Title
Statistical Models for Cognitive Social Structures

Permalink
https://escholarship.org/uc/item/19t22037

Author
Shao, Kanghong

Publication Date
2018

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA

Los Angeles

Statistical Models for Cognitive Social Structures

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Science in Statistics

by

Kanghong Shao

2018
Cognitive social structures (CSS) is an area in social network research that has enduring importance but lacks flexible models. In this paper we consider statistical models for CSS systems where we observe a three-dimensional binary array of relational ties characterized by the “sender” of the relation, the “receiver” of the relation, and the “perceiver” of the relation from the ”sender” to the “receiver”. Such systems have been represented as networks by Krackhardt [1]. Durante, Dunson and Vogelstein proposed a flexible Bayesian nonparametric approach [2] to model the population distribution of network valued data, in which the joint distribution of the edges probabilities is defined through a mixture model that reduces dimensionality and incorporates information within each mixture component based on latent space models. Inspired by this work, we modify the model to characterize cognitive social structures by adding a parameter for cognitive error. As a case study, we apply our model
to Krackhardt’s network data of 21 managers in a high-tech machine manufacturing firm. The results show distinct effects of cognitive error and illustrate that our model is capable of characterizing cognitive social structures. The results also motivate future improvement on transitivity and triangle relations.
The thesis of Kanghong Shao is approved.

Tao Gao

Peter M. Bentler

Mark Stephen Handcock, Committee Chair

University of California, Los Angeles

2018
# TABLE OF CONTENTS

1 Introduction ......................................................... 1

2 Overview ......................................................... 4

2.1 Mixture Modeling of Populations of Networks .................. 4

3 Methodology ..................................................... 7

3.1 Model ......................................................... 7

3.2 Prior Specification ............................................ 9

3.3 Posterior Computation ....................................... 9

3.3.1 Posterior Computation When There Is No Cognitive Error .. 10

3.3.2 Posterior Computation When There May Be Cognitive Error . 14

4 Application to Cognitive Networks among High-Tech Managers .... 27

4.1 Data ............................................................ 27

4.2 Goodness-of-Fit Checks ....................................... 28

4.3 Results ........................................................ 30

5 Discussion ....................................................... 35

References .......................................................... 38
LIST OF FIGURES

4.1 Posterior Predictive Distribution for $\theta = 0$ ........................................ 29
4.2 Posterior Predictive Distribution for $\theta \neq 0$ ........................................ 29
4.3 Posterior Distribution for $\theta$ with a prior centered at zero $\mu_\theta = 0$ ........ 31
4.4 Posterior Distribution for $\theta$ with a prior centered at $3 \mu_\theta = 3.$ ........... 31
4.5 Posterior Tie Probabilities $\bar{\pi}_t$ Removing the Cognitive Error for $\theta = 0$ . 34
4.6 Posterior Tie Probabilities $\bar{\pi}_t$ Removing the Cognitive Error for $\theta \neq 0$ . 34
LIST OF TABLES

4.1 Tie Densities between Ego-Alter Pairs and Alter-Alter Pairs .............. 32
4.2 Posterior Component Possibilities ............................................. 32
4.3 Predictive Summary Measures .................................................. 33
4.4 Observed Summary Measures .................................................. 33
CHAPTER 1

Introduction

Cognitive social structures (CSS) has been of enduring importance in social network research. This is an inevitable issue because network representations of a group’s social structure are often derived from information provided by each of the members of the group [4]. While social network analysis (SNA) describes the actual configurations of ties of individuals, CSS instead focuses on the patterns of interactions as perceived by individuals [3]. David Krackhardt [1] formalized the definition of cognitive social structures and provided an empirical example of 21 managers from a small high-tech, machine manufacturing firm to demonstrate the features of cognitive social structures and his approach to modeling it. A cognitive social structure can be represented as $R_{i,j,k}$, where $i$ is the ”sender” of the relation, $j$ is the ”receiver” of the relation, and $k$ is the ”perceiver” of the relation from $i$ to $j$. And $R_{i,j,k} = 1$ would be interpreted as meaning that person $k$ thinks that person $i$ forms a directed tie with $j$. By paying attention to the cognitive structures of social networks, CSS has broaden and deepen the field of social networks. It intends to reveal underlying cognitive schemas for social relationships [1] and provides with a different explanation of the constructs that social network surveys want to measure [3]. Furthermore, the study of cognition level helps us better understand how individuals’ surrounding social networks affect
their outcomes by emphasizing individuals’ perceptions of their networks [1]. Cognitive error is of great importance in studying CSS. As discussed in [4], reported network representations of a group’s social structure may be deceptive because there are systematic individual biases and knowledge differences (cognitive errors). These biases suggest an egocentric view of the social structure as perceived from each individual’s vantage point. However, we notice that CSS lacks flexible statistical models. It might be largely due to practical data collection problems since getting a respondent to voluntarily record his/her perception of every \((i, j)\) dyad for each relation is a very difficult task [1].

On the other hand, the modeling of population of networks has become a more popular research area in recent years because of the increased scale of network data. Durante, Dunson and Vogelstein proposed a flexible Bayesian nonparametric approach for modeling the population distribution of network valued data via a mixture model [2]. Although the model reduces dimensionality and can flexibly characterize a broad variety of generative mechanism, it lacks the ability to model the population of networks with CSS. In particular, they assume that the edge probabilities conditional on the mixture component \(h\) are independent and identically distributed. But the model does not specify the difference between Ego-Alter pairs (pairs between ego and its alters) and Alter-Alter pairs (pairs between ego’s alters), in which the cognitive error might occur. In this paper, we proposed a modified model based on their model but allows different distributions of the ensemble due to cognition error, which makes our model a more general case and their model a special case of ours. In order to model cognitive social structure, we add a new parameter \(\theta\) to capture the cognitive error in individual perceptions. We focus on a simple and interpretable cognition structure where
each ego has clear perception of their own ties and has poor perception of ties between their alters due to an individual’s lack of knowledge and information processing limits of human mind.

In the remainder of the paper, we overview the model from [2], describe the details of our proposed model, apply the model to the data collected by David Krackhardt in [1] and discuss its advantages, weakness and possible future improvements.
CHAPTER 2

Overview

Firstly we make an overview of the model proposed by Durante, Dunson and Vogelstein in [2]. They try to model the population distribution of network valued data by a flexible Bayesian nonparametric method. The priors, likelihoods and posterior computations are explicitly formulated with details in Section 3.2 and 3.3.1. Based on latent space representations [6], they define the joint distribution of the edge probability by a mixture model such that the edge probabilities conditional on the component $h$ are independent and identically distributed. The full model is reviewed below.

