J. Appl. Math. & Informatics Vol. **40***(2022), No. 3 - 4, pp. 515 - 530*
https://doi.org/10.14317/jami.2022.515

CONVERGENCE OF A GENERALIZED BELIEF PROPAGATION ALGORITHM FOR BIOLOGICAL NETWORKS†

SANG-MOK CHOO, YOUNG-HEE KIM*

Abstract. A factor graph and belief propagation can be used for finding stochastic values of link weights in biological networks. However it is not easy to follow the process of use and so we presented the process with a toy network of three nodes in our prior work. We extend this work more generally and present numerical example for a network of 100 nodes..

AMS Mathematics Subject Classification : 65F08, 68Q07.
Key words and phrases : Factor graph, belief propagation, convergence, biological networks, probability distribution, Banach's fixed point theorem.

1. Introduction

It is imporant to understand the dynamics of interactions of genes or proteins in biological systems ([1]). The dynamics can be described by mathematical models such as differential equation model and Boolean model ([2], [3]). The models have played an important role in this area. However it is not easy to determine parameters in the models ([4]).

The authors of the papers [5] and [6] provided a framework to find parameters in differential equation model by using experimental data, where the parameters are the weights of links in a prior knowledge network(PKN) described by system of ordinary differential equations. They consider each weight a discrete random variable and find its probability mass function(PMF) by using a 'belief propagation(BP) algorithm' on a factor graph ([7], [8], [9], [10], [11], [12]).

Given PKN and experimental data, a stochastic network can be obtained by application of this algorithm. However it is difficult to follow each step in the algorithm and the convergence of iterative schemes in the algorithm was

Received November 25, 2021. Revised March 7, 2022. Accepted March 21, 2022.
*Corresponding author.
†This work was conducted during the sabbatical leave with support from Kwangwoon University in 2021.

not shown. So, we explained the steps with a toy network of three nodes and presented a sufficient condition for the convergence in [13].

In this paper, we extend our results in [13] for networks without restriction on the number of nodes and present a new sufficient condition for the convergence of the general network based on a Banach fixed-point theorem ([14]). Numerical examples are given to illustrate the convergence and application of a network of 100 nodes.

2. Preliminaries

We consider a network which has measured nodes x_i $(1 \le i \le M)$ and drug nodes x_j $(M+1 \le j \le N)$, where x_j has outgoing link $w_{j,i}$ to x_i and no incoming link to any node. Each measured node has an outgoing link to each other measured node and incoming links from each other nodes. Two treatments are assumed to be given to the network, which are called the 1^{st} and 2^{nd} perturbations. Symbols x_i^v $(1 \le i \le N,\, v = 1, 2)$ denote $\log_2\left(x_i^{v,\text{after}}/x_i^{v,\text{before}}\right)$, where $x_i^{v,\text{before}}$ and $x_i^{v,\text{after}}$ are the concentrations of x_i at steady state before and after the v^{th} perturbation, respectively. The dynamics of the given situation is modeled as in [13] and steady state value $x_i^{v,s}$ of $x_{i,v}$ becomes

$$x_i^{v,s} = \begin{cases} \phi\left(\sum_{j=1, j\neq i}^{N} w_{i,j} x_j^v\right) & (1 \le i \le M) \\ x_i^v & (M+1 \le i \le N) \end{cases}, \tag{1}$$

where $\phi(x) = \tanh(x)$ and $w_{i,j}$ is a discrete random variable with PMF

$$P(w_{i,j} = w)\ (1 \le i \le M,\, 1 \le j \le N,\, i \neq j), w \in \{-1, 0, 1\}. \tag{2}$$

To find an approximation of the PMF is our goal using a factor graph and BP.

3. System of equations for marginal PMFs

A large and low cost between simulated and experimental values can be related to low and high probabilities of models, respectively ([5]). So, the joint PMF of all weights W is defined as follows.

$$P(W) = \frac{1}{Z} \exp(-\text{Cost}) \tag{3}$$

and the cost function is defined by

$$\text{Cost} = \beta \sum_{i=1}^{N} \sum_{v=1}^{2} (x_i^{v,s} - x_i^v)^2 + \lambda \sum_{i=1}^{M} \sum_{j=1, j\neq i}^{N} \delta(w_{i,j}), \tag{4}$$

where Z, β, λ are the constants, $\delta(w_{i,j})$ is a penalty function such that $\delta(w_{i,j} = 0) = 0$ and $\delta(w_{i,j} = \pm 1) = 1$. Substituting (1) into (4) gives

$$\text{Cost} = \beta \sum_{i=1}^{M} \sum_{v=1}^{2} \left\{ \phi\Bigg(\sum_{j=1, j\neq i}^{N} w_{i,j} x_j^v \Bigg) - x_i^v \right\}^2 + \lambda \sum_{i=1}^{M} \sum_{j=1, j\neq i}^{N} \delta(w_{i,j}). \tag{5}$$

Substituting (5) into (3) gives a factorization of $P(W)$ with probabilities:

$$\begin{aligned} P(W) &= \frac{1}{Z}\exp\left[\sum_{i=1}^{M}\left(\begin{array}{l} -\beta\sum\limits_{\upsilon=1}^{2}\left\{\phi\left(\sum\limits_{j=1,j\neq i}^{N} w_{i,j}x_j^{\upsilon}\right)-x_i^{\upsilon}\right\}^2 \\ -\lambda\sum\limits_{j=1,j\neq i}^{N}\delta\left(w_{i,j}\right)\end{array}\right)\right] \\ &= \prod_{i=1}^{M}\frac{1}{Z_i}e^{-\lambda\sum\limits_{j=1,j\neq i}^{N}\delta(w_{i,j})}\prod_{\upsilon=1}^{2}\exp\left[-\beta\left\{\phi\left(\sum_{j=1,j\neq i}^{N} w_{i,j}x_j^{\upsilon}\right)-x_i^{\upsilon}\right\}^2\right] \\ &\equiv \prod_{i=1}^{M}P\left(W_{i*}\right), \end{aligned}$$

where W_{i^*} denotes weight incoming to x_i and $Z=\prod_{i=1}^{M}Z_i$. Since each PMF in (2) can be calculated as the marginal PMF of $P\left(W_{i^*}\right)$, we calculate PMF $P\left(w_{1,2}\right)$ instead of $P\left(w_{i,j}\right)$:

$$\begin{aligned} P\left(w_{1,2}\right) &= \sum_{\{W_{1^*}\}-\{w_{1,2}\}} P\left(W_{1^*}\right) \equiv \sum\nolimits_2 P\left(W_{1^*}\right) \\ &= \frac{1}{Z_1}\sum\nolimits_2 e^{-\lambda\sum\limits_{j=2}^{N}\delta(w_{1,j})}\prod_{\upsilon=1}^{2}e^{-\beta\left\{\phi\left(\sum\limits_{j=2}^{N}w_{1,j}x_j^{\upsilon}\right)-x_1^{\upsilon}\right\}^2}. \end{aligned} \tag{6}$$

It is not efficient to calculate the exact marginal in (6) for large N. So a factor graph and BP are used for inferring approximate the marginal. From now on, we explain the complicate multi-step process in [5] in the following three steps.

Step 1. Introduction of a factor graph and BP.

