Dynamic Additive and Multiplicative Effects (DAME) Model with Application to the United Nations Voting Behaviors

Bomin Kim[1], Xiaoyue Niu[1], David Hunter[1], and Xun Cao[2]

[1]Department of Statistics, The Pennsylvania State University
[2]Department of Political Science, The Pennsylvania State University

Abstract

In this paper, we introduce the dynamic version of a Gaussian additive and multiplicative effects (AME) model—a statistical regression model for symmetric discrete-time networks that are correlated over time. Our model extends the additive and multiplicative latent factor network model of Hoff (2009) and Minhas et al. (2016a) by incorporating the temporal correlation structure into the prior specifications of the parameters. The temporal evolution of the network is modeled through a Gaussian process (GP) as in Durante and Dunson (2013), whereas we estimate the unknown covariance structure from the dataset. We analyze United Nations voting data from 1983 to 2014 (Voeten et al., 2016) and show the effectiveness of our model at inferring the dyadic dependence structure among the international voting behaviors as well as allowing for a changing number of nodes over time. Overall, the DAME model shows significantly better fit to the dataset compared to the alternative approaches. Moreover, after controlling for other covariates such as geographic distances and bilateral trade between countries, the model-estimated latent positions and the movement of those positions reveal interesting and meaningful foreign policy positions and alliances of various countries.

1 Introduction

In recent decades, social network analysis has been well-established and is widely used in a variety of applications, ranging from friendship or collaboration networks to disease transmission. Because a network naturally evolves over time, there has been growing need for methods of modeling networks that change over time. A number of models have been suggested that are extensions of static network models, such as the temporal exponential random graph model (TERGM) (Hanneke et al., 2010) and the dynamic stochastic blockmodel (Xu and Hero III, 2013), or new models for network dynamics, such as the stochastic actor oriented model (SAOM) (Snijders et al., 2010). However, most previous studies on dynamic network modeling have given little consideration to the unobserved latent space (Hoff et al., 2002)——the structure of the network that is not explained through the use of exogenous node and dyad covariates. We therefore extend existing latent factor models (Hoff, 2005, 2009; Hoff et al., 2014; Minhas et al., 2016a) and develop the dynamic additive and multiplicative effects model (DAME) to model symmetric discrete-time networks, with emphasis on the latent structures——unmeasured attributes of nodes for tie formation in networks and their evolution over time.

Since Hoff et al. (2002) developed a class of models where the probability of a relation between actors depends on the positions of individuals in an unobserved "social space," there have been several versions of the latent projection model in which the probability of a tie beteween nodes i and j is determined by the two nodes' latent vectors using $f(v_i, v_j) = \frac{v_i' v_j}{|v_j|}$, gradually providing additional explicability and flexibility. Hoff (2005) first introduced the symmetric multiplicative interaction effect ($v_i' v_j$) into the network generalized linear model, in order to capture third-order dependence patterns——often described by the three features, transitivity, balance, and clusterability (Hoff, 2005). After modifying the parameterization of multiplicative effects into the form of an eigendecomposition $u_i^T \Lambda u_j$, Hoff (2008) claimed that this "latent eigenmodel" is able to represent

a wide array of patterns in the data, due to the fact that the eigenmodel provides an unrestricted low-rank approximation to the symmetric relational data. Hoff (2009) extended the framework to model asymmetric social networks using the singular value decomposition $u_i^T D v_j$, and finally, Hoff et al. (2014) and Minhas et al. (2016a) combined the additive and multiplicative effects to model the second-order (or reciprocity) and third-order dependencies and estabilished the AME regression model for a dyadic response data y_{ij} in which

$$
\begin{aligned}
&y_{ij} = g(\theta_{ij}),\\
&\theta_{ij} = \boldsymbol{\beta}^T \boldsymbol{X}_{ij} + e_{ij},\\
&e_{ij} = a_i + b_j + \epsilon_{ij} + \alpha(\boldsymbol{u}_i, \boldsymbol{v}_j), \text{ where}\\
&\alpha(\boldsymbol{u}_i, \boldsymbol{v}_j) = \boldsymbol{u}_i^T \boldsymbol{D} \boldsymbol{v}_j
\end{aligned}
$$

and $g(\theta_{ij})$ is a link function to model various type of relational matrix, such as continuous, binary, ordinal, or rank-based responses.