2.1 Mixture Modeling of Populations of Networks

First we develop a notation for populations of networks. We focus throughout on undirected networks. For an undirected network, $A$, over $V$ nodes, let $l = V(j - 1) - j(j - 1)/2 + i - j$ index the tie variable $A_{i,j}$, $i > j$. This is clearly a 1-1 as the column index $j$ is the $\min_k\{k : l < Vk - k(k+1)/2 - k - 1\}$ and $i$ is $l - V(j - 1) - j(j - 1)/2 - j$. Let $c(l) = \{(i,j) : j = \min_k\{k : l < Vk - k(k+1)/2 - k - 1\}, i = l - V(j - 1) - j(j - 1)/2 - j\}$. From now on let $l$ index the tie variables in this way. Also if $M$ is a $V \times V$ matrix, let $\text{vec}(M) = \{M_{c(l)} : l = 1, \ldots, V(V - 1)/2\}$. 
We further define a multivariate network-valued random variable $\mathcal{L}(A) = (A_{21}, A_{31}, \ldots, A_{V1}, A_{32}, \ldots, A_{V2}, \ldots, A_{V(V-1)})^T$ with binary entries $\mathcal{L}(A)_l \in 0, 1$, for $l = 1, \ldots, V(V-1)/2$. $\mathcal{L}(A)$ can be interpreted as a categorical random variable with each category being one of the possible network configurations, hence the probability mass function $p_{\mathcal{L}(A)}(a)$ is the probability of the random variable $\mathcal{L}(A)$ being the network configuration $a$. To fix notation:

$$p_{\mathcal{L}(A)}(a) = pr\{\mathcal{L}(A) = a\} = \sum_{h=1}^{H} \nu(h) \prod_{l=1}^{L} (\pi_l^{(h)})^{a_l} (1 - \pi_l^{(h)})^{1-a_l}$$ (2.1)

$$\pi_l^{(h)} = \frac{1}{1 + e^{-(Z + D^{(h)})}}$$ (2.2)

$$D^{(h)} = \text{vec}\left(\mathcal{L}(X^{(h)}) \Lambda^{(h)} X^{(h)\top}\right)$$ (2.3)

where $Z = \text{vec}(\{Z_{c(l)}\}_{l=1}^{L})$ is the similarity vector, $D^{(h)} = \text{vec}(\{D^{(h)}_{c(l)}\}_{l=1}^{L})$ is the component-specific deviation, and $L = V(V-1)/2$. $\nu(h) \in (0, 1)$ is defined as the probability of to be in mixture component $h$ and $\pi_l^{(h)} \in (0, 1)$ is the probability of an edge forms between the $l$th pair of nodes in mixture component $h$. $X^{(h)}$ is a $V \times R$ matrix with each row $X^{(h)}_l$ representing the $R$ latent coordinates of each node $v$ in mixture component $h$. $\Lambda^{(h)}$ is defined as a $R \times R$ diagonal matrix with each diagonal element $\lambda^{(h)}_r \geq 0$ the weight of the $r$th dimension in
mixture component $h$ for defining the component-specific deviation $D^{(h)}$. 
CHAPTER 3

Methodology

The model in Section 2.1 is for a population of independent and identically distributed networks. In this section we enhance the model to represent cognitive social structures. We write down a model for each network in the CSS and modify (2.2) by adding a cognition level vector $C_{ih}$ to the component-specific edge probability vector (see equations (3.2) to (3.4)).

3.1 Model

The cognitive network for person $i$ is modeled as:

$$p_{\mathcal{L}(A_i)}(a) = pr\{\mathcal{L}(A_i) = a\} = \sum_{h=1}^{H} \nu(h) \prod_{l=1}^{L} (\pi_{il}^{(h)})^{a_l} (1 - \pi_{il}^{(h)})^{1-a_l} \quad (3.1)$$

$$\pi_{il}^{(h)} = \frac{1}{1 + e^{-\psi_{il}^{(h)}}} \quad (3.2)$$
\[ \psi_{il}^{(h)} = Z_l + C_{il} + D_l^{(h)} \]  

(3.3)

\[ D_l^{(h)} = \mathcal{L}(X^{(h)} \Lambda^{(h)} X^{(h)T})_l \]  

(3.4)

where \( Z = \text{vec}(\{Z_{c(l)}\}_{l=1}^L) \) is the similarity vector, \( D^{(h)} \text{vec}(\{D_{c(l)}^{(h)}\}_{l=1}^L) \) is the component-specific deviation, and \( L = V(V - 1)/2 \). where \( Z_l \) is the similarity vector for \( l \)th pair of nodes, \( D_l^{(h)} \) is the component-specific deviation, and \( C_{il} \) is the cognition level for the \( l \)th pair of people in the reported network of the \( i \)th individual (ego). Different cognitive perspectives are represented by mixture components. Cognitive error is quantified by the cognition level.

While complex forms of cognition are straightforward to model, here we focus on a simple and interpretable form of cognition where each ego has clear perception of their own ties and has poor perception of ties between their alters. Specifically, we assume they have clear perception of ties between some alters and also have false perception of a proportion of alter-to-alter ties. Here we assume that the cognition is independent and identically distributed across pairs of alter nodes. For simplicity we assume it is the same for each ego. These restrictions can be lifted. Under this model

\[ C_{il} = \begin{cases} 
0 & \text{one of the nodes in pair } c(l)isi \\
-\theta & \text{otherwise}
\end{cases} \]
3.2 Prior Specification

The similarity vector $Z$ is assumed to have a prior distribution given by:

$$Z \sim N_L(\mu_Z, \Sigma_Z)$$  \hspace{1cm} (3.5)

$$\mu_Z \in \mathcal{R}^L$$  \hspace{1cm} (3.6)

$$\Sigma_Z = \text{diag}(\sigma_1^2, \ldots, \sigma_L^2)$$  \hspace{1cm} (3.7)

For the case considered above, the cognition vectors $C_{il}$ are parametrized by $\theta$, which is assumed to have a prior distribution given by:

$$\theta \sim N(\mu_\theta, \sigma_\theta^2)$$  \hspace{1cm} (3.8)

$$\mu_\theta \in \mathcal{R}, \quad \sigma_\theta^2 > 0$$  \hspace{1cm} (3.9)

3.3 Posterior Computation

First we verify the method in [2], which corresponds to our situation where there is no cognitive error, that is, $\theta = 0$ so that $C_l = 0$, which means $\psi^{(h)}_{il} = Z_l + D_l^{(h)}$. Then we extend the algorithm to deal with our model, for which it is possible that $\theta \neq 0$. 

9
3.3.1 Posterior Computation When There Is No Cognitive Error

From Step [3] of Algorithm 1 in [2], we have:

$$\omega_i^{(h)}|\sim \text{PG}(n_h, \psi_i^{(h)})$$

Hence Step [4] of Algorithm 1 to sample from the full conditional distribution of $Z$ can be derived using variants of results in [2]. The likelihood of the aggregation of $l$th tie variables, $Y_l^{(h)}$ is

$$L_{lh}(\psi_l^{(h)}) = \left\{\exp(\psi_l^{(h)})\right\}^{Y_l^{(h)}} \over 1 + \exp(\psi_l^{(h)}) .$$

Theorem 1 in [5] states: Let $p(\omega)$ denote the density of the random $\omega \sim \text{PG}(b, 0)$ and $b > 0$. Then the following integral identity holds for all $a \in \mathbb{R}$:

$$\frac{(e^\psi)^a}{(1 + e^\psi)^b} = 2^{-b} e^{\psi k} \int_0^{\infty} e^{-\omega^2/2} p(\omega) d\omega$$

where $k = a - b/2$ and the conditional distribution of $\omega$ given $\psi$ is:

$$p(\omega|\psi) = \frac{e^{-\omega^2/2} p(\omega)}{\int_0^{\infty} e^{-\omega^2/2} p(\omega) d\omega}$$

Based on Theorem 1, we can write the likelihood contribution of the aggregation of $l$th tie variables in component $h$, $Y_l^{(h)}$, as:

$$L_{lh}(\psi) = \exp(\kappa_l^{(h)} \psi_l^{(h)}) \int_0^{\infty} e^{-\frac{\omega_l^{(h)} \psi_l^{(h)}^2}{2}} p(\omega_l^{(h)}) d\omega_l^{(h)},$$
so that the conditional likelihood contribution of the aggregation of $l$th tie variables in component $h$, $Y_{l}^{(h)}$, given the latent variable $\omega_{l}^{(h)} \sim \text{PG}(n_{h}, 0) \sim \text{PG}(n_{h}, \psi_{l}^{(h)})$ is

$$L_{lh}(\psi_{l}^{(h)} | \omega_{l}^{(h)}) = \exp\{\kappa_{l}^{(h)}\psi_{l}^{(h)} - \omega_{l}^{(h)}\psi_{l}^{(h)2}/2\},$$

where $\kappa_{l}^{(h)} = Y_{l}^{(h)} - \frac{n_{h}}{2}$.

