ij}.$$

Therefore $b_{ij}^{(n)}(+,-) \to 0$ or $b_{ij}^{(n)}(-,+) \to 0$ holds; we assume $b_{ij}^{(n)}(+,-) \to 0$ without loss of generality. Now we have

$$b_i^{(n)}(+) = b_{ij}^{(n)}(+,-) + b_{ij}^{(n)}(+,+) \to 0.$$

In this situation, the following claim holds.

Claim. *Let $k \in N_i$. In the limit of $n \to \infty$,*

$$\sum_{x_i, x_k = \pm 1} x_i \log \frac{b_{ik}^{(n)}(x_i, x_k)}{b_i^{(n)}(x_i)} = \log \left[\frac{b_{ik}^{(n)}(+,+)b_{ik}^{(n)}(+,-)b_i^{(n)}(-)^2}{b_{ik}^{(n)}(-,+)b_{ik}^{(n)}(-,-)b_i^{(n)}(+)^2} \right] \tag{E.2}$$

converges to a finite value.

proof of claim. From $b_i^{(n)}(+) \to 0$, we have

$$b_{ik}^{(n)}(+,-), b_{ik}^{(n)}(+,+) \longrightarrow 0 \quad \text{and} \quad b_i^{(n)}(-) \to 1.$$

Case 1: $b_{ik}^{(n)}(-,+) \longrightarrow b_{ik}^*(-,+) \neq 0$ and $b_{ik}^{(n)}(-,-) \longrightarrow b_{ik}^*(-,-) \neq 0$.
In the same way as (E.2),

$$\nabla F(q_n)_{ik} = \frac{1}{4} \log \frac{b_{ik}^{(n)}(+,+)b_{ik}^{(n)}(-,-)}{b_{ik}^{(n)}(+,-)b_{ik}^{(n)}(-,+)} \longrightarrow \eta_{ik}.$$

7

Therefore

$$\frac{b_{ik}^{(n)}(+,+)}{b_{ik}^{(n)}(+,-)} \longrightarrow {}^{\exists}r \neq 0.$$

Then we see that (E.2) converges to a finite value.

Case 2: $b_{ik}^{(n)}(-,+) \longrightarrow 1$ and $b_{ik}^{(n)}(-,-) \longrightarrow 0$.
Similar to the case 1, we have

$$\frac{b_{ik}^{(n)}(+,+)b_{ik}^{(n)}(-,-)}{b_{ik}^{(n)}(+,-)} \longrightarrow {}^{\exists}r \neq 0.$$

Therefore $\frac{b_{ik}^{(n)}(+,-)}{b_{ik}^{(n)}(+,+)} \to 0$. This implies that $\frac{b_i^{(n)}(+)}{b_{ik}^{(n)}(+,+)} \to 1$. Then we see that (E.2) converges to a finite value.

Case 3: $b_{ik}^{(n)}(-,+) \longrightarrow 0$ and $b_{ik}^{(n)}(-,-) \longrightarrow 1$.
Same as the case 2. $\qquad\square$

Now let us get back to the proof of lemma 1. We rewrite (B.1) as

$$\nabla F(q_n)_i = \frac{1}{2}\log b_i^{(n)}(+) - \frac{1}{2}\log b_i^{(n)}(-) + \frac{1}{4}\sum_{k \in N_i}\sum_{x_i, x_k = \pm 1} x_i \log \frac{b_{ik}^{(n)}(x_i, x_k)}{b_i^{(n)}(x_i)}$$

From (E.1), this value converges to ξ_i. The second and third terms in (E.2) converges to a finite value, while the first value converges to infinite. This is a contradiction. $\qquad\square$

E.3 Detailed proof of theorem 7

Theorem 7. *If* $\det \nabla^2 F(q) \neq 0$ *for all* $q \in (\nabla F)^{-1}(0)$ *then*

$$\sum_{q: \nabla F(q) = 0} \operatorname{sgn}\left(\det \nabla^2 F(q)\right) = 1, \quad \text{where } \operatorname{sgn}(x) := \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x < 0. \end{cases} \tag{E.3}$$

We call each summand,which is $+1$ *or* -1, *index of* F *at* q.

Proof. Define a map $\Phi : L(G) \to \mathbb{R}^{N+M}$ by

$$\Phi(q)_i = (1 - d_i)\frac{1}{2}\sum_{x_i = \pm 1} x_i \log b_i(x_i) + \frac{1}{4}\sum_{k \in N_i}\sum_{x_i, x_k = \pm 1} x_i \log b_{ik}(x_i, x_k), \tag{E.4}$$

$$\Phi(q)_{ij} = \frac{1}{4}\sum_{x_i, x_j = \pm 1} x_i x_j \log b_{ij}(x_i, x_j), \tag{E.5}$$

where $b_{ij}(x_i, x_j)$ and $b_i(x_i)$ are given by $q = \{m_i, \chi_{ij}\} \in L(G)$. Therefore, we have $\nabla F = \Phi - \binom{h}{J}$ and $\nabla \Phi = \nabla^2 F$. Then following claim holds.

