Research Article

Convergence of a Belief Propagation Algorithm for Biological Networks

Sang-Mok Choo [ID][1] **and Young-Hee Kim** [ID][2]

[1]*Department of Mathematics, University of Ulsan, Ulsan 44610, Republic of Korea*
[2]*Ingenium College of Liberal Arts, Kwangwoon University, Seoul 01897, Republic of Korea*

Correspondence should be addressed to Young-Hee Kim; yhkim@kw.ac.kr

Constructing network models of biological systems is important for effective understanding and control of the biological systems. For the construction of biological networks, a stochastic approach for link weights has been recently developed by using experimental data and belief propagation on a factor graph. The link weights were variable nodes of the factor graph and determined from their marginal probability mass functions which were approximated by using an iterative scheme. However, there is no convergence analysis of the iterative scheme. In this paper, at first, we present a detailed explanation of the complicated multistep process step by step with a network of small size and artificial experimental data, and then we show a sufficient condition for the convergence of the iterative scheme. Numerical examples are given to illustrate the whole process and to verify our result.

1. Introduction

Systems of mathematical equations have been used for modeling interactions of genes or proteins in biological systems. The modeling consists of two parts: one is to construct a network of nodes and links, where nodes usually represent genes or proteins and links denote interactions of nodes. The second is to build a set of mathematical formulas for governing the network dynamics, which might be expressed as a system of differential equations or Boolean logic functions [1, 2]. Modeling of a biological network has been focused on the construction of a directed network for a given biological system and its mathematical formulas, with which biological phenomena and mechanisms have been explained and controlled [3] and, furthermore, modeling of a biological network is a useful tool for understanding and explaining the distribution of the biological communities too [4]. However, there are few researches on how to determine an optimal set of nodes and links which are key components of networks for biological phenomena.

There could be multiple networks for explaining a given biological system. So, it would be meaningful to use a probabilistic modeling approach which could yield multiple networks for the biological system. One approach was recently developed in [5, 6] for the construction of network models of a melanoma cell line (SK-MEL-133), the most serious skin cancer, in which each link weight was considered as a discrete random variable and its marginal probability mass function (PMF) was constructed by using a factor graph and a belief propagation (BP) algorithm. The marginal PMFs were used to choose multiple weights of high-probability values and then to construct multiple networks for the biological system. The initial step of the approach was to determine a prior knowledge network (PKN) for melanoma and its mathematical formula structure. Nodes were chosen for representing some drugs (drug nodes) for melanoma or proteins (measured nodes) which take important roles in melanoma. The drug nodes had outgoing links to each measured node without incoming links. Each measured node had an outgoing link to each other measured node and an incoming link from each other measured node. Each link weight in the mathematical formulas was not determined in the initial step. Next, each protein concentration was measured at steady state before and after treatment of the

drugs, which produced experimental data for determining the link weights. Then, the joint PMF of link weights was defined as the Boltzmann/Gibbs form [7] for the factorization of the joint PMF. The form contained a cost function of the experimental data and the simulated values of measured nodes at steady state, where the simulated values were not constants for undetermined link weights. So, the factorization of the joint PMF directly led to the factor graph which is a bipartite graph consisting of variable nodes (link weights in the network) and factor nodes (factors in the factorization). Finally, the marginal PMFs were approximated with an iterative scheme constructed from belief (message) propagation on the factor graph.

In general, computation time could be high to directly compute marginal distributions from a joint probability distribution. So, factor graphs and BP have been applied to infer marginals in diverse problems [8–12]. It is known that a marginal obtained from applying BP becomes the true marginal for an acyclic factor graph and BP could be successfully applied to a cyclic graph [13–18]. There have been many researches on conditions for the convergence of beliefs, which can be applied for computing marginals from a Gaussian joint probability density function (PDF) [19, 20]. However, the convergence conditions cannot be applied to the convergence of beliefs (messages) in [5, 6], since the joint random variable is discrete. And the authors [5, 6] did not analyze the convergence. As far as we know, there is no framework for the inference of marginal PMFs based on both BP and perturbation data except the approach in [5, 6]. For the use of the stochastic approach [5] in the construction of networks based on BP as well as perturbation data, a sufficient condition for the convergence of messages is necessary. Under such a sufficient condition marginal PMFs can be identified and, as a result, multiple network models for the given biological system can be constructed.

In this paper, with our simple network, we present a detailed explanation of the long and intricate process of constructing a system of equations with messages from variable nodes to factor nodes, and vice versa for the approximation of marginal PMFs. Since approximate marginal PMFs are defined with the solution of the system of equations, we identify recursive relations for the solution of the system of equations and construct iterative schemes for solving the recursive relations. Finally we show a sufficient condition for the convergence of the schemes by using a Banach fixed-point theorem (see, e.g., [21, 22]). Numerical examples are given to illustrate the process in Section 4 and to verify our results.

2. Preliminaries

For a clear explanation of the convergence of the iterative schemes mentioned in the Introduction section, we make a simple network of three nodes and four links (see Figure 1(a)), which is assumed to be a PKN. Since the simple network has key components necessary for determining approximate marginal PMFs, the simple network can be extended to any network, including the biological network in [5]. So, our convergence analysis in the Results section can be applied for any network. The network has two measured nodes (x_1, x_2)

and one drug node (x_3). We consider two treatments as the 1st and 2nd perturbations. Symbols x_i^v ($v = 1, 2$, $i = 1, 2, 3$) in Figure 1(b) denote $\log_2(x_i^{v,\text{after}}/x_i^{v,\text{before}})$, 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 vth perturbation, respectively. Since each drug node has an outgoing link to each measured node and no incoming link to any node, drug node x_3 has two outgoing links to x_1 and x_2 with weights $w_{1,3}$ and $w_{2,3}$, respectively. Each measured node has an outgoing link to each other measured node and incoming links from each other nodes, and so, measured node x_1 has one outgoing link to x_2 with weight $w_{2,1}$ and two incoming links from x_2 and x_3 with weights $w_{1,2}$ and $w_{1,3}$, respectively. Similarly, measured node x_2 has one outgoing link to x_1 with weight $w_{1,2}$ and two incoming links from x_1 and x_3 with weights $w_{2,1}$ and $w_{2,3}$, respectively.

As in [5], the dynamics of the given situation is modeled with the mathematical formulas as in Figure 1(c). Simulated values $x_i^{v,s}$ at steady state under the vth perturbation are assumed to be the form

$$x_i^{v,s} = \phi\left(\sum_{j=1, j\neq i}^{3} w_{i,j} x_j^v\right) \quad (i = 1, 2), \ x_3^{v,s} = x_3^v, \quad (1)$$

where $\phi(x) = \tanh(x)$. From now on, weights $w_{i,j}$ are considered as discrete random variables with three values w (-1, 0 and 1) and their PMFs are denoted as follows: for $w \in \{-1, 0, 1\}$

$$
\begin{aligned}
&P\left(w_{1,2} = w\right), \\
&P\left(w_{1,3} = w\right), \\
&P\left(w_{2,1} = w\right), \\
&P\left(w_{2,3} = w\right),
\end{aligned}
\quad (2)
$$

which are approximately calculated in the Results section by using a factor graph and BP based on the factor graph. In [3], discretization over the 3 weight values led to a high rate of convergence and did not capture appropriate uncertainty in a probability distribution and then 11 weight values were used. However, our goals in this paper are to find a convergence condition for messages and to explain the complicated multistep process in [3] step by step by using a prior knowledge network of small size. So, the three weight values are enough for our goals. For the extension of our results to networks of large size and real perturbation data, some notations and mathematical expressions used in our paper might be rather cumbersome in later sections.

3. Results

In this section, we present a system of equations for each marginal PMF, iterative schemes for the solutions and a sufficient condition for the convergence of the iterative schemes.

3.1. System of Equations for Marginal PMFs. To define a joint PMF of link weights based on the experimental data we
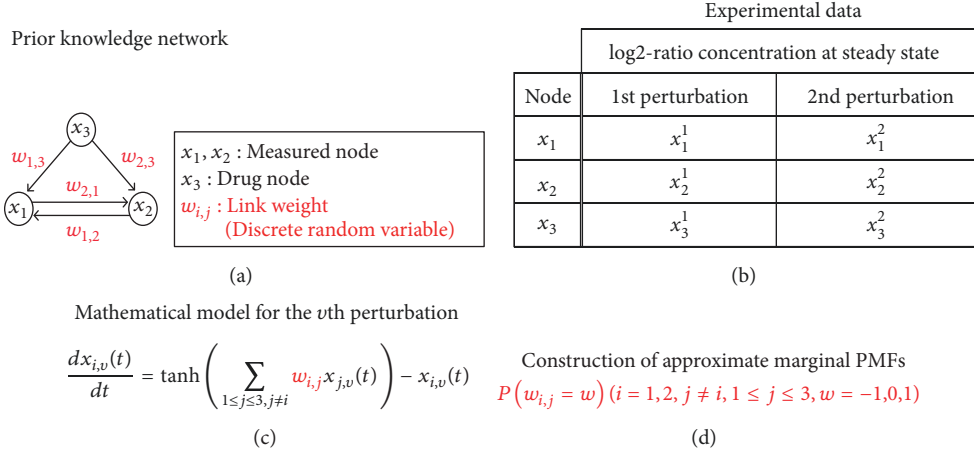
Experimental data

Prior knowledge network

(a)

Node	log2-ratio concentration at steady state	
	1st perturbation	2nd perturbation
x_1	x_1^1	x_1^2
x_2	x_2^1	x_2^2
x_3	x_3^1	x_3^2

(b)

Mathematical model for the vth perturbation

$$\frac{dx_{i,v}(t)}{dt} = \tanh\left(\sum_{1 \le j \le 3, j \ne i} w_{i,j}x_{j,v}(t)\right) - x_{i,v}(t)$$

Construction of approximate marginal PMFs

$$P\left(w_{i,j} = w\right)(i = 1, 2, j \ne i, 1 \le j \le 3, w = -1,0,1)$$

(c)

(d)

FIGURE 1: Given information and our goal. (a) A prior knowledge network of two measured nodes x_1 and x_2, and one drug node x_3. The symbol $w_{i,j}$ denotes the link weight from x_i to x_j, which is a discrete random variable. (b) Experimental data. The 1st and 2nd perturbations mean the 1st and 2nd experiments. Symbol x_i^v ($i = 1, 2, 3$, $v = 1, 2$) denotes the log 2-ratio concentration of x_i at steady state before and after the vth perturbation. (c) Differential equation model for the vth perturbation. Symbol $x_{i,v}(t)$ denotes the relative change of x_i at time t in the vth perturbation. (d) Construction of approximate marginal PMFs from a joint PMF. Set $\{w \mid w = -1, 0, 1\}$ denotes the domain of each marginal PMF.

consider the sum of squared errors between simulated values $x_i^{v,s}$ and experimental values x_i^v

$$\sum_{i=1}^{3}\sum_{v=1}^{2}\left(x_i^{v,s} - x_i^v\right)^2 \tag{3}$$

as well as the penalty

$$\sum_{i=1}^{3}\sum_{j=1,j\ne i}^{3}\delta\left(w_{i,j}\right), \quad \delta\left(w_{i,j} = w\right) = \begin{cases} 0 & (w = 0) \\ 1 & (w \ne 0), \end{cases} \tag{4}$$

which reflects the property that the total number of links tends not to be large. Note that $w_{3,j} = 0$ ($j = 1, 2$) due to no incoming link to drug node x_3. Then the cost function is defined as

$$\text{Cost} = \beta\sum_{i=1}^{3}\sum_{v=1}^{2}\left(x_i^{v,s} - x_i^v\right)^2 + \lambda\sum_{i=1}^{3}\sum_{j=1,j\ne i}^{3}\delta\left(w_{i,j}\right), \tag{5}$$

where the error and penalty are weighted by positive constants β and λ, respectively. Using (1), the cost is written as

$$\text{Cost} = \beta\sum_{i=1}^{2}\sum_{v=1}^{2}\left\{\phi\left(\sum_{j=1,j\ne i}^{3}w_{i,j}x_j^v\right) - x_i^v\right\}^2$$

$$+ \lambda\sum_{i=1}^{2}\sum_{j=1,j\ne i}^{3}\delta\left(w_{i,j}\right). \tag{6}$$