So for all the observations, the conditional posterior probability of $\psi$ is given by:

$$p(\psi \mid \omega, Y, \ldots) \propto p(\psi) \prod_{h=1}^{H} \prod_{l=1}^{L} L_{lh}(\psi_{l}^{(h)} | \omega_{l}^{(h)})$$

$$= p(\psi) \prod_{h=1}^{H} \prod_{l=1}^{L} \exp\{\kappa_{l}^{(h)}\psi_{l}^{(h)} - \omega_{l}^{(h)}\psi_{l}^{(h)2}/2\}$$

$$\propto p(\psi) \prod_{h=1}^{H} \prod_{l=1}^{L} \exp\{-\frac{\omega_{l}^{(h)}\psi_{l}^{(h)2}}{2} - \kappa_{l}^{(h)}\psi_{l}^{(h)} + \frac{\kappa_{l}^{(h)2}}{2\omega_{l}^{(h)}}\}$$

$$= p(\psi) \prod_{h=1}^{H} \prod_{l=1}^{L} \exp\{-\frac{\omega_{l}^{(h)}}{2}(\psi_{l}^{(h)2} - \frac{2\kappa_{l}^{(h)}\psi_{l}^{(h)}}{\omega_{l}^{(h)}} + \frac{\kappa_{l}^{(h)2}}{\omega_{l}^{(h)2}})\}$$

$$= p(\psi) \prod_{h=1}^{H} \prod_{l=1}^{L} \exp\{-\frac{\omega_{l}^{(h)}}{2}(\psi_{l}^{(h)2} - \frac{\kappa_{l}^{(h)}}{\omega_{l}^{(h)}\psi_{l}^{(h)}}\}$$

$$\propto p(\psi) \exp\left\{-\frac{1}{2}(q - X*\psi)^{T}\Omega(q - X*\psi)\right\},$$

where

$$q = \text{vec}\{\frac{\kappa_{l}^{(h)}}{\omega_{l}^{(h)}}\}_{l=1}^{L} , h=1 \}^{H}$$
\[
\Omega = \text{diag}\left(\text{vec}(\{\omega_i^{(h)}\}_{i=1}^{LH}_{h=1})\right)
\]

and \(X^*\) is the \(LH \times L\) design matrix with \((i,j)\)th element 1 if, and only if, \([i/H] = j\).

Note that \(X^* = 1_H \otimes I_L\). Here, for matrices with \(L\) rows and \(H\) columns we define \(\text{vec}(B) = \{B_{c(l)} : l = 1, \ldots, LH\}\) and \(c(l) = \{(i,j) : j = \min\{k : l < kL\}, i = l - (j - 1)L\}\).

In [2], \(\psi^{(h)}_l = Z_l + D_l^{(h)}\), so we have \(X^*\psi = X^*Z + X^*D\), where \(X^*D\) is a constant vector.

Hence the full conditional distribution for \(Z\) is:

\[
p(Z|\omega, Y, \ldots) \propto p(Z) \exp\left\{ -\frac{1}{2} (q^* - X^*Z)^T \Omega (q^* - X^*Z) \right\},
\]

where the working response is:

\[
q^* = q - X^*D
\]

We can denote the prior density of \(Z\) as:

\[
p(Z) \propto \exp\left\{ -\frac{1}{2} (Z - \mu_Z)^T \Sigma_Z^{-1} (Z - \mu_Z) \right\}
\]

Hence, the conditional posterior probability of \(Z\) can be expressed as:

\[
p(Z \mid \omega, Y, \ldots) \propto p(Z) \exp\left\{ -\frac{1}{2} (q^* - X^*Z)^T \Omega (q^* - X^*Z) \right\} = \exp\left\{ -\frac{1}{2} (Z - \mu_Z)^T \Sigma_Z^{-1} (Z - \mu_Z) \right\} \exp\left\{ -\frac{1}{2} (q^* - X^*Z)^T \Omega (q^* - X^*Z) \right\}
\]
The quadratic terms in the exponential can be re-arranged as a quadratic form in $Z - \tilde{\mu}_Z$:

$$(q^* - X^*Z)^T \Omega (q^* - X^*Z) + (Z - \mu_z)^T \Sigma_Z^{-1} (Z - \mu_z)$$

$$= (Z - \tilde{\mu}_Z)^T (X^T \Omega X^* + \Sigma^{-1}_Z)(Z - \tilde{\mu}_Z) + q^T \Omega q^* - \tilde{\mu}^T_Z (X^T \Omega X^* + \Sigma^{-1}_Z) \tilde{\mu}_Z + \mu_z \Sigma^{-1}_Z \mu_z$$

$$= (Z - \tilde{\mu}_Z)^T \tilde{\Sigma}^{-1}_Z (Z - \tilde{\mu}_Z) + q^T \Omega q^* - \tilde{\mu}^T_Z \tilde{\Sigma}^{-1}_Z \tilde{\mu}_Z + \mu_z \Sigma^{-1}_Z \mu_z$$

where

$$\tilde{\Sigma}_Z = (X^T \Omega X^* + \Sigma^{-1}_Z)^{-1}$$

and

$$\tilde{\mu}_Z = \tilde{\Sigma}^{-1}_Z (X^T \Omega q^* + \Sigma^{-1}_Z \mu_z)$$

So the conditional posterior distribution can be expressed as a normal distribution times a constant that only depends on data:

$$p(Z | \omega, Y, \ldots) \propto \exp \left\{ -\frac{1}{2} (Z - \tilde{\mu}_Z)^T \tilde{\Sigma}_Z^{-1} (Z - \tilde{\mu}_Z) \right\}$$

Hence the conditional posterior distribution of $Z$ is Gaussian with mean $\tilde{\mu}_Z$ and diagonal covariance matrix $\tilde{\Sigma}_Z$. Explicitly,

$$Z \mid - \sim N_L(\tilde{\mu}_Z, \tilde{\Sigma}_Z),$$

with
\[
\tilde{\mu}_Z = \tilde{\Sigma}_Z^{-1}(X^* \kappa^* + \Sigma_Z^{-1} \mu_Z)
\]
and
\[
\kappa^* = \Omega q^* = \Omega(q - X^* D) = \text{vec}(\{\kappa_{i}^{(h)} - D_{i}^{(h)} \omega_{i}^{(h)}\}_{i=1;h=1}^{L:H}).
\]

So we see that \(\tilde{\Sigma}_Z\) is diagonal with diagonal elements \(\tilde{\sigma}_{Zi}^2\) where
\[
\tilde{\sigma}_{Zi}^2 = \left[\sum_{h=1}^{H} \omega_{i}^{(h)} + \sigma_{Zi}^{-2}\right]^{-1}
\]

And \(\tilde{\mu}_Z\) has elements \(\tilde{\mu}_{Zi}\) where
\[
\tilde{\mu}_{Zi} = \tilde{\sigma}_{Zi}^2 \left[\sum_{h=1}^{H} \kappa_{i}^{* (h)} + \sigma_{Zi}^{-2} \mu_{Zi}\right]
\]
\[
= \tilde{\sigma}_{Zi}^2 \left[\sum_{h=1}^{H} [Y_{i}^{(h)} - \frac{n_{h}}{2} - \omega_{i}^{(h)} \mathcal{L}(X^{(h)} \Lambda^{(h)} X^{(h)\top})] + \sigma_{Zi}^{-2} \mu_{Zi}\right]
\]

These verify Step [4] of Algorithm 1 in [2].

