Joint Modeling of Longitudinal Relational Data and Exogenous Variables

Rajarshi Guhaniyogi and Abel Rodriguez

October 23, 2017

Abstract

A fundamental aspect of relational data, such as data from social network along with the attributes of its constituent actors, is the possibility of dependence between network and the attributes over time. This article proposes a time varying stochastic framework that jointly models co-evolution of the network and the attributes over time. To be more specific, we propose time varying stochastic latent factor models with shared latent parameters in modeling the network and the actor attributes. Our model derives multiple advantages over the existing literature. Unlike the popular co-evolution models, the proposed framework is flexible enough to allow dynamic actor attributes to be measured in both ordinal and continuous scale. It specifically provides model based assessment of the set of predictors jointly influencing relation between nodes. Additionally, the model is easy to compute and readily yields inference and prediction for missing link between nodes. We employ our model framework to study co-evolution of international relations between 22 countries and the country specific indicators over a period of 11 years.

Keywords: latent factor model, nodal attribute, ROC curve, social network, spike and slab prior, systemic dimensions.

1 Introduction

Understanding the coevolution of relational and nodal attributes is a common problem in fields as diverse as public health (Christakis and Fowler, 2007; Fowler and Christakis, 2008), finance (Kalyagin et al., 2014) and genomics (Butland et al., 2005). In these types of applications, data consists of two parts: a time series of dyadic relationship among a common set of actors, which is encoded in a sociomatrix $\boldsymbol{Y}(t) = \{y_{i,j}(t) : i, j = 1, .., n; t \in \mathbb{N}\}$, and a collection of time series of attributes associated with the actors $\boldsymbol{Z}(t) = \{z_{i,k}(t) : i = 1, \ldots, n; k = 1, \ldots, p\}$, where p is the number of attributes and n is the number of actors in the network.

Broadly speaking, the literature approaches the study of the association between nodal and network attributes through two different approaches. One of these approaches focuses on modeling the structure of the network conditional of the nodal attributes. The goal in this case is to understand how social relationships are formed based on attributes of individuals, a process known as "selection". The other approach consists of models of the nodal attributes and their association conditional on the network structure. These models are employed to understand how relationships affect attributes of the individuals in a network, a process referred to as "influence" or "contagion".

Typically, models of selection are built by regressing $y_{ij}(t)$ on node- or dyad-specific regressors using Exponential Random Graph Models (ERGMs) (Holland and Leinhardt, 1981; Robins et al., 2007) or mixed-effects generalized linear models (Wasserman and Anderson, 1987; Holland et al., 1983; Hoff et al., 2002; Hoff, 2005). ERGMs, also known as p^* models, express the conditional density proportional to an exponential function of a few summary statistics of the network and function of nodal covariates. Though naturally appealing, ERGMs are computationally expensive and weak at capturing local features of a network resulting in less than satisfactory performances in many real world problems (Snijders, 2002; Handcock et al., 2003). Alternatively, latent variable models regress the network (or a function of it) on covariates and/or unobserved latent variables. These models are

computationally efficient and useful in modeling transitivity and homophily. On the other hand, models of contagion are usually constructed by regressing nodal attributes against those of other nodes in their network (for example, see Christakis and Fowler, 2007; Fowler and Christakis, 2008; Shoham et al., 2015 and references therein). Common methodological approaches include simultaneous autoregressive (SAR) models (e.g., see Lin, 2010) and threshold models (e.g., see Watts and Dodds, 2009).

Determining whether selection or contagion are at play (i.e., the direction of the causal relationship) is typically a difficult problem (Doreian, 2001). Instead, we focus on jointly modeling the co-evolution of network and nodal attributes through shared latent variables. The goal of our model is twofold. First, we are interested in developing tests of association between structural features of the network and individual nodal attributes. In particular, although our approach does not allow us to distinguish between contagion and selection, it does allow us to carry out tests of independence between the network and the nodal attributes. Second, we are interested in developing predictive models that can be used to jointly predict both future links and future nodal attributes. Joint models of network and nodal attributes have started to receive increasing attention over the last few years. In a static setting, Fosdick and Hoff (2015) recently proposed an extension of the bilinear model of Hoff (2005) in which the nodal attributes and the latent factors used to explain transitivity in the network are jointly modeled using a multivariate normal distribution. In their model, testing for association between the network and nodal attributes reduces to testing whether the cross-covariance matrix between latent factors and the nodal attributes is the null matrix using a likelihood ratio test. On the other hand, Durante et al. (2017) proposes joint modeling of a binary/categorical response and a network using latent variable tensor factorization of the joint probability model. This framework finds its application in clustering networks into multiple groups and thus has a different focus than ours. In a dynamic setting, De la Haye et al. (2010) proposed time varying joint models for network and attributes when the attributes are binary or categorical, while Niezink and Snijders (2016)

extend the framework to accommodate continuous nodal attributes. In both cases inference is carried out using the method of moments implemented using a Robbins-Monro stochastic approximation algorithm.

In this paper we propose a fully Bayesian approach to inference, testing and prediction for co-evolving networks and nodal attributes. The development of this model is motivated by the study of the relationship between international relationships and country-specific economic performance. Our approach is related to, but distinct from, the one presented in Fosdick and Hoff (2015). In addition to accommodating both discrete and continuous attributes and considering the more general case of time series data, our approach uses a common set of latent factors to explain network transitivity and covariation among attributes and network structure, and provides a fully Bayesian test of association that can be used to study individual nodal attributes. When the nodal attributes are assumed to follow conditional Gaussian distribution, our model can be interpreted as a dynamic version of the model presented in Fosdick and Hoff (2015), but with a structured (and more parsimonious) prior on the covariance matrix between the latent traits and the nodal attributes. When the attributes are non-Gaussian, modeling through shared latent traits allows us to include them in the model in a straightforward fashion, something that the formulation in Fosdick and Hoff (2015) does not permit. Shared latent factors have been recently employed, among others, by Rodríguez and Moser (2015) to jointly model voting outcomes and abstentions in roll-call data in the U.S. Congress. Our approach captures the dynamics of the system using autoregressive priors for the shared latent parameters, in an approach reminiscent to Sarkar and Moore (2006), Durante and Dunson (2014) and Sewell and Chen (2015). We refer to our proposed joint modeling framework as Joint Latent Factor Model (JLAFAC).

The remainder of the article flows as follows. Section 2 describes the joint modeling framework with latent factors and highlights some features of the proposed joint model. This section details out prior distribution on the model parameters and latent factors. Section 3 explains posterior computation and explains how MCMC samples are used for link and

attribute prediction and assessments of association between structural features of the network and individual nodal attributes. Sections 4 and 5 demonstrate performance of the proposed framework along with the competing models in simulation studies and in a financial trading network data respectively. Finally, Section 6 concludes the article. Details of MCMC updates are described in the Appendix.

2 Model formulation

2.1 A joint model for nodal attributes and networks

Assume that a group of n actors is followed over time, and let $\boldsymbol{Y}(t) = [y_{i,j}(t)]$ denote the $n \times n$ binary matrix capturing dyadic interactions between these actors at time t, and $\boldsymbol{Z}(t) = [z_{i,k}(t)]$ be the $n \times p$ matrix of continuous or discrete attributes for these actors at time t. Although we concentrate here in binary directed relationships between actors, the model can be easily extended to continuous and other categorical relationships, as well as to undirected relationships (please see Section 6). The network and actor attributes are observed at a finite number of time points $t_1 < t_2 < \cdots < t_T$ resulting in realizations $\boldsymbol{Y}(t_1), \ldots, \boldsymbol{Y}(t_T)$ of the stochastic network $\boldsymbol{Y}(t)$, and $\boldsymbol{Z}(t_1), \ldots, \boldsymbol{Z}(t_T)$ of the stochastic attributes $\boldsymbol{Z}(t)$.

We propose conditionally random effect models for $\boldsymbol{Y}(t)$ and $\boldsymbol{Z}(t)$ with shared latent factors to accommodate co-evolution. To be more precise, for the entries of the sociomatrix $\boldsymbol{Y}(t)$ we consider a bilinear model,

$$y_{i,j}(t) \sim \mathsf{Ber}\left(\theta_{i,j}(t)\right), \qquad \theta_{i,j}(t) = G\left(\mu(t) + \sum_{r=1}^{R} \lambda_r u_{i,r}(t) v_{j,r}(t)\right), \tag{1}$$

where $\boldsymbol{u}_i(t) = (u_{i,1}(t), \ldots, u_{i,R}(t))'$ and $\boldsymbol{v}_j(t) = (v_{j,1}(t), \ldots, v_{j,R}(t))'$ are both R-dimensional time-varying latent variables, and G is an appropriate link function. In the sequel we focus our discussion on the probit link where $G(\cdot)$ corresponds to the cumulative distribution

function of the standard normal distribution, but more general links such as the logistic can be easily accommodated with relatively minor changes to our computational approach.

Equation (1) corresponds to the bilinear model proposed in Hoff (2009). The time varying latent vectors $\boldsymbol{u}_1(t), \ldots, \boldsymbol{u}_n(t)$ and $\boldsymbol{v}_1(t), \ldots, \boldsymbol{v}_n(t)$ capture transitivity and reciprocity in the dyadic relationship (Hoff, 2009). The eigenvalues $\lambda_1, \ldots, \lambda_R$ play a crucial role in determining the specific form of that relationship, in the sense that similar values $u_{i,r}(t)$, $u_{j,r}(t)$ and $v_{i,r}(t)$, $v_{j,r}(t)$ contribute positively or negatively to the relationships $i \to j$ and $j \to i$, depending on whether $\lambda_r > 0$ or $\lambda_r < 0$. Furthermore, note that if $\lambda_r = 0$ then the rth dimension of the latent variables has no impact on the network structure.

The time varying attributes attached to each node can either be continuous, binary or categorical. To model them, we propose to use a set of conditionally independent generalized linear models where

$$E\left\{z_{i,k}(t)\right\} = H_k\left(\eta_k(t) + \sum_{r=1}^{R} \psi_{k,r}\alpha_{k,r}(t)u_{i,r}(t) + \sum_{r=1}^{R} \xi_{k,r}\beta_{k,r}(t)v_{i,r}(t)\right), \tag{2}$$

$\boldsymbol{\alpha}_k(t) = (\alpha_{k,1}(t), \ldots, \alpha_{k,R}(t))'$ and $\boldsymbol{\beta}_k(t) = (\beta_{k,1}(t), \ldots, \beta_{k,R}(t))'$ are $R \times 1$ vectors, and $H_k(\cdot)$ is an appropriate link function for the kth outcome of interest. For example, when $z_{i,k}(t)$ corresponds to a continuous attribute we might set

$$z_{i,k}(t) = \eta_k(t) + \sum_{r=1}^{R} \psi_{k,r}\alpha_{k,r}(t)u_{i,r}(t) + \sum_{r=1}^{R} \xi_{k,r}\beta_{k,r}(t)v_{i,r}(t) + \zeta_{i,k}(t) \tag{3}$$

where $\zeta_{i,k}(t)$ corresponds to realizations from a white-noise process with variance ϕ_k^{-1}.