Claim. *The sets* $\Phi^{-1}(\binom{h}{J}), \Phi^{-1}(0) \subset L(G)$ *are finite and*

$$\sum_{q \in \Phi^{-1}\binom{h}{J}} \operatorname{sgn}(\det \nabla \Phi(q)) = \sum_{q \in \Phi^{-1}(0)} \operatorname{sgn}(\det \nabla \Phi(q)), \tag{E.6}$$

holds.

Before the proof of this claim, we prove theorem 7 under the claim.

From (E.4) and (E.5), it is easy to see that $\Phi(q) = 0 \Leftrightarrow q = \{m_i = 0, \chi_{ij} = 0\}$. At this point, we can easily check that $\nabla \Phi = \nabla^2 F$ is a positive definite matrix. Therefore the right hand side of (E.6) is equal to one. The left hand side of (E.6) is equal to the left hand side of (E.3), because $q \in \Phi^{-1}(\binom{h}{J}) \Leftrightarrow \nabla F(q) = 0$. Then the assertion of theorem 7 is proved. $\qquad\square$

8
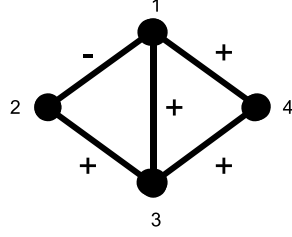
Figure F.1: The graph G.
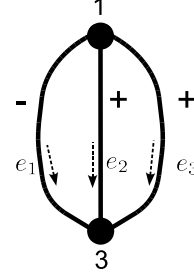
Figure F.2: The graph $\hat{G}$.

Proof of the claim. First, we prove that $\Phi^{-1}\left(\binom{h}{J}\right) = (\nabla F)^{-1}(0)$ is a finite set. If not, we can choose a sequence $\{q_n\}$ of distinct points from this set. Let $\overline{L(G)}$ be the closure of $L(G)$. Since $\overline{L(G)}$ is compact, we can choose a subsequence that converges to some point $q_* \in \overline{L(G)}$. From lemma 1, $q_* \in L(G)$ and $\nabla F(q_*) = 0$ hold. By the assumption in theorem 7, we have $\det \nabla^2 F(q_*) \neq 0$. This implies that $\nabla F(q) \neq 0$ in some neighborhood of q_*. This is a contradiction because $q_n \to q_*$.

Secondly, we prove the equality (E.6) using lemma 2. Define a sequence of compact convex sets $C_n := \left\{ q \in L(G) \mid \sum_{ij \in E} \sum_{x_i, x_j} - \log b_{ij} \leq n \right\}$, which increasingly converges to $L(G)$. Since $\Phi^{-1}(0)$ and $\Phi^{-1}\binom{h}{J}$ are finite, they are included in C_n for sufficiently large n. Take $K > 0$ and $\epsilon > 0$ to satisfy $K - \epsilon > \left\| \binom{h}{J} \right\|$. From lemma 1, we see that $\Phi(\partial C_n) \cap B_0(K) = \phi$ for sufficiently large n. Let n_o be such a large number. Let $\Pi_\epsilon : \mathbb{R}^{N+M} \to B_0(K)$ be a smooth map that is identity on $B_0(K - \epsilon)$, monotonically increasing on $\|x\|$, and $\Pi_\epsilon(x) = \frac{K}{\|x\|} x$ for $\|x\| \geq K$. Then we obtain a composition map $\tilde{\Phi} := \Pi_\epsilon \circ \Phi : C_{n_0} \to B_0(K)$ that satisfy $\tilde{\Phi}(\partial C_{n_0}) \subset \partial B_0(K)$. By definition, we have $\Phi^{-1}(0) = \tilde{\Phi}^{-1}(0)$ and $\Phi^{-1}\binom{h}{J} = \tilde{\Phi}^{-1}\binom{h}{J}$. Therefore, both 0 and $\binom{h}{J}$ are regular values of $\tilde{\Phi}$. From lemma 2, we have

$$\sum_{q \in \tilde{\Phi}^{-1}\binom{h}{J}} \mathrm{sgn}(\det \nabla \tilde{\Phi}(q)) = \sum_{q \in \tilde{\Phi}^{-1}(0)} \mathrm{sgn}(\det \nabla \tilde{\Phi}(q)).$$

Then, the assertion of the claim is proved. $\qquad\square$

F Proof of corollary 4

F.1 Detailed proof of example 2

In this subsection we prove the assertion of corollary 4 for the graph of example 2, which is displayed in figure F.1. The $+$ and $-$ signs represent that of two body interactions.