Using the factorization in (6), the factor nodes F_1^{υ} $(\upsilon=1,2)$ are defined as

$$F_1^{\upsilon}\left(W_{1^*}\right)=\exp\left[-\beta\left\{\phi\left(\sum_{j=2}^{N}w_{1,j}x_j^{\upsilon}\right)-x_1^{\upsilon}\right\}^2\right] \tag{7}$$

and then (6) becomes

$$P\left(w_{1,2}\right)=\frac{1}{Z_1}\sum\nolimits_2 e^{-\lambda\sum\limits_{j=2}^{N}\delta(w_{1,j})}\prod_{\upsilon=1}^{2}F_1^{\upsilon}\left(W_{1^*}\right),$$

which gives the factor graph of $N-1$ variable nodes $\left(w_{1,2},\ldots,w_{1,N}\right)$ and two factor nodes (F_1^1, F_1^2). Following BP on the factor graph, the message $P^{\upsilon}(w_{1,2})$ from the variable node $w_{1,2}$ to the factor node $F_1^{\upsilon}(W_{1^*})$ is defined as

$$P^{\upsilon}\left(w_{1,2}\right)=\frac{1}{Z_{1,2}^{\upsilon}}e^{-\lambda\delta(w_{1,2})}\prod_{\mu=1,\mu\neq\upsilon}^{2}\rho^{\mu}\left(w_{1,2}\right), \tag{8}$$

where $Z^{\upsilon}_{1,2}$ is the normalization constant of the probability $P^{\upsilon}(w_{1,2})$ and the message $\rho^{\upsilon}(w_{1,2})$ from $F^{\upsilon}_1(W_{1^*})$ to $w_{1,2}$ is defined as

$$\rho^{\upsilon}(w_{1,2}) = \sum\nolimits_2 \left\{ F^{\upsilon}_1(W_{1^*}) \prod_{k=3}^{N} P^{\upsilon}(w_{1,k}) \right\}, \tag{9}$$

where symbol $\sum_2$ is defined in (6). Using BP, the marginal PMF $P(w_{1,2})$ can be approximated as

$$P(w_{1,2}) = \frac{1}{Z_{1,2}} e^{-\lambda\delta(w_{1,2})} \prod_{\upsilon=1}^{2} \rho^{\upsilon}(w_{1,2}), \tag{10}$$

where $Z_{1,2}$ is the normalization constant of $P(w_{1,2})$. By the definitions (8) and (9), the message $P^{\upsilon}(w_{1,2})$ corresponds to an approximation of $P(w_{1,2})$ depending on the υ^{th} perturbation and $\rho^{\upsilon}(w_{1,2})$ corresponds to a factor of $P^{\upsilon}(w_{1,2})$.

Step 2. Approximation of the summation (9).

The process of the approximation used in [5] can be divided into two parts: the first is to change multiple summations into a single summation with a new random variable and the second is to change the summation into an integral.

Part A. Note that ρ^{υ} in (9) is a function of $w_{1,2}$. Therefore all random variables in $F^{\upsilon}_1(w_{1^*})$ in (9) can be divided into two type of random variables: one is $w_{1,2}$ and the other is

$$s^{\upsilon}_{1,2} = \sum_{\xi \neq 1,2}^{N} w_{1,\xi} x^{\upsilon}_{\xi}, \tag{11}$$

which is a linear combination of random variables $w_{1,\xi}$. Then $F^{\upsilon}_1(W_{1^*})$ in (7) can be written as

$$F^{\upsilon}_1\left(s^{\upsilon}_{1,2}, w_{1,2}\right) = \exp\left[-\beta\left\{\phi\left(s^{\upsilon}_{1,2} + w_{1,2}x^{\upsilon}_2\right) - x^{\upsilon}_1\right\}^2\right], \tag{12}$$

which is a function of random variables $s^{\upsilon}_{1,2}$ and $w_{1,2}$. Substituting (12) into (9) gives

$$\rho^{\upsilon}(w_{1,2}) = \sum\nolimits_2 \left\{ F^{\upsilon}_1\left(s^{\upsilon}_{1,2}, w_{1,2}\right) \prod_{1<\ell\leq N,\, \ell\neq 2} P^{\upsilon}(w_{1,\ell}) \right\}. \tag{13}$$

For some positive integer m, letting

$$\left\{ \sum_{1<\xi\leq N,\, \xi\neq 2} \tilde{w}_{1,\xi} x^{\upsilon}_{\xi} \,\middle|\, \tilde{w}_{1,\xi} \in \{-1,0,1\}, 1<\xi\leq N, \xi\neq 2 \right\} = \left\{ \tilde{s}^{\upsilon}_{1,2,k} \,\middle|\, 1\leq k\leq m \right\},$$

the message ρ^{υ} in (13) becomes

$$\begin{aligned}\rho^{\upsilon}\left(w_{1,2}\right) &= \sum_{k=1}^{m} \sum_{\substack{w_{1,\xi}=\tilde{w}_{1,\xi}\ (2<\xi\leq N) \\ \text{such that} \sum_{2<\xi\leq N} \tilde{w}_{1,\xi} x_{\xi}^{\upsilon}=\tilde{s}_{1,2,k}^{\upsilon}}} \left\{ F_1^{\upsilon}\left(s_{1,j}^{\upsilon}, w_{1,j}\right) \prod_{2<\ell\leq N} P^{\upsilon}\left(w_{1,\ell}\right) \right\} \\ &= \sum_{k=1}^{m} F_1^{\upsilon}\left(\tilde{s}_{1,2,k}^{\upsilon}, w_{1,2}\right) \left\{ \sum_{\substack{w_{1,\xi}=\tilde{w}_{1,\xi}\ (2<\xi\leq N) \\ \text{such that} \sum_{2<\xi\leq N} \tilde{w}_{1,\xi} x_{\xi}^{\upsilon}=\tilde{s}_{1,2,k}^{\upsilon}}} \prod_{2<\ell\leq N} P^{\upsilon}\left(w_{1,\ell}\right) \right\}.\end{aligned} \tag{14}$$

Since

$$\begin{aligned}&\sum_{k=1}^{m} \left\{ \sum_{\substack{w_{1,\xi}=\tilde{w}_{1,\xi}\ (2<\xi\leq N) \\ \text{such that} \sum_{2<\xi\leq N} \tilde{w}_{1,\xi} x_{\xi}^{\upsilon}=\tilde{s}_{1,2,k}^{\upsilon}}} \prod_{2<\ell\leq N} P^{\upsilon}\left(w_{1,\ell}\right) \right\} \\ &\quad = \sum_{\substack{w_{1,\xi} \\ (2<\xi\leq N)}} \prod_{2<\ell\leq N} P^{\upsilon}\left(w_{1,\ell}\right) = \prod_{2<\ell\leq N} \left\{ \sum_{w_{1,\ell}} P^{\upsilon}\left(w_{1,\ell}\right) \right\} = 1,\end{aligned}$$

the following can be a PMF of $s_{1,2}^{\upsilon}$ for $1 \leq k \leq m$

$$P_s^{\upsilon}\left(s_{1,2}^{\upsilon} = \tilde{s}_{1,2,k}^{\upsilon}\right) = \sum_{\substack{w_{1,\xi}=\tilde{w}_{1,\xi}\ (2<\xi\leq N) \\ \text{such that} \sum_{2<\xi\leq N} \tilde{w}_{1,\xi} x_{\xi}^{\upsilon}=\tilde{s}_{1,2,k}^{\upsilon}}} \prod_{2<\ell\leq N} P^{\upsilon}\left(w_{1,\ell}\right). \tag{15}$$

Substituting (15) into (14) gives

$$\rho^{\upsilon}\left(w_{1,2}\right) = \sum_{s_{1,2}^{\upsilon}} F_1^{\upsilon}\left(s_{1,2}^{\upsilon}, w_{1,2}\right) P_s^{\upsilon}\left(s_{1,2}^{\upsilon}\right). \tag{16}$$