Posteriori probability (joint PMF) $P(W) = P(w_{1,2}, w_{1,3}, w_{2,1}, w_{2,3})$ is defined as

$$P(W) = \frac{1}{Z}\exp(-\text{Cost}), \tag{7}$$

where Z is the constant ensuring that the sum of the probabilities over all model configurations is equal to one. Let $Z = Z_1 Z_2$, where Z_1 and Z_2 are the constants ensuring that the sums of probabilities $P(w_{1,2}, w_{1,3})$ and $P(w_{2,1}, w_{2,3})$ over all model configurations are equal to one, respectively. Substituting the cost (6) into $P(W)$ gives

$$P(W)$$

$$= \frac{1}{Z}\exp\left[\sum_{i=1}^{2}\left(-\beta\sum_{v=1}^{2}\left\{\phi\left(\sum_{j=1,j\ne i}^{3}w_{i,j}x_j^v\right)\right.\right.\right.$$

$$\left.\left.\left. - x_i^v\right\}^2 - \lambda\sum_{j=1,j\ne i}^{3}\delta\left(w_{i,j}\right)\right)\right]$$

$$= \prod_{i=1}^{2} \left(\frac{1}{Z_i} \exp\left\{ -\lambda \sum_{j=1,j\neq i}^{3} \delta\left(w_{i,j}\right) \right\} \prod_{v=1}^{2} \exp\left[-\beta \left\{ \phi\left(\sum_{j=1,j\neq i}^{3} w_{i,j} x_j^v \right) - x_i^v \right\}^2 \right] \right)$$

$$\equiv P\left(w_{1,2}, w_{1,3}\right) P\left(w_{2,1}, w_{2,3}\right),$$

$$(8)$$

which implies that $P(w_{1,2}, w_{1,3})$ and $P(w_{2,1}, w_{2,3})$ can be independently determined by using a similar approach. So, we consider $P(w_{1,2}, w_{1,3})$ to show the approximation of marginal PMF $P(w_{1,2})$:

$$P\left(w_{1,2}\right) = \sum_{w_{1,3}} P\left(w_{1,2}, w_{1,3}\right)$$

$$(9)$$

$$= \frac{1}{Z_1} \sum_{w_{1,3}} e^{-\lambda \sum_{j=2}^{3} \delta(w_{1,j})} \prod_{v=1}^{2} e^{-\beta\{\phi(\sum_{j=2}^{3} w_{1,j} x_j^v) - x_1^v\}^2}.$$

It is not efficient to calculate the exact marginal with the enumeration in the cases where the numbers of nodes in a network or experimental perturbations become large. Therefore, the process to approximate the marginal is presented step by step by using a factor graph and BP on the factor graph.

Step 1 (introduction of a factor graph and BP on the factor graph). Using the factorization in (9), factor nodes F_1^v ($v = 1, 2$) are defined as

$$F_1^v\left(w_{1,2}, w_{1,3}\right) = \exp\left[-\beta \left\{ \phi\left(\sum_{j=2}^{3} w_{1,j} x_j^v \right) - x_1^v \right\}^2 \right] \quad (10)$$

and then marginal $P(w_{1,2})$ in (9) can be written as follows:

$$P\left(w_{1,2}\right) = \frac{1}{Z_1} \sum_{w_{1,3}} e^{-\lambda \sum_{j=2}^{3} \delta(w_{1,j})} \prod_{v=1}^{2} F_1^v\left(w_{1,2}, w_{1,3}\right), \quad (11)$$

which yields the factor graph of two variable nodes $(w_{1,2}, w_{1,3})$ and two factor nodes (F_1^1, F_1^2) as in Figure 2. Following BP on the factor graph, message $P^v(w_{1,j})$ from variable node $w_{1,j}$ to factor node $F_1^v(w_{1,2}, w_{1,3})$ is defined as

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

$$(12)$$

$$\left(v = 1, 2, \ j = 2, 3\right),$$

where $Z_{1,j}^v$ is the normalization constant ensuring that the sum of the probabilities over all model configurations is equal to one and $\rho^\mu(w_{1,j})$ is the message from $F_1^\mu(w_{1,2}, w_{1,3})$ to $w_{1,j}$ defined as

$$\rho^\mu\left(w_{1,j}\right)$$

$$(13)$$

$$= \sum_{\{w_{1,2}, w_{1,3}\} - \{w_{1,j}\}} F_1^\mu\left(w_{1,2}, w_{1,3}\right) \prod_{2\leq\ell\leq 3,\ell\neq j}^{3} P^\mu\left(w_{1,\ell}\right).$$

Using message $\rho^v(w_{1,j})$, marginal PMF $P(w_{1,j})$ in (11) can be approximated as

$$P\left(w_{1,j}\right) = \frac{1}{Z_{1,j}} e^{-\lambda\delta(w_{1,j})} \prod_{v=1}^{2} \rho^v\left(w_{1,j}\right), \quad (14)$$

where $Z_{1,j}$ is the normalization constant ensuring that the sum of the probabilities over all model configurations is equal to one.

Step 2 (approximation of the summation (13)). 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 by introducing a new random variable. The next is to change the single summation into an integral.

Step 2-1. Since ρ^v in (13) is a function of $w_{1,j}$, all random variables in F_1^v $(w_{1,2}, \ldots, w_{1,n})$, which are in (13), can be divided into two type of random variables: one is $w_{1,j}$ and the other is $s_{1,j}^v = \sum_{\xi\neq 1,j}^{n} w_{1,\xi} x_\xi^v$ ($n = 3$), which becomes the linear combination of random variables $w_{1,\xi}$ in the cases where the number n of nodes becomes large. Using the definitions of $s_{1,j}^v$ and F_1^v $(w_{1,2}, \ldots, w_{1,n})$ in (10), we can write F_1^v in (13) as

$$F_1^v\left(s_{1,j}^v, w_{1,j}\right) = \exp\left\{ -\beta\left(\phi\left(s_{1,j}^v + w_{1,j} x_j^v \right) - x_1^v \right)^2 \right\}, \quad (15)$$

which is a function of random variables $s_{1,j}^v$ and $w_{1,j}$. Then (13) becomes

$$\rho^v\left(w_{1,j}\right)$$

$$= \sum_{w_{1,\xi}(1<\xi\leq n,\xi\neq j)} \left\{ F_1^v\left(s_{1,j}^v, w_{1,j}\right) \prod_{2\leq\ell\leq n,\ell\neq j} P^v\left(w_{1,\ell}\right) \right\}. \quad (16)$$

Letting

$$\left\{ \sum_{1<\xi\leq n,\xi\neq j} \widetilde{w}_{1,\xi} x_\xi^v \mid \widetilde{w}_{1,\xi} \in \{-1, 0, 1\}, \ 1 < \xi \leq n, \ \xi \right.$$

$$(17)$$

$$\left. \neq j \right\} = \left\{ \widetilde{s}_{1,j,k}^v \mid 1 \leq k \leq m \right\}$$
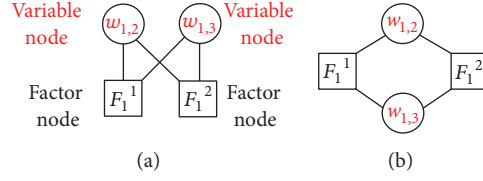
FIGURE 2: Two representations of the factor graph: bipartite graph (a) and cyclic graph (b). The factor graph contains a cycle. The connection of factor node F_1^1 to variable nodes $w_{1,2}$ and $w_{1,3}$ denotes the dependence of F_1^1 on $w_{1,2}$ and $w_{1,3}$ since F_1^1 is a function of $w_{1,2}$ and $w_{1,3}$.

for some positive integer m, message ρ^v in (16) becomes

$$\rho^v\left(w_{1,j}\right) = \sum_{k=1}^{m} \sum_{\substack{w_{1,\xi}=\tilde{w}_{1,\xi}(1<\xi\leq n,\xi\neq j) \\ \sum_{1<\xi\leq n,\xi\neq j} \tilde{w}_{1,\xi}x_\xi^v=\tilde{s}_{1,j,k}^v}} \left\{ F_1^v\left(s_{1,j}^v,w_{1,j}\right) \right.$$

$$\left. \cdot \prod_{2\leq\ell\leq n,\ell\neq j} P^v\left(w_{1,\ell}\right) \right\} = \sum_{k=1}^{m} F_1^v\left(\tilde{s}_{1,j,k}^v,w_{1,j}\right) \quad (18)$$

$$\cdot \left\{ \sum_{\substack{w_{1,\xi}=\tilde{w}_{1,\xi}(1<\xi\leq n,\xi\neq j) \\ \sum_{1<\xi\leq n,\xi\neq j} \tilde{w}_{1,\xi}x_\xi^v=\tilde{s}_{1,j,k}^v}} \prod_{2\leq\ell\leq n,\ell\neq j} P^v\left(w_{1,\ell}\right) \right\}.$$

Note that

$$\sum_{k=1}^{m} \left\{ \sum_{\substack{w_{1,\xi}=\tilde{w}_{1,\xi}(1<\xi\leq n,\xi\neq j) \\ \sum_{1<\xi\leq n,\xi\neq j} \tilde{w}_{1,\xi}x_\xi^v=\tilde{s}_{1,j,k}^v}} \prod_{2\leq\ell\leq n,\ell\neq j} P^v\left(w_{1,\ell}\right) \right\}$$

$$= \sum_{\substack{w_{1,\xi} \\ (1<\xi\leq n,\xi\neq j)}} \prod_{2\leq\ell\leq n,\ell\neq j} P^v\left(w_{1,\ell}\right) \quad (19)$$

$$= \prod_{2\leq\ell\leq n,\ell\neq j} \left\{ \sum_{w_{1,\ell}} P^v\left(w_{1,\ell}\right) \right\} = 1,$$

which implies that the following can be a PMF of $s_{1,j}^v$

$$P_s^v\left(s_{1,j}^v = \tilde{s}_{1,j,k}^v\right) = \sum_{\substack{w_{1,\xi}=\tilde{w}_{1,\xi}(1<\xi\leq n,\xi\neq j) \\ \sum_{1<\xi\leq n,\xi\neq j} \tilde{w}_{1,\xi}x_\xi^v=\tilde{s}_{1,j,k}^v}} \prod_{2\leq\ell\leq n,\ell\neq j} P^v\left(w_{1,\ell}\right) \quad (20)$$

$$\left(1 \leq k \leq m\right).$$

So, message ρ^v in (18) can be written as a single summation

$$\rho^v\left(w_{1,j}\right) = \sum_{s_{1,j}^v} F_1^v\left(s_{1,j}^v,w_{1,j}\right) P_s^v\left(s_{1,j}^v\right). \quad (21)$$

Step 2-2. Note that $s_{1,j}^v$ is a sum of random variables $w_{1,\xi}$ ($1 < \xi \leq n$, $\xi \neq j$), which are assumed to be independent. Even though $s_{1,j}^v$ is not identically distributed, the authors [5] invoked the central limit theorem to approximate the PMF of $s_{1,j}^v$ as a Gaussian with reference to [23], 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 [24]) is satisfied, we think that such a condition might be implicitly assumed in [24]. So, the approximate PMF of $s_{1,j}^v$ becomes

$$P_s^v\left(s_{1,j}^v\right) = \frac{1}{\sqrt{2\pi\Delta_{1,j}^v}} \exp\left[-\frac{\left(s_{1,j}^v - \overline{s_{1,j}^v}\right)^2}{2\Delta_{1,j}^v} \right], \quad (22)$$

where average $\overline{s_{1,j}^v}$ and variance $\Delta_{1,j}^v$ of $s_{1,j}^v$ are defined as

$$\overline{s_{1,j}^v} = E\left(s_{1,j}^v\right) = E\left(\sum_{\ell\neq 1,j}^{3} w_{1,\ell}x_\ell^v \right) = \sum_{\ell\neq 1,j}^{3} \overline{w_{1,\ell}}x_\ell^v, \quad (23)$$

$$\Delta_{1,j}^v = V\left(s_{1,j}^v\right) = V\left(\sum_{\ell\neq 1,j}^{3} w_{1,\ell}x_\ell^v \right)$$

$$= \sum_{\ell\neq 1,j}^{3} V\left(w_{1,\ell}\right)\left(x_\ell^v\right)^2 \quad (24)$$

$$= \sum_{\ell\neq 1,j}^{3} \left(\overline{w_{1,\ell}^2} - \overline{w_{1,\ell}}^2\right)\left(x_\ell^v\right)^2.$$