### 3.3.2 Posterior Computation When There May Be Cognitive Error

In the case that \(\psi_{il}^{(h)} = Z_{i} + C_{il} + D_{i}^{(h)}\), networks in the same component \(h\) are still independent conditional on the component-specific edge probability vector \(\pi^{(h)}\). However, they are not identically distributed due to the cognitive error. Suppose \(G^{(h)} = \{i : G_{i} = h\}\) and \(Y^{(h)} = \{A_{i} : G_{i} = h\}\), the set of networks conditionally assigned to component \(h\). Recall
that \( c(l) \) are the ordered indices of the pair of nodes in \( l \)th pair of nodes. Then

\[
YE_l^{(h)} = \{ A_i[uv] : \{u,v\} \cap c(l) \neq \emptyset, \ i \in G^{(h)}, \ v = 1,\ldots, V, \ V \geq u > v \}
\]

is the set of tie variables of type \( l \) reported by an ego-to-alter and

\[
YA_l^{(h)} = \{ A_i[uv] : \{u,v\} \cap c(l) = \emptyset, \ i \in G^{(h)}, \ v = 1,\ldots, V, \ V \geq u > v \}
\]

is the set of tie variables of type \( l \) reported as alter-to-alter. According to the model

\[
YE_l^{(h)} | Z, X^{(h)}, \lambda(h), Z, \omega \sim \text{Binom}(n_{lh}^E, \pi_l^{E(h)})
\]

where

\[
\pi_l^{E(h)} = \frac{1}{1 + e^{-\left(Z_l + D_l^{(h)}\right)}} \tag{3.10}
\]

independently for \( l = 1,\ldots, L \) and \( h = 1,\ldots, H \), with \( n_{lh}^E \) the number of tie variables in \( YE_l^{(h)} \). Similarly,

\[
YA_l^{(h)} | Z, X^{(h)}, \lambda(h), Z, \omega \sim \text{Binom}(n_{lh}^A, \pi_l^{A(h)})
\]

where

\[
\pi_l^{A(h)} = \frac{1}{1 + e^{-\left(Z_l - \theta + D_l^{(h)}\right)}} \tag{3.11}
\]
independently for \( l = 1, \ldots, L \) and \( h = 1, \ldots, H \), with \( n_{lh}^A \) the number of tie variables in \( YA_l^{(h)} \).

We need to sample from the full conditional distribution of \( Z \). To do this we will show that Steps [3] and [4] in the above algorithm can be modified by considering two type of observed tie variables. We replace Step [3] in Algorithm 1 with:

**[3’] Data Augmentation Step via Pólya-gamma Variables**

Generate the Pólya-gamma augmented data for \( \omega_i^{E(h)} \) and \( \omega_i^{A(h)} \), respectively, from their full conditional Pólya-gamma distributions

\[
\omega_i^{E(h)} | - \sim PG(n_{lh}^E, \psi_i^{E(h)})
\]

\[
\omega_i^{A(h)} | - \sim PG(n_{lh}^A, \psi_i^{A(h)})
\]

where

\[
\psi_i^{E(h)} = Z_l + D_i^{(h)}
\]

\[
\psi_i^{A(h)} = Z_l - \theta + D_i^{(h)}
\]

(3.12)

We now need to consider the likelihood contribution of each network \( A_i \) and decompose it to the two types of tie variables, those subject to cognitive error and those not. The likelihood of the aggregation of the \( l \)th tie variables is now

\[
L_{lh}(\psi_i^{E(h)}, \psi_i^{A(h)}) = \frac{\{\exp(\psi_i^{E(h)})\}^{YE_i^{(h)}}}{1 + \exp(\psi_i^{E(h)})} \cdot \frac{\{\exp(\psi_i^{A(h)})\}^{YA_i^{(h)}}}{1 + \exp(\psi_i^{A(h)})}.
\]
As for the algorithm without cognitive error, this can be written given the latent variables $\omega^E_l$ and $\omega^A_l$ as

$$L_{th}(\psi^E_l, \psi^A_l | \omega^E_l, \omega^A_l) = \exp\{\kappa^E_l \psi^E_l - \omega^E_l \psi^E_l^2 / 2 + \kappa^A_l \psi^A_l - \omega^A_l \psi^A_l^2 / 2\}$$

Combining all the observations, the conditional posterior probability of $\psi = (\{\psi^E_l\}_{l=1:h=1}, \{\psi^A_l\}_{l=1:h=1})$ is given by:

$$p(\psi | \omega, Y, \ldots) \propto p(\psi) \prod_{h=1}^H \prod_{l=1}^L L_{th}(\psi^E_l, \psi^A_l | \omega^E_l, \omega^A_l)$$

$$= p(\psi) \prod_{h=1}^H \prod_{l=1}^L \exp\{\kappa^E_l \psi^E_l - \omega^E_l \psi^E_l^2 / 2\} \cdot \exp\{\kappa^A_l \psi^A_l - \omega^A_l \psi^A_l^2 / 2\}$$

$$= p(\psi) \prod_{h=1}^H \prod_{l=1}^L \exp\{-\omega^E_l \left(\psi^E_l - \frac{\kappa^E_l}{\omega^E_l}\right)^2\} \cdot \exp\{-\omega^A_l \left(\psi^A_l - \frac{\kappa^A_l}{\omega^A_l}\right)^2\}$$

$$\propto p(\psi) \exp\left\{-\frac{1}{2} (q - X^\top \psi)^T \Omega (q - X^\top \psi) \right\},$$

where

$$\kappa^E_l = Y_{l}^{E(h)} - \frac{n_{lh}^E}{2} \quad \kappa^A_l = Y_{l}^{A(h)} - \frac{n_{lh}^A}{2}$$

$$q = \text{vec}\left(\{\kappa^E_l/\omega^E_l\}_{l=1:h=1}, \{\kappa^A_l/\omega^A_l\}_{l=1:h=1}\right)$$

$$\Omega = \text{diag}\left(\text{vec}(\omega^E_l)_{l=1:h=1}, \text{vec}(\omega^A_l)_{l=1:h=1}\right)$$

and $X^\top$ is the $2LH \times L$ design matrix $[X^T, X^T]^T$.
Hence, using (3.12), the full conditional distribution for $Z$ is:

$$p(Z \mid \omega, A, \ldots) \propto p(Z) \exp \left\{ -\frac{1}{2}(q^\dagger - X^\dagger Z)^T \Omega(q^\dagger - X^\dagger Z) \right\},$$

where the working response is:

$$q^\dagger = \text{vec}(\{\kappa^{E(h)}_l / \omega^{E(h)}_l - D^{(h)}_{l,h=1}^L\} \cup \{\kappa^{A(h)}_l / \omega^{A(h)}_l + \theta - D^{(h)}_{l,h=1}^L\})$$