Part B. The single summation (16) can be changed into an integral in this part. Note that $s_{1,2}^{\upsilon}$ defined in (11) is a sum of random variables $w_{1,\xi}$ $(2 < \xi \leq N)$, which are assumed to be independent. Even though $w_{1,\xi}$ are not identically distributed, Braunstein et al. [5] invoked the central limit theorem to approximate the PMF of $s_{1,2}^{\upsilon}$ as a Gaussian with reference to [16], where there was no explicit justification for the application of this theorem. Since sums of independent random variables converge in distribution to the standard normal as long as some condition (e.g., the Lindeberg Condition [17]) is satisfied, we think that such a condition might be implicitly assumed in [5]. So the approximate PMF of $s_{1,2}^{\upsilon}$ becomes

$$P_s^{\upsilon}\left(s_{1,2}^{\upsilon}\right) = \frac{1}{\sqrt{2\pi\Delta_{1,2}^{\upsilon}}} \exp\left[-\frac{\left(s_{1,2}^{\upsilon} - \overline{s_{1,2}^{\upsilon}}\right)^2}{2\Delta_{1,2}^{\upsilon}}\right], \tag{17}$$

where $\overline{s^{\upsilon}_{1,2}}$ and $\Delta^{\upsilon}_{1,2}$ are average and variance of $s^{\upsilon}_{1,2}$, respectively:

$$\begin{aligned} \overline{s^{\upsilon}_{1,2}} &= E\left(s^{\upsilon}_{1,2}\right) = E\left(\sum_{\ell\neq 1,2}^{N} w_{1,\ell}x^{\upsilon}_{\ell}\right) = \sum_{\ell\neq 1,2}^{N} E\left(w_{1,\ell}\right)x^{\upsilon}_{\ell} \\ &= \sum_{\ell\neq 1,2}^{N}\left\{\sum_{w} wP^{\upsilon}\left(w_{1,\ell}=w\right)\right\}x^{\upsilon}_{\ell}, \end{aligned} \tag{18}$$

$$\begin{aligned} \Delta^{\upsilon}_{1,2} &= V\left(s^{\upsilon}_{1,2}\right) = V\left(\sum_{\ell\neq 1,2}^{N} w_{1,\ell}x^{\upsilon}_{\ell}\right) = \sum_{\ell\neq 1,2}^{N} V\left(w_{1,\ell}\right)\left(x^{\upsilon}_{\ell}\right)^2 \\ &= \sum_{\ell\neq 1,2}^{N}\left[E\left({w_{1,\ell}}^2\right) - \left\{E\left(w_{1,\ell}\right)\right\}^2\right]\left(x^{\upsilon}_{\ell}\right)^2 \\ &= \sum_{\ell\neq 1,2}^{N}\left[\begin{array}{c} \left\{\sum_{w} w^2P^{\upsilon}\left(w_{1,\ell}=w\right)\right\} \\ -\left\{\left\{\sum_{w} wP^{\upsilon}\left(w_{1,\ell}=w\right)\right\}\right\}^2 \end{array}\right]\left(x^{\upsilon}_{\ell}\right)^2. \end{aligned} \tag{19}$$

Then the sum over configurations $\{w_{1,\ell}\,|\,2<\ell\leq N\}$ in (16) is approximated with a Gaussian integration as follows:

$$\rho^{\upsilon}\left(w_{1,2}\right) \approx \int_{-\infty}^{\infty} F^{\upsilon}_{1}\left(s^{\upsilon}_{1,2}, w_{1,2}\right)P^{\upsilon}_{s}\left(s^{\upsilon}_{1,2}\right)ds^{\upsilon}_{1,2}. \tag{20}$$

Step 3. Approximation of the improper integral (20).

To approximate the improper integral (20), the error $\phi\left(s^{\upsilon}_{1,2} + w_{1,2}x^{\upsilon}_{2}\right) - x^{\upsilon}_{1}$ in (12) is linearized in $s^{\upsilon}_{1,2}$ with respect to the maximization of the fitness in (12). Note that the equality

$$\phi\left(s^{\upsilon}_{1,2} + w_{1,2}x^{\upsilon}_{2}\right) - x^{\upsilon}_{1} = 0$$

can be written as

$$\phi^{-1}(x^{\upsilon}_{1}) - w_{1,2}x^{\upsilon}_{2} - s^{\upsilon}_{1,2} = 0$$

under the assumption that

the experimental data x^{υ}_{1} is contained in the codomain of ϕ. (21)

Then error $\phi\left(s^{\upsilon}_{1,2} + w_{1,2}x^{\upsilon}_{2}\right) - x^{\upsilon}_{1}$ in (12) is approximated by

$$\phi^{-1}(x^{\upsilon}_{1}) - w_{1,2}x^{\upsilon}_{2} - s^{\upsilon}_{1,2},$$

which gives

$$\begin{aligned} F^{\upsilon}_{1}\left(s^{\upsilon}_{1,2}, w_{1,2}\right) &= \exp\left[-\beta\left\{\phi\left(s^{\upsilon}_{1,2} + w_{1,2}x^{\upsilon}_{2}\right) - x^{\upsilon}_{1}\right\}^2\right] \\ &\approx \exp\left[-\beta\left\{\phi^{-1}(x^{\upsilon}_{1}) - w_{1,2}x^{\upsilon}_{2} - s^{\upsilon}_{1,2}\right\}^2\right]. \end{aligned} \tag{22}$$

By (17) and (22), the improper integral (20) becomes

$$\begin{aligned}\rho^{\upsilon}\left(w_{1,2}\right) &\approx \int_{-\infty}^{\infty} F_1^{\upsilon}\left(s_{1,2}^{\upsilon}, w_{1,2}\right) P_s^{\upsilon}\left(s_{1,2}^{\upsilon}\right) ds_{1,2}^{\upsilon} \\ &\approx \int_{-\infty}^{\infty} \left\{ \begin{array}{l} \exp\left[-\beta\left\{\phi^{-1}(x_1^{\upsilon}) - w_{1,2}x_2^{\upsilon} - s_{1,2}^{\upsilon}\right\}^2\right] \\ \cdot\frac{1}{\sqrt{2\pi\Delta_{1,2}^{\upsilon}}} \exp\left[-\frac{\left(s_{1,2}^{\upsilon}-\overline{s_{1,2}^{\upsilon}}\right)^2}{2\Delta_{1,2}^{\upsilon}}\right] \end{array} \right\} ds_{1,2}^{\upsilon} \\ &= \frac{1}{\left(1+2\beta\Delta_{1,2}^{\upsilon}\right)^{1/2}} \exp\left[-\beta\frac{\left\{\phi^{-1}(x_1^{\upsilon}) - w_{1,2}x_2^{\upsilon} - \overline{s_{1,2}^{\upsilon}}\right\}^2}{1+2\beta\Delta_{1,2}^{\upsilon}}\right],\end{aligned}$$

where the last equality is obtained by both the identity

$$\begin{aligned}&-\beta\left\{\phi^{-1}(x_1^{\upsilon}) - w_{1,2}x_2^{\upsilon} - s_{1,2}^{\upsilon}\right\}^2 - \frac{\left(s_{1,2}^{\upsilon}-\overline{s_{1,2}^{\upsilon}}\right)^2}{2\Delta_{1,2}^{\upsilon}} \\ &\quad = -\frac{1}{2}\frac{\left(1+2\beta\Delta_{1,2}^{\upsilon}\right)\left[s_{1,2}^{\upsilon} - \frac{\overline{s_{1,2}^{\upsilon}}+\beta\left\{\phi^{-1}(x_1^{\upsilon})-w_{1,2}x_2^{\upsilon}\right\}2\Delta_{1,2}^{\upsilon}}{1+2\beta\Delta_{1,2}^{\upsilon}}\right]^2}{\Delta_{1,2}^{\upsilon}} \\ &\qquad -\beta\frac{\left\{\phi^{-1}(x_1^{\upsilon}) - w_{1,2}x_2^{\upsilon} - \overline{s_{1,2}^{\upsilon}}\right\}^2}{1+2\beta\Delta_{1,2}^{\upsilon}}\end{aligned}$$

and the property that the integral of a PDF over its domain is equal to 1. Therefore $P(w_{1,2} = w)$ in (10) can be obtained by solving the system of equations

$$\rho^{\upsilon}\left(w_{1,2}=w\right) = \frac{1}{\left(1+2\beta\Delta_{1,2}^{\upsilon}\right)^{1/2}} \exp\left[-\beta\frac{\left\{\phi^{-1}(x_1^{\upsilon}) - wx_2^{\upsilon} - \overline{s_{1,2}^{\upsilon}}\right\}^2}{1+2\beta\Delta_{1,2}^{\upsilon}}\right], \quad (23)$$

$$P^{\upsilon}\left(w_{1,2}=w\right) = \frac{1}{Z_{1,2}^{\upsilon}} e^{-\lambda\delta(w_{1,2}=w)} \rho^{3-\upsilon}\left(w_{1,2}=w\right), \quad (24)$$

where $\overline{s_{1,2}^{\upsilon}}$ and $\Delta_{1,2}^{\upsilon}$ are in (18) and (19) with $\upsilon = 1, 2$ and $w = -1, 0, 1$.