Symbols $\overline{w_{1,\ell}}$ and $\overline{w_{1,\ell}^2}$ denote the averages of $w_{1,\ell}$ and $(w_{1,\ell})^2$, respectively:

$$\overline{w_{1,\ell}} = \sum_w w P^v\left(w_{1,\ell} = w\right),$$

$$\overline{w_{1,\ell}^2} = \sum_w w^2 P^v\left(w_{1,\ell} = w\right) \quad (25)$$

and then the sum over configurations $\{w_{1,\ell} \mid 2 \leq \ell \leq 3, \ell \neq j\}$ in (19) is approximated with a Gaussian integration:

$$\rho^v\left(w_{1,j}\right) \approx \int_{-\infty}^{\infty} F_1^v\left(s_{1,j}^v,w_{1,j}\right) P_s^v\left(s_{1,j}^v\right) ds_{1,j}^v. \quad (26)$$

Step 3 (approximation of the improper integral (26)). For the calculation of the improper integral (26), error $\phi(s_{1,j}^v + w_{1,j}x_j^v) - x_j^v$ is linearized with respect to the maximization of the fitness in (15). Note that the equality

$$\phi\left(s_{1,j}^v + w_{1,j}x_j^v\right) - x_1^v = 0, \quad (27)$$

can be written as

$$\phi^{-1}\left(x_1^v\right) - w_{1,j}x_j^v - s_{1,j}^v = 0, \tag{28}$$

under the assumption that

experimental data x_1^v are contained in the codomain of ϕ. (29)

Then error $\phi(s_{1,j}^v + w_{1,j}x_j^v) - x_1^v$ in (15) is approximated by

$$\phi^{-1}\left(x_1^v\right) - w_{1,j}x_j^v - s_{1,j}^v, \tag{30}$$

which yields

$$F_1^v\left(s_{1,j}^v, w_{1,j}\right) = \exp\left\{-\beta\left(\phi\left(s_{1,j}^v + w_{1,j}x_j^v\right) - x_1^v\right)^2\right\}$$
$$\approx \exp\left\{-\beta\left(\phi^{-1}\left(x_1^v\right) - w_{1,j}x_j^v - s_{1,j}^v\right)^2\right\}. \tag{31}$$

$$-\beta\left(\phi^{-1}\left(x_1^v\right) - w_{1,j}x_j^v - s_{1,j}^v\right)^2$$
$$-\frac{\left(s_{1,j}^v - \overline{s_{1,j}^v}\right)^2}{2\Delta_{1,j}^v} = -\frac{1}{2}\frac{\left(1 + 2\beta\Delta_{1,j}^v\right)\left[s_{1,j}^v - \left(\overline{s_{1,j}^v} + \beta\left\{\phi^{-1}\left(x_1^v\right) - w_{1,j}x_j^v\right\}2\Delta_{1,j}^v\right)/\left(1 + 2\beta\Delta_{1,j}^v\right)\right]^2}{\Delta_{1,j}^v} \tag{33}$$
$$-\beta\frac{\left\{\phi^{-1}\left(x_1^v\right) - w_{1,j}x_j^v - \overline{s_{1,j}^v}\right\}^2}{1 + 2\beta\Delta_{1,j}^v}$$

and the property of a PDF (the integral of a PDF over its domain is equal to 1). Therefore adopting the final approximation of ρ^v in [5], approximate marginal PMF $P(w_{1,j} = w)$ in (14) can be obtained by solving the system of equations

$$\rho^v\left(w_{1,j} = w\right) = \exp\left[-\beta\frac{\left\{\phi^{-1}\left(x_1^v\right) - wx_j^v - \overline{s_{1,j}^v}\right\}^2}{1 + 2\beta\Delta_{1,j}^v}\right], \tag{34}$$

$$P^v\left(w_{1,j} = w\right) = \frac{1}{Z_{1,j}^v}e^{-\lambda\delta(w_{1,j}=w)}\rho^{3-v}\left(w_{1,j} = w\right), \tag{35}$$

where $v = 1, 2$, $j = 2, 3$, $w = -1, 0, 1$, $\overline{s_{1,j}^v}$ and $\Delta_{1,j}^v$ are in (23)–(25).

Remark 1. Since the approximation (34) was used in [5] and our goal in this paper is to find a convergence condition for the algorithm in [5], we also use (34) instead of

$$\rho^v\left(w_{1,j}\right) = \frac{1}{\left(1 + 2\beta\Delta_{1,j}^v\right)^{1/2}}$$
$$\cdot \exp\left[-\beta\frac{\left\{\phi^{-1}\left(x_1^v\right) - w_{1,j}x_j^v - \overline{s_{1,j}^v}\right\}^2}{1 + 2\beta\Delta_{1,j}^v}\right]. \tag{36}$$

Hence we have

$$\rho^v\left(w_{1,j}\right) \approx \int_{-\infty}^{\infty}F_1^v\left(s_{1,j}^v, w_{1,j}\right)P_s^v\left(s_{1,j}^v\right)ds_{1,j}^v$$
$$\approx \int_{-\infty}^{\infty}\exp\left\{-\beta\left(\phi^{-1}\left(x_1^v\right) - w_{1,j}x_j^v - s_{1,j}^v\right)^2\right\}$$
$$\cdot \frac{1}{\sqrt{2\pi\Delta_{1,j}^v}}\exp\left[-\frac{\left(s_{1,j}^v - \overline{s_{1,j}^v}\right)^2}{2\Delta_{1,j}^v}\right]ds_{1,j}^v \tag{32}$$
$$= \frac{1}{\left(1 + 2\beta\Delta_{1,j}^v\right)^{1/2}}$$
$$\cdot \exp\left[-\beta\frac{\left\{\phi^{-1}\left(x_1^v\right) - w_{1,j}x_j^v - \overline{s_{1,j}^v}\right\}^2}{1 + 2\beta\Delta_{1,j}^v}\right],$$

where the last equality is obtained both by the identity

Even when using (36) instead of (34), the convergence analysis in the later subsections could be applied (we do not show it in this paper).

Remark 2. Similarly approximate marginal PMF $P(w_{2,j})$ can be obtained by replacing subscripts 1 and $j = 2, 3$, related with node x_1, in $w_{1,j}$, $\overline{s_{1,j}^v}$, $\Delta_{1,j}^v$, and $w_{1,\ell}$ in (34), (35), and (23)–(25) with 2 and $j = 1, 3$, related with node x_2: for $v = 1, 2$, $j = 1, 3$:

$$P\left(w_{2,j}\right) = \frac{1}{Z_{2,j}}e^{-\lambda\delta(w_{2,j})}\prod_{v=1}^{2}\rho^v\left(w_{2,j}\right), \tag{37}$$

$$\rho^v\left(w_{2,j}\right) = \exp\left[-\beta\frac{\left\{\phi^{-1}\left(x_2^v\right) - w_{2,j}x_j^v - \overline{s_{2,j}^v}\right\}^2}{1 + 2\beta\Delta_{2,j}^v}\right], \tag{38}$$

$$P^v\left(w_{2,j}\right) = \frac{1}{Z_{2,j}^v}e^{-\lambda\delta(w_{2,j})}\rho^{3-v}\left(w_{2,j}\right), \tag{39}$$

where $Z_{2,j}^v$ is the normalization constant ensuring that the sum of the probabilities over all model configurations is equal to one and

$$\overline{s_{2,j}^v} = \sum_{\ell \neq 2,j}^{3} \left\{ \sum_w w P^v \left(w_{2,\ell} = w \right) \right\} x_\ell^v, \tag{40}$$

$$\Delta_{2,j}^v = \sum_{\ell \neq 2,j}^{3} \left\{ \left(\sum_w w^2 P^v \left(w_{2,\ell} = w \right) \right) \right.$$
$$\left. - \left(\sum_w w P^v \left(w_{2,\ell} = w \right) \right)^2 \right\} \left(x_\ell^v \right)^2 . \tag{41}$$

3.2. Iterative Schemes for Solving the System of Equations. In this subsection, we construct sequences $\{\rho_{1,j,n}^v(w)\}$ and $\{P_{1,j,n}^v(w)\}$ ($v = 1, 2, \ j = 2, 3, \ w = -1, 0, 1$) by using (34) and (35). In the next subsection we present a sufficient condition for the convergence of the two sequences, which implies that the two limits satisfy (34) and (35). So, the limit of $\{\rho_{1,j,n}^v(w)\}$ becomes value $\rho^v(w_{1,j} = w)$, leading to the construction of approximate marginal PMF $P(w_{1,j})$ in (14). We assume that

$$\text{initial terms } \rho_{1,j,0}^v(w) \text{ of } \left\{ \rho_{1,j,n}^v(w) \right\} \text{ are given as positive numbers} \tag{42}$$

and initial terms of $\{P_{1,j,n}^v(w)\}$ are defined as

$$P_{1,j,0}^v(w) = \frac{1}{Z_{1,j,0}^v} e^{-\lambda \delta(w)} \rho_{1,j,0}^{3-v}(w), \tag{43}$$

where $Z_{1,j,0}^v$ is the normalization constant ensuring that the sum of the probabilities over all model configurations is equal to one. The 1st iterations $\rho_{1,j,1}^v(w)$ and $P_{1,j,1}^v(w)$ are defined similarly to (34) and (35) and so we need to define $\overline{s_{1,j,0}^v}$ and $\Delta_{1,j,0}^v$ by using (23) and (24):

$$\overline{s_{1,j,0}^v} = \sum_{\ell \neq 1,j}^{3} \overline{w_{1,\ell,0}} x_\ell^v = \sum_{\ell \neq 1,j}^{3} \left\{ \sum_{w \in \{-1,0,1\}} w P_{1,\ell,0}^v(w) \right\} x_\ell^v$$
$$= \sum_{\ell \neq 1,j}^{3} \left\{ -P_{1,\ell,0}^v(-1) + P_{1,\ell,0}^v(1) \right\} x_\ell^v$$
$$= \begin{cases} \left(-P_{1,3,0}^v(-1) + P_{1,3,0}^v(1) \right) x_3^v & (j = 2) \\ \left(-P_{1,2,0}^v(-1) + P_{1,2,0}^v(1) \right) x_2^v & (j = 3) \end{cases} \tag{44}$$
$$= \frac{e^{-\lambda}}{Z_{1,5-j,0}^v} \left\{ -\rho_{1,5-j,0}^{3-v}(-1) + \rho_{1,5-j,0}^{3-v}(1) \right\} x_{5-j}^v,$$

and

$$\Delta_{1,j,0}^v = \sum_{\ell \neq 1,j}^{3} \left(\overline{w_{1,\ell,0}^2} - \overline{w_{1,\ell,0}}^2 \right) \left(x_\ell^v \right)^2$$

$$= \sum_{\ell \neq 1,j}^{3} \left\{ \left(\sum_w w^2 P_{1,\ell,0}^v(w) \right) - \left(\sum_w w P_{1,\ell,0}^v(w) \right)^2 \right\}$$
$$\cdot \left(x_\ell^v \right)^2$$

$$= \sum_{\ell \neq 1,j}^{3} \left\{ \left(P_{1,\ell,0}^v(-1) + P_{1,\ell,0}^v(1) \right) - \left(-P_{1,\ell,0}^v(-1) + P_{1,\ell,0}^v(1) \right)^2 \right\} \tag{45}$$
$$\cdot \left(x_\ell^v \right)^2$$

$$= \frac{e^{-\lambda}}{Z_{1,5-j,0}^v}$$

$$\cdot \left[\left\{ \rho_{1,5-j,0}^{3-v}(-1) + \rho_{1,5-j,0}^{3-v}(1) \right\} - \frac{e^{-\lambda}}{Z_{1,5-j,0}^v} \left\{ -\rho_{1,5-j,0}^{3-v}(-1) + \rho_{1,5-j,0}^{3-v}(1) \right\} \right]^2 \left(x_{5-j}^v \right)^2 .$$