Note that the latent factors $\boldsymbol{u}_1(t), \ldots, \boldsymbol{u}_n(t)$ and $\boldsymbol{v}_1(t), \ldots, \boldsymbol{v}_n(t)$ we introduced in (1) reappear as predictors in (2). The static coefficients $\psi_{k,r}$ and $\xi_{k,r}$ capture the "average" effect of the kth component of the latent factors on the corresponding nodal attribute, and the time varying coefficients $\alpha_{k,r}(t)$ and $\beta_{k,r}(t)$ control how those baseline effects vary over time. Hence, $\boldsymbol{u}_1(t), \ldots, \boldsymbol{u}_n(t)$ and $\boldsymbol{v}_1(t), \ldots, \boldsymbol{v}_n(t)$ not only determine the structure of the network but also control the level of association between nodal attributes, as well as the co-

evolution between the network and the nodal attributes. Furthermore, just like in the case of λ_r, note that setting $\psi_{k,r} = 0$ and $\xi_{k,r} = 0$ simultaneously implies that the rth dimension of the latent factors has no effect in the evolution of $z_{i,k}(t)$.

Although the previous formulation does not incorporate known covariates that might impact the evolution structure of the network or the nodal attributes, extending it in this direction is trivial by including additional regression terms in (1) and (2).

2.2 Modeling the evolution of the nodal attributes and networks

In order to borrow information across time, the stochastic processes $\{\mu_k(t) : t \in R^+\}$, $\{u_{i,r}(t) : t \in R^+\}$, $\{v_{i,r}(t) : t \in R^+\}$, $\{\eta_k(t) : t \in R^+\}$, $\{\alpha_{k,r}(t) : t \in R^+\}$ and $\{\beta_{k,r}(t) : t \in R^+\}$ for $i, j = 1, \ldots, n$, $k = 1, \ldots, p$ and $r = 1, \ldots, R$ are modeled using independent stationary Ornstein-Uhlenbeck processes. For ease of presentation we assume that the times $t_1, t_2, \ldots, t_T$ at which the data is observed are equally spaced, in which case an Euler discretization of these Ornstein-Uhlenbeck processes leads to first-order autoregressive priors as described in the next few paragraphs. This discretization can be easily modified in the case of irregularly spaced observations to provide a time-consistent model.

Let $u_{i,r}(t_l) = u_{i,r,l}$ and $v_{i,r}(t_l) = v_{i,r,l}$ for all $l = 1, \ldots, T$. Assuming that $t_1, t_2, \ldots, t_T$ are equally spaced, the Ornstein-Uhlenbeck prior implies that

$$u_{i,r,l} = \rho_1 u_{i,r,l-1} + \epsilon^u_{i,r,l}, \qquad \epsilon^u_{i,r,l} \sim \mathsf{N}\left(0, \sigma_u^2\right),$$
$$v_{i,r,l} = \rho_1 v_{i,r,l-1} + \epsilon^v_{i,r,l}, \qquad \epsilon^v_{i,r,l} \sim \mathsf{N}\left(0, \sigma_v^2\right),$$

where $|\rho_1| < 1$ is the autocorrelation coefficient. The initial states $u_{i,r,0}$ and $v_{i,r,0}$ are assigned the respective stationary distribution, $u_{i,r,0} \sim \mathsf{N}\left(0, \frac{\sigma_u^2}{1-\rho_1^2}\right)$ and $v_{i,r,0} \sim \mathsf{N}\left(0, \frac{\sigma_v^2}{1-\rho_1^2}\right)$, which implies that the marginal prior distribution for $u_{i,r,l}$ and $v_{i,r,l}$ have a zero mean and the same constant variance at any l. Furthermore, setting $\rho_1 = 0$ leads to independent priors at every point in time. Note that we have elected to use the same autocorrelation coefficient

ρ_1 for all latent traits, which substantially reduces the number of parameters. However, this assumption implies the we expect the mean reversion times of the different variables to be around the same order of magnitude. We believe that this assumption is reasonable in most applications (and, in particular, in the illustration discussed in Section 5), but if necessary it could be relaxed.

We use a similar argument for each of the other four sets of dynamic parameters in our model. More specifically, letting $\alpha_{k,r}(t_l) = \alpha_{k,r,l}$, $\beta_{k,r}(t_l) = \beta_{k,r,l}$, $\mu(t_l) = \mu_l$ and $\eta_k(t_l) = \eta_{k,l}$, we have

$$\begin{aligned}
\alpha_{k,r,l} &= \rho_2 \alpha_{k,r,l-1} + \epsilon^{\alpha}_{k,r,l}, & \epsilon^{\alpha}_{k,r,l} &\sim \mathsf{N}(0, \sigma^2_{\alpha}), \\
\beta_{k,r,l} &= \rho_2 \beta_{k,r,l-1} + \epsilon^{\beta}_{k,r,l}, & \epsilon^{\beta}_{k,r,l} &\sim \mathsf{N}(0, \sigma^2_{\beta}), \\
\eta_{k,l} &= \rho_3 \eta_{k,l-1} + \epsilon^{\eta}_{k,l}, & \epsilon^{\eta}_{k,l} &\sim \mathsf{N}(0, \sigma^2_{\eta}), \\
\mu_l &= \rho_4 \mu_{l-1} + \epsilon^{\mu}_{l}, & \epsilon^{\mu}_{l} &\sim \mathsf{N}(0, \sigma^2_{\mu}),
\end{aligned}$$

where, as before, the processes are assumed to start at their respective stationary distribution, $\alpha_{k,r,0} \sim \mathsf{N}\left(0, \frac{\sigma^2_{\alpha}}{1-\rho_2^2}\right)$, $\beta_{k,r,0} \sim \mathsf{N}\left(0, \frac{\sigma^2_{\beta}}{1-\rho_2^2}\right)$, $\eta_{k,0} \sim \mathsf{N}\left(0, \frac{\sigma^2_{\eta}}{1-\rho_3^2}\right)$ and $\mu_0 \sim \mathsf{N}\left(0, \frac{\sigma^2_{\mu}}{1-\rho_4^2}\right)$. As before, ρ_2, ρ_3, ρ_4 are autocorrelation coefficients with $|\rho_2| < 1, |\rho_3| < 1$ and $|\rho_4| < 1$ to ensure stationarity.

Without loss of generality, σ_u^2 and σ_v^2 are both fixed at 1. The rest of the variance parameters $\sigma^2_{\alpha}, \sigma^2_{\beta}, \sigma^2_{\eta}, \sigma^2_{\mu}$ and ϕ_k^{-1}, $k = 1, \ldots, p$ follow proper inverse gamma distributions with infinite means, $\mathsf{IGam}(1, 2)$. Finally, lag parameters $\rho_1, \rho_2, \rho_3, \rho_4$ are assumed to follow a uniform distribution on the $[-1, 1]$ interval.

2.3 Dimension selection and association testing

As mentioned in Section 3.2, the coefficients $\lambda_1, \ldots, \lambda_R$ can potentially be used to determine which components of the latent factor vector affect the structure of the network. To exploit

this property of the model we propose to use independent Gaussian mixture priors,

$$\lambda_r \mid \gamma_r \sim \begin{cases} \mathsf{N}\left(0, \frac{1}{\tau_r}\right) & \gamma_r = 1 \\ \mathsf{N}\left(0, \frac{1}{v_0 \tau_r}\right) & \gamma_r = 0 \end{cases} \tag{4}$$

where $\gamma_r \mid \pi_\lambda \sim \mathsf{Ber}(\pi_\lambda)$, $\pi_\lambda \sim \mathsf{Beta}(a_\lambda, b_\lambda)$, and $\tau_r \sim \mathsf{Gam}\left(q^{3(r-1)}, q^{2(r-1)}\right)$, $q > 1$. This prior specification on τ_r ensures fast decay of the precision parameter $\frac{1}{\tau_r}$, so as to safeguard the implied prior on the $\theta_{i,j}(t)$s from degenerating as R becomes large. In particular, $\sum_{r=1}^{R} E\left\{1/\tau_r\right\} = \sum_{r=1}^{R} q^{2(r-1)} / \left\{q^{3(r-1)} - 1\right\}$, so that $\sum_{r=1}^{R} E\{1/\tau_r\}$ converges as $R \to \infty$. In our experiments we work with $q = 1.5$. On the other hand, allowing the prior inclusion probability π_λ to be random allows us to automatically adjust for multiple comparisons (Scott et al., 2010).

Note that, as $v_0 \to \infty$, the mixture component associated with $\gamma_r = 0$ converges to a Dirac delta function at 0, δ_0. Though using a degenerate measure at zero as one component of mixture distribution is suitable for selecting unimportant dimensions, we instead use a large value of v_0 in (4) to mimic the effect of δ_0. Sensitivity of the procedure to the choice of v_0 is explored and discussion is offered in Section 4.

We use a similar approach to identify components of the latent factors that do not affect the different nodal attributes. In particular, we let

$$\psi_{k,r} \mid \omega_{k,r} \sim \begin{cases} \mathsf{N}\left(0, \frac{1}{\nu_{k,r}}\right) & \omega_{k,r} = 1 \\ \mathsf{N}\left(0, \frac{1}{v_0 \nu_{k,r}}\right) & \omega_{k,r} = 0 \end{cases}, \qquad \xi_{k,r} \mid \varsigma_{k,r} \sim \begin{cases} \mathsf{N}\left(0, \frac{1}{\kappa_{k,r}}\right) & \varsigma_{k,r} = 1 \\ \mathsf{N}\left(0, \frac{1}{v_0 \kappa_{k,r}}\right) & \varsigma_{k,r} = 0 \end{cases}, \tag{5}$$

where

$$\begin{aligned} &\omega_{k,r} \mid \pi_{\psi,k} \sim \mathsf{Ber}(\pi_{\psi,k}), \qquad \pi_{\psi,k} \sim \mathsf{Beta}(a_\psi, b_\psi), \qquad \nu_{k,r} \sim \mathsf{Gam}\left(q^{3(r-1)}, q^{2(r-1)}\right), \\ &\varsigma_{k,r} \mid \pi_{\xi,k} \sim \mathsf{Ber}(\pi_{\xi,k}), \qquad \pi_{\xi,k} \sim \mathsf{Beta}(a_\xi, b_\xi), \qquad \kappa_{k,r} \sim \mathsf{Gam}\left(q^{3(r-1)}, q^{2(r-1)}\right). \end{aligned}$$

The indicator variables $\gamma_1, \ldots, \gamma_R$, $\omega_{1,1}, \ldots, \omega_{p,R}$ and $\varsigma_{1,1}, \ldots, \varsigma_{p,R}$ can be used to investigate the pattern of association between the network structure and the nodal attributes, as well as to estimate the effective dimension of the latent space (see Section 3.2 for more details, as well as for a discussion on the elicitation of the parameters a_λ, b_λ, a_ψ, b_ψ, a_ξ and b_ξ).