Remark 3.1. Similarly, approximate marginal PMF $P(w_{i,j})$ can be obtained as follows.

$$\begin{aligned}P\left(w_{i,j}\right) &= \frac{1}{Z_{i,j}} e^{-\lambda\delta(w_{i,j})} \prod_{\upsilon=1}^{2} \rho^{\upsilon}\left(w_{i,j}\right), \\ \rho^{\upsilon}\left(w_{i,j}\right) &= \frac{1}{\left(1+2\beta\Delta_{i,j}^{\upsilon}\right)^{1/2}} \exp\left[-\beta\frac{\left\{\phi^{-1}(x_i^{\upsilon}) - w_{i,j}x_j^{\upsilon} - \overline{s_{i,j}^{\upsilon}}\right\}^2}{1+2\beta\Delta_{i,j}^{\upsilon}}\right], \\ P^{\upsilon}\left(w_{i,j}\right) &= \frac{1}{Z_{i,j}^{\upsilon}} e^{-\lambda\delta(w_{i,j})} \rho^{3-\upsilon}\left(w_{i,j}\right),\end{aligned}$$

where $Z^{\upsilon}_{i,j}$ is the normalization constant of probability $P^{\upsilon}(w_{i,j})$ and equations (18), (19) give

$$\begin{aligned}\overline{s^{\upsilon}_{i,j}} &= \sum\nolimits_{\ell\neq i,j}^{N}\left\{\sum\nolimits_{w} wP^{\upsilon}(w_{i,\ell}=w)\right\}x^{\upsilon}_{\ell}\\ &= \sum\nolimits_{\ell\neq i,j}^{N}\frac{e^{-\lambda}}{Z^{\upsilon}_{i,\ell}}\left\{-\rho^{3-\upsilon}(-1)+\rho^{3-\upsilon}(1)\right\}x^{\upsilon}_{\ell}, \qquad (25)\end{aligned}$$

$$\begin{aligned}\Delta^{\upsilon}_{i,j} &= \sum\nolimits_{\ell\neq i,j}^{N}\left\{\left(\sum\nolimits_{w} w^2P^{\upsilon}(w)\right)-\left(\sum\nolimits_{w} wP^{\upsilon}(w)\right)^2\right\}(x^{\upsilon}_{\ell})^2\\ &= \sum\nolimits_{\ell\neq i,j}^{N}\left[\left\{P^{\upsilon}(-1)+P^{\upsilon}(1)\right\}-\left\{-P^{\upsilon}(-1)+P^{\upsilon}(1)\right\}^2\right](x^{\upsilon}_{\ell})^2\\ &= \sum_{\ell\neq i,j}^{N}\frac{e^{-\lambda}}{Z^{\upsilon}_{i,\ell}}\left[\begin{array}{l}\left\{\rho^{3-\upsilon}(-1)+\rho^{3-\upsilon}(1)\right\}\\ -\frac{e^{-\lambda}}{Z^{\upsilon}_{i,\ell}}\left\{-\rho^{3-\upsilon}(-1)+\rho^{3-\upsilon}(1)\right\}^2\end{array}\right](x^{\upsilon}_{\ell})^2. \qquad (26)\end{aligned}$$

4. Iteration method for marginal PMFs

In this section, we present iterative schemes for solving the equations (23), (24) and show a sufficient condition for the convergence of the schemes.

4.1. Iterative schemes for solving the system of equations.

We construct sequences $\{\rho^{\upsilon}_{1,2,n}(w)\}$ and $\{P^{\upsilon}_{1,2,n}(w)\}$ using (23) and (24), which limits satisfy (23) and (24). So the limit of $\{\rho^{\upsilon}_{1,2,n}(w)\}$ becomes value $\rho^{\upsilon}(w_{1,2}=w)$, leading to the construction of approximate $P(w_{1,2})$ in (10). Assume that

$$\text{initial terms } \rho^{\upsilon}_{1,2,0}(w) \text{ are given as positive numbers} \qquad (27)$$

and initial terms of $\{P^{\upsilon}_{1,2,n}(w)\}$ are defined as

$$P^{\upsilon}_{1,2,0}(w)=\frac{1}{Z^{\upsilon}_{1,2,0}}e^{-\lambda\delta(w)}\rho^{3-\upsilon}_{1,2,0}(w), \qquad (28)$$

where $Z^{\upsilon}_{1,2,0}$ is the normalization constant. The first iterations $\rho^{\upsilon}_{1,2,1}(w)$ and $P^{\upsilon}_{1,2,1}(w)$ are defined similarly to (23) and (24). So, equation (18) gives the definition of $\overline{s^{\upsilon}_{1,2,0}}$ as follows.

$$\begin{aligned}\overline{s^{\upsilon}_{1,2,0}} &= \sum\nolimits_{\ell\neq 1,2}^{N}\left\{-P^{\upsilon}_{1,\ell,0}(-1)+P^{\upsilon}_{1,\ell,0}(1)\right\}x^{\upsilon}_{\ell}\\ &= \sum\nolimits_{\ell\neq 1,2}^{N}\frac{e^{-\lambda}}{Z^{\upsilon}_{1,\ell,0}}\left\{-\rho^{3-\upsilon}_{1,\ell,0}(-1)+\rho^{3-\upsilon}_{1,\ell,0}(1)\right\}x^{\upsilon}_{\ell}.\end{aligned}$$

And equation (19) gives the definition of $\Delta^{\upsilon}_{1,2,0}$

$$\Delta^{\upsilon}_{1,2,0} = \sum\nolimits_{\ell\neq 1,2}^{N}\left[\begin{array}{l}\left\{P^{\upsilon}_{1,\ell,0}(-1)+P^{\upsilon}_{1,\ell,0}(1)\right\}\\ -\left\{-P^{\upsilon}_{1,\ell,0}(-1)+P^{\upsilon}_{1,\ell,0}(1)\right\}^2\end{array}\right](x^{\upsilon}_{\ell})^2$$

$$= \sum_{\ell \neq 1,2}^{N} \frac{e^{-\lambda}}{Z^{v}_{1,\ell,0}} \left[\begin{array}{c} \left\{ \rho^{3-v}_{1,\ell,0}(-1) + \rho^{3-v}_{1,\ell,0}(1) \right\} \\ -\frac{e^{-\lambda}}{Z^{v}_{1,\ell,0}} \left\{ -\rho^{3-v}_{1,\ell,0}(-1) + \rho^{3-v}_{1,\ell,0}(1) \right\}^2 \end{array} \right] (x^{v}_{\ell})^2.$$