Hence, the conditional posterior probability of $Z$ can be expressed as:

$$p(Z \mid \omega, Y, \ldots) \propto p(Z) \exp \left\{ -\frac{1}{2}(Z - \mu_Z)^T \Sigma^{-1}_Z (Z - \mu_Z) \right\} \exp \left\{ -\frac{1}{2}(q^\dagger - X^\dagger Z)^T \Omega(q^\dagger - X^\dagger Z) \right\}$$

As before, the quadratic terms in the exponential can be re-arranged as a quadratic form in $Z - \tilde{\mu}_Z$:

$$(q^\dagger - X^\dagger Z)^T \Omega(q^\dagger - X^\dagger Z) + (Z - \mu_Z)^T \Sigma^{-1}_Z (Z - \mu_Z)$$

$$= (Z - \tilde{\mu}_Z)^T (X^T \Omega X + \Sigma^{-1}_Z)(Z - \tilde{\mu}_Z) + q^T \Omega q^\dagger - \tilde{\mu}_Z^T (X^T \Omega X^\dagger + \Sigma^{-1}_Z)\mu_Z + \mu_Z \Sigma^{-1}_Z \mu_Z$$

$$= (Z - \tilde{\mu}_Z)^T \tilde{\Sigma}_Z^{-1} (Z - \tilde{\mu}_Z) + q^T \Omega q^\dagger - \tilde{\mu}_Z^T \tilde{\Sigma}_Z^{-1} \tilde{\mu}_Z + \mu_Z \Sigma^{-1}_Z \mu_Z$$

where

$$\tilde{\Sigma}_Z = (X^T \Omega X^\dagger + \Sigma^{-1}_Z)^{-1}$$
\[ \tilde{\mu}_Z = \tilde{\Sigma}_Z (X^T \Omega q^\dagger + \Sigma^{-1}_Z \mu_Z) \]

Hence, as before, the conditional posterior distribution of \( Z \) is Gaussian with mean \( \tilde{\mu}_Z \) and diagonal covariance matrix \( \tilde{\Sigma}_Z \). Explicitly,

\[ Z \mid - \sim N_L(\tilde{\mu}_Z, \tilde{\Sigma}_Z), \]

with

\[ \tilde{\mu}_Z = \tilde{\Sigma}_Z (X^T \kappa^\dagger + \Sigma^{-1}_Z \mu_Z) \]

and

\[ \kappa^\dagger = \Omega q^\dagger = \text{vec}(\{\kappa_l^{E(h)} - \omega_l^{E(h)} D_l^{(h)} \}_l=1;h=1, \{\kappa_l^{A(h)} + \omega_l^{A(h)} \theta - \omega_l^{A(h)} D_l^{(h)} \}_l=1;h=1). \]

So we see that \( \tilde{\Sigma}_Z \) is diagonal with diagonal elements \( \tilde{\sigma}^2_{Z_l} \) where

\[ \tilde{\sigma}^2_{Z_l} = \left[ \sum_{h=1}^{H} \omega_l^{E(h)} + \sum_{h=1}^{H} \omega_l^{A(h)} + \sigma_{Z_l}^{-2} \right]^{-1} = \left[ \sum_{h=1}^{H} \omega_l^{(h)} + \sigma_{Z_l}^{-2} \right]^{-1} \]

where \( \omega_l^{(h)} = \omega_l^{E(h)} + \omega_l^{A(h)} \) and \( \tilde{\mu}_Z \) has elements \( \tilde{\mu}_{Z_l} \) where

\[ \tilde{\mu}_{Z_l} = \tilde{\sigma}^2_{Z_l} \left[ \sum_{h=1}^{H} \kappa_l^{(h)} + \sigma_{Z_l}^{-2} \mu_{Z_l} \right] \]
Similarly, the full conditional distribution for $\theta$ can be derived. From (3.12), the likelihood of the aggregation of the $l$th tie variables is now

$$L_{lh}(\theta) = \frac{\exp(\psi_{i}^{A(h)} \lambda_{h}^{(h)}) Y_{l}^{(h)}}{1 + \exp(\psi_{i}^{A(h)})}.$$ 

Combining all the observations, the conditional posterior probability of $\theta$ is given by:

$$p(\theta \mid \omega, Y, Z, D, \ldots) \propto p(\theta) \prod_{h=1}^{H} \prod_{l=1}^{L} L_{lh}(\theta|\omega_{i}^{A(h)})$$ 

$$= p(\theta) \prod_{h=1}^{H} \prod_{l=1}^{L} \exp\left\{ \kappa_{i}^{A(h)} \psi_{i}^{(h)} - \omega_{i}^{A(h)} \psi_{i}^{(h)}^{2}/2 \right\}$$ 

$$= p(\theta) \prod_{h=1}^{H} \prod_{l=1}^{L} \exp\left\{ -\frac{\omega_{i}^{A(h)} \lambda_{h}^{(h)}}{2} \left( \frac{\kappa_{i}^{A(h)}}{\omega_{i}^{A(h)}} - Z_{l} - D_{l}^{(h)} + \theta \right)^{2} \right\}$$ 

$$\propto p(\theta) \exp\left\{ -\frac{1}{2} (q^{\dagger} - X^{\dagger} \theta)^{T} \Omega (q^{\dagger} - X^{\dagger} \theta) \right\},$$
where

\[ \kappa_l^{A(h)} = Y A_l^{(h)} - \frac{n_l^{A}}{2} \]

\[ q^\dagger = \text{vec}(\{\kappa_l^{A(h)}/\omega_l^{A(h)} - Z_l - D_l^{(h)}\}_{l=1, h=1}^{L: H}) \]

\[ \Omega = \text{diag}(\text{vec}(\{\omega_l^{A(h)}\}_{l=1, h=1}^{L: H})) \]

and \( X^\dagger \) is the \( LH \times 1 \) design matrix with elements all -1.

Similarly as for \( Z \), the conditional posterior distribution of \( \theta \) is Gaussian with mean \( \tilde{\mu}_\theta \)
and diagonal covariance matrix \( \tilde{\Sigma}_\theta \). Explicitly,

\[ \theta \mid - \sim N(\tilde{\mu}_\theta, \tilde{\sigma}_\theta^2), \]

with

\[ \tilde{\sigma}_\theta^2 = (X^\dagger^T \Omega X^\dagger + \sigma_\theta^{-2})^{-1} \]

and

\[ \tilde{\mu}_\theta = \tilde{\sigma}_\theta^2 (X^\dagger^T \kappa^\dagger + \sigma_\theta^{-2} \mu_\theta) \]

and

\[ \kappa^\dagger = \Omega q^\dagger = \text{vec}(\{\kappa_l^{A(h)} - \omega_l^{A(h)} Z_l - \omega_l^{A(h)} D_l^{(h)}\}_{l=1, h=1}^{L: H}). \]

So we see that

\[ \tilde{\sigma}_\theta^2 = \left[ \sum_{l=1}^{L} \sum_{h=1}^{H} \omega_l^{A(h)} + \sigma_\theta^{-2} \right]^{-1} \]
and
\[
\tilde{\mu}_\theta = \tilde{\sigma}_\theta^2 \left[ -\sum_{l=1}^L \sum_{h=1}^H \kappa^{(h)}_l + \sigma_\theta^{-2} \mu_\theta \right]
\]
\[
= \tilde{\sigma}_\theta^2 \left[ -\sum_{l=1}^L \sum_{h=1}^H [Y A_l^{(h)} - \frac{n_{lh}^A}{2} - \omega_l^{A(h)} Z_l - \omega_l^{A(h)} L(X^{(h)} \Lambda^{(h)} X^{(h)T})_l] + \sigma_\theta^{-2} \mu_\theta \right]
\]