3 Posterior Inference

Under a Bayesian framework, parameter estimation can be achieved via Markov chain Monte Carlo (MCMC) algorithms, in which posterior distributions for the unknown quantities are approximated with empirical distributions of samples from a Markov chain. To streamline computation, we follow Albert and Chib (1993) and introduce latent variables $w_{i,j,l}$ for the network data such that $w_{i,j,l} > 0$ if $y_{i,j}(t_l) = 1$ and $w_{i,j,l} < 0$ otherwise. Equation (1) can now be written in terms of $w_{i,j,l}$ as

$$w_{i,j,l} = \mu_l + \sum_{r=1}^{R} \lambda_r u_{i,r,l} v_{j,r,l} + \epsilon_{i,j,l}, \qquad \epsilon_{i,j,l} \sim N(0,1). \tag{6}$$

For the nodal attributes we follow a similar approach. In particular, for nodal attributes that follow Gaussian distributions no data augmentation is required. For binary or ordinal attributes we follow a similar latent variable augmentation. This approach leads to updates that mostly use Gibbs sampling steps (see Appendix A for details).

Note that the unknown parameters $\boldsymbol{u}_{i,l} = (u_{i,1,l}, ..., u_{i,R,l})'$, $\boldsymbol{v}_{i,l} = (v_{i,1,l}, ..., v_{i,R,l})'$, $\boldsymbol{\alpha}_{k,l} = (\alpha_{k,1,l}, ..., \alpha_{k,R,l})'$, $\boldsymbol{\beta}_{k,l} = (\beta_{k,1,l}, ..., \beta_{k,R,l})'$, $\lambda_1, \ldots, \lambda_R$, $\psi_{k,1}, \ldots, \psi_{k,R}$ and $\xi_{k,1}, \ldots, \xi_{k,R}$ are not individually identifiable. In particular, the latent space is invariant to rotations and rescalings. Hence, in our analysis we focus on performing inferences on identifiable functions of the model parameters.

3.1 Link and nodal attribute prediction

We can approximate the posterior probability of a directed dyad from node k_1 to node k_2 at time t_l as an average of M post burn-in, suitably thinned, MCMC samples as

$$P\left(y_{k_1,k_2}(t_l)=1 \mid \text{data}\right) \approx \frac{1}{M}\sum_{s=1}^{M} G\left(\mu_l^{(s)} + \sum_{r=1}^{R} \lambda_r^{(s)} u_{k_1,r,l}^{(s)} v_{k_2,r,l}^{(s)}\right).$$

where the superscript (s) denotes the s-th post burn-in MCMC sample for a parameter after suitable thinning. Note that this estimated link probability can be used to infer missing links within observed networks (under the additional assumption of ignorable missingness), or to predict the structure of the network at a future time. To decide whether a directed dyad between nodes k_1 and k_2 is present, one can choose a cut-off $c \in (0,1)$ so that $\theta_{k_1,k_2}(t_l) > c$ implies a directed link. By varying c we can construct a receiver characteristic curve for our network prediction algorithm.

A similar approach can be used to predict the value of nodal attributes. For example, for the purpose of point prediction,

$$E\left\{z_{i,k}(t_l) \mid \text{data}\right\} \approx \frac{1}{M}\sum_{s=1}^{M} H_k\left(\eta_{k,l}^{(s)} + \sum_{r=1}^{R} \psi_{k,r}^{(s)} \alpha_{k,r,l}^{(s)} u_{i,r,l}^{(s)} + \sum_{r=1}^{R} \xi_{k,r}^{(s)} \beta_{k,r,l}^{(s)} v_{i,r,l}^{(s)}\right).$$

3.2 Association tests

The R dimensions of the latent space can be divided into four groups: *systemic* dimensions, which influence both the network structure and at least one of the nodal attributes, two groups of *idiosyncratic* dimensions, one group that is associated only with the network structure, and a second group that solely impacts the association between nodal attributes, and a set of inactive dimensions that have no effect on either of the outcomes of interest. More specifically, a given dimension r is systemic if and only if it significantly affects the structure of the network (i.e., $\gamma_r = 1$) and it also affects at least one of the nodal attributes (i.e., if there is at least one k for which $\omega_{k,r} = 1$ or $\varsigma_{k,r} = 1$, or both). Hence, q_S, the number

of systemic dimensions in the latent space is given by

$$q_S = \sum_{r=1}^{R} \gamma_r \left\{ 1 - \prod_{k=1}^{p} \left(1 - \omega_{k,r}\right) \left(1 - \varsigma_{k,r}\right) \right\}.$$

Note that q_S is an identifiable quantity even if the individual elements are not. Furthermore, $q_S = 0$ if and only if the sequence of networks and the time series of nodal attributes are mutually independent. Hence,

$$P\left(q_S = 0 \mid \text{data}\right) \approx \frac{1}{M} \sum_{s=1}^{M} I\left(q_S^{(s)} = 0\right),$$

where $I(\cdot)$ denotes the indicator function. $P\left(q_S = 0 \mid \text{data}\right)$ provides us with a mechanism to test for global association between the network structure and nodal attributes, with high values of this probability indicating that these are independent. This test is a Bayesian alternative to the likelihood ratio test discussed in Fosdick and Hoff (2015) that also takes into account the dynamic nature of our data.

The general association test we just described can be slightly modified to investigate whether the network structure is associated with a particular nodal attribute. For this purpose, define

$$q_{S,k} = \sum_{r=1}^{R} \gamma_r \left\{ 1 - \left(1 - \omega_{k,r}\right) \left(1 - \varsigma_{k,r}\right) \right\}.$$

Again, an estimate of $P(q_{S,k} = 0 \mid \text{data})$ can be obtained from the samples generated by the MCMC algorithm. High values for this probability indicate that the network structure and the kth nodal attribute are (marginally) independent from each other.

A similar approach can be used to define the number of idiosyncratic dimensions associated only with the network structure,

$$q_N = \sum_{r=1}^{R} \gamma_r \prod_{k=1}^{p} (1 - \omega_{k,r})(1 - \varsigma_{k,r}),$$

and the number of idiosyncratic dimensions associated with the nodal attributes,

$$q_A = \sum_{r=1}^{R}(1-\gamma_r)\left\{1-\prod_{k=1}^{p}(1-\omega_{k,r})(1-\varsigma_{k,r})\right\}.$$

Note that if $q_A = 0$, then all correlations among the nodal attributes is explained by some of the same factors that explain the network structure. Furthermore, although the model allows the dimension of the latent space to be as high as R, the effective dimensionally of the space is $R^* = q_S + q_N + q_A \leq R$, with the number of inactive dimensions being $R - q_S - q_N - q_A$. Hence, a posteriori our model is able to learn the dimension of the latent space.

3.2.1 Association tests and hyperparameter selection

The prior distribution on the summaries q_S, q_N and q_A depends critically on the hyperparmeters a_λ, b_λ, a_ψ, b_ψ, a_ξ and b_ξ, which must be carefully chosen. For example, a priori, $q_N \mid \Upsilon_N \sim \mathsf{Bin}(R, \Upsilon_N)$, where $\Upsilon_N = \pi_\lambda \prod_{k=1}^{p}(1-\pi_{\psi,k})(1-\pi_{\varsigma,k})$ is a random variable with expectation,

$$\mathrm{E}\left\{\Upsilon_N\right\} = \frac{a_\lambda}{a_\lambda + b_\lambda}\left(\frac{a_\psi}{a_\psi + b_\psi}\right)^p\left(\frac{a_\xi}{a_\xi + b_\xi}\right)^p.$$

For even moderate p, the usual choice of $a_\psi = b_\psi = a_\xi = b_\xi = 1$ (i.e., using uniform distributions for the inclusion probabilities) will lead to a prior distribution on q_N that is heavily skewed towards 0. Such prior distribution would typically be unappealing since it would make it difficult to identify components that are idiosyncratic to the network, and therefore lead to an overestimation of the number of systemic components.

To address this issue, in our own data analysis we set $a_\lambda = b_\lambda = 1$, which ensures that the $\sum_{r=1}^{R}\gamma_r$ follows a uniform distribution on $\{0, \ldots, R\}$, and then set the values of a_ψ, b_ψ, a_ξ and b_ξ so that $a_\psi = a_\xi$, $b_\psi = b_\xi$ and the marginal distributions for q_S, q_N and q_A are (approximately) identical.

4 Simulation study

We evaluate the performance of our model using four simulated datasets, each generated under a different scenario. The purpose of this simulation study is threefold: (a) to assess predictive performance of JLAFAC in terms of link and nodal attribute predictions, (b) to investigate the ability of our model to identify dependence and independence relationships in the data, and (c) to assess the impact of hyperparameters (and, in particular, of the constant v_0 used in our variable selection prior) on model performance. For each of the four datasets we fitted JLAFAC using $R = 5$ latent dimensions. All posterior inferences are based on 40,000 samples from the MCMC iterations obtained after a burn-in period of 10,000 iterations and thinning the chain every 10 samples.