So the 1^{st} iteration is defined as

$$\rho^{v}_{1,2,1}(w) = \frac{1}{\left(1+2\beta\Delta^{v}_{1,2,0}\right)^{1/2}} \exp\left[-\beta \frac{\left\{ \phi^{-1}(x^{v}_{1}) - w x^{v}_{2} - \overline{s^{v}_{1,2,0}} \right\}^2}{1+2\beta\Delta^{v}_{1,2,0}} \right],$$
$$P^{v}_{1,2,1}(w) = \frac{1}{Z^{v}_{1,2,1}} e^{-\lambda\delta(w)} \rho^{3-v}_{1,2,1}(w),$$

where $Z^{v}_{1,2,1}$ is the normalization constant. Similarly the $(n+1)^{th}$ iteration is defined as

$$\rho^{v}_{1,2,n+1}(w) = \Phi^{v}_{1,2,w}\left(\rho^{3-v}_{1,2^*,n}(-1), \rho^{3-v}_{1,2^*,n}(1) \right), \tag{29}$$
$$P^{v}_{1,2,n+1}(w) = \frac{1}{Z^{v}_{1,2,n+1}} e^{-\lambda\delta(w)} \rho^{3-v}_{1,2,n+1}(w) \ \ (n \geq 0), \tag{30}$$

where $Z^{v}_{1,2,n+1}$ is the normalization constant. Here the function $\Phi^{v}_{1,2,w}$ of $\rho^{3-v}_{1,2^*,n}(-1)$ and $\rho^{3-v}_{1,2^*,n}(1)$ is defined as

$$\Phi^{v}_{1,2,w}\left(\rho^{3-v}_{1,2^*,n}(-1), \rho^{3-v}_{1,2^*,n}(1) \right) = \frac{1}{\left(1+2\beta\Delta^{v}_{1,2,n}\right)^{1/2}}$$
$$\times \exp\left[-\beta \frac{\left\{ \phi^{-1}(x^{v}_{1}) - w x^{v}_{2} - \sum_{\ell \neq 1,2}^{N} \frac{e^{-\lambda}}{Z^{v}_{1,\ell,n}} \left(-\rho^{3-v}_{1,\ell,n}(-1) + \rho^{3-v}_{1,\ell,n}(1) \right) x^{v}_{\ell} \right\}^2}{1+2\beta\Delta^{v}_{1,2,n}} \right] \tag{31}$$

and $\rho^{3-v}_{1,2^*,n}$ denotes $(\rho^{3-v}_{1,3,n}, \cdots, \rho^{3-v}_{1,N,n})$. Therefore the schemes consist of (27)–(31) under the assumption (21).

Remark 4.1. Note that sequence $\{\rho^{v}_{1,2,n}(w)\}$ in the recursive relation (29) contains no $P^{v}_{1,2,n}(w)$. And similarly $\{\rho^{v}_{1,j,n}(w)\}$ and $\{P^{v}_{1,j,n}(w)\}$ are defined. In the next subsection, we present a sufficient condition for the convergence of sequence $\{\rho^{v}_{1,j,n}(w)\}$ without using $\{P^{v}_{1,j,n}(w)\}$ and, as a result, the limit of $\{\rho^{v}_{1,j,n}(w)\}$ is used to define the message $\rho^{v}(w_{1,j} = w)$.

4.2. A sufficient condition for the convergence of the iterative schemes. Replacing subscript $(1,2)$ in (27)–(31) with $(1,j)$ gives the iterative scheme for

$\rho^{v}(w_{1,j}=w)$. Let $\mathbf{X}_1^{(n)}$ be a vector in R^{6N-6} $(N\geq 2)$ defined by

$$\mathbf{X}_1^{(n)} = \begin{pmatrix} X_{1,1}^{(n)}, & X_{1,2}^{(n)}, & X_{1,3}^{(n)}, \cdots, \\ X_{1,3N-5}^{(n)}, X_{1,3N-4}^{(n)}, X_{1,3N-3}^{(n)}, \\ X_{1,3N-2}^{(n)}, X_{1,3N-1}^{(n)}, X_{1,3N}^{(n)}, \cdots, \\ X_{1,6N-8}^{(n)}, X_{1,6N-7}^{(n)}, X_{1,6N-6}^{(n)} \end{pmatrix} = \begin{pmatrix} \rho_{1,2,n}^{1}(-1), \rho_{1,2,n}^{1}(0), \rho_{1,2,n}^{1}(1), \cdots, \\ \rho_{1,N,n}^{1}(-1), \rho_{1,N,n}^{1}(0), \rho_{1,N,n}^{1}(1), \\ \rho_{1,2,n}^{2}(-1), \rho_{1,2,n}^{2}(0), \rho_{1,2,n}^{2}(1), \cdots, \\ \rho_{1,N,n}^{2}(-1), \rho_{1,N,n}^{2}(0), \rho_{1,N,n}^{2}(1) \end{pmatrix}$$

and $\mathbf{\Phi}_1$ be a function from R^{6N-6} to R^{6N-6} defined by

$$\mathbf{\Phi}_1 = \begin{pmatrix} \Phi_{1,1}, & \Phi_{1,2}, & \Phi_{1,3}, \cdots, \\ \Phi_{1,3N-5}, \Phi_{1,3N-4}, \Phi_{1,3N-3}, \\ \Phi_{1,3N-2}, \Phi_{1,3N-1}, \Phi_{1,3N}, \cdots, \\ \Phi_{1,6N-8}, \Phi_{1,6N-7}, \Phi_{1,6N-6} \end{pmatrix} = \begin{pmatrix} \Phi_{1,2,-1}^{1}, \Phi_{1,2,0}^{1}, \Phi_{1,2,1}^{1}, \cdots, \\ \Phi_{1,N,-1}^{1}, \Phi_{1,N,0}^{1}, \Phi_{1,N,1}^{1}, \\ \Phi_{1,2,-1}^{2}, \Phi_{1,2,0}^{2}, \Phi_{1,2,1}^{2}, \cdots, \\ \Phi_{1,N,-1}^{2}, \Phi_{1,N,0}^{2}, \Phi_{1,N,1}^{2} \end{pmatrix},$$

where the subscript 1 of $\mathbf{X}_1$ and $\mathbf{\Phi}_1$ represents node x_1 and the definition of $\Phi_{1,k}$ follows that of equation (31). For example, $\Phi_{1,2,-1}^{1}$ is defined as follows.

$$\Phi_{1,2,-1}^{1}(\mathbf{X}) = \frac{1}{\left(1+2\beta\Delta_{1,2,\mathbf{X}}^{1}\right)^{1/2}} \exp\left[-\beta\frac{\left\{\phi^{-1}(x_1^1) - wx_2^1 - \overline{s_{1,2,\mathbf{X}}^{1}}\right\}^2}{1+2\beta\Delta_{1,2,\mathbf{X}}^{1}}\right],$$

where $\overline{s_{1,2,\mathbf{X}}^{1}}$ and $\Delta_{1,2,\mathbf{X}}^{1}$ are defined by following equtions (25) and (26):

$$\overline{s_{1,2,\mathbf{X}}^{1}} = \frac{e^{-\lambda}}{Z_{1,3,\mathbf{X}}^{1}}(-X_{3N+1}+X_{3N+3})\,x_3^1 + \cdots + \frac{e^{-\lambda}}{Z_{1,N,\mathbf{X}}^{1}}(-X_{6N-8}+X_{6N-6})\,x_N^1,$$

$$\begin{aligned}\Delta_{1,2,\mathbf{X}}^{1} = {}& \frac{e^{-\lambda}}{Z_{1,3,\mathbf{X}}^{1}}\left\{(X_{3N+1}+X_{3N+3}) - \frac{e^{-\lambda}}{Z_{1,3,\mathbf{X}}^{1}}(-X_{3N+1}+X_{3N+3})^2\right\}\left(x_3^1\right)^2 \\ &+ \cdots + \frac{e^{-\lambda}}{Z_{1,N,\mathbf{X}}^{1}}\left\{(X_{6N-8}+X_{6N-6}) - \frac{e^{-\lambda}}{Z_{1,N,\mathbf{X}}^{1}}(-X_{6N-8}+X_{6N-6})^2\right\}\left(x_N^1\right)^2,\end{aligned}$$

where $Z_{1,j,\mathbf{X}}^{1}$ is defined by following the definition of $Z_{1,j,n}^{1}$ such as

$$Z_{1,3,\mathbf{X}}^{1} = e^{-\lambda}X_{3N+1} + X_{3N+2} + e^{-\lambda}X_{3N+3}.$$

The $(n+1)^{th}$ iteration in (29) is written as

$$\mathbf{X}_1^{(n+1)} = \mathbf{\Phi}_1\left(\mathbf{X}_1^{(n)}\right). \tag{32}$$

We use Banach fixed-point theorem [14] for the convergence of the sequence (32) to prove Theorem 4.4, which is our main result.