So, the 1st iteration is defined as

$$\rho_{1,j,1}^{v}(w) = \exp\left[-\beta\frac{\left\{\phi^{-1}(x_1^v) - wx_j^v - \overline{s_{1,j,0}^v}\right\}^2}{1 + 2\beta\Delta_{1,j,0}^v}\right],$$

$$P_{1,j,1}^{v}(w) = \frac{1}{Z_{1,j,1}^{v}}e^{-\lambda\delta(w)}\rho_{1,j,1}^{3-v}(w), \tag{46}$$

where $Z_{1,j,1}^{v}$ is the normalization constant ensuring that the sum of the probabilities over all model configurations is equal to one. Similarly the $(n+1)^{th}$ iteration is defined as

$$\rho_{1,j,n+1}^{v}(w) = \Phi_{1,j,w}^{v}\left(\rho_{1,5-j,n}^{3-v}(-1), \rho_{1,5-j,n}^{3-v}(1)\right) \tag{47}$$

$$(n \geq 0),$$

$$P_{1,j,n+1}^{v}(w) = \frac{1}{Z_{1,j,n+1}^{v}}e^{-\lambda\delta(w)}\rho_{1,j,n+1}^{3-v}(w), \tag{48}$$

for $v = 1,2$, $j = 2,3$ and $w = -1,0,1$. Here $Z_{1,j,n+1}^{v}$ is the normalization constant ensuring that the sum of the probabilities over all model configurations is equal to one and

$$\Phi_{1,j,w}^{v}\left(\rho_{1,5-j,n}^{3-v}(-1), \rho_{1,5-j,n}^{3-v}(1)\right)$$

$$= \exp\left[-\beta\frac{\left\{\phi^{-1}(x_1^v) - wx_j^v - \left(e^{-\lambda}/Z_{1,5-j,n}^{v}\right)\left(-\rho_{1,5-j,n}^{3-v}(-1) + \rho_{1,5-j,n}^{3-v}(1)\right)x_{5-j}^v\right\}^2}{1 + 2\beta\left(e^{-\lambda}/Z_{1,5-j,n}^{v}\right)\left\{\left(\rho_{1,5-j,n}^{3-v}(-1) + \rho_{1,5-j,n}^{3-v}(1)\right) - \left(e^{-\lambda}/Z_{1,5-j,n}^{v}\right)\left(-\rho_{1,5-j,n}^{3-v}(-1) + \rho_{1,5-j,n}^{3-v}(1)\right)^2\right\}\left(x_{5-j}^v\right)^2}\right]. \tag{49}$$

Therefore the desired iterative schemes consist of (42), (43), (47), (48), and (49) under the assumption (29).

Remark 3. Since $P_{1,j,n+1}^{v}(w_{1,j})$ defined in (48) is a PMF due to (47) and (49), we have

$$0 < \frac{e^{-\lambda}}{Z_{1,5-j,n}^{v}}\rho_{1,5-j,n}^{3-v}(-1) < 1,$$

$$0 < \frac{e^{-\lambda}}{Z_{1,5-j,n}^{v}}\rho_{1,5-j,n}^{3-v}(1) < 1, \tag{50}$$

which gives the positivity of the following term in (49)

$$\rho_{1,5-j,n}^{3-v}(-1) + \rho_{1,5-j,n}^{3-v}(1)$$

$$- \frac{e^{-\lambda}}{Z_{1,5-j,n}^{v}}\left(-\rho_{1,5-j,n}^{3-v}(-1) + \rho_{1,5-j,n}^{3-v}(1)\right)^2 > 0. \tag{51}$$

So, we obtain a lower bound of $\rho_{1,j,n+1}^{v}(w)$

$$\rho_{1,j,n+1}^{v}(w)$$

$$\geq \exp\left[-\beta\left\{\left|\phi^{-1}(x_1^v)\right| + \left|x_j^v\right| + 2\left|x_{5-j}^v\right|\right\}^2\right]. \tag{52}$$

Therefore we obtain an upper bound

$$\frac{e^{-\lambda}}{Z_{1,j,n}^{v}} = \frac{e^{-\lambda}}{e^{-\lambda}\rho_{1,j,n}^{v}(-1) + \rho_{1,j,n}^{v}(0) + e^{-\lambda}\rho_{1,j,n}^{v}(1)}$$

$$\leq e^{-\lambda}U_{j,\beta,\lambda}^{v}, \tag{53}$$

$$U_{j,\beta,\lambda}^{v}$$

$$= \frac{1}{1 + 2e^{-\lambda}}\exp\left[\beta\left\{\left|\phi^{-1}(x_1^v)\right| + \left|x_j^v\right| + 2\left|x_{5-j}^v\right|\right\}^2\right],$$

which is used in the proof of Lemma 8.

Remark 4. Note that sequence $\{\rho_{1,j,n}^{v}(w)\}$ in the recursive relation (47) contains no $P_{1,j,n}^{v}(w)$. Then, in the next subsection, we present a sufficient condition for the convergence of sequence $\{\rho_{1,j,n}^{v}(w)\}$ without using $\{P_{1,j,n}^{v}(w)\}$ and, as a result, the limit of $\{\rho_{1,j,n}^{v}(w)\}$ is used to define the message $\rho^{v}(w_{1,j} = w)$.

Remark 5. Similarly, to define messages $\rho^{v}(w_{2,j} = w)$ ($v = 1,2$, $j = 1,3$, $w = -1,0,1$), we use the following sequences:

$$\rho_{2,j,n+1}^{v}(w) = \Phi_{2,j,w}^{v}\left(\rho_{2,4-j,n}^{3-v}(-1), \rho_{2,4-j,n}^{3-v}(1)\right) \tag{54}$$

$$(n \geq 0),$$

$$P_{2,j,n+1}^{v}(w) = \frac{1}{Z_{2,j,n+1}^{v}}e^{-\lambda\delta(w)}\rho_{2,j,n+1}^{3-v}(w), \tag{55}$$

where $Z_{2,j,n+1}^{v}$ is the normalization constant ensuring that the sum of the probabilities over all model configurations is equal to one and

$$\Phi_{2,j,w}^{v}\left(\rho_{2,4-j,n}^{3-v}(-1),\rho_{2,4-j,n}^{3-v}(1)\right)$$

$$= \exp\left[-\beta\frac{\left\{\phi^{-1}(x_{2}^{v})-wx_{j}^{v}-\left(e^{-\lambda}/Z_{2,4-j,n}^{v}\right)\left(-\rho_{2,4-j,n}^{3-v}(-1)+\rho_{2,4-j,n}^{3-v}(1)\right)x_{4-j}^{v}\right\}^{2}}{1+2\beta\left(e^{-\lambda}/Z_{2,4-j,n}^{v}\right)\left\{\left(\rho_{2,4-j,n}^{3-v}(-1)+\rho_{2,4-j,n}^{3-v}(1)\right)-\left(e^{-\lambda}/Z_{2,4-j,n}^{v}\right)\left(-\rho_{2,4-j,n}^{3-v}(-1)+\rho_{2,4-j,n}^{3-v}(1)\right)^{2}\right\}\left(x_{4-j}^{v}\right)^{2}}\right]. \tag{56}$$

3.3. A Sufficient Condition for the Convergence of the Iterative Schemes. Let

$$\mathbf{X}_1^{(n)} = \begin{pmatrix} X_{1,1}^{(n)}, X_{1,2}^{(n)}, X_{1,3}^{(n)}, \\ X_{1,4}^{(n)}, X_{1,5}^{(n)}, X_{1,6}^{(n)}, \\ X_{1,7}^{(n)}, X_{1,8}^{(n)}, X_{1,9}^{(n)}, \\ X_{1,10}^{(n)}, X_{1,11}^{(n)}, X_{1,12}^{(n)} \end{pmatrix} \tag{57}$$

$$= \begin{pmatrix} \rho_{1,2,n}^{1}(-1),\rho_{1,2,n}^{1}(0),\rho_{1,2,n}^{1}(1), \\ \rho_{1,3,n}^{1}(-1),\rho_{1,3,n}^{1}(0),\rho_{1,3,n}^{1}(1), \\ \rho_{1,2,n}^{2}(-1),\rho_{1,2,n}^{2}(0),\rho_{1,2,n}^{2}(1), \\ \rho_{1,3,n}^{2}(-1),\rho_{1,3,n}^{2}(0),\rho_{1,3,n}^{2}(1) \end{pmatrix} \in R^{12}$$

and

$$\Phi_1 = \begin{pmatrix} \Phi_{1,1}, \Phi_{1,2}, \Phi_{1,3}, \\ \Phi_{1,4}, \Phi_{1,5}, \Phi_{1,6}, \\ \Phi_{1,7}, \Phi_{1,8}, \Phi_{1,9}, \\ \Phi_{1,10}, \Phi_{1,11}, \Phi_{1,12} \end{pmatrix}$$

$$= \begin{pmatrix} \Phi_{1,2,-1}^{1}, \Phi_{1,2,0}^{1}, \Phi_{1,2,1}^{1}, \\ \Phi_{1,3,-1}^{1}, \Phi_{1,3,0}^{1}, \Phi_{1,3,1}^{1}, \\ \Phi_{1,2,-1}^{2}, \Phi_{1,2,0}^{2}, \Phi_{1,2,1}^{2}, \\ \Phi_{1,3,-1}^{2}, \Phi_{1,3,0}^{2}, \Phi_{1,3,1}^{2} \end{pmatrix}, \tag{58}$$

where the subscript 1 of $\mathbf{X}_1$ and Φ_1 represents node x_1. The $(n+1)^{th}$ iteration in (47) is written as

$$\mathbf{X}_1^{(n+1)} = \Phi_1\left(\mathbf{X}_1^{(n)}\right). \tag{59}$$

It follows from (47), (57), and (58) that $\Phi_{1,\theta}$ ($1 \le \theta \le 12$) are functions of two independent variables: for $\mathbf{X} \in R^{12}$

$$\Phi_1(\mathbf{X})$$

$$= \begin{pmatrix} \Phi_{1,1}(X_{10},X_{12}),\Phi_{1,2}(X_{10},X_{12}),\Phi_{1,3}(X_{10},X_{12}), \\ \Phi_{1,4}(X_{7},X_{9}),\ \Phi_{1,5}(X_{7},X_{9}),\ \Phi_{1,6}(X_{7},X_{9}), \\ \Phi_{1,7}(X_{4},X_{6}),\ \Phi_{1,8}(X_{4},X_{6}),\ \Phi_{1,9}(X_{4},X_{6}), \\ \Phi_{1,10}(X_{1},X_{3}),\ \Phi_{1,11}(X_{1},X_{3}),\ \Phi_{1,12}(X_{1},X_{3}) \end{pmatrix}. \tag{60}$$

We use Banach fixed-point theorem [21] for the convergence of the sequence (59) to prove Theorem 9, which is our main result.

Theorem 6. *Let D be a closed subset of R^m for a positive integer m. If a function $\Psi : D \longrightarrow D$ satisfies that for a constant $k \in (0,1)$ and all $\mathbf{x}, \mathbf{y}$ in D*

$$\|\Psi(\mathbf{x}) - \Psi(\mathbf{y})\| \le k\|\mathbf{x} - \mathbf{y}\|, \tag{61}$$

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

For the application of Theorem 6 to the proof of our main result we need the following lemmas.