In our first simulation scenario the true dimension of the latent space is $R^* = 5$, and the number of attributes is $p = 5$. The true vector of eigenvalues takes the form $\boldsymbol{\lambda} = (0, 0, 0, \lambda_4, \lambda_5)$, while the matrices of selection coefficients take the form

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_{1,1} & \psi_{1,2} & \psi_{1,3} & 0 & 0 \\ \psi_{2,1} & \psi_{2,2} & \psi_{2,3} & \psi_{2,4} & 0 \\ 0 & 0 & 0 & \psi_{3,4} & \psi_{3,5} \\ 0 & 0 & \psi_{4,3} & 0 & 0 \\ \psi_{5,1} & \psi_{5,2} & \psi_{5,3} & \psi_{5,4} & \psi_{5,5} \end{pmatrix}, \qquad \boldsymbol{\Xi} = \begin{pmatrix} \xi_{1,1} & \xi_{1,2} & \xi_{1,3} & 0 & 0 \\ \xi_{2,1} & \xi_{2,2} & \xi_{2,3} & \xi_{2,4} & 0 \\ 0 & 0 & 0 & \xi_{3,4} & \xi_{3,5} \\ 0 & 0 & \xi_{4,3} & 0 & 0 \\ \xi_{5,1} & \xi_{5,2} & \xi_{5,3} & \xi_{5,4} & \xi_{5,5} \end{pmatrix},$$

Note that, in this case, we have $q_S = 2$, $q_N = 0$ and $q_A = 3$. For our second scenario we again set the true dimension of the latent space to $R^* = 5$, and the number of attributes is $p = 5$. However, in this case, the true vector of eigenvalues takes the form $\boldsymbol{\lambda} = (0, 0, 0, 0, \lambda_5)$

and the selection matrices are of the form

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_{1,1} & \psi_{1,2} & 0 & 0 & 0 \\ \psi_{2,1} & \psi_{2,2} & 0 & 0 & 0 \\ 0 & 0 & \psi_{3,3} & \psi_{3,4} & 0 \\ 0 & \psi_{4,2} & \psi_{4,3} & 0 & 0 \\ \psi_{5,1} & \psi_{5,2} & 0 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{\Xi} = \begin{pmatrix} \xi_{1,1} & \xi_{1,2} & 0 & 0 & 0 \\ \xi_{2,1} & \xi_{2,2} & 0 & 0 & 0 \\ 0 & 0 & \xi_{3,3} & \xi_{3,4} & 0 \\ 0 & 0 & \xi_{4,3} & \xi_{4,4} & 0 \\ \xi_{5,1} & \xi_{5,2} & \xi_{5,3} & \xi_{5,4} & 0 \end{pmatrix},$$

In this case $q_S = 0$, $q_N = 1$ and $q_A = 4$, i.e., the network and the nodal attributes are independent in this dataset. Note that the second simulation study allows structural mismatch in terms of positioning of nonzero entries between $\boldsymbol{\Psi}$ and $\boldsymbol{\Xi}$. We set up the third simulation study to allow a mismatch between the fitted dimension and the true dimension of latent variables. Specifically, we set the true dimension of the latent space to $R^* = 3$ and the number of attributes to $p = 5$. In this case the true vector of eigenvalues is $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ and we set

$$\boldsymbol{\Psi} = \begin{pmatrix} 0 & \psi_{1,2} & \psi_{1,3} \\ 0 & \psi_{2,2} & \psi_{2,3} \\ 0 & \psi_{3,2} & \psi_{3,3} \\ 0 & \psi_{4,2} & \psi_{4,3} \\ 0 & \psi_{5,2} & \psi_{5,3} \end{pmatrix}, \qquad \boldsymbol{\Xi} = \begin{pmatrix} \xi_{1,1} & \xi_{1,2} & 0 \\ \xi_{2,1} & \xi_{2,2} & 0 \\ \xi_{3,1} & \xi_{3,2} & 0 \\ \xi_{4,1} & \xi_{4,2} & 0 \\ \xi_{5,1} & \xi_{5,2} & 0 \end{pmatrix},$$

so that $q_S = 3$, $q_N = 0$ and $q_A = 0$. Finally, for the fourth simulation scenario the true dimension of the latent space is $R^* = 5$, but we increase the number of attributes to $p = 10$.

ρ_1	σ_u^2	σ_v^2	ρ_2	σ_α^2	σ_β^2	ρ_3	σ_η^2	ρ_4	σ_μ^2
0.2	1.0	1.0	0.4	1	0.5	0.1	0.25	0.3	0.25

Table 1: True value of the hyperparameters used to generate data for our simulation study.

The true vector of eigenvalues now has the form $\boldsymbol{\lambda} = (0, 0, 0, \lambda_4, \lambda_5)$ and we set

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_{1,1} & \psi_{1,2} & 0 & 0 & 0 \\ \psi_{2,1} & \psi_{2,2} & 0 & 0 & 0 \\ 0 & 0 & \psi_{3,3} & \psi_{3,4} & 0 \\ 0 & 0 & \psi_{4,3} & \psi_{4,4} & 0 \\ \psi_{5,1} & \psi_{5,2} & \psi_{5,3} & \psi_{5,4} & \psi_{5,5} \\ \psi_{6,1} & \psi_{6,2} & \psi_{6,3} & \psi_{6,4} & \psi_{6,5} \\ \psi_{7,1} & \psi_{7,2} & \psi_{7,3} & \psi_{7,4} & \psi_{7,5} \\ \psi_{8,1} & \psi_{8,2} & \psi_{8,3} & \psi_{8,4} & \psi_{8,5} \\ \psi_{9,1} & \psi_{9,2} & \psi_{9,3} & \psi_{9,4} & \psi_{9,5} \\ \psi_{10,1} & \psi_{10,2} & \psi_{10,3} & \psi_{10,4} & \psi_{10,5} \end{pmatrix}, \quad \boldsymbol{\Xi} = \begin{pmatrix} \xi_{1,1} & \xi_{1,2} & 0 & 0 & 0 \\ \xi_{2,1} & \xi_{2,2} & 0 & 0 & 0 \\ 0 & 0 & \xi_{3,3} & \xi_{3,4} & 0 \\ 0 & 0 & \xi_{4,3} & \xi_{4,4} & 0 \\ \xi_{5,1} & \xi_{5,2} & \xi_{5,3} & \xi_{5,4} & \xi_{5,5} \\ \xi_{6,1} & \xi_{6,2} & \xi_{6,3} & \xi_{6,4} & \xi_{6,5} \\ \xi_{7,1} & \xi_{7,2} & \xi_{7,3} & \xi_{7,4} & \xi_{7,5} \\ \xi_{8,1} & \xi_{8,2} & \xi_{8,3} & \xi_{8,4} & \xi_{8,5} \\ \xi_{9,1} & \xi_{9,2} & \xi_{9,3} & \xi_{9,4} & \xi_{9,5} \\ \xi_{10,1} & \xi_{10,2} & \xi_{10,3} & \xi_{10,4} & \xi_{10,5} \end{pmatrix},$$

In this scenario we have $q_S = 2$, $q_N = 0$ and $q_A = 3$. In all four scenarios we work with $n = 40$ nodes and $T = 10$ time points, and generate the non-zero entries of $\boldsymbol{\lambda}$, $\boldsymbol{\Psi}$ and $\boldsymbol{\Xi}$ from normal distributions with means 2, 2 and -3 and variances 2, 0.5 and 0.5, respectively. On the other hand, the true value of the random processes $\{\eta_k(t)\}$, $\{\alpha_{k,r}(t)\}$, $\{\beta_{k,r}(t)\}$, $\mu(t)$, $\{u_{i,r}(t)\}$ and $\{v_{i,r}(t)\}$ are generated using the hyperparaters presented in Table 1. Finally, all the attribute values are generated from Gaussian distributions according to Equation (3), with variances $\phi_1^{-1}, .., \phi_p^{-1}$ drawn i.i.d from $U(0.1, 0.5)$.

4.1 Link and attribute prediction

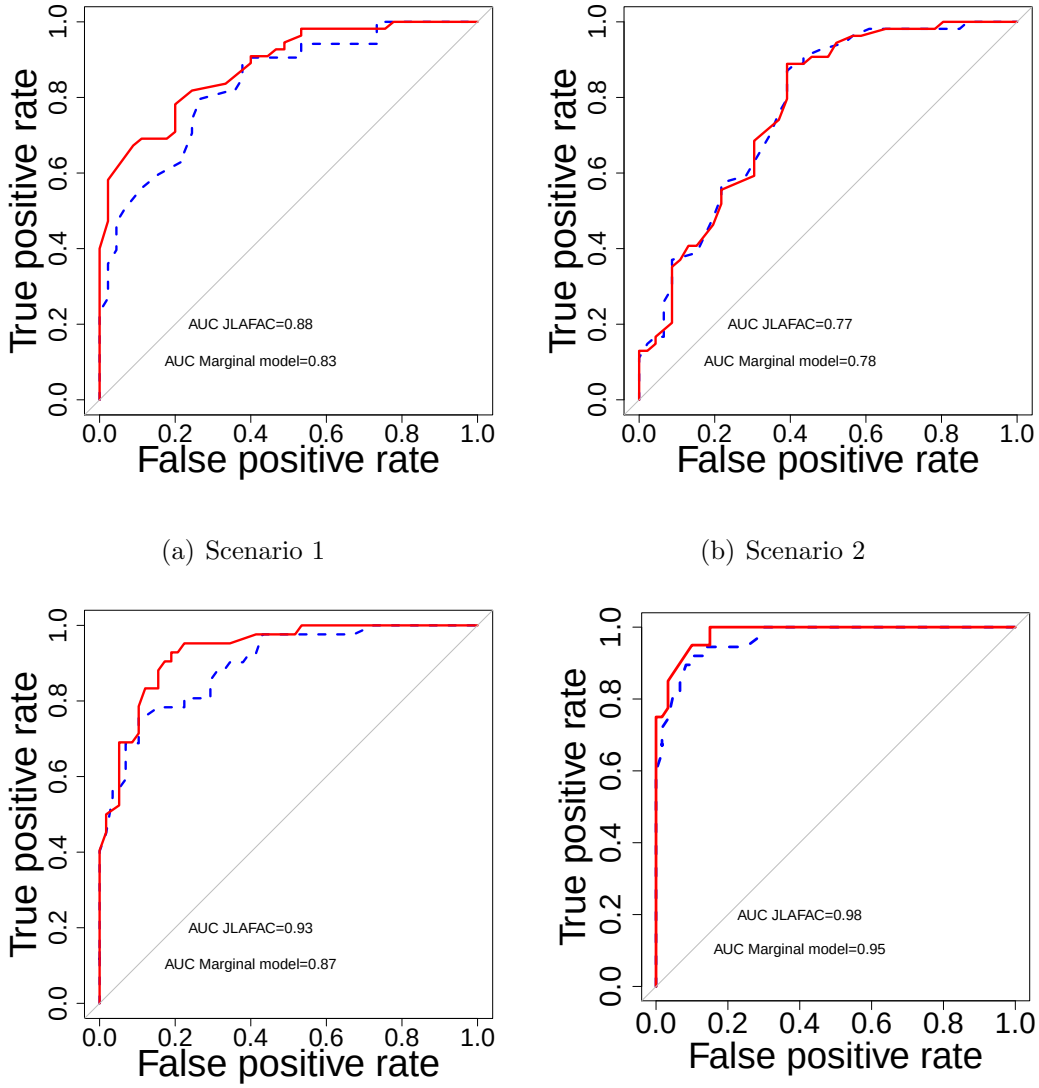
In order to evaluate the ability of the model to predict links, we carry out an out-of-sample crossvalidation exercise. More specifically, we randomly select 400 dyads to hold-out as a

validation set and then estimate our model treating these dyads as if they were missing at random. Figure 1 shows the ROC curve and associated area under the ROC curve (AUC) of JLAFAC for the observations in the validation set. In order to assess the value of jointly modeling nodal attributes and network features, we also present the results generated by fitting only the components of the model associated with the network data, referred to as the marginal model. It is evident that in simulations with $q_S > 0$ (cases 1, 3 and 4), joint modeling has clear advantages in terms of predicting missing links. This can be attributed to the fact that the shared latent factors in such cases are estimated based on both the relational and nodal attribute data, leading to more accurate estimation. The biggest gains in predictive performance arise in scenarios 1 and 3 (the predictive performance of the marginal model in the case of scenario 4 is already quite good, which might help explain why the gains are relatively minor in this case). On the other hand, the predictive performance of both models is essentially identical when $q_S = 0$, as would be expected.