These lead to modifications of Algorithm 1 of [2], replacing Step [4] and adding an additional Step. Step [4] is replaced with:

[4'] Update the Shared Similarity Vector

Sample the shared similarity vector \( Z \) from its Gaussian full conditional

\[
Z \mid \sim N_L(\tilde{\mu}_Z, \tilde{\Sigma}_Z),
\]

with \( \tilde{\Sigma}_Z \) diagonal with diagonal elements

\[
\tilde{\sigma}_Z^2 = \left[ \sum_{h=1}^H \frac{\omega_l^{A(h)} A_l^{(h)}}{2} - \omega_l^{A(h)} L(X^{(h)} \Lambda^{(h)} X^{(h)T})_l] + \sigma_\theta^{-2} \mu_\theta \right]
\]

and \( \tilde{\mu}_Z \) has \( l \)th element:

\[
\tilde{\sigma}_Z^2 \left[ \sum_{h=1}^H [Y_l^{(h)} - \frac{n_{lh}}{2} + \theta \omega_l^{A(h)} - \omega_l^{A(h)} L(X^{(h)} \Lambda^{(h)} X^{(h)T})_l] + \sigma_\theta^{-2} \mu_\theta \right]
\]

We also add an additional Step to update the cognitive error level, \( \theta \):

[4’'] Update the Cognition Error Level
Sample the cognitive error level $\theta$ from its Gaussian full conditional

$$
\theta \mid \sim N(\tilde{\mu}_\theta, \tilde{\sigma}^2_\theta),
$$

with

$$
\tilde{\sigma}^2_\theta = \left[ \sum_{l=1}^L \sum_{h=1}^H \omega_l^{A(h)} + \sigma^{-2}_\theta \right]^{-1}
$$

and

$$
\tilde{\mu}_\theta = \tilde{\sigma}^2_\theta \left[ - \sum_{l=1}^L \sum_{h=1}^H \left[ Y\Lambda_i^{(h)} - \frac{J_l^{(h)}}{2} - \omega_l^{A(h)} z_l - \omega_i^{A(h)} L(X^{(h)}\Lambda^{(h)}X^{(h)T})_{l,l} \right] + \sigma^{-2}_\theta \mu_\theta \right]
$$

In addition, we also make modification of Step [5] and replace it with [5']. To maintain conjugacy in sampling the component-specific deviations, we reparameterize the model to update $\bar{X}^{(h)} = X^{(h)}\Lambda^{(h)1/2}$ and $\Lambda^{(h)}, h = 1, \ldots, H$, so $D^{(h)} = L(\bar{X}^{(h)}\bar{X}^{(h)T})$.

**[5'] Update the Component-specific Weighted Latent Coordinates**

Denote by $\bar{X}_{-v}^{(h)}$ the $(V - 1) \times R$ matrix, obtained by removing the $v$th row in $\bar{X}^{(h)}$.

For each component, $h$, we block-sample each row of $\bar{X}^{(h)}$, that is $\bar{X}_v^{(h)T} = (\bar{X}_{v1}^{(h)}, \ldots, \bar{X}_{vR}^{(h)})$, conditionally on all the other parameters and $\bar{X}_{-v}^{(h)}$, for $v = 1, \ldots, V$.

We use a similar extension of Step [5] in [2] as developed in Section 3.3.2. Explicitly, let

$$
YE_{(v)}^{(h)} = \{ A_i[c(l)] : v \in c(l), i \in c(l), i \in G^{(h)}, l = 1, \ldots, L \}$$
be the set of tie variables reported by an ego-to-alter case where either the alter or ego is \( v \).

Similarly, let

\[
YA^{(h)}_{(v)} = \{ A_i[c(l)] : v \in c(l), i \not\in c(l), i \in G^{(h)}, l = 1, \ldots, L \}
\]

be the set of tie variables reported as alter-to-alter case where \( v \) is one of the alters. According to the model

\[
YE^{(h)}_{(v)} | Z, X^{(h)}, \lambda(h), Z, \omega \sim \text{Binom}(n_{vh}^E, \pi_{(v)}^{E(h)})
\]

where

\[
\pi_{(v)}^{E(h)} = \frac{1}{1 + e^{-\left(Z_{(v)} + X_{-v}^{(h)} \bar{X}_v^{(h)}\right)}}
\]  

(3.13)

independently for \( v = 1, \ldots, V \) and \( h = 1, \ldots, H \), with \( n_{vh}^E \) the number of tie variables in \( YE^{(h)}_{(v)} \). Similarly,

\[
YA^{(h)}_{(v)} | Z, X^{(h)}, \lambda(h), Z, \omega \sim \text{Binom}(n_{vh}^A, \pi_{(v)}^{A(h)})
\]

where

\[
\pi_{(v)}^{A(h)} = \frac{1}{1 + e^{-\left(Z_{(v)} - \theta + X_{-v}^{(h)} \bar{X}_v^{(h)}\right)}}
\]  

(3.14)

independently for \( v = 1, \ldots, V \) and \( h = 1, \ldots, H \), with \( n_{vh}^A \) the number of tie variables in \( YA^{(h)}_{(v)} \).

Following a similar augmentation by Pólya-gamma variables as in Step [3’], we can write
the full conditional distribution for $\bar{X}_v^{(h)}$ as:

$$p(X_v^{(h)} | \omega, A, X_v^{-(h)}, \ldots) \propto p(X_v^{(h)}) \exp\left\{-\frac{1}{2}(q^\dagger - X_v^{(h)})^T \Omega_v^{(h)}(q^\dagger - X_v^{(h)})\right\},$$

where the working response is:

$$q^\dagger = \text{vec}\left(\left\{\kappa_{(v)}^{E(h)} / \omega_{(v)}^{E(h)} - Z_{(v)}^{(h)}\right\}_{v=1}, \left\{\kappa_{(v)}^{A(h)} / \omega_{(v)}^{A(h)} - Z_{(v)}^{(h)} + \theta\right\}_{v=1}\right)$$

$$\kappa_{(v)}^{E(h)} = Y_{(v)}^{E(h)} - \frac{n_{vh}^E}{2} \quad \kappa_{(v)}^{A(h)} = Y_{(v)}^{A(h)} - \frac{n_{vh}^A}{2}$$

$$\Omega_v^{(h)} = \text{diag}\left(\left\{\omega_{(v)}^{E(h)}\right\}_{v=1}, \left\{\omega_{(v)}^{A(h)}\right\}_{v=1}\right)$$

and $\left\{\omega_{(v)}^{A(h)}\right\}_{v=1}$ are the augmenting Pólya-gamma variables and the $2VH \times R$ design matrix $X^\dagger$ is $[\bar{X}_v^{-(h)^T}, \bar{X}_v^{-(h)^T}]^T$. The prior for $\bar{X}_v^{(h)}$ is mean-zero Gaussian with covariance matrix $\Lambda^{(h)}$.