Lemma 7. *Assume that experimental data x_1^v are contained in the codomain of ϕ. Assume that positive constants β and λ satisfy inequalities $1 - 2\beta(e^{-\lambda}U_{j,\beta,\lambda}^{v})^2(x_j^v)^2 > 0$ for $v = 1, 2, j = 2, 3$ and $U_{j,\beta,\lambda}^{v}$ defined in (53). Then Φ_1 defined in (49) and (57)–(60) becomes a function from domain $[0,1]^{12}$ to codomain $[0,1]^{12}$.*

Proof. It follows from (49) and (57)–(60) that for $(\theta, w) = (1,-1), (2,0), (3,1)$

$$\Phi_{1,\theta}(\mathbf{X}) = \Phi_{1,2,w}^{1}(X_{10},X_{12}),$$

$$\mathbf{X} = (X_1,\ldots,X_{12}) \in [0,1]^{12} \tag{62}$$

and

$$\Phi_{1,2,w}^{1}(X_{10},X_{12}) = \exp\left[-\beta\frac{\left\{\phi^{-1}(x_{1}^{1})-wx_{2}^{1}-\left(e^{-\lambda}/Z_{1,3,n}^{1}\right)\left(-X_{10}+X_{12}\right)x_{3}^{1}\right\}^{2}}{1+2\beta\left(e^{-\lambda}/Z_{1,3,n}^{1}\right)\left(\left(X_{10}+X_{12}\right)-\left(e^{-\lambda}/Z_{1,3,n}^{1}\right)\left(-X_{10}+X_{12}\right)^{2}\right)\left(x_{3}^{1}\right)^{2}}\right]. \tag{63}$$

We have that for $\beta > 0$, $\lambda > 0$ and $\{X_{10}, X_{12}\} \subset [0, 1]$

$$1 + 2\beta \frac{e^{-\lambda}}{Z_{1,3,n}^1} \left((X_{10} + X_{12}) - \frac{e^{-\lambda}}{Z_{1,3,n}^1} (-X_{10} + X_{12})^2 \right)$$
$$\cdot \left(x_3^1 \right)^2 \geq 1 - 2\beta \left(e^{-\lambda} U_{3,\beta,\lambda}^1 \right)^2 \left(x_3^1 \right)^2 > 0, \tag{64}$$

and for $\beta > 0$

$$- \beta \left\{ \phi^{-1} \left(x_1^1 \right) - w x_2^1 - \frac{e^{-\lambda}}{Z_{1,3,n}^1} (-X_{10} + X_{12}) x_3^1 \right\}^2 \tag{65}$$
$$\leq 0,$$

and so, we obtain

$$0 < \Phi_{1,\theta} (\mathbf{X}) \leq 1 \quad (1 \leq \theta \leq 3). \tag{66}$$

Similarly, under the same condition, we can obtain

$$0 < \Phi_{1,\theta} (\mathbf{X}) \leq 1 \quad (4 \leq \theta \leq 12), \tag{67}$$

which are desired results. $\qquad\square$

Lemma 8. *Assume that experimental data x_1^v are contained in the codomain of ϕ. Assume that positive constants β and λ satisfy inequalities $1 - 2\beta (e^{-\lambda} U_{j,\beta,\lambda}^v)^2 (x_j^v)^2 > 0$ for $v = 1, 2, j = 2, 3$ and $U_{j,\beta,\lambda}^v$ defined in (53). Then for $\mathbf{\Phi}_1$ defined in (49) and (57)–(60) and $\{\mathbf{X}, \mathbf{Y}\} \subset [0, 1]^{12}$*

$$\|\mathbf{\Phi}_1 (\mathbf{X}) - \mathbf{\Phi}_1 (\mathbf{Y})\|$$
$$\leq \sqrt{6} \max \left\{ C_{j,\beta,\lambda}^v \mid v = 1, 2, \; j = 2, 3 \right\} \beta e^{-\lambda} \|\mathbf{X} - \mathbf{Y}\|, \tag{68}$$

where

$$C_{j,\beta,\lambda}^v = \left[2 \left\{ \left| \phi^{-1} \left(x_1^v \right) \right| + \left| x_j^v \right| + e^{-\lambda} U_{5-j,\beta,\lambda}^v \left| x_{5-j}^v \right| \right\} \right.$$
$$\cdot U_{5-j,\beta,\lambda}^v \left| x_{5-j}^v \right|$$
$$\times \left[1 + 2\beta e^{-\lambda} U_{5-j,\beta,\lambda}^v \left(2 + e^{-\lambda} U_{5-j,\beta,\lambda}^v \right) \left(x_{5-j}^v \right)^2 \right]$$
$$+ \left\{ \left| \phi^{-1} \left(x_1^v \right) \right| + \left| x_j^v \right| + e^{-\lambda} U_{5-j,\beta,\lambda}^v \left| x_{5-j}^v \right| \right\}^2$$
$$\times 2\beta U_{5-j,\beta,\lambda}^v \left(1 + 2e^{-\lambda} U_{5-j,\beta,\lambda}^v \right) \left(x_{5-j}^v \right)^2 \right]$$
$$\times \left\{ 1 - 2\beta \left(e^{-\lambda} U_{5-j,\beta,\lambda}^v \right)^2 \left(x_{5-j}^v \right)^2 \right\}^{-2}. \tag{69}$$

Proof. Using (62) and (63), we have that for $\{\mathbf{X}, \mathbf{Y}\} \subset [0, 1]^{12}$

$$\Phi_{1,1} (\mathbf{X}) - \Phi_{1,1} (\mathbf{Y}) = \Phi_{1,2,-1}^1 (X_{10}, X_{12}) \tag{70}$$
$$- \Phi_{1,2,-1}^1 (Y_{10}, Y_{12})$$
$$= \exp \left[f (X_{10}, X_{12}) \right] \tag{71}$$
$$- \exp \left[f (Y_{10}, Y_{12}) \right],$$

where $\{X_{10}, X_{12}, Y_{10}, Y_{12}\} \subset [0, 1]$ and function $f : [0, 1]^2 \longrightarrow [0, 1]^2$ is defined as follows: for $\zeta, \eta \in [0, 1]$

$$f (\zeta, \eta) = -\beta$$
$$\cdot \frac{\left\{ \phi^{-1} \left(x_1^1 \right) - (-1) x_2^1 - \left(e^{-\lambda}/Z_{1,3,n}^1 \right) (-\zeta + \eta) x_3^1 \right\}^2}{1 + 2\beta \left(e^{-\lambda}/Z_{1,3,n}^1 \right) \left((\zeta + \eta) - \left(e^{-\lambda}/Z_{1,3,n}^1 \right) (-\zeta + \eta)^2 \right) \left(x_3^1 \right)^2}. \tag{72}$$

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

$$\exp \left[f (X_{10}, X_{12}) \right] - \exp \left[f (Y_{10}, Y_{12}) \right]$$
$$= \left\{ \nabla \exp \left[f \left\{ (1 - c) (X_{10}, X_{12}) + c (Y_{10}, Y_{12}) \right\} \right] \right\} \tag{73}$$
$$\cdot (X_{10} - Y_{10}, X_{12} - Y_{12}).$$

Note that

$$\left| \frac{\partial}{\partial \zeta} \exp \left[f (\zeta, \eta) \right] \right| \leq \left| \frac{\partial}{\partial \zeta} f (\zeta, \eta) \right|, \tag{74}$$

where

$$\frac{\partial}{\partial \zeta} f (\zeta, \eta) = -\beta \frac{\left[\begin{array}{l} 2 \left\{ \phi^{-1} \left(x_1^1 \right) + x_2^1 - \left(e^{-\lambda}/Z_{1,3,n}^1 \right) (-\zeta + \eta) x_3^1 \right\} \left(e^{-\lambda}/Z_{1,3,n}^1 \right) x_3^1 \\ \times \left[1 + 2\beta \left(e^{-\lambda}/Z_{1,3,n}^1 \right) \left((\zeta + \eta) - \left(e^{-\lambda}/Z_{1,3,n}^1 \right) (-\zeta + \eta)^2 \right) \left(x_3^1 \right)^2 \right] \\ - \left\{ \phi^{-1} \left(x_1^1 \right) + x_2^1 - \left(e^{-\lambda}/Z_{1,3,n}^1 \right) (-\zeta + \eta) x_3^1 \right\}^2 \\ \times 2\beta \left(e^{-\lambda}/Z_{1,3,n}^1 \right) \left(1 - 2 \left(e^{-\lambda}/Z_{1,3,n}^1 \right) (-\zeta + \eta) (-1) \right) \left(x_3^1 \right)^2 \end{array} \right]}{\left[1 + 2\beta \left(e^{-\lambda}/Z_{1,3,n}^1 \right) \left((\zeta + \eta) - \left(e^{-\lambda}/Z_{1,3,n}^1 \right) (-\zeta + \eta)^2 \right) \left(x_3^1 \right)^2 \right]^2}. \tag{75}$$

Using $\{\zeta, \eta\} \subset [0, 1]$, we have

$$\left|\frac{\partial}{\partial \zeta} f(\zeta, \eta)\right| \leq \beta \frac{\begin{bmatrix} 2\left\{\left|\phi^{-1}\left(x_1^1\right)\right| + \left|x_2^1\right| + \left(e^{-\lambda}/Z_{1,3,n}^1\right)\left|x_3^1\right|\right\}\left(e^{-\lambda}/Z_{1,3,n}^1\right)\left|x_3^1\right| \\ \times \left[1 + 2\beta\left(e^{-\lambda}/Z_{1,3,n}^1\right)\left(2 + e^{-\lambda}/Z_{1,3,n}^1\right)\left(x_3^1\right)^2\right] \\ + \left\{\left|\phi^{-1}\left(x_1^1\right)\right| + \left|x_2^1\right| + \left(e^{-\lambda}/Z_{1,3,n}^1\right)\left|x_3^1\right|\right\}^2 \\ \times 2\beta\left(e^{-\lambda}/Z_{1,3,n}^1\right)\left(1 + 2\left(e^{-\lambda}/Z_{1,3,n}^1\right)\right)\left(x_3^1\right)^2 \end{bmatrix}}{\left[1 + 2\beta\left(e^{-\lambda}/Z_{1,3,n}^1\right)\left((\zeta + \eta) - \left(e^{-\lambda}/Z_{1,3,n}^1\right)(-\zeta + \eta)^2\right)\left(x_3^1\right)^2\right]^2}. \tag{76}$$

Using the given condition in this lemma and (53), we have the positive lower bound

$$1 + 2\beta \frac{e^{-\lambda}}{Z_{1,3,n}^1}\left((\zeta + \eta) - \frac{e^{-\lambda}}{Z_{1,3,n}^1}(-\zeta + \eta)^2\right)\left(x_3^1\right)^2$$

$$\geq 1 - 2\beta\left(e^{-\lambda}U_{3,\beta,\lambda}^1\right)^2\left(x_3^1\right)^2 > 0, \tag{77}$$

which gives that

$$\left|\frac{\partial}{\partial \zeta} f(\zeta, \eta)\right| \leq \beta e^{-\lambda}$$

$$\cdot \frac{\begin{bmatrix} 2\left\{\left|\phi^{-1}\left(x_1^1\right)\right| + \left|x_2^1\right| + e^{-\lambda}U_{3,\beta,\lambda}^1\left|x_3^1\right|\right\}U_{3,\beta,\lambda}^1\left|x_3^1\right| \\ \times \left[1 + 2\beta e^{-\lambda}U_{3,\beta,\lambda}^1\left(2 + e^{-\lambda}U_{3,\beta,\lambda}^1\right)\left(x_3^1\right)^2\right] \\ + \left\{\left|\phi^{-1}\left(x_1^1\right)\right| + \left|x_2^1\right| + e^{-\lambda}U_{3,\beta,\lambda}^1\left|x_3^1\right|\right\}^2 \\ \times 2\beta U_{3,\beta,\lambda}^1\left(1 + 2e^{-\lambda}U_{3,\beta,\lambda}^1\right)\left(x_3^1\right)^2 \end{bmatrix}}{\left[1 - 2\beta\left(e^{-\lambda}U_{3,\beta,\lambda}^1\right)^2\left(x_3^1\right)^2\right]^2}$$

$$= \beta e^{-\lambda} C_{2,\beta,\lambda}^1. \tag{78}$$

Similarly,

$$\left|\frac{\partial}{\partial \eta} f(\zeta, \eta)\right| \leq \beta e^{-\lambda} C_{2,\beta,\lambda}^1. \tag{79}$$

Substituting (73), (74), (78), and (79) into (70), we have

$$\left|\Phi_{1,1}(\mathbf{X}) - \Phi_{1,1}(\mathbf{Y})\right| = \left|\Phi_{1,2,-1}^1(\mathbf{X}) - \Phi_{1,2,-1}^1(\mathbf{Y})\right|$$

$$= \left|\exp\left[f\left(X_{10}, X_{12}\right)\right] - \exp\left[f\left(Y_{10}, Y_{12}\right)\right]\right| \tag{80}$$

$$\leq C_{2,\beta,\lambda}^1 \beta e^{-\lambda}\left\{\left|X_{10} - Y_{10}\right| + \left|X_{12} - Y_{12}\right|\right\}.$$

Following the calculations for obtaining (80), we can obtain that for the 1st perturbation

$$\left|\Phi_{1,\theta}(\mathbf{X}) - \Phi_{1,\theta}(\mathbf{Y})\right|$$

$$\leq C_{2,\beta,\lambda}^1 \beta e^{-\lambda}\left\{\left|X_{10} - Y_{10}\right| + \left|X_{12} - Y_{12}\right|\right\}$$

$$(\theta = 2, 3),$$

$$\left|\Phi_{1,\theta}(\mathbf{X}) - \Phi_{1,\theta}(\mathbf{Y})\right|$$

$$\leq C_{3,\beta,\lambda}^1 \beta e^{-\lambda}\left\{\left|X_7 - Y_7\right| + \left|X_9 - Y_9\right|\right\}$$

$$(\theta = 4, 5, 6),$$

$$\tag{81}$$

where $C_{3,\beta,\lambda}^1$ is defined in Lemma 8.