We also carry out prediction of attribute values based on both JLAFAC and a dynamic factor model that ignores the relational information. Table 2 presents the root mean squared error (RMSE) values obtained under each model and for each scenario. Again, the results suggest that jointly modeling both sources of information leads to much improved predictions of nodal attributes when $q_S > 0$, and that those advantages disappear in scenario 2 where $q_S = 0$ and therefore the nodal attributes evolve independently from the relational data. Note, however, that unlike in the dyad prediction problem, the largest improvement is seen in scenario 4.

4.2 Association tests

Prior and posterior distributions for q_S, q_N, q_A and $R^* = q_S + q_N + q_A$ under JLAFAC for each of the four simulation scenarios are recorded in Table 3. In scenarios 1 and 2 the model is capable of identifying both the right dimension of the latent space and its breakdown into systemic and idiosyncratic components. In particular, the model is able to correctly assess



(a) Scenario 1

(b) Scenario 2

(c) Scenario 3

(d) Scenario 4

Figure 1: ROC curves and corresponding AUC values for our out-of-sample crossvalidation in four different simulation scenarios. Continuous curves correspond to the results under our JLAFAC model, while the dashed curves come from a marginal model that does not incorporate the attribute data.

	RMSE	
Scenario	JLAFAC	Dynamic Factor
1	1.23	1.88
2	1.05	1.10
3	0.34	0.47
4	0.35	1.45

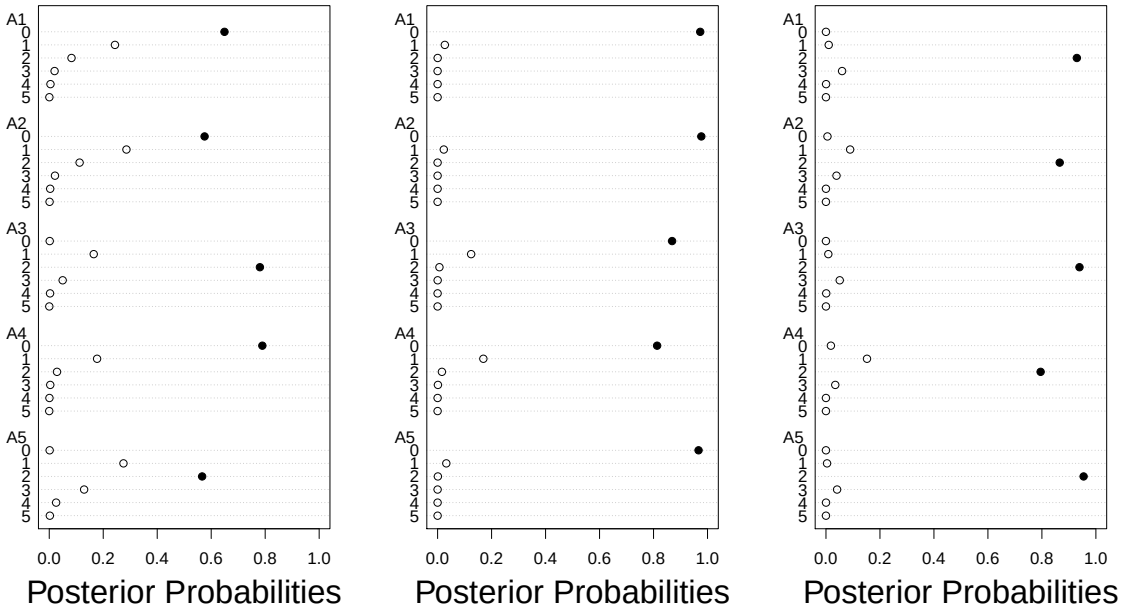
Table 2: Root mean squared error (RMSE) for the two competitors in simulation cases 1 to 4. RMSE is calculated as the square root of the mean absolute squared deviation of predicted and true values of all the predictors.

the independence of network and attribute data in scenario 2, and the lack of independence in scenario 1. In scenario 3 the model is again able to identify the right dimension of the latent space (which, in this case, is smaller than the maximum dimension allowed by our model), but somewhat underestimates the dimension of the systemic component while overestimating the number of idiosyncratic dimensions associated with the network. The performance of the model deteriorates somewhat for scenario 4. In this case the model correctly identifies the dimension of the systemic component of the space, but tends to overestimate the number of idiosyncratic dimensions associated with the network structure while underestimating the number of idiosyncratic dimensions associated with the attributes.

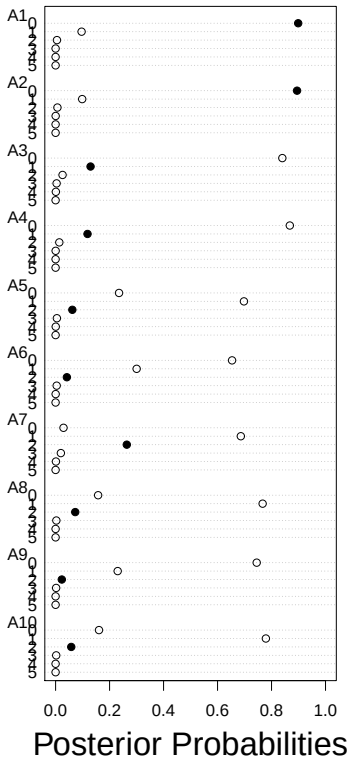
Figure 2 shows the posterior distributions for $q_{S,1}, \ldots, q_{S,p}$ on each of the four simulation scenarios. Again, the model is capable of correctly identifying the number of systemic components associated with each attribute for scenarios 1, 2 and 3, but struggles in the case of scenario 4. In particular, in scenario 4 the model correctly captures the fact that no systemic dimension is associated with attributes 1 and 2, but tends to underestimate the number of systemic components associated with attributes 3 to 10.

4.3 Sensitivity Analysis

All of the results presented above were obtained after fixing $v_0 = 100$. We checked the sensitivity of the results to this choice by re-running the model with $v_0 = 1,000$ and $v_0 = 10,000$ and obtained similar results in every case.



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3



(d) Scenario 4

Figure 2: Posterior distribution for $q_{S,1}, \ldots, q_{S,p}$ under each of our four simulation scenarios. Filled bullets indicate the true value of the summary. Ak denotes the kth attribute.

		Scenario 1		Scenario 2		Scenario 3		Scenario 4	
		Prior	Post	Prior	Post	Prior	Post	Prior	Post
q_S	0	0.31	0.00	0.31	0.64	0.31	0.00	0.31	0.00
	1	0.28	0.22	0.28	0.30	0.28	0.00	0.28	0.36
	2	0.21	0.63	0.21	0.04	0.21	0.79	0.21	0.45
	3	0.13	0.12	0.13	0.00	0.13	0.19	0.13	0.15
	4	0.05	0.01	0.05	0.00	0.05	0.01	0.05	0.01
	5	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
q_N	0	0.37	0.79	0.37	0.23	0.37	0.11	0.37	0.41
	1	0.30	0.17	0.30	0.74	0.30	0.61	0.30	0.50
	2	0.19	0.02	0.19	0.02	0.19	0.21	0.19	0.07
	3	0.09	0.00	0.09	0.00	0.09	0.04	0.09	0.00
	4	0.03	0.00	0.03	0.00	0.03	0.00	0.03	0.00
	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_A	0	0.31	0.12	0.31	0.00	0.31	0.80	0.31	0.27
	1	0.27	0.26	0.27	0.01	0.27	0.18	0.27	0.42
	2	0.21	0.29	0.21	0.12	0.21	0.01	0.21	0.23
	3	0.12	0.30	0.12	0.44	0.12	0.00	0.12	0.06
	4	0.05	0.00	0.05	0.41	0.05	0.00	0.05	0.00
	5	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
R^*	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	1	0.03	0.00	0.03	0.00	0.03	0.00	0.03	0.00
	2	0.09	0.07	0.09	0.00	0.09	0.01	0.09	0.12
	3	0.19	0.22	0.19	0.05	0.19	0.50	0.20	0.34
	4	0.29	0.28	0.29	0.41	0.29	0.37	0.28	0.36
	5	0.37	0.42	0.37	0.52	0.37	0.12	0.37	0.16

Table 3: Prior and posterior distributions on q_S, q_N, q_A and $R^* = q_S + q_N + q_A$ under JLAFAC for the four datasets in our simulation study. Grey backgrounds indicate the true value under each scenario.

5 Analysis of Longitudinal International Relationship Data with Country Specific Indicators of Economic Performance

In this section we investigate the association between international conflicts and various economic indicators in a set of 22 countries. The list of countries includes all five permanent members of the United Nations security council (China, France, Rusia, USA and the UK), a

few members of the G-20 group (Australia, Germany, India, Japan, South Korea, Turkey), as well as various countries from the middle East (Egypt, Iran, Iraq, Israel, Lebanon, Palestine, Sudan), eastern Europe (Ukraine, Georgia) and south Asia (Afghanistan, Pakistan).