It is enough to check that $\det(I - \mathcal{B}\mathcal{M}) > 0$ for arbitrary $0 \leq \beta_{13}, \beta_{23}, \beta_{14}, \beta_{34} < 1$ and $-1 < \beta_{12} \leq 0$. The graph $\hat{G}$ in figure F.2 is obtained by erasing vertices 2 and 4 in G. To compute $\det(I - \mathcal{B}\mathcal{M})$, it is enough to consider $\hat{G}$. In fact

$$\det(I - \mathcal{B}\mathcal{M}) = \zeta_G(\boldsymbol{\beta})^{-1}$$
$$= \prod_{\mathfrak{p} \in P} (1 - g(\mathfrak{p})) \tag{F.1}$$
$$= \prod_{\hat{\mathfrak{p}} \in \hat{P}} (1 - g(\hat{\mathfrak{p}})) \tag{F.2}$$
$$= \zeta_{\hat{G}}(\hat{\boldsymbol{\beta}})^{-1} = \det(I - \hat{\mathcal{B}}\hat{\mathcal{M}}),$$

where $\hat{\beta}_{e_1} := \beta_{12}\beta_{23}$, $\hat{\beta}_{e_2} := \beta_{13}$, $\hat{\beta}_{e_3} := \beta_{14}\beta_{34}$ and $\hat{\beta}_{e_i} = \hat{\beta}_{\bar{e}_i}$. The equality between (F.1) and (F.2) is obtained by the one to one correspondence between prime cycles of G and $\hat{G}$.

9
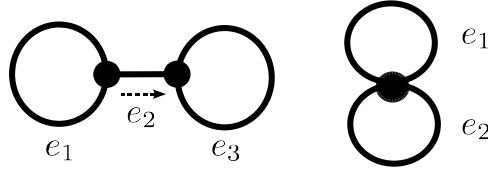
Figure F.3: Two other types of graphs.
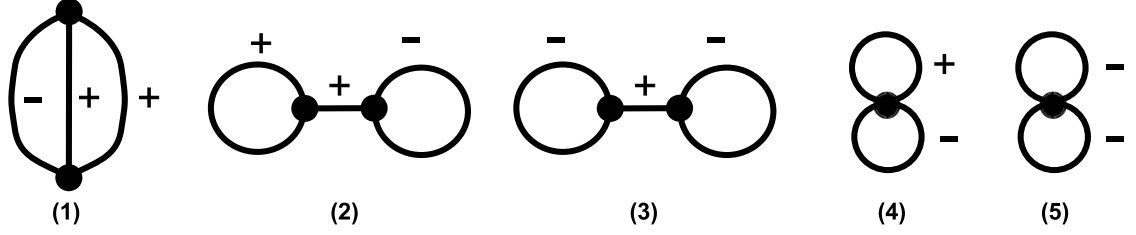
Figure F.4: List of interaction types.

By definition, we have

$$\hat{\mathcal{B}}\hat{\mathcal{M}} = \begin{bmatrix} 0 & 0 & 0 & 0 & \hat{\beta}_{e_1} & \hat{\beta}_{e_1} \\ 0 & 0 & 0 & \hat{\beta}_{e_2} & 0 & \hat{\beta}_{e_2} \\ 0 & 0 & 0 & \hat{\beta}_{e_3} & \hat{\beta}_{e_3} & 0 \\ 0 & \hat{\beta}_{e_1} & \hat{\beta}_{e_1} & 0 & 0 & 0 \\ \hat{\beta}_{e_2} & 0 & \hat{\beta}_{e_2} & 0 & 0 & 0 \\ \hat{\beta}_{e_3} & \hat{\beta}_{e_3} & 0 & 0 & 0 & 0 \end{bmatrix},$$

where the rows and columns are indexed by $e_1, e_2, e_3, \bar{e}_1, \bar{e}_2$ and $\bar{e}_3$. Then the determinant is

$$\det(I - \hat{\mathcal{B}}\hat{\mathcal{M}}) = \det\left[I - \begin{pmatrix} 0 & \hat{\beta}_{e_1} & \hat{\beta}_{e_1} \\ \hat{\beta}_{e_2} & 0 & \hat{\beta}_{e_2} \\ \hat{\beta}_{e_3} & \hat{\beta}_{e_3} & 0 \end{pmatrix}\right] \det\left[I + \begin{pmatrix} 0 & \hat{\beta}_{e_1} & \hat{\beta}_{e_1} \\ \hat{\beta}_{e_2} & 0 & \hat{\beta}_{e_2} \\ \hat{\beta}_{e_3} & \hat{\beta}_{e_3} & 0 \end{pmatrix}\right]$$
$$= (1 - \hat{\beta}_{e_1}\hat{\beta}_{e_2} - \hat{\beta}_{e_1}\hat{\beta}_{e_3} - \hat{\beta}_{e_2}\hat{\beta}_{e_3} - 2\hat{\beta}_{e_1}\hat{\beta}_{e_2}\hat{\beta}_{e_3})$$
$$(1 - \hat{\beta}_{e_1}\hat{\beta}_{e_2} - \hat{\beta}_{e_1}\hat{\beta}_{e_3} - \hat{\beta}_{e_2}\hat{\beta}_{e_3} + 2\hat{\beta}_{e_1}\hat{\beta}_{e_2}\hat{\beta}_{e_3}).$$