Similarly we can have that for the 2nd perturbation

$$\left|\Phi_{1,\theta}(\mathbf{X}) - \Phi_{1,\theta}(\mathbf{Y})\right|$$

$$\leq C_{2,\beta,\lambda}^2 \beta e^{-\lambda}\left\{\left|X_4 - Y_4\right| + \left|X_6 - Y_6\right|\right\}$$

$$(\theta = 7, 8, 9),$$

$$\left|\Phi_{1,\theta}(\mathbf{X}) - \Phi_{1,\theta}(\mathbf{Y})\right|$$

$$\leq C_{3,\beta,\lambda}^2 \beta e^{-\lambda}\left\{\left|X_1 - Y_1\right| + \left|X_3 - Y_3\right|\right\}$$

$$(\theta = 10, 11, 12),$$

$$\tag{82}$$

where $C_{j,\beta,\lambda}^2$ $(j = 2, 3)$ are defined in Lemma 8. Therefore, it follows from (80)–(82) that

$$\left\|\mathbf{\Phi}_1(\mathbf{X}) - \mathbf{\Phi}_1(\mathbf{Y})\right\|^2 = \sum_{\theta=1}^{12}\left(\Phi_{1,\theta}(\mathbf{X}) - \Phi_{1,\theta}(\mathbf{Y})\right)^2$$

$$\leq 6 \max_{\substack{\upsilon=1,2 \\ j=2,3}}\left(C_{j,\beta,\lambda}^\upsilon\right)^2 \beta^2 e^{-2\lambda}\|\mathbf{X} - \mathbf{Y}\|^2, \tag{83}$$

which gives the desired result. □

Using Theorem 6 and Lemmas 7 and 8, we can obtain our main result.

Theorem 9. *Assume that the experimental data x_1^v ($v = 1, 2$) are contained in the codomain of ϕ. Suppose that positive constants β and λ satisfy*

$$1 - 2\beta \left(e^{-\lambda} U_{j,\beta,\lambda}^1\right)^2 \left(x_j^v\right)^2 > 0 \tag{84}$$

and

$$\sqrt{6} \max \left\{C_{j,\beta,\lambda}^v \mid v = 1, 2, \ j = 2, 3\right\} \beta e^{-\lambda} < 1, \tag{85}$$

for $v = 1, 2$, $j = 2, 3$, $U_{j,\beta,\lambda}^v$ defined in (53) and $C_{j,\beta,\lambda}^v$ defined in Lemma 8. Then sequence $\mathbf{X}_1^{(n+1)} = \mathbf{\Phi}_1(\mathbf{X}_1^{(n)})$ ($n \geq 0$) converges for any $\mathbf{X}_1^{(0)} \in [0, 1]^{12}$.

Proof. Let D be the closed subset $[0, 1]^{12}$ of R^{12} and

$$k = \sqrt{6} \max \left\{C_{j,\beta,\lambda}^v \mid v = 1, 2, \ j = 2, 3\right\} \beta e^{-\lambda}. \tag{86}$$

Then Lemmas 7 and 8 give that $\mathbf{\Phi}_1 : D \longrightarrow D$ and

$$\left\|\mathbf{\Phi}_1(\mathbf{X}) - \mathbf{\Phi}_1(\mathbf{Y})\right\| \leq k \left\|\mathbf{X} - \mathbf{Y}\right\|. \tag{87}$$

Therefore we can apply Theorem 6 and so we obtain a unique fixed point $\mathbf{X}_1^* \in D$ such that $\mathbf{\Phi}_1(\mathbf{X}_1^*) = \mathbf{X}_1^*$, which implies the convergence of sequence $\mathbf{X}_1^{(n+1)} = \mathbf{\Phi}_1(\mathbf{X}_1^{(n)})$ ($n \geq 0$) for any $\mathbf{X}_1^{(0)} \in [0, 1]^{12}$ with limit $\mathbf{X}_1^*$. □

Remark 10. We define

$$\begin{pmatrix} \rho^1\left(w_{1,2} = -1\right), \rho^1\left(w_{1,2} = 0\right), \rho^1\left(w_{1,2} = 1\right), \\ \rho^1\left(w_{1,3} = -1\right), \rho^1\left(w_{1,3} = 0\right), \rho^1\left(w_{1,3} = 1\right), \\ \rho^2\left(w_{1,2} = -1\right), \rho^2\left(w_{1,2} = 0\right), \rho^2\left(w_{1,2} = 1\right), \\ \rho^2\left(w_{1,3} = -1\right), \rho^2\left(w_{1,3} = 0\right), \rho^2\left(w_{1,3} = 1\right) \end{pmatrix} \tag{88}$$

$$= \begin{pmatrix} \rho^1_{1,2,-1}, \rho^1_{1,2,0}, \rho^1_{1,2,1}, \\ \rho^1_{1,3,-1}, \rho^1_{1,3,0}, \rho^1_{1,3,1}, \\ \rho^2_{1,2,-1}, \rho^2_{1,2,0}, \rho^2_{1,2,1}, \\ \rho^2_{1,3,-1}, \rho^2_{1,3,0}, \rho^2_{1,3,1} \end{pmatrix},$$

where the vector in the right side denotes limit $\mathbf{X}_1^*$ in Theorem 9. Since $\{\mathbf{X}_1^{(n)}\}$ in (57)–(59) converges, $\{P_{1,j,n}^v(w)\}$ in (48) also converges to limit $e^{-\lambda\delta(w)}\rho^{3-v}_{1,j,w}$. Therefore, $\rho^v(w_{1,j})$ in (88) and $P^v(w_{1,j}) = e^{-\lambda\delta(w_{1,j})}\rho^{3-v}_{1,j,w_{1,j}}$ satisfy the system of equations in (34) and (35), which give approximate marginal PMFs $P(w_{1,j}) = (1/Z_{1,j})e^{-\lambda\delta(w_{1,j})}\prod_{v=1}^2 \rho^v(w_{1,j})$ in (14) as follows:

$$P\left(w_{1,j} = w\right) = \frac{1}{Z_{1,j}} e^{-\lambda\delta(w_{1,j}=w)} \prod_{v=1}^2 \rho^v_{1,j,w} \tag{89}$$

$$\left(j = 2, 3, \ w = -1, 0, 1\right).$$

Remark 11. When using the sequences in (54)–(56) instead of the sequences in (47)–(49), we can obtain the limits of the sequences in (54)–(56) under conditions similar to

those in Theorem 9: assume that experimental data x_2^v ($v = 1, 2$) are contained in the codomain of ϕ. Let $U_{j,\beta,\lambda}^v = (1/(1 + 2e^{-\lambda})) \exp[\beta\{|\phi^{-1}(x_2^v)| + |x_j^v| + 2|x_{4-j}^v|\}^2]$ ($j = 1, 3$). Suppose that positive constants β and λ satisfy that $1 - 2\beta(e^{-\lambda}U_{j,\beta,\lambda}^1)^2(x_j^v)^2 > 0$ and

$$\sqrt{6} \max \left\{C_{j,\beta,\lambda}^v \mid v = 1, 2, \ j = 1, 3\right\} \beta e^{-\lambda} < 1, \tag{90}$$

where

$$\begin{aligned} C_{j,\beta,\lambda}^v &= \left[2 \left\{\left|\phi^{-1}\left(x_2^v\right)\right| + \left|x_j^v\right| + e^{-\lambda} U_{4-j,\beta,\lambda}^v \left|x_{4-j}^v\right|\right\}\right. \\ &\quad \cdot U_{4-j,\beta,\lambda}^v \left|x_{4-j}^v\right| \\ &\quad \times \left[1 + 2\beta e^{-\lambda} U_{4-j,\beta,\lambda}^v \left(2 + e^{-\lambda} U_{4-j,\beta,\lambda}^v\right) \left(x_{4-j}^v\right)^2\right] \\ &\quad + \left\{\left|\phi^{-1}\left(x_2^v\right)\right| + \left|x_j^v\right| + e^{-\lambda} U_{4-j,\beta,\lambda}^v \left|x_{4-j}^v\right|\right\}^2 \\ &\quad \times 2\beta U_{4-j,\beta,\lambda}^v \left(1 + 2e^{-\lambda} U_{4-j,\beta,\lambda}^v\right) \left(x_{4-j}^v\right)^2\right] \times \Big\{1 \\ &\quad - 2\beta \left(e^{-\lambda} U_{4-j,\beta,\lambda}^1\right)^2 \left(x_{4-j}^v\right)^2\Big\}^{-2}. \end{aligned} \tag{91}$$

We can define the messages as in Remark 4

$$\begin{pmatrix} \rho^1\left(w_{2,1} = -1\right), \rho^1\left(w_{2,1} = 0\right), \rho^1\left(w_{2,1} = 1\right), \\ \rho^1\left(w_{2,3} = -1\right), \rho^1\left(w_{2,3} = 0\right), \rho^1\left(w_{2,3} = 1\right), \\ \rho^2\left(w_{2,1} = -1\right), \rho^2\left(w_{2,1} = 0\right), \rho^2\left(w_{2,1} = 1\right), \\ \rho^2\left(w_{2,3} = -1\right), \rho^2\left(w_{2,3} = 0\right), \rho^2\left(w_{2,3} = 1\right) \end{pmatrix} \tag{92}$$

$$= \begin{pmatrix} \rho^1_{2,1,-1}, \rho^1_{2,1,0}, \rho^1_{2,1,1}, \\ \rho^1_{2,3,-1}, \rho^1_{2,3,0}, \rho^1_{2,3,1}, \\ \rho^2_{2,1,-1}, \rho^2_{2,1,0}, \rho^2_{2,1,1}, \\ \rho^2_{2,3,-1}, \rho^2_{2,3,0}, \rho^2_{2,3,1} \end{pmatrix}$$

and then obtain the approximate marginal PMFs

$$P\left(w_{2,j} = w\right) = \frac{1}{Z_{2,j}} e^{-\lambda\delta(w_{2,j}=w)} \prod_{v=1}^2 \rho^v_{2,j,w} \tag{93}$$

$$\left(j = 1, 3, \ w = -1, 0, 1\right).$$

4. Numerical Examples

In this section, we verify our main result and show its application.