Hence Step [5'] consists of $H$ sub-steps, for $h = 1, \ldots, H$, updating $\bar{X}_v^{(h)}$ from the full conditional distribution

$$\bar{X}_v^{(h)} | - \sim N\left(\mu_{\bar{X}_v^{(h)}}, \Sigma_{\bar{X}_v^{(h)}}\right)$$

with

$$\mu_{\bar{X}_v^{(h)}} = \left(\bar{X}_v^{(h)^T} \Omega_v^{(h)} \bar{X}_v^{(h)} + \Lambda^{(h)^{-1}}\right)^{-1} \eta_v^{(h)}$$

and

$$\Sigma_{\bar{X}_v^{(h)}} = \left(\bar{X}_v^{(h)^T} \Omega_v^{(h)} \bar{X}_v^{(h)} + \Lambda^{(h)^{-1}}\right)^{-1}$$

where

$$\eta_v^{(h)} = \bar{X}_v^{(h)^T} \left(\left.Y_{(v)}^{(h)} - 1_{V-1}n_{vh}/2 + \theta \Omega_v^{A(h)} - \Omega_v^{(h)} Z_{(v)}\right)\right)$$

25
where

\[ \Omega^{A(h)}_{(v)} = \text{diag}\left(\{\omega^{A(h)}_{(v)}\}^V_{v=1}\right) \]

\[ \Omega^{(h)}_{(v)} = \text{diag}\left(\{\omega^{E(h)}_{(v)}\}^V_{v=1}, \{\omega^{A(h)}_{(v)}\}^V_{v=1}\right) \]

and

\[ Y^{(h)}_{(v)} = Y^{E(h)}_{(v)} + Y^{A(h)}_{(v)} \]

\[ n_{vh} = n^{E}_{vh} + n^{A}_{vh} \]
CHAPTER 4

Application to Cognitive Networks among High-Tech Managers

4.1 Data

We consider data collected by David Krackhardt from questionnaires answered by 21 managers in a small high-tech, machine manufacturing firm. The social relation asked in the questionnaire was ”Who is a friend of X?”. Each person indicated not only his or her own friendship relationships, but also the relations he or she perceived among all other managers, generating a full $21 \times 21 \times 21$ directed binary array. Other manager attributes including Department, Level, Age, and Tenure are also included in the data. To test our model, we only look at the friendship relationships among the managers. Moreover, the original network data is directed. While a directed model is straightforward we focus on on the undirected version to compare the results to [2]. We take as a relationship between two people if ego indicates that both are friends of each other. Explicitly, we define a tie if and only if the reports are mutual, i.e. $A[i,j] = A[j,i] = 1$. This leads to 21 symmetric adjacency matrices with binary values $\{0, 1\}$.

The average density of ties in the networks is 0.0544. As we are interested in the cognitive

27
social structure, we further compute the average densities for the reported Alter-Alter ties and Ego-Alter ties. The density of Alter-Alter ties is 0.0414, while the density of Ego-Alter ties is 0.179. The density in the latter case is significantly higher than that the density in the first case, indicating the existence of some sort of cognitive error. The empirical cognitive error level is 1.617, computed by taking the difference between the logit of the two densities.

4.2 Goodness-of-Fit Checks

To evaluate the performance of our statistical model in recovering the generative mechanism underlying the observed data, we use the same posterior predictive checks as in [2]. In particular, we first simulate networks from their posterior distribution obtained from our model and then compute a wide set of network summary measures of interest, obtaining samples from the posterior predictive distributions associated with these measures. To compare the goodness-of-fit performances of our model and the original model, we give \( \theta \) different prior specifications. For the case when there is cognitive error, \( \theta \neq 0 \), we specify the prior to be \( \mu_\theta = 0 \) and \( \sigma^2_\theta = 10 \). For the case where \( \theta = 0 \), we let the prior to be \( \mu_\theta = 0 \) and \( \sigma^2_\theta = 10^{-9} \), which empirically restricts \( \theta = 0 \) and make our model exactly the same as the original model in [2].

Figure 4.1 and Figure 4.2 show the the posterior predictive distributions for different summary measures with \( \theta = 0 \) and with \( \theta \neq 0 \), respectively. We observe that the density graphs of the posterior distribution under our model are similar to those under the original model where \( \theta = 0 \) but sharper, centering more heavily at clusters of observed data and having smaller variances than the density graphs of the original model. In summary, the
posterior predictive checks show that our modified model has a slightly better performance of
in recovering the generative mechanism underlying the observed CSS data than the original
model.

Figure 4.1: Posterior Predictive Distribution for $\theta = 0$

Figure 4.2: Posterior Predictive Distribution for $\theta \neq 0$
4.3 Results

As a case study, we analyzed David Krackhardt’s manager data under our model. Since the managers worked in the same firm, we treat the networks as samples from the same population. For prior specification, we tried different prior means for $Z$ such as $\mu_Z = 0$ and $\mu_Z = -3$. We notice that different prior means of $Z$ do have influences on the results. We analyze the results for $\mu_Z = -3$ because $Z$ is the overall odds and $\mu_Z = -3$ means a sparse network without cognitive error. In addition, we also tried $\mu_\theta = 0$ and $\mu_\theta = 3$ and the posterior distributions for $\theta$ are shown in Figure 4.3 and Figure 4.4 respectively.

First of all, we checked whether the results indicates the presence of cognitive error, i.e. $\theta \neq 0$. We look at the the posterior density graph of $\theta$ with the prior mean $\mu_\theta = 0$ as shown in Figure 4.3. The posterior distribution of $\theta$ has a mean of 3.042 with a standard deviation of 0.260, which indicates that $\theta$ is significantly greater than 0. This provides strong evidence of cognitive error in the observed CSS data. In fact, the model cognitive error is about twice as much as the naive cognitive error level of 1.617. The naive cognitive error is computed from the observed data by taking the logit of the difference between the densities of reported Alter-Alter ties and reported Ego-Alter ties.

Figure 4.4 below is the posterior density graph for the prior mean $\mu_\theta = 3$. The density graph for $\mu_\theta = 3$ looks almost the same as that for $\mu_\theta = 0$. It implies that the effect of the prior mean is very small because the data sends such a strong signal about $\theta$.

Secondly, we check to see whether our model is able to recover the patterns of ties between Ego-Alter pairs and Alter-Alter pairs. Table 4.1 shows that the tie density of Ego-Alter pairs in predictive networks are little higher than the observed ones, and Alter-Alter tie densities
Figure 4.3: Posterior Distribution for $\theta$ with a prior centered at zero $\mu_\theta = 0$

Figure 4.4: Posterior Distribution for $\theta$ with a prior centered at 3 $\mu_\theta = 3$. 
Predictive Tie Densities

<table>
<thead>
<tr>
<th>Type</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ego-Alter Pair</td>
<td>0.197</td>
</tr>
<tr>
<td>Alter-Alter Pair</td>
<td>0.0410</td>
</tr>
</tbody>
</table>

Observed Tie Densities

<table>
<thead>
<tr>
<th>Type</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ego-Alter Pair</td>
<td>0.179</td>
</tr>
<tr>
<td>Alter-Alter Pair</td>
<td>0.0414</td>
</tr>
</tbody>
</table>

Table 4.1: Tie Densities between Ego-Alter Pairs and Alter-Alter Pairs

are very similar. The results demonstrates that our model recovers the cognitive social structure (Ego-Alter relation and Alter-Alter relation) pretty well in terms of tie densities.

Since $p_{L(A)}(a)$ is defined via a mixture model, we empirically set the maximum number of components $H = 5$ as there are only 21 networks in total. The posterior component possibilities are computed from the simulated networks, as shown at the upper part in Table 4.2. It seems that $H = 3$ as the probabilities for component 4 and 5 are relatively smaller. Notice that the cutoff is arbitrary, but practically $H = 3$ is a reasonable choice for this case. We also set the maximum number of components $H = 3$. The result is shown at the lower part in Table 4.2, which stays consistent with our choice.