Since $-1 < \hat{\beta}_{e_1} \leq 0$ and $0 \leq \hat{\beta}_{e_2}, \hat{\beta}_{e_3} < 1$, we conclude that this is positive.

F.2 Other cases

There are two operations on graphs that do not change the set of prime cycles. The first one is adding or removing a vertex of degree two on any edge. The second one is adding or removing an edge with a vertex of degree one. With these two operations, all graphs that have two linearly independent cycles are reduced to three types of graphs. The first type is in figure F.2. The other types are in figure F.3.

Up to equivalence of interactions, all types of signs of two body interactions are listed in figure F.4 except for the attractive case. We check the uniqueness for each case in order.

Case (1): Proved in example 2.

10

Case (2): In this case,

$$
\mathcal{BM} = \begin{bmatrix}
\beta_{e_1} & 0 & 0 & 0 & \beta_{e_1} & 0 \\
\beta_{e_2} & 0 & 0 & \beta_{e_2} & 0 & 0 \\
0 & \beta_{e_3} & \beta_{e_3} & 0 & 0 & 0 \\
0 & 0 & 0 & \beta_{e_1} & \beta_{e_1} & 0 \\
0 & 0 & \beta_{e_2} & 0 & 0 & \beta_{e_2} \\
0 & \beta_{e_3} & 0 & 0 & 0 & \beta_{e_3}
\end{bmatrix},
$$

where rows and columns are labeled by $e_1, e_2, e_3, \bar{e}_1, \bar{e}_2$ and $\bar{e}_3$. Then the determinant is

$$
\det(I - \mathcal{BM}) = (1 - \beta_{e_1})(1 - \beta_{e_3})(1 - \beta_{e_1} - \beta_{e_3} + \beta_{e_1}\beta_{e_3} - 4\beta_{e_1}\beta_{e_2}^2\beta_{e_3}). \tag{F.3}
$$

This is positive when $0 \le \beta_{e_1}, \beta_{e_2} < 1$ and $-1 < \beta_{e_3} \le 0$.

Case (3): The determinant (F.3) is also positive when $0 \le \beta_{e_2} < 1$ and $-1 < \beta_{e_1}, \beta_{e_3} \le 0$.

Case (4): In this case,

$$
\mathcal{BM} = \begin{bmatrix}
\beta_{e_1} & \beta_{e_1} & 0 & \beta_{e_1} \\
\beta_{e_2} & \beta_{e_2} & \beta_{e_2} & 0 \\
0 & \beta_{e_1} & \beta_{e_1} & \beta_{e_1} \\
\beta_{e_2} & 0 & \beta_{e_2} & \beta_{e_2}
\end{bmatrix},
$$

where rows and columns are labeled by $e_1, e_2, \bar{e}_1$ and $\bar{e}_2$. Then we have

$$
\det(I - \mathcal{BM}) = (1 - \beta_{e_1})(1 - \beta_{e_2})(1 - \beta_{e_1} - \beta_{e_2} - 3\beta_{e_1}\beta_{e_2}). \tag{F.4}
$$

This is positive when $0 \le \beta_{e_1} < 1$ and $-1 < \beta_{e_2} \le 0$.

Case (5): The determinant (F.4) is positive when $-1 < \beta_{e_1}, \beta_{e_2} \le 0$.

References

[S1] C.D. Godsil and G. Royle. *Algebraic Graph Theory*. Springer, 2001.

[S2] K. Hashimoto. On zeta and L-functions of finite graphs. *Internat. J. Math*, 1(4):381–396, 1990.

[S3] M. Kotani and T. Sunada. Zeta functions of finite graphs. *J. Math. Sci. Univ. Tokyo*, 7(1):7–25, 2000.

[S4] S. Northshield. A note on the zeta function of a graph. *Journal of Combinatorial Theory, Series B*, 74(2):408–410, 1998.

[S5] C. Furtlehner, J.M. Lasgouttes, and A. De La Fortelle. Belief propagation and Bethe approximation for traffic prediction. *INRIA RR-6144, Arxiv preprint physics/0703159*, 2007.

11