Theorem 4.1. *Let D be a closed subset of R^m for a positive integer m. If a function $\Psi : D \to D$ satisfies that for a constant $k \in (0,1)$ and all $\mathbf{x}, \mathbf{y}$ in D*

$$\|\Psi(\mathbf{x}) - \Psi(\mathbf{y})\| \le k\|\mathbf{x} - \mathbf{y}\|,$$

then there exists a unique fixed point $\mathbf{x}^* \in D$ such that $\Psi(\mathbf{x}^*) = \mathbf{x}^*$, which is the limit of sequence $\mathbf{x}^{(n+1)} = \Psi\left(\mathbf{x}^{(n)}\right)$ for any $\mathbf{x}^{(0)} \in D$.

Since each function $\Phi_{1,k}$ of $\mathbf{\Phi}_1$ is defined by equation (31), the codomain of $\Phi_{1,k}$ is $[0,1]$, we assume that the following lemma holds.

Lemma 4.2. *Assume that experimental data are contained in the codomain of ϕ. Then there exists a closed bounded domain $\mathcal{D} \subset R^{6N-6}$ such that $\mathbf{\Phi}_1$ in (32) becomes a function from domain $\mathcal{D}$ to codomain $\mathcal{D}$.*

Lemma 4.3. *Assume that experimental data are contained in the codomain of ϕ. Let $\mathcal{D}$ be in Lemma 4.2. Then for $\mathbf{\Phi}_1$ defined in (32) and $\{\mathbf{X}, \mathbf{Y}\} \subset \mathcal{D}$,*

$$\|\mathbf{\Phi}_1(\mathbf{X}) - \mathbf{\Phi}_1(\mathbf{Y})\| \le \beta M_\beta \|\mathbf{X} - \mathbf{Y}\|,$$

where

$$M_\beta = \max_{\mathbf{X} \subset \mathcal{D}, i, j} \left| \frac{\partial \Delta^v_{1,j,\mathbf{X}}}{\partial X_i} \right| + \max_{\mathbf{X} \subset \mathcal{D}, i, j} \left| \frac{\partial}{\partial X_i} \frac{\left\{ \phi^{-1}(x^v_1) - w x^v_j - \overline{s^v_{1,j,\mathbf{X}}} \right\}^2}{1 + 2\beta \Delta^v_{1,j,\mathbf{X}}} \right|$$

Proof. Using (31), we have that for $\{\mathbf{X}, \mathbf{Y}\} \subset \mathcal{D}$

$$\begin{aligned} \Phi_{1,1}(\mathbf{X}) - \Phi_{1,1}(\mathbf{Y}) &= \Phi^1_{1,2,-1}(\mathbf{X}) - \Phi^1_{1,2,-1}(\mathbf{Y}) \\ &= f(\mathbf{X}) \exp\left[g(\mathbf{X})\right] - f(\mathbf{Y}) \exp\left[g(\mathbf{Y})\right], \end{aligned} \tag{33}$$

where functions f and g are defined as follows:

$$\begin{aligned} f(\mathbf{X}) &= \left(1 + 2\beta \Delta^1_{1,2,\mathbf{X}}\right)^{-1/2}, \\ g(\mathbf{X}) &= -\beta \frac{\left\{ \phi^{-1}(x^1_1) - w x^1_2 - \overline{s^1_{1,2,\mathbf{X}}} \right\}^2}{1 + 2\beta \Delta^1_{1,2,\mathbf{X}}}. \end{aligned}$$

Due to the mean value theorem there exists a constant c in (0,1) such that

$$|f(\mathbf{X}) - f(\mathbf{Y})| \le \|\nabla f((1-c)\mathbf{X} + c\mathbf{Y})\| \|\mathbf{X} - \mathbf{Y}\|. \tag{34}$$

Using the following property

$$\left| \frac{\partial f(\mathbf{X})}{\partial X_{3N+1}} \right| = \left| -\frac{1}{2} f(\mathbf{X})^3 2\beta \frac{\partial \Delta^1_{1,2,\mathbf{X}}}{\partial X_{3N+1}} \right| \le \beta \left| \frac{\partial \Delta^1_{1,2,\mathbf{X}}}{\partial X_{3N+1}} \right|,$$

equation (34) becomes

$$|f(\mathbf{X}) - f(\mathbf{Y})| \le \left(\beta \max_{\mathbf{X} \subset \mathcal{D}, i} \left| \frac{\partial \Delta^1_{1,2,\mathbf{X}}}{\partial X_i} \right| \right) \|\mathbf{X} - \mathbf{Y}\|. \tag{35}$$

There exists a constant c in (0,1) such that

$$|\exp\left[g\left(\mathbf{X}\right)\right]-\exp\left[g\left(\mathbf{Y}\right)\right]|\leq\|\nabla\exp[g((1-c)\mathbf{X}+c\mathbf{Y})]\|\|\mathbf{X}-\mathbf{Y}\|. \tag{36}$$

Since $0<\exp[g(\mathbf{X})]\leq 1$, we have

$$\begin{aligned}\left|\frac{\partial}{\partial X_{3N+1}}\exp\left[g\left(\mathbf{X}\right)\right]\right| &\leq \left|\frac{\partial}{\partial X_{3N+1}}g\left(\mathbf{X}\right)\right| \\ &\leq \beta\left|\frac{\partial}{\partial X_{3N+1}}\left(-\beta\frac{\left\{\phi^{-1}(x_1^1)-wx_2^1-\overline{s_{1,2,\mathbf{X}}^1}\right\}^2}{1+2\beta\Delta_{1,2,\mathbf{X}}^1}\right)\right|.\end{aligned}$$

Then (36) becomes

$$|\exp[g(\mathbf{X})]-\exp[g(\mathbf{Y})]|\leq\left(\beta\max_{\mathbf{X}\subset\mathcal{D},i}\left|\frac{\partial}{\partial X_i}\frac{\left\{\phi^{-1}(x_1^1)-wx_2^1-\overline{s_{1,2,\mathbf{X}}^1}\right\}^2}{1+2\beta\Delta_{1,2,\mathbf{X}}^1}\right|\right)\|\mathbf{X}-\mathbf{Y}\|. \tag{37}$$

Substituting (35) and (37) into (33) gives

$$|\Phi_{1,1}\left(\mathbf{X}\right)-\Phi_{1,1}\left(\mathbf{Y}\right)|\leq\beta\left(\begin{array}{c}\max_{\mathbf{X}\subset\mathcal{D},i}\left|\frac{\partial\Delta_{1,2,\mathbf{X}}^1}{\partial X_i}\right| \\ +\max_{\mathbf{X}\subset\mathcal{D},i}\left|\frac{\partial}{\partial X_i}\frac{\left\{\phi^{-1}(x_1^1)-wx_2^1-\overline{s_{1,2,\mathbf{X}}^1}\right\}^2}{1+2\beta\Delta_{1,2,\mathbf{X}}^1}\right|\end{array}\right)\|\mathbf{X}-\mathbf{Y}\|,$$

which gives the desired result. □

Using Theorem 4.1, Lemmas 4.2 and 4.3, we can obtain our main result.