Annual networks of interactions among countries between 2004 and 2014 were defined by linking two countries if there was at least one "positive verbal action" during the first week of the year (Goldstein, 1992). Examples of positive verbal actions include granting diplomatic recognition and positive verbal support in international forums. These data on international relations is available from `http://www.stat.washington.edu/people/pdhoff/Code` and was previously studied in Hoff (2015). In addition to the relational data, we consider the evolution of three country-specific economic indicators, which are modeled using a conditionally Gaussian model: the growth rate of Gross Domestic Product (GDP) Per Capita, the level of National Savings as a percentage of GDP, and the amount of International Trade as a percentage of GDP. Data on these indicators are available from the World Development Indicators Database, which is published annually by the World Bank. For this dataset we estimate JLAFAC using $R = 5$ latent dimensions. All inferences presented below are based on $50,000$ samples from our MCMC algorithm obtained after a burn-in period of $10,000$ samples and thinning the chain every 10 iterations.

As in Section 4, we first evaluate the ability of JLAFAC to make out-of-sample predictions. In this case we hold out 500 randomly-chosen dyads, whose values are predicted using the rest of the data. Figure 3 presents the ROC curve and the AUC value under both JLAFAC and a dynamic network model that ignores the attribute data (this comparison is similar to the one we carried out in Section 4). As in the simulation study, it is clear that jointly modeling the relational and nodal attribute data provides some improvement in the predictions, although the improvement in this case is relatively minor.

Next we investigate the association structure in the data. Table 4 presents the prior and posterior distributions for q_S, q_N, q_A and $R^* = R - q_S - q_N - q_A$. It is evident from the Table that the data favors the use of 4 latent dimensions. Furthermore, the model provides strong

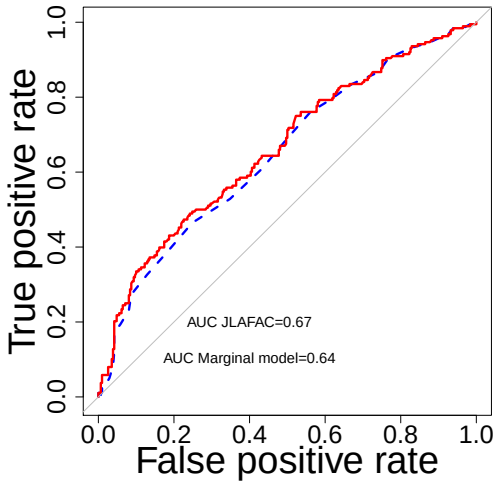


Figure 3: ROC curves and corresponding AUC values for JLAFAC and a dynamic network model that ignores information about network attributes in the international relations data. Solid and dashed curves represent ROC curves for JLAFAC and dynamic network model without network attributes respectively.

evidence for the presence of at least one systemic component ($P(q_S > 0 \mid \text{data}) \approx 0.92$), suggesting that the network structure and the nodal attributes are indeed associated. To investigate this association in more detail we present in Figure 4 the posterior distributions for $q_{S,1}$, $q_{S,2}$ and $q_{S,3}$, where subscripts 1,2,3 correspond to growth of GDP per capita, ratio of trade to GDP and ratio of savings to GDP respectively. Note that while there is strong evidence that the ratio of trade to GDP is associated with the structure of international relations ($P(q_{S,2} > 0 \mid \text{data}) \approx 0.90$), the evidence for an association with GDP growth per capita and national savings is weak to non-existant ($P(q_{S,1} > 0 \mid \text{data}) \approx 0.50$ and $P(q_{S,3} > 0 \mid \text{data}) \approx 0.14$, respectively). A review of the literature suggests that our results are consistent with previous findings. Indeed, there is a well established relationship between international trade and conflict, which was first established in the classic paper by Polachek (1980) (see also Reuveny and Kang, 1996, Morrow, 1999 and Reuveny, 2000). Similarly, there is some evidence in the literature for an association between GDP growth and international conflict (e.g., see Rodrik, 1999, Anderton et al., 2003, Miguel et al., 2004 and Polachek and

Sevastianova, 2012), although the evidence is highly disputed and the direction of the causal effect often unclear. On the other hand, as far as we could find, an association between international conflict and the level aggregate national savings has not been proposed or reported in the literature.

	q_S		q_N		q_A		R^*	
	Prior	Posterior	Prior	Posterior	Prior	Posterior	Prior	Posterior
0	0.28	0.08	0.42	0.61	0.29	0.11	0.00	0.00
1	0.26	0.52	0.30	0.28	0.26	0.20	0.02	0.00
2	0.21	0.27	0.17	0.07	0.20	0.27	0.07	0.04
3	0.14	0.09	0.07	0.01	0.14	0.26	0.17	0.18
4	0.07	0.01	0.02	0.00	0.07	0.13	0.29	0.41
5	0.02	0.00	0.00	0.00	0.02	0.00	0.42	0.34

Table 4: Prior and posterior distributions for q_S, q_N, q_A and $R^* = R - q_S - q_N - q_A$ for the international relations data.

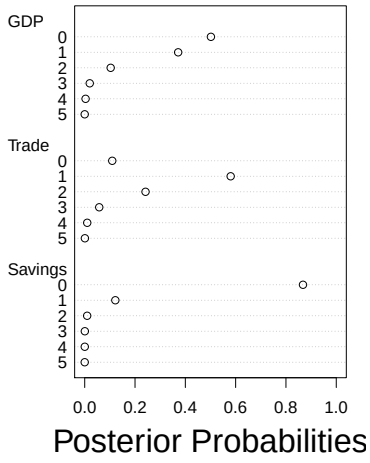


Figure 4: Posterior distributions for $q_{S,1}$, $q_{S,2}$ and $q_{S,3}$ for the international relations data.

6 Conclusion

This article introduces the idea of jointly modeling of network and associated nodal attributes over time. Our proposed framework relied on modeling the network and nodal attributes

jointly through latent factor representations, with some latent factors shared across both models to introduce dependence of directional dyad between nodes on nodal attributes. Evolution of both network and nodal attributes over time is modeled by allowing the latent factors to vary over time, and inference is carried out using a Bayesian approach.

Because of the application that motivated our work we have assumed that the network process $\boldsymbol{Y}(t)$ is directed and binary. The model can easily be reformulated to accommodate undirected networks by replacing equations (1) and (2) with

$$y_{i,j}(t) \sim \mathsf{Ber}\left(\theta_{i,j}(t)\right), \qquad \theta_{i,j}(t) = G\left(\mu(t) + \sum_{r=1}^{R} \lambda_r u_{i,r}(t) u_{j,r}(t)\right),$$

and

$$E\left\{z_{i,k}(t)\right\} = H_k\left(\eta_k(t) + \sum_{r=1}^{R} \psi_{k,r}\alpha_{k,r}(t) u_{i,r}(t)\right),$$

respectively. Similarly, situation in which $y_{i,j}(t)$ belong to other members of the exponential family can be easily accommodated through appropriate generalized linear factor analysis models.

One important extension of the proposed model would be to scale it for larger datasets with millions of individuals connected under social networking and observed over a large number of time points. Computationally convenient time series models are to be used to specify correlations across time for these models. It is also interesting to extend our methodology for online social networks. Some of these constitutes our current work.

A MCMC algorithm

This section details out the MCMC algorithm for the proposed JLAFAC model. It is important to emphasize that the algorithm is presented assuming $G(\cdot)$ as the probit link and $H_k(\cdot)$'s as identity links, i.e. the nodal attributes are assumed to follow Gaussian distributions.

Let, $\boldsymbol{\eta}_k = (\eta_{k,1}, \ldots, \eta_{k,T})'$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_T)'$, $\boldsymbol{Z}_{i,k} = (z_{i,k}(t_1), ..., z_{i,k}(t_T))'$, and $\boldsymbol{w}_{i,j} = (w_{i,j,1}, ..., w_{i,j,T})'$, for $k = 1, \ldots, p$ and $i = 1, .., n$. Further assume that $\boldsymbol{D}_1, \boldsymbol{D}_2, \boldsymbol{D}_3, \boldsymbol{D}_4$ are the four $T \times T$ matrix with the (s, s')-th entry of $\boldsymbol{D}_j$ is given by $\boldsymbol{D}_{j,ss'} = \rho_j^{|s-s'|}$, $j = 1, 2, 3, 4$. Let $\boldsymbol{\Lambda} = diag(\lambda_1, \ldots, \lambda_R)$, $\boldsymbol{\Psi}_k = diag(\psi_{k,1}, \ldots, \psi_{k,R})$ and $\boldsymbol{\Xi}_k = diag(\xi_{k,1}, \ldots, \xi_{k,R})$, $k = 1, \ldots, p$. Further define $\tilde{\boldsymbol{uv}}_{i,j} = (\boldsymbol{u}'_{i,1}\boldsymbol{\Lambda}\boldsymbol{v}_{j,1}, \ldots, \boldsymbol{u}'_{i,T}\boldsymbol{\Lambda}\boldsymbol{v}_{j,T})'$ $\underline{\boldsymbol{uv}}_{i,j} = (\boldsymbol{u}'_{i,1}\boldsymbol{v}_{j,1}, \ldots, \boldsymbol{u}'_{i,T}\boldsymbol{v}_{j,T})'$, $\underline{\boldsymbol{u\alpha}}_{i,k} = (\boldsymbol{u}'_{i,1}\boldsymbol{\alpha}_{k,1}, \ldots, \boldsymbol{u}'_{i,T}\boldsymbol{\alpha}_{k,T})'$, $\underline{\boldsymbol{v\beta}}_{ik} = (\boldsymbol{v}'_{i,1}\boldsymbol{\beta}_{k,1}, \ldots, \boldsymbol{v}'_{i,T}\boldsymbol{\beta}_{k,T})'$, $\boldsymbol{w}_{i,j} = (w_{i,j,1}, \ldots, w_{i,j,T})'$, $\boldsymbol{\alpha}_k = (\boldsymbol{\alpha}'_{k,1}, .., \boldsymbol{\alpha}'_{k,T})'$, $\boldsymbol{\beta}_k = (\boldsymbol{\beta}'_{k,1}, .., \boldsymbol{\beta}'_{k,T})'$, $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_R)'$, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_R)'$, $\boldsymbol{\nu}_k = (\nu_{k,1}, \ldots, \nu_{k,R})'$, $\boldsymbol{\omega}_k = (\omega_{k,1}, \ldots, \omega_{k,R})'$, $\boldsymbol{\kappa}_k = (\kappa_{k,1}, \ldots, \kappa_{k,R})'$, $\boldsymbol{\varsigma}_k = (\varsigma_{k,1}, \ldots, \varsigma_{k,R})'$ for all $i, j = 1, ..., n$ and $k = 1, ..., p$. The full conditionals for the parameters are given as following.