<table>
<thead>
<tr>
<th>Component</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.577</td>
</tr>
<tr>
<td>2</td>
<td>0.222</td>
</tr>
<tr>
<td>3</td>
<td>0.117</td>
</tr>
<tr>
<td>4</td>
<td>0.0594</td>
</tr>
<tr>
<td>5</td>
<td>0.0249</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Component</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.622</td>
</tr>
<tr>
<td>2</td>
<td>0.282</td>
</tr>
<tr>
<td>3</td>
<td>0.0958</td>
</tr>
</tbody>
</table>

Table 4.2: Posterior Component Possibilities
We already analyzed the posterior density graphs of network summary measures in Section 4.2. In addition to that, we compare the average summary measures of simulated networks to the observed ones, see Table 4.3 and Table 4.4. Most of the summary measures are recovered well, except that average triangle frequency and average transitivity are lower than the observed values. The reduction in triangle frequency and transitivity can also be observed in the plots in Figure 4.1 and Figure 4.2, in which their density graphs are centered around 0 while the observed values scatter. This indicates the lack of ability for our model to model the transitivity and triangle relations in networks and implies potential improvement.

Figure 4.5 and Figure 4.6 each shows the summaries of the posterior distribution for the tie probabilities $\bar{\pi}_t$ removing the cognitive error $\theta$ in the reporting of ties. We consider the two models (with $\theta = 0$ and $\theta \neq 0$, respectively). From the plots, we find that the tie probabilities removing the cognitive error for $\theta \neq 0$ are much higher than those for $\theta = 0$. 

### Table 4.3: Predictive Summary Measures

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>0.0578</td>
</tr>
<tr>
<td>Triangle Frequency</td>
<td>0.00141</td>
</tr>
<tr>
<td>Transitivity</td>
<td>0.136</td>
</tr>
<tr>
<td>Average Eigencentrality</td>
<td>0.215</td>
</tr>
<tr>
<td>Average Path Length</td>
<td>2.085</td>
</tr>
<tr>
<td>Average Degree</td>
<td>1.157</td>
</tr>
<tr>
<td>SE Degree</td>
<td>1.271</td>
</tr>
</tbody>
</table>

### Table 4.4: Observed Summary Measures

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>0.0544</td>
</tr>
<tr>
<td>Triangle Frequency</td>
<td>0.00226</td>
</tr>
<tr>
<td>Transitivity</td>
<td>0.235</td>
</tr>
<tr>
<td>Average Eigencentrality</td>
<td>0.224</td>
</tr>
<tr>
<td>Average Path Length</td>
<td>2.013</td>
</tr>
<tr>
<td>Average Degree</td>
<td>1.088</td>
</tr>
<tr>
<td>SE Degree</td>
<td>1.300</td>
</tr>
</tbody>
</table>
Of course removing $\theta$ will not affect the tie probabilities much for $\theta = 0$ case, but the overall tie probabilities are greatly increased by removing the $\theta$ for $\theta \neq 0$ case. The results show a very strong effect of $\theta$ in tie probabilities and confirm our finding that the cognitive error $\theta \neq 0$. It implies that the probability of forming ties is much higher when the cognitive error is taken into account and the probability of reporting ties between Ego-Alter pair (without $\theta$) is much higher than those of Alter-Alter pair (with $\theta$).

![Figure 4.5](image1)

(a) Posterior Mean  
(b) Posterior 2.5% Quantile  
(c) Posterior 97.5% Quantile

Figure 4.5: Posterior Tie Probabilities $\bar{\pi}_l$ Removing the Cognitive Error for $\theta = 0$

![Figure 4.6](image2)

(a) Posterior Mean  
(b) Posterior 2.5% Quantile  
(c) Posterior 97.5% Quantile

Figure 4.6: Posterior Tie Probabilities $\bar{\pi}_l$ Removing the Cognitive Error for $\theta \neq 0$
CHAPTER 5

Discussion

Cognitive social structure is one of the areas in social network research that is of enduring importance but lacks flexible statistical models. From the viewpoint of CSS, network representations of a group’s social structure are often derived from the information reported by each of the members of the group. On the other hand, Durante, Dunson and Vogelstein have propose in [2] a flexible Bayesian nonparametric approach for modeling the population distribution of network valued data. Based on latent space representations, they defines the joint distribution of the edge probability by a mixture model such that the edge probabilities conditional on the component $h$ are independent and identically distributed. It successfully reduces dimensionality and incorporates information within each mixture component $h$. Although their method can flexibly characterize various generative mechanisms, it lacks the ability to model the population of networks with CSS.

Inspired by their work and its strength in flexibility, we construct a statistical model for cognitive social structures by adding a cognitive error level parameter $\theta$ to their model to represent the cognitive error of Ego person $k$ perceiving the social relations between Alter persons $i$ and $j$. Specifically, we assume Ego person $k$ has clear perception of their own ties (Ego-Alter ties) and has poor perception of ties between their Alters (Alter-Alter ties),
represented by the cognitive error level $\theta$. By construct, we expect the estimated $\theta$ to be a positive number if there is any cognitive error that leads to fewer ties between Alter-Alter pairs perceived by the Ego.

As a case study to assess the performance of our method in characterizing the generative mechanism underlying CSS, we apply the model to the manager data collected by David Krackhardt. We first evaluate the goodness-of-fit of our model through posterior predictive checks. The density graphs of the posterior distribution of summary measures under our model are similar to those under the original model but center more heavily at clusters of observed data and have smaller variances. Our model has a slightly better performance in recovering the generative mechanism underlying the observed CSS data than the original model. In the case study, we try different prior means for $Z$ and notice that they have influences on the results. For our case we analyze the results for $\mu_Z = -3$ as it means a sparse network without cognitive error. We also try different different prior means for $\theta$ but the effect is very small. The results from our case study strongly support the presence of cognitive error. The posterior distribution for $\theta$ centers at 3.042 with a standard deviation of 0.260, which indicates that the estimated $\theta$ is significantly greater than 0. We confirm that our model is capable of recovering the patterns of ties between Ego-Alter pairs and Alter-Alter pairs as the tie densities of Ego-Alter and Alter-Alter pairs in predictive networks are close to the observed ones. We also investigate the flexibility of our model by checking the posterior probabilities of the mixture components. There appears to be three main groups of individuals among the 21 managers according to our results. Furthermore, we observe that most of the average summary measures of simulated networks are close to the observed
ones, expect for average triangle frequency and average transitivity. This and the results from Goodness-of-Fit Checks together show that our model is able to recover the mechanism underlying the observed CSS data well. Finally, we plot the posterior summaries of tie probabilities removing the cognitive error $\theta$ for $\theta \neq 0$ and without $\theta$. Not surprisingly, we find that $\theta$ has a very strong effect of $\theta$ in tie probabilities and confirm that the cognitive error $\theta \neq 0$. To conclude, the results illustrate that our model is capable of characterizing cognitive social structures via a flexible Bayesian nonparametric approach.

This work has another contribution to the literature, as we have fully derived and verified the method in [2] and its posterior computations steps for Gibbs sampler. Moreover, we have also explicitly derived the posterior distributions and its posterior computation steps for Gibbs sampler for our modified model. Therefore this work can be of great use to develop other modifications of the mixture model of population and their posterior computation algorithms. We also notice in the results that our model does not recover triangle frequency and transitivity well. This implies that our model lacks the ability to model the transitivity and triangle relations in CSS, which motivates future improvement about them.
REFERENCES