Theorem 4.4. *Assume that the experimental data x_1^υ $(\upsilon=1,2)$ are contained in the codomain of ϕ. Let $\mathcal{D}$ be in Lemma 4.2. Suppose that positive constants β and λ satisfy*

$$\beta\left(\max_{\mathbf{X}\subset\mathcal{D},i,j}\left|\frac{\partial\Delta_{1,j,\mathbf{X}}^\upsilon}{\partial X_i}\right|+\max_{\mathbf{X}\subset\mathcal{D},i,j}\left|\frac{\partial}{\partial X_i}\frac{\left\{\phi^{-1}(x_1^\upsilon)-wx_j^\upsilon-\overline{s_{1,j,\mathbf{X}}^\upsilon}\right\}^2}{1+2\beta\Delta_{1,j,\mathbf{X}}^\upsilon}\right|\right)<1.$$

Then sequence $\mathbf{X}_1^{(n+1)}=\mathbf{\Phi}_1\left(\mathbf{X}_1^{(n)}\right)$ converges for any $\mathbf{X}_1^{(0)}\subset\mathcal{D}$.

Remark 4.2. Using the limit $\rho_{1,j,w}^\upsilon$ of $\{\rho_{1,j,n}^\upsilon\left(w\right)\}$, we can obtain approximate marginal PMFs

$$P\left(w_{1,j}=w\right)=\frac{1}{Z_{1,j}}e^{-\lambda\delta(w_{1,j}=w)}\prod_{\upsilon=1}^{2}\rho_{1,j,w}^\upsilon\ (2\leq j\leq N,\ w=-1,0,1).$$

5. Numerical examples

In order to show the convergence of sequences $\{\rho^{v}_{1,j,n}(w)\}$ for $N = 100$ and $2 \le j \le N$, we randomly generate artificial data for x^{ν}_{i} $(1 \le i \le N,\ \nu = 1,2)$ in the open interval $(-1,1)$, the codomain of $\phi(x) = \tanh(x)$, and so the condition on the experimental data is satisfied.

We set $(\beta,\lambda) = (0.1,1)$ and simulate the following system of equations

$$\begin{aligned}
\rho^{v}_{1,j,n+1}(w) &= \frac{1}{\left(1+2\beta\Delta^{v}_{1,j,n}\right)^{1/2}} \exp\left[-\beta\frac{\left\{\phi^{-1}(x^{v}_{1}) - wx^{v}_{j} - \overline{s^{v}_{1,j,n}}\right\}^{2}}{1+2\beta\Delta^{v}_{1,j,n}}\right],\\
\overline{s^{v}_{1,j,n}} &= \sum\nolimits_{\ell\neq 1,j}^{N} \frac{e^{-\lambda}}{Z^{v}_{1,\ell,n}}\left\{-\rho^{3-v}_{1,\ell,n}(-1) + \rho^{3-v}_{1,\ell,n}(1)\right\} x^{v}_{\ell},\\
\Delta^{v}_{1,j,n} &= \sum_{\ell\neq 1,j}^{N} \frac{e^{-\lambda}}{Z^{v}_{1,\ell,n}}\left[\begin{array}{l}\left\{\rho^{3-v}_{1,\ell,n}(-1) + \rho^{3-v}_{1,\ell,n}(1)\right\}\\ -\frac{e^{-\lambda}}{Z^{v}_{1,\ell,n}}\left\{-\rho^{3-v}_{1,\ell,n}(-1) + \rho^{3-v}_{1,\ell,n}(1)\right\}^{2}\end{array}\right](x^{v}_{\ell})^{2},\\
Z^{v}_{1,\ell,n} &= e^{-\lambda}\rho^{3-v}_{1,\ell,n}(-1) + \rho^{3-v}_{1,\ell,n}(0) + e^{-\lambda}\rho^{3-v}_{1,\ell,n}(1),
\end{aligned}$$

where the initial values $\rho^{v}_{1,j,0}(w)$ are randomly generated in (0,1) and the other initial values of $Z^{v}_{1,j,n}$, $\overline{s^{v}_{1,j,n}}$ and $\Delta^{v}_{1,j,n}$ are defined by replacing $\rho^{v}_{1,j,n}(w)$ in the definition of $Z^{v}_{1,j,n}$, $\overline{s^{v}_{1,j,n}}$ and $\Delta^{v}_{1,j,n}$ with $\rho^{v}_{1,j,0}(w)$. The convergence is measured by using the difference of the consecutive terms in each sequence $\{\rho^{v}_{1,j,n}(w)\}$. Figure 1 shows that $\{\rho^{v}_{1,2,4}(w)\}$ and $\{\rho^{v}_{1,2,5}(w)\}$ are very close, which implies the convergence of $\{\rho^{v}_{1,2,n}(w)\}$.

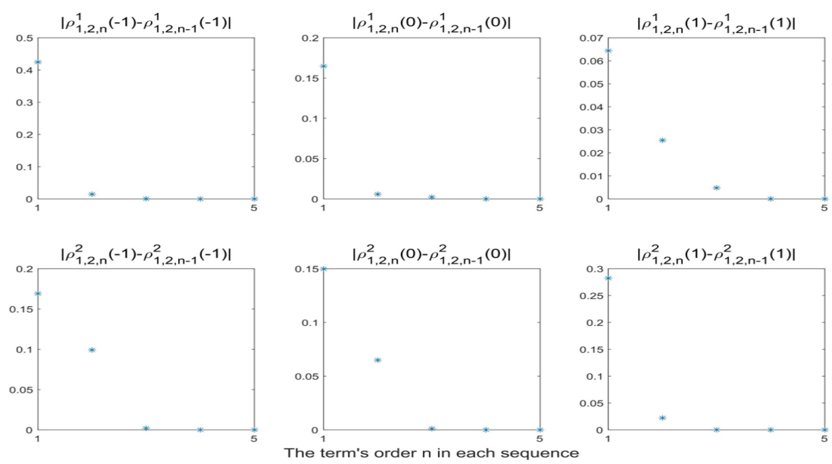

FIGURE 1. Absolute values of $\{\rho^{v}_{1,2,n}(w)\} - \{\rho^{v}_{1,2,n-1}(w)\}$ $(n = 1,2,3,4,5, w = -1,0,1)$ and the last value is 2.8365e-07.

Figure 2 shows that $\{\rho^v_{1,j,10}(w)\}$ and $\{\rho^v_{1,j,9}(w)\}$ $(2 \leq j \leq N)$ are very close, which implies the convergence of $\{\rho^v_{1,j,n}(w)\}$.

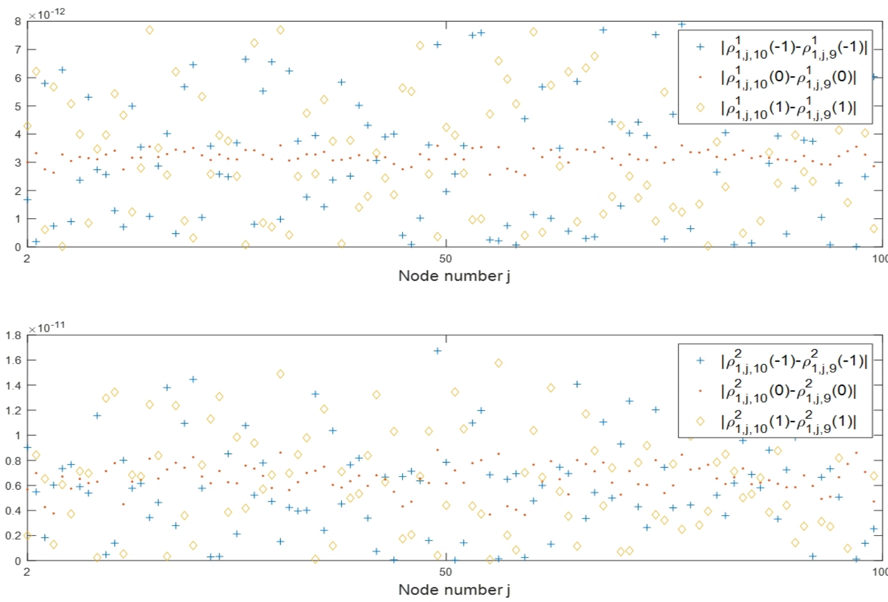

FIGURE 2. Absolute values of $\{\rho^v_{1,j,10}(w)\} - \{\rho^v_{1,j,9}(w)\}$ $(2 \leq j \leq 100, w = -1, 0, 1)$ and the last value is 6.7597e-12.