Although the latent space model is widely used for cross-sectional networks, networks at a specific point in time, the latent modeling of dynamic networks is relatively new in the statistics lieterature. The latent distance model has been applied to dynamic networks by many authors (Sarkar and Moore, 2005; Sarkar et al., 2007; Sewell and Chen, 2015, 2016; Friel et al., 2016), while a comparatively small literature employs the AME framework to develop dynamic network models. Ward and Hoff, 2007 first introduced the concept of dynamic latent factors and this idea was expanded by Ward et al. (2013) to analyze bilateral trade using the generalized bi-linear mixed effect model (Hoff, 2005). This model allowed time-varying edge covariate parameters as well as latent factors, however, it did not involve any dynamic information in the dataset. There also exists a series of papers for modeling a tensor representation of network or multi-way array (Hoff, 2011; Hoff et al., 2011; Hoff, 2015; Minhas et al., 2016b), but this framework can be considered as larger class of models since the applications are not limited to longitudinal networks.

Focusing on the temporal aspect of networks, Durante and Dunson (2013) proposed the dynamic latent space model for binary symmetric matrices, which assumes that the latent factors are evolving in continuous time via Gaussian processes (GP), with the model formulation

$$
\begin{aligned}
&y_{ij,t} | \pi_{ij}(t) \sim \text{Bern}(\pi_{ij}(t)),\\
&\pi_{ij}(t) = \frac{1}{1 + e^{-s_{ij}(t)}},\\
&s_{ij}(t) = \mu(t) + x_i(t)' x_j(t)
\end{aligned}
$$

for $i < j$, with $x_{ih}(\cdot) \sim \text{GP}(0, \tau_h^{-1} c_X)$ and $\mu(\cdot) \sim \text{GP}(0, c_\mu)$, where $x_i(t) = [x_{i1}(t), ..., x_{iH}(t)]'$ for $i = 1, ..., V$ are the latent vectors of node i, together with the baseline $\mu(t)$, τ_h^{-1} for $h = 1, ..., H$ is a shrinkage parameter with a gamma prior, and c_μ and c_x are the squared exponential functions (i.e. $c_X(t, t') = \exp(-\kappa_X ||t - t'||_2^2)$ and $c_\mu(t, t') = \exp(-\kappa_\mu ||t - t'||_2^2)$). This approach is based on nonparametric Bayesian inference and has the strong advantage of learning the number of latent dimensions H in the model (Bhattacharya and Dunson, 2011). Although the model has been successfully applied to different types of longitudinal networks (Durante and Dunson, 2014a,b), it lacks several benefits of the AME model. First, model (1) does not include the additive effects a_i and b_j, which can capture significant heterogeneity in activity levels across nodes. Second, model (1) uses the term $x_i(t)' x_j(t)$ to represent multiplicative latent factor effects. However, the form $u_i(t)^T D(t) u_j(t)$ used in the AME model allows the flexibility of negative eigenvalues $d_r(t)$. Depending on the sign of $d_r(t)$, similar values of the latent coordinates $u_{ir}(t)$ and $u_{jr}(t)$ will contribute positively or negatively to the relationship between i and j in the specific latent dimension r. According to Hoff (2008), the parameterization of $u_i(t)^T D(t) u_j(t)$ can represent both positive or negative transitivity in varying degrees; on the other hand, the parameterization of $v_i(t)^T v_j(t)$ is not able to explain negative transitivity or stochastic equivalence, where nodes with the same or similar latent vectors do not have trong relationships with one another.

We propose a new model, the dynamic additive and multiplicative effects (DAME) model, combining the advantages of the AME model and the dynamic latent space model; we use the same formulation as AME model, while incorporating the time-varying prior structures of Durante and Dunson (2013). In addition, the DAME employs two innovations. First, we learn the temporal correlation of networks by estimating Gaussian process length parameters. Instead of using fixed covariance structure (Durante and Dunson, 2013, 2014b), this does not require any initial guess on correlations and further enables efficient estimation on the fixed and random effect parameters. Second, in order to increase flexibility and accuracy of the model, the DAME allows the number of nodes to change over time by allowing a special case of missing values, which are referred to as "structural zeros" in this paper. In what follows, we present the model formulation of the DAME, describing how we take advantage of temporal correlation; derive the sampling equations for hierarchical Bayesian inference with the DAME; present simulation