Step 1 (generation of artificial experimental data). We randomly generate 100 sets of artificial experimental data, where each set consists of the six values of x_j^v ($v = 1, 2$, $j = 1, 2, 3$) as in Figure 3(a). Every randomly generated experimental data has a value contained in the open interval $(-1, 1)$, the codomain of $\phi(x) = \tanh(x)$, and so the condition on the experimental data in Theorem 9 is satisfied.
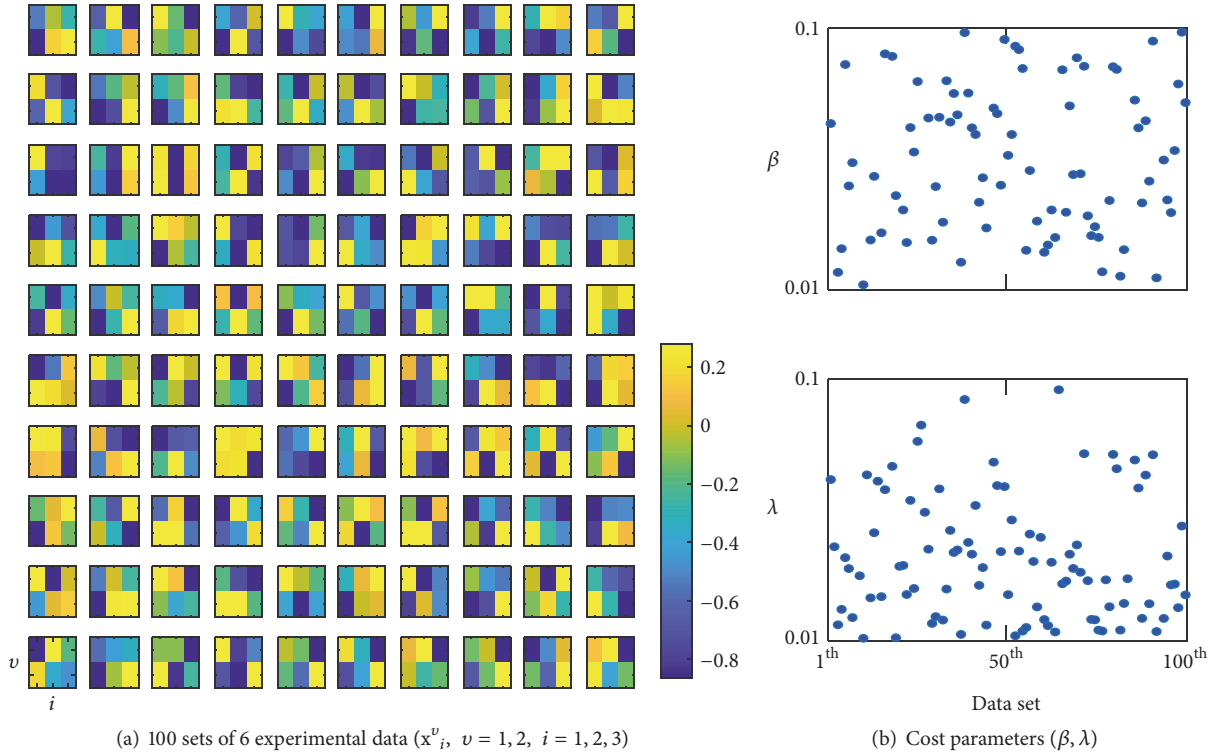
(a) 100 sets of 6 experimental data (x^v_i, $v = 1, 2$, $i = 1, 2, 3$)

(b) Cost parameters (β, λ)

FIGURE 3: Experimental data and cost parameters. (a) Each of 100 color matrices denotes the values of 6 artificial experimental data x^v_i ($v = 1, 2$, $i = 1, 2, 3$). Values of the x- and y-axis in each color matrix denote the values of i and v, respectively, where the color in the entry (i, v) denotes the value of x^v_i. (b) A value on the x-axis denotes an experimental data set in Figure 3(a). The data set at entry (i, j) in Figure 3(a) is represented as the $10(i - 1) + j^{th}$ data set in Figure 3(b). There are few values of β and λ greater than 0.1 and we do not show them here.

Step 2 (determination of cost parameters). We find a pair of positive cost parameters (β, λ) as in Figure 3(b) for each experimental data set, which satisfy the condition on (β, λ) in Theorem 9.

Step 3 (generation of values of initial terms in sequences). We randomly generate initial values of 24 terms $\rho^v_{i,j,0}(w)$ ($v = 1, 2$, $i = 1, 2$, $j \neq i$, $1 \leq j \leq 3$, $w = -1, 0, 1$) for each experimental data set, which satisfy the condition on initial values in Theorem 9 (see Figure 4).

Step 4 (convergence of the sequences). We obtain the convergence of all sequences $\rho^v_{i,j,n}(w)$ ($v = 1, 2$, $i = 1, 2$, $j \neq i$, $1 \leq j \leq 3$, $w = -1, 0, 1$) for each experimental data set in Figure 3(a), (β, λ) in Figure 3(b), and initial values in Figure 4, where every sequence converges within 14 iterations (see Figure 5). Due to the recursive relations in (47), (48), (49), (53), (54), (55), and (56), we can divide all sequences into four groups ($\{\rho^1_{1,2,n}(w), \rho^2_{1,3,n}(w)\}$, $\{\rho^1_{1,3,n}(w), \rho^2_{1,2,n}(w)\}$, $\{\rho^1_{2,1,n}(w), \rho^2_{2,3,n}(w)\}$, and $\{\rho^1_{2,3,n}(w), \rho^2_{2,1,n}(w)\}$) and show their convergences as in Figure 5.

Step 5 (approximation of marginal PMFs). Substituting the limits (Figure 6(c)) of the sequences for the 1st experimental data set (Figure 6(a)) and its corresponding cost parameters (β, λ) (Figure 6(a)) into (89) and (93), we obtain the approximate marginal PMFs $P(w_{1,2} = w)$, $P(w_{1,3} = w)$, $P(w_{2,1} = w)$,

and $P(w_{2,3} = w)$ as in Figure 6(d), where the maximum value of each PMF is marked as red.

Step 6 (construction of multiple networks and their mathematical formulas from the approximate marginal PMFs). We can use the maximum values of four link weights to make a network and its corresponding mathematical formulas (Figure 6(d), first grey box), which could be considered as the most possible network model of the artificial biological system. In addition, since value $P(w_{2,1} = 0)$ is almost equal to value $P(w_{2,1} = 1)$, we can also consider the network with $w_{2,1} = 0$ and its mathematical formula as in the second grey box in Figure 6(d). We present two networks and their mathematical formulas in the two grey boxes to show that there could be multiple networks for explaining the artificial biological system by using the approach in [5]. Similarly, in cases where a given PKN has more nodes and discrete values of link weights, we could obtain multiple networks and their corresponding mathematical formulas, where probability $P(w_{i,j} = w)$ is high.

5. Conclusions and Future Work

In this paper, using a simple network with artificial experimental data, we have presented a detailed explanation of the whole process for determining approximate marginal PMFs of link weights in a prior knowledge network based on BP

(a) 100 sets of 12 initial values ($\rho_{1,j,0}^{v}(w)$, $v = 1, 2$, $j = 2, 3$, $w = -1, 0, 1$)

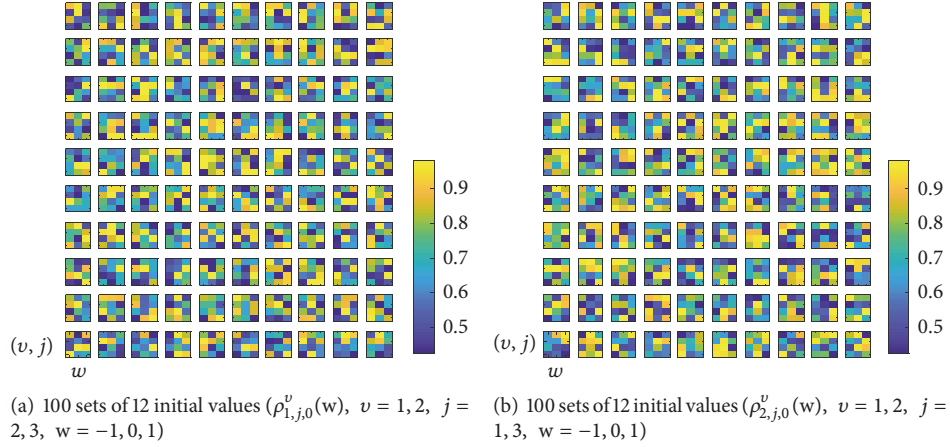(b) 100 sets of 12 initial values ($\rho_{2,j,0}^{v}(w)$, $v = 1, 2$, $j = 1, 3$, $w = -1, 0, 1$)

FIGURE 4: Initial values for the sequences. Due to the recursive relations in (47), (48), (54), and (55), it is enough to independently consider the convergences of the two types of the sequences $\rho_{1,j,n}^{v}(w)$ ($v = 1, 2$, $j = 2, 3$, $w = -1, 0, 1$) and $\rho_{2,j,n}^{v}(w)$ ($v = 1, 2$, $j = 1, 3$, $w = -1, 0, 1$). (a) Each set contains values of 12 initial terms $\rho_{1,j,0}^{v}(w)$ ($v = 1, 2$, $j = 2, 3$, $w = -1, 0, 1$) for each experimental data set. (b) Each set contains values of 12 initial terms $\rho_{2,j,0}^{v}(w)$ ($v = 1, 2$, $j = 1, 3$, $w = -1, 0, 1$) for each experimental data set.

(a)

(b)

(c)

(d)
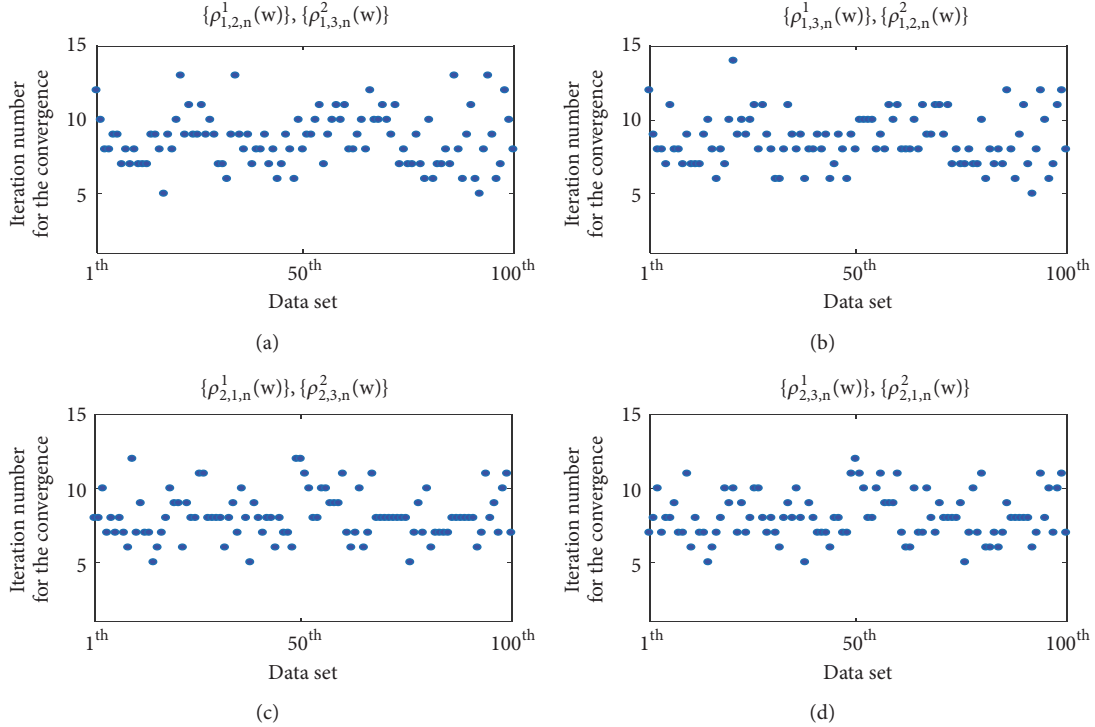
FIGURE 5: Iteration number for the convergence. (a) The system of recursive relations for $\rho_{1,2,n}^{v}(w)$ and $\rho_{1,3,n}^{v}(w)$ ($v = 1, 2$, $w = -1, 0, 1$) converges within 10 iterations for tolerance 10^{-10}. Symbols and data in Figures 5(b), 5(c), and 5(d) have similar meanings as in Figure 5(a).

on the factor graph. For the factorization of the joint PMF of link weights, the Boltzmann/Gibbs form has been used. And we have shown a sufficient condition for the convergence of beliefs, which leads to determining the approximate marginal PMFs. Since our simple PKN and experimental data have key components necessary for the application of BP in determining approximate marginal PMFs, our results can be extended to any PKN with perturbation data, including the PKN in [5], which is our future study.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

(a)

		Node			Cost parameters	
		x1	x2	x3	β	λ
Experiment	1st	0.7059	0.6807	-0.4108	0.0226	0.0170
	2nd	0.4996	0.5448	0.6262		

(b)

12 initial values for the sequences $\{\rho^{v}_{1,j,n}(w)\}$ ($v=1,2$, $j=2,3$, $w=-1,0,1$)