In order to show the application of PMF $P(w_{1,j} = w)$ we consider node x_j an activation node to x_1 if $P(w_{1,j} = 1) > P(w_{1,j} = -1)$. Similarly, a node x_j an inhibition node to x_i if $P(w_{1,j} = 1) < P(w_{1,j} = -1)$. As in Figure 3, nodes x_2 and x_{99} are activation and inhibition nodes to x_1, respectively, where the height of each line at x_j denotes its probability. Even if the heights are similar, we can select top 10 activation and inhibition nodes to x_1 among 99 nodes $x_j(2 \leq j \leq N)$.

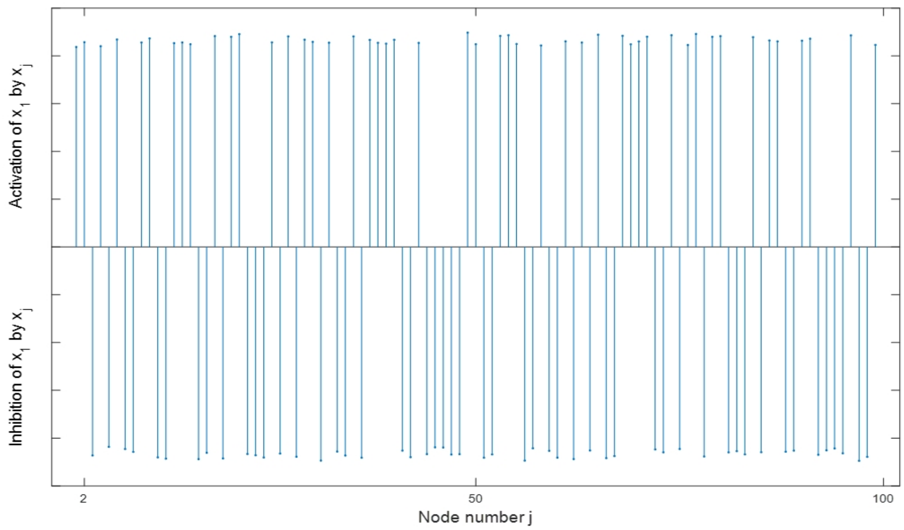

FIGURE 3. Activation and inhibition nodes x_j $(2 \leq j \leq 100)$ to x_1. Node x_2 is an activation node to x_1 and the height of line from x_2 denotes probability $P\left(w_{1,2}=1\right)$. Node x_{99} is an inhibition node to x_1 and the height of line from x_{99} denotes $P\left(w_{1,99}=-1\right)$.

6. Conclusions

In this paper we extend our results in [13] to a network of N nodes. We present the process to define approximate PMFs of link weights in the network based on BP on the factor graph, where the PMFs can be calculated by solving system of equations. However the system cannot be solved analytically. To find the solution of the system we construct sequences of which limits are the solution and find a sufficient condition for the convergence of the sequences. The construction of the sequences is more general than that in our prior work. We show the convergence numerically and an application of the PMFs.

References

1. J.D.J. Hand, *Understanding biological functions through molecular networks*, Cell research **18** (2008), 224-237.
2. W. Materi and D.S. Wishart, *Computational systems biology in cancer: modeling methods and applications*, Gene Regulation and Systems Biology **1** (2007), 91-110.
3. H.F. Fumia and M.L. Martins,, *Boolean network model for cancer pathways: predicting carcinogenesis and targeted therapy outcomes*, PLoS One **8** (2013), e69008.
4. R. Barbuti, R. Gori et al., *A survey of gene regulatory networks modelling methods: from differential equations, to Boolean and qualitative bioinspired models* Journal of Membrane Computing **2** (2020), 1-20.

5. E.J. Molinelli, A. Korkut et al., *Perturbation biology: inferring signaling networks in cellular systems*, PLoS Computational Biology **9** (2013), Article ID e1003290.
6. A. Korkut, W. Wang et al., *Perturbation biology nominates upstream–downstream drug combinations in RAF inhibitor resistant melanoma cells*, Elife **4** (2015), Article ID e04640.
7. F.R. Kschischang, B.J. Frey and H.A. Loeliger, *Factor graphs and the sum-product algorithm*, IEEE Transactions on information theory **47** (2001), 498-519.
8. T. Heskes, *Stable fixed points of loopy belief propagation are local minima of the bethe free energy*, In Advances in neural information processing systems (2003), 359-366.
9. H.A. Loeliger, *An introduction to factor graphs*, IEEE Signal Processing Magazine **21** (2004), 28-41.
10. C.M. Bishop, *Pattern recognition and machine learning*, springer, New York, 2006.
11. H.A. Loeliger, J. Dauwels et al., *The factor graph approach to model-based signal processing*, Proceedings of the IEEE **95** (2007), 1295-1322.
12. K.P. Murphy, *Machine learning: a probabilistic perspective*, MIT press, Cambridge, MA, 2012.
13. S.-M. Choo and Y.-H. Kim, *Convergence of a Belief Propagation Algorithm for Biological Network*, Discrete Dyn. Nat. Soc. **2019** (2019), Art.ID 9362179.
14. M.A. Khamsi and A.K. William, *An introduction to metric spaces and fixed point theory*, J. Wiley and Sons, New York, 2011.
15. A. Braustein, R. Mezard and R. Zecchina, *Survey propagation: An algorithm for satisfiablility*, Random structures and algorithms **27** (2005), 201-226.
16. A. Braunstein et al., *Inference algorithm for gene networks: a statical mechanics analysis*, J. Statistical Mechanics: Theory and Experiment **2008** (2008), Art.ID P12001.
17. D.R. Hunter, *Notes for a Graduate-Level Course in Asymptotics for Statics*, Penn State University, Penn, USA, 2014.
18. J. Pearl, *Probabilistic Reasoning in Intelligent systems: Networks of Plausible Inference*, Morgan Kaufmann, San Francisco, CA, 1988.
19. Q. Su and Y.C. Wu, *On Convergence Conditions of Gaussian Belief Propagation*, IEEE Trans. Signal Processing **63** (2015), 1144-1155.
20. J.S. Yedidia, W.T. Freeman, and Y. Weiss, *Understanding belief propagation and its generalizations*, Exploring artificial intelligence in the new millennium **8** (2003), 236-239.

Sang-Mok Choo received M.Sc. and Ph.D. from Seoul National University. He is currently a professor at University of Ulsan since 2001. His research interests are numerical analysis, system biology and neural networks.

Department of Mathematics, University of Ulsan, Ulsan 44610, Korea.
e-mail: smchoo@ulsan.ac.kr

Young-Hee Kim received M.Sc. and Ph.D. from Yonsei University. Since 2003 she has been at Kwangwoon University. Her research interests are numerical analysis, information system and biological networks.

Ingenium College of Liberal Arts-Mathematics, Kwangwoon University, Seoul 01897, Korea.
e-mail: yhkim@kw.ac.kr