- $w_{i,j,l}|- \sim Truncated - Normal(\mu_l + \boldsymbol{u}'_{i,l}\boldsymbol{\Lambda}\boldsymbol{v}_{j,l}, 1, 0, \infty)$, if $y_{i,j,l} = y_{i,j}(t_l) = 1$
  $w_{i,j,l}|- \sim Truncated - Normal(\mu_l + \boldsymbol{u}'_{i,l}\boldsymbol{\Lambda}\boldsymbol{v}_{j,l}, 1, -\infty, 0)$, if $y_{i,j,l} = y_{i,j}(t_l) = 0$.
- $\boldsymbol{\mu}|- \sim N(\boldsymbol{A}_\mu, \boldsymbol{\Sigma}_\mu)$, $\boldsymbol{\Sigma}_\mu = \left(n^2\boldsymbol{I}_T + \boldsymbol{D}_4^{-1}\frac{(1-\rho_4^2)}{\sigma_\mu^2}\right)^{-1}$, $\boldsymbol{A}_\mu = \boldsymbol{\Sigma}_\mu \sum_{i,j=1}^n (\boldsymbol{w}_{i,j} - \tilde{\boldsymbol{uv}}_{i,j})$.
- $\sigma_\alpha^2|- \sim IG(1 + TRp/2, 2 + (1-\rho_2^2)\sum_{k=1}^p \boldsymbol{\alpha}'_k(\boldsymbol{D}_2^{-1} \otimes \boldsymbol{I}_R)\boldsymbol{\alpha}_k/2)$.
- $\sigma_\beta^2|- \sim IG(1 + TRp/2, 2 + (1-\rho_2^2)\sum_{k=1}^p \boldsymbol{\beta}'_k(\boldsymbol{D}_2^{-1} \otimes \boldsymbol{I}_R)\boldsymbol{\beta}_k/2)$.
- $\sigma_\eta^2|- \sim IG(1 + Tp/2, 2 + (1-\rho_3^2)\sum_{k=1}^p \boldsymbol{\eta}'_k\boldsymbol{D}_3^{-1}\boldsymbol{\eta}_k/2)$
- $\sigma_\mu^2|- \sim IG(1 + T/2, 2 + (1-\rho_4^2)\boldsymbol{\mu}'\boldsymbol{D}_4^{-1}\boldsymbol{\mu}/2)$
- $\tau_r|- \sim Gamma(0.5 + q^{3(r-1)}, q^{2(r-1)} + (\lambda_r^2\gamma_r + \lambda_r^2 v_0(1-\gamma_r))/2)$, $r = 1, \ldots, R$.
- $\pi_\lambda|- \sim Beta(\sum_{r=1}^R I(\gamma_r = 1) + a_\lambda, R + b_\lambda - \sum_{r=1}^R I(\gamma_r = 1))$
- $\kappa_{k,r}|- \sim Gamma(0.5 + q^{3(r-1)}, q^{2(r-1)} + (\xi_{k,r}^2\varsigma_{k,r} + \xi_{k,r}^2 v_0(1-\varsigma_{k,r}))/2)$, $r = 1, \ldots, R$.
- $\nu_{k,r}|- \sim Gamma(0.5 + q^{3(r-1)}, q^{2(r-1)} + (\psi_{k,r}^2\omega_{k,r} + \psi_{k,r}^2 v_0(1-\omega_{k,r}))/2)$, $r = 1, \ldots, R$.
- $\pi_{\psi,k}|- \sim Beta(\sum_{r=1}^R I(\omega_{k,r} = 1) + a, R + b - \sum_{r=1}^R I(\omega_{k,r} = 1))$
- $\pi_{\xi,k}|- \sim Beta(\sum_{r=1}^R I(\varsigma_{k,r} = 1) + a, R + b - \sum_{r=1}^R I(\varsigma_{k,r} = 1))$

- $\gamma_r|- \sim Ber(\tilde{\pi}_r)$, $\tilde{\pi}_r = \frac{\pi_\lambda N\left(\lambda_r|0,\frac{1}{\tau_r}\right)}{\pi_\lambda N\left(\lambda_r|0,\frac{1}{\tau_r}\right)+(1-\pi_\lambda)N\left(\lambda_r|0,\frac{1}{v_0\tau_r}\right)}$.

- $\varsigma_{k,r}|- \sim Ber(\tilde{\pi}_{\xi,k,r})$, $\tilde{\pi}_{\xi,k,r} = \frac{\pi_{\xi,k} N\left(\xi_{k,r}|0,\frac{1}{\kappa_{k,r}}\right)}{\pi_{\xi,k} N\left(\xi_{k,r}|0,\frac{1}{\kappa_{k,r}}\right)+(1-\pi_{\xi,k})N\left(\xi_{k,r}|0,\frac{1}{v_0\kappa_{k,r}}\right)}$.

- $\omega_{k,r}|- \sim Ber(\tilde{\pi}_{\psi,k,r})$, $\tilde{\pi}_{\psi,k,r} = \frac{\pi_{\psi,k} N\left(\psi_{k,r}|0,\frac{1}{\nu_{k,r}}\right)}{\pi_{\psi,k} N\left(\psi_{k,r}|0,\frac{1}{\nu_{k,r}}\right)+(1-\pi_{\psi,k})N\left(\psi_{k,r}|0,\frac{1}{v_0\nu_{k,r}}\right)}$.

- $diag(\boldsymbol{\Lambda})|- \sim N(\boldsymbol{A}_\lambda, \boldsymbol{\Sigma}_\lambda)$, where $\boldsymbol{\Sigma}_\lambda = \left(\sum_{i,j=1}^n \underline{\boldsymbol{uv}}_{i,j} + \boldsymbol{H}\right)^{-1}$,
  $\boldsymbol{H} = diag(\boldsymbol{\tau}'\boldsymbol{\gamma} + \boldsymbol{\tau}'(1-\boldsymbol{\gamma})v_0)$ and $\boldsymbol{A}_\lambda = \boldsymbol{\Sigma}_\lambda \sum_{i,j=1}^n \underline{\boldsymbol{uv}}_{i,j}(\boldsymbol{w}_{i,j} - \boldsymbol{\mu})$.

- $diag(\boldsymbol{\Psi}_k)|- \sim N(\boldsymbol{A}_{\boldsymbol{\Psi},k}, \boldsymbol{\Sigma}_{\boldsymbol{\Psi},k})$, where $\boldsymbol{\Sigma}_{\boldsymbol{\Psi},k} = \left(\sum_{i=1}^n \underline{\boldsymbol{u\alpha}}_{i,k}\phi_k + \boldsymbol{H}_2\right)^{-1}$,
  $\boldsymbol{H}_2 = diag(\boldsymbol{\nu}'_k\boldsymbol{\omega}_k + \boldsymbol{\nu}'_k(1-\boldsymbol{\omega}_k)v_0)$ and $\boldsymbol{A}_{\boldsymbol{\Psi},k} = \boldsymbol{\Sigma}_{\boldsymbol{\Psi},k} \sum_{i=1}^n \underline{\boldsymbol{u\alpha}}_{i,k}(\boldsymbol{Z}_{i,k} - \boldsymbol{\eta}_k - \underline{\boldsymbol{v\beta}}_{i,k} diag(\boldsymbol{\Xi}_k))$.

- $diag(\boldsymbol{\Xi}_k)|- \sim N(\boldsymbol{A}_{\boldsymbol{\Xi},k}, \boldsymbol{\Sigma}_{\boldsymbol{\Xi},k})$, where $\boldsymbol{\Sigma}_{\boldsymbol{\Xi},k} = \left(\sum_{i=1}^n \underline{\boldsymbol{v\beta}}_{i,k}\phi_k + \boldsymbol{H}_3\right)^{-1}$,
  $\boldsymbol{H}_3 = diag(\boldsymbol{\kappa}'_k\boldsymbol{\varsigma}_k + \boldsymbol{\kappa}'_k(1-\boldsymbol{\varsigma}_k)v_0)$ and $\boldsymbol{A}_{\boldsymbol{\Xi},k} = \boldsymbol{\Sigma}_{\boldsymbol{\Xi},k} \sum_{i=1}^n \underline{\boldsymbol{v\beta}}_{i,k}(\boldsymbol{Z}_{i,k} - \boldsymbol{\eta}_k - \underline{\boldsymbol{u\alpha}}_{i,k} diag(\boldsymbol{\Psi}_k))$.

- $\boldsymbol{\alpha}_k = (\boldsymbol{\alpha}'_{k,1}, \ldots, \boldsymbol{\alpha}'_{k,T})'|- \sim N(\boldsymbol{A}_{\boldsymbol{\alpha},k}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha},k})$,
  where $\boldsymbol{\Sigma}_{\boldsymbol{\alpha},k} = \left(\sum_{i=1}^n \boldsymbol{D}'_{u\boldsymbol{\Psi}_k,i}\boldsymbol{D}_{u\boldsymbol{\Psi}_k,i}\phi_k + \frac{((1-\rho_2^2)\boldsymbol{D}_2^{-1}\otimes \boldsymbol{I}_R)}{\sigma_\alpha^2}\right)^{-1}$,
  $\boldsymbol{D}_{u\boldsymbol{\Psi}_k,i} = Block{-}diag(\boldsymbol{u}'_{i,1}\boldsymbol{\Psi}_k, \ldots, \boldsymbol{u}'_{i,T}\boldsymbol{\Psi}_k)$, $\boldsymbol{D}_{v\boldsymbol{\Xi}_k,i} = Block{-}diag(\boldsymbol{v}'_{i,1}\boldsymbol{\Xi}_k, \ldots, \boldsymbol{v}'_{i,T}\boldsymbol{\Xi}_k)$,
  $\boldsymbol{A}_{\boldsymbol{\alpha},k} = \boldsymbol{\Sigma}_{\boldsymbol{\alpha},k} \sum_{i=1}^n \boldsymbol{D}'_{u\boldsymbol{\Psi}_k,i}(\boldsymbol{Z}_{i,k} - \boldsymbol{\eta}_k - \boldsymbol{D}_{v\boldsymbol{\Xi}_k,i}\boldsymbol{\beta}_k)$.