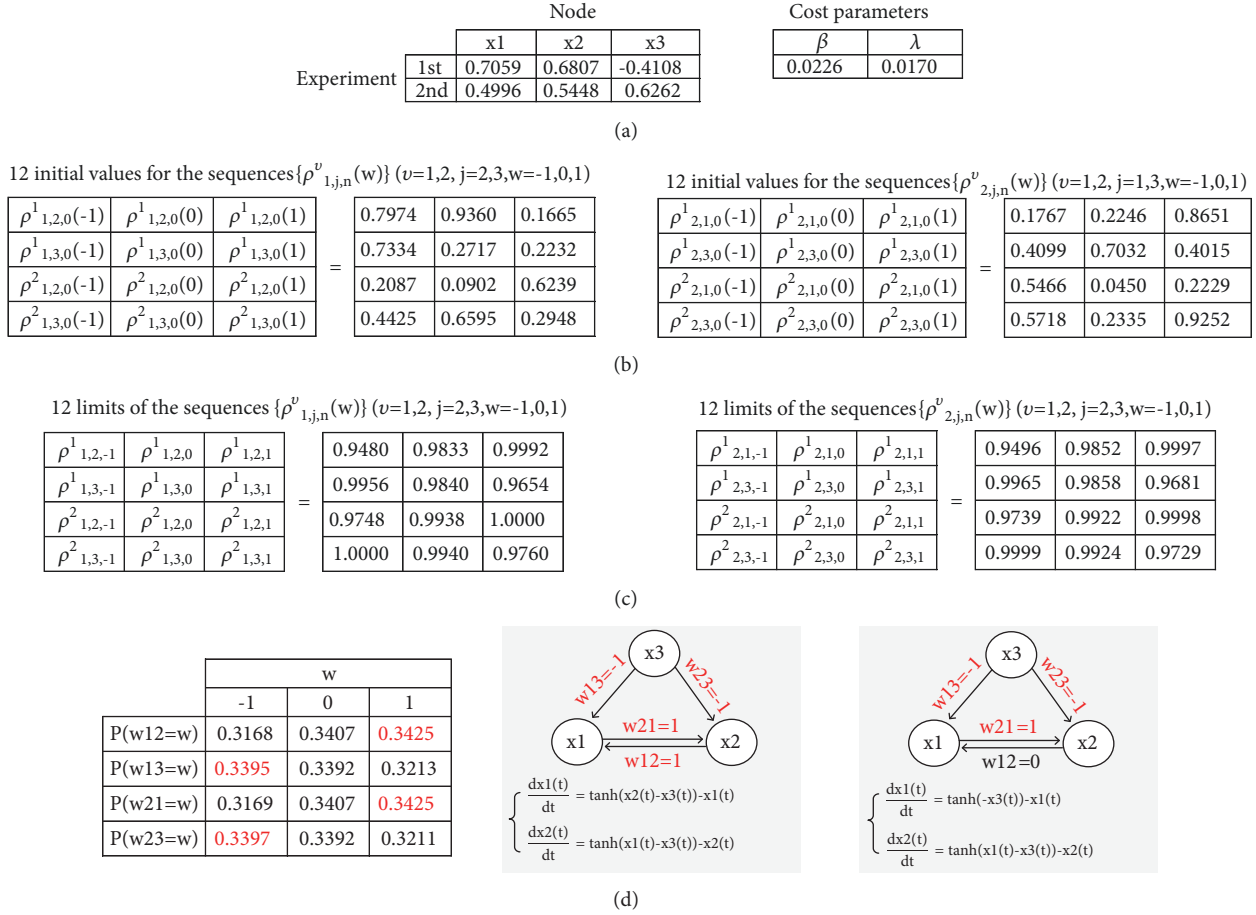|						
$\rho^1_{1,2,0}(-1)$	$\rho^1_{1,2,0}(0)$	$\rho^1_{1,2,0}(1)$		0.7974	0.9360	0.1665
$\rho^1_{1,3,0}(-1)$	$\rho^1_{1,3,0}(0)$	$\rho^1_{1,3,0}(1)$	$=$	0.7334	0.2717	0.2232
$\rho^2_{1,2,0}(-1)$	$\rho^2_{1,2,0}(0)$	$\rho^2_{1,2,0}(1)$		0.2087	0.0902	0.6239
$\rho^2_{1,3,0}(-1)$	$\rho^2_{1,3,0}(0)$	$\rho^2_{1,3,0}(1)$		0.4425	0.6595	0.2948

12 initial values for the sequences $\{\rho^{v}_{2,j,n}(w)\}$ ($v=1,2$, $j=1,3$, $w=-1,0,1$)

$\rho^1_{2,1,0}(-1)$	$\rho^1_{2,1,0}(0)$	$\rho^1_{2,1,0}(1)$		0.1767	0.2246	0.8651
$\rho^1_{2,3,0}(-1)$	$\rho^1_{2,3,0}(0)$	$\rho^1_{2,3,0}(1)$	$=$	0.4099	0.7032	0.4015
$\rho^2_{2,1,0}(-1)$	$\rho^2_{2,1,0}(0)$	$\rho^2_{2,1,0}(1)$		0.5466	0.0450	0.2229
$\rho^2_{2,3,0}(-1)$	$\rho^2_{2,3,0}(0)$	$\rho^2_{2,3,0}(1)$		0.5718	0.2335	0.9252

(c)

12 limits of the sequences $\{\rho^{v}_{1,j,n}(w)\}$ ($v=1,2$, $j=2,3$, $w=-1,0,1$)

$\rho^1_{1,2,-1}$	$\rho^1_{1,2,0}$	$\rho^1_{1,2,1}$		0.9480	0.9833	0.9992
$\rho^1_{1,3,-1}$	$\rho^1_{1,3,0}$	$\rho^1_{1,3,1}$	$=$	0.9956	0.9840	0.9654
$\rho^2_{1,2,-1}$	$\rho^2_{1,2,0}$	$\rho^2_{1,2,1}$		0.9748	0.9938	1.0000
$\rho^2_{1,3,-1}$	$\rho^2_{1,3,0}$	$\rho^2_{1,3,1}$		1.0000	0.9940	0.9760

12 limits of the sequences $\{\rho^{v}_{2,j,n}(w)\}$ ($v=1,2$, $j=2,3$, $w=-1,0,1$)

$\rho^1_{2,1,-1}$	$\rho^1_{2,1,0}$	$\rho^1_{2,1,1}$		0.9496	0.9852	0.9997
$\rho^1_{2,3,-1}$	$\rho^1_{2,3,0}$	$\rho^1_{2,3,1}$	$=$	0.9965	0.9858	0.9681
$\rho^2_{2,1,-1}$	$\rho^2_{2,1,0}$	$\rho^2_{2,1,1}$		0.9739	0.9922	0.9998
$\rho^2_{2,3,-1}$	$\rho^2_{2,3,0}$	$\rho^2_{2,3,1}$		0.9999	0.9924	0.9729

(d)

	w		
	-1	0	1
$P(w12=w)$	0.3168	0.3407	0.3425
$P(w13=w)$	0.3395	0.3392	0.3213
$P(w21=w)$	0.3169	0.3407	0.3425
$P(w23=w)$	0.3397	0.3392	0.3211

First network ($w13=-1$, $w23=-1$, $w21=1$, $w12=1$):

$$\begin{cases} \dfrac{dx1(t)}{dt} = \tanh(x2(t)-x3(t))-x1(t) \\ \dfrac{dx2(t)}{dt} = \tanh(x1(t)-x3(t))-x2(t) \end{cases}$$

Second network ($w13=-1$, $w23=-1$, $w21=1$, $w12=0$):

$$\begin{cases} \dfrac{dx1(t)}{dt} = \tanh(-x3(t))-x1(t) \\ \dfrac{dx2(t)}{dt} = \tanh(x1(t)-x3(t))-x2(t) \end{cases}$$

FIGURE 6: Approximate marginal PMFs. (a) The experimental data and cost parameters (β, λ) are the 61th data set and its corresponding values of (β, λ) in Figures 3(a) and 3(b), respectively. (b) 12 initial values in the left and right denote the values in the 61th initial data set in Figures 4(a) and 4(b), respectively. (c) 12 limit values in the top and bottom are obtained for Figures 6(a) and 6(b). (d) The approximate marginal PMFs are obtained from (89) and (93). The red numbers denote the maximum values of the marginal PMFs. The first network and its mathematical model in the right denote those in Figures 1(a) and 1(c) with link weights of maximum values. Replacing $w12 = 1$ in the first network with $w12 = 0$ gives the second network and its mathematical model.

Authors' Contributions

All authors carried out the main results and completed the corresponding proof. All authors read and approved the final manuscript.

Acknowledgments

References

[1] W. Materi and D. S. Wishart, "Computational systems biology in cancer: modeling methods and applications," *Gene Regulation and Systems Biology*, vol. 1, pp. 91–110, 2007.

[2] H. F. Fumiã and M. L. Martins, "Boolean Network model for cancer pathways: predicting carcinogenesis and targeted therapy outcomes," *PLoS ONE*, vol. 8, no. 7, p. e69008, 2013.

[3] A. Barbagallo and M. A. Ragusa, "On Lagrange duality theory for dynamics vaccination games," *Ricerche di Matematica*, vol. 67, no. 2, pp. 969–982, 2018.

[4] R. Costa and P. Pavone, "Diachronic biodiversity analysis of a metropolitan area in the Mediterranean region," *Acta Horticulturae*, Article ID 1215_8, 2018.

[5] E. J. Molinelli, A. Korkut, W. Wang et al., "Perturbation biology: inferring signaling networks in cellular systems," *PLoS Computational Biology*, vol. 9, no. 12, Article ID e1003290, 2013.

[6] A. Korkut, W. Wang et al., "Perturbation biology nominates upstream-downstream drug combinations in RAF inhibitor resistant melanoma cells," *Elife*, vol. 4, Article ID e04640, 2015.

[7] L. D. Landau, L. P. Pitaevskii, and E. M. Lifshits, *Statistical Physics: Transl. from the Russian by J. B. Sykes and M. J. Kearsley*, Pergamon Press, 3rd, Rev. and Enl. edition, 1980.

[8] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.

[9] H. Loeliger, "An Introduction to factor graphs," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 28–41, 2004.

[10] H. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, "The factor graph approach to model-based signal processing," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, 2007.

[11] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, NY, USA, 2006.

[12] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT press, 2012.

[13] R. G. Gallager, "Low-density parity-check codes," *IEEE Transactions on Information Theory*, vol. 8, pp. 21–28, 1962.

[14] R. G. Gallager, *Low Density Parity Check Codes*, vol. 21 of *Research Monograph Series*, 1963.

[15] D. J. MacKay and R. M. Neal, "Good codes based on very sparse matrices," in *Proceedings of the IMA International Conference on Cryptography and Coding*, pp. 100–111, Springer, Berlin, Germany, 1995.

[16] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng, "Turbo decoding as an instance of Pearl's 'belief propagation' algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 140–152, 1998.

[17] T. Heskes, "Stable fixed points of loopy belief propagation are local minima of the bethe free energy," in *Advances in Neural Information Processing Systems*, pp. 359–366, 2003.

[18] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Understanding belief propagation and its generalizations," *Exploring Artificial Intelligence in the New Millennium*, vol. 8, pp. 236–239, 2003.

[19] Q. Su and Y. Wu, "On convergence conditions of gaussian belief propagation," *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1144–1155, 2015.

[20] Z. Zhang and M. Fu, "On convergence rate of the gaussian belief propagation algorithm for markov networks," https://arxiv.org/abs/1903.02658.

[21] M. A. Khamsi and A. K. William, *An Introduction to Metric Spaces And Fixed Point Theory*, vol. 53, John Wiley & Sons, 2011.

[22] M. A. Ragusa, "Local Hölder regularity for solutions of elliptic systems," *Duke Mathematical Journal*, vol. 113, no. 2, pp. 385–397, 2002.

[23] A. Braunstein, A. Pagnani, M. Weigt, and R. Zecchina, "Inference algorithms for gene networks: a statistical mechanics analysis," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, Article ID P12001, 2008.

[24] D. R. Hunter, *Notes for a Graduate-Level Course in Asymptotics for Statisticians*, Penn State University, Pennsylvania, Penn, USA, 2014.