- $\boldsymbol{\beta}_k = (\boldsymbol{\beta}'_{k,1}, \ldots, \boldsymbol{\beta}'_{k,T})'|- \sim N(\boldsymbol{A}_{\boldsymbol{\beta},k}, \boldsymbol{\Sigma}_{\boldsymbol{\beta},k})$,
  where $\boldsymbol{\Sigma}_{\boldsymbol{\beta},k} = \left(\sum_{i=1}^n \boldsymbol{D}'_{v\boldsymbol{\Xi}_k,i}\boldsymbol{D}_{v\boldsymbol{\Xi}_k,i}\phi_k + \frac{((1-\rho_2^2)\boldsymbol{D}_2^{-1}\otimes \boldsymbol{I}_R)}{\sigma_\beta^2}\right)^{-1}$,
  $\boldsymbol{A}_{\boldsymbol{\beta},k} = \boldsymbol{\Sigma}_{\boldsymbol{\beta}k} \sum_{i=1}^n \boldsymbol{D}'_{v\boldsymbol{\Xi}_k,i}(\boldsymbol{Z}_{i,k} - \boldsymbol{\alpha}_k - \boldsymbol{D}_{u\boldsymbol{\Psi}_k,i}\boldsymbol{\alpha}_k)$.

- $\phi_k|- \sim Gamma\left(c + Tn/2, d + \frac{1}{2}\sum_{i=1}^n ||\boldsymbol{Z}_{i,k} - \boldsymbol{\eta}_k - \boldsymbol{D}_{u\boldsymbol{\Psi}_k,i}\boldsymbol{\alpha}_k - \boldsymbol{D}_{v\boldsymbol{\Xi}_k,i}\boldsymbol{\beta}_k||^2\right)$

- Let $\boldsymbol{v}_j = (\boldsymbol{v}'_{j,1}, \ldots, \boldsymbol{v}'_{j,T})'$, $\boldsymbol{D}_{u\boldsymbol{\Lambda},i} = Block - diag(\boldsymbol{u}'_{i,1}\boldsymbol{\Lambda}, \ldots, \boldsymbol{u}'_{i,T}\boldsymbol{\Lambda})$, $\boldsymbol{D}_{\alpha\boldsymbol{\Psi},k} = Block - diag(\boldsymbol{\alpha}'_{k,1}\boldsymbol{\Psi}_k, \ldots, \boldsymbol{\alpha}'_{k,T}\boldsymbol{\Psi}_k)$ and $\boldsymbol{D}_{\beta\boldsymbol{\Xi},k} = Block{-}diag(\boldsymbol{\beta}'_{k,1}\boldsymbol{\Xi}_k, \ldots, \boldsymbol{\beta}'_{k,T}\boldsymbol{\Xi}_k)$. Then $\boldsymbol{v}_j|- \sim N(\boldsymbol{A}_v, \boldsymbol{\Sigma}_v)$, where
  $\boldsymbol{\Sigma}_v = \left(\sum_{i=1}^n \boldsymbol{D}'_{u\boldsymbol{\Lambda},i}\boldsymbol{D}_{u\boldsymbol{\Lambda},i} + \sum_{k=1}^p \boldsymbol{D}'_{\beta\boldsymbol{\Psi},k}\boldsymbol{D}_{\beta\boldsymbol{\Psi},k}\phi_k + ((1-\rho_1^2)\boldsymbol{D}_1^{-1}\otimes \boldsymbol{I}_R)\right)^{-1}$,
  $\boldsymbol{A}_v = \boldsymbol{\Sigma}_v\left[\sum_{i=1}^n \boldsymbol{D}'_{u\boldsymbol{\Lambda},i}(\boldsymbol{w}_{ij} - \boldsymbol{\mu}) + \sum_{k=1}^p \boldsymbol{D}'_{\beta\boldsymbol{\Xi},k}(\boldsymbol{Z}_{j,k} - \boldsymbol{D}_{\alpha\boldsymbol{\Psi},k}\boldsymbol{u}_j)\right]$

- $\boldsymbol{u}_i|- \sim N\left(\boldsymbol{A}_u, \boldsymbol{\Sigma}_u\right)$, where
  $\boldsymbol{\Sigma}_u = \left(\sum_{j=1}^{n} \boldsymbol{D}'_{v\boldsymbol{\Lambda},j}\boldsymbol{D}_{v\boldsymbol{\Lambda},j} + \sum_{k=1}^{p} \boldsymbol{D}'_{\alpha\boldsymbol{\Psi},k}\boldsymbol{D}_{\alpha\boldsymbol{\Psi},k}\phi_k + ((1-\rho_1^2)\boldsymbol{D}_1^{-1} \otimes \boldsymbol{I}_R)\right)^{-1}$,
  $\boldsymbol{A}_u = \boldsymbol{\Sigma}_u \left[\sum_{j=1}^{n} \boldsymbol{D}'_{v\boldsymbol{\Lambda},j}\left(\boldsymbol{w}_{i,j} - \boldsymbol{\mu}\right) + \sum_{k=1}^{p} \boldsymbol{D}'_{\alpha\boldsymbol{\Psi},k}\left(\boldsymbol{Z}_{i,k} - \boldsymbol{D}_{\beta\boldsymbol{\Xi},k}\boldsymbol{v}_i\right)\right]$.

- $\rho_1, \rho_2, \rho_3, \rho_4$ do not admit closed form full conditional distributions and are updated using the simple random walk Metropolis algorithm within the Gibbs sampler.

References

Albert, J. H. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association 88*(422), 669–679.

Anderton, C. H., N. Beck, J. R. Carter, H. Dorussen, E. Gartzke, R. Gissinger, N. Gleditsch, H. Hegre, J. S. Levy, Q. Li, et al. (2003). *Globalization and armed conflict*. Rowman & Littlefield Publishers.

Butland, G., J. M. Peregrín-Alvarez, J. Li, W. Yang, X. Yang, V. Canadien, A. Starostine, D. Richards, B. Beattie, N. Krogan, et al. (2005). Interaction network containing conserved and essential protein complexes in escherichia coli. *Nature 433*(7025), 531–537.

Christakis, N. A. and J. H. Fowler (2007). The spread of obesity in a large social network over 32 years. *n engl j med 2007*(357), 370–379.

De la Haye, K., G. Robins, P. Mohr, and C. Wilson (2010). Obesity-related behaviors in adolescent friendship networks. *Social Networks 32*(3), 161–167.

Doreian, P. (2001). Causality in social network analysis. *Sociological Methods & Research 30*(1), 81–114.

Durante, D. and D. B. Dunson (2014). Nonparametric bayes dynamic modelling of relational data. *Biometrika 101*(4), 883–898.

Durante, D., D. B. Dunson, et al. (2017). Bayesian inference and testing of group differences in brain networks. *Bayesian Analysis*.

Fosdick, B. K. and P. D. Hoff (2015). Testing and modeling dependencies between a network and nodal attributes. *Journal of the American Statistical Association 110*(511), 1047–1056.

Fowler, J. H. and N. A. Christakis (2008). Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the framingham heart study. *Bmj 337*, a2338.

Goldstein, J. S. (1992). A conflict-cooperation scale for weis events data. *Journal of Conflict Resolution 36*(2), 369–385.

Handcock, M. S., G. Robins, T. Snijders, J. Moody, and J. Besag (2003). Assessing degeneracy in statistical models of social networks. Technical report, Working paper.

Hoff, P. D. (2005). Bilinear mixed-effects models for dyadic data. *Journal of the american Statistical association 100*(469), 286–295.

Hoff, P. D. (2009). Multiplicative latent factor models for description and prediction of social networks. *Computational and mathematical organization theory 15*(4), 261.

Hoff, P. D. (2015). Multilinear tensor regression for longitudinal relational data. *The annals of applied statistics 9*(3), 1169.

Hoff, P. D., A. E. Raftery, and M. S. Handcock (2002). Latent space approaches to social network analysis. *Journal of the american Statistical association 97*(460), 1090–1098.

Holland, P. W., K. B. Laskey, and S. Leinhardt (1983). Stochastic blockmodels: First steps. *Social networks 5*(2), 109–137.

Holland, P. W. and S. Leinhardt (1981). An exponential family of probability distributions for directed graphs. *Journal of the american Statistical association 76*(373), 33–50.

Kalyagin, V., A. Koldanov, P. Koldanov, P. Pardalos, and V. Zamaraev (2014). Measures of uncertainty in market network analysis. *Physica A: Statistical Mechanics and its Applications 413*, 59–70.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics 28*(4), 825–860.

Miguel, E., S. Satyanath, and E. Sergenti (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of political Economy 112*(4), 725–753.

Morrow, J. D. (1999). How could trade affect conflict? *Journal of Peace Research 36*(4), 481–489.

Niezink, N. M. K. and T. A. B. Snijders (2016). Co-evolution of social networks and continuous actor attributes.

Polachek, S. W. (1980). Conflict and trade. *Journal of Conflict resolution 24*(1), 55–78.

Polachek, S. W. and D. Sevastianova (2012). Does conflict disrupt growth? evidence of the relationship between political instability and national economic performance. *The Journal of International Trade & Economic Development 21*(3), 361–388.

Reuveny, R. (2000). The trade and conflict debate: A survey of theory, evidence and future research. *Peace Economics, Peace Science and Public Policy 6*(1).

Reuveny, R. and H. Kang (1996). International trade, political conflict/cooperation, and granger causality. *American Journal of Political Science*, 943–970.

Robins, G., P. Pattison, Y. Kalish, and D. Lusher (2007). An introduction to exponential random graph (p*) models for social networks. *Social networks 29*(2), 173–191.

Rodríguez, A. and S. Moser (2015). Measuring and accounting for strategic abstentions in the us senate, 1989–2012. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 64*(5), 779–797.

Rodrik, D. (1999). Where did all the growth go? external shocks, social conflict, and growth collapses. *Journal of economic growth 4*(4), 385–412.

Sarkar, P. and A. W. Moore (2006). Dynamic social network analysis using latent space models. In *Advances in Neural Information Processing Systems*, pp. 1145–1152.

Scott, J. G., J. O. Berger, et al. (2010). Bayes and empirical-bayes multiplicity adjustment in the variable-selection problem. *The Annals of Statistics 38*(5), 2587–2619.

Sewell, D. K. and Y. Chen (2015). Latent space models for dynamic networks. *Journal of the American Statistical Association 110*(512), 1646–1657.

Shoham, D. A., R. Hammond, H. Rahmandad, Y. Wang, and P. Hovmand (2015). Modeling social norms and social influence in obesity. *Current epidemiology reports 2*(1), 71–79.

Snijders, T. A. (2002). Markov chain monte carlo estimation of exponential random graph models. *Journal of Social Structure 3*(2), 1–40.

Wasserman, S. and C. Anderson (1987). Stochastic a posteriori blockmodels: Construction and assessment. *Social Networks 9*(1), 1–36.

Watts, D. J. and P. Dodds (2009). Threshold models of social influence. *The Oxford handbook of analytical sociology*, 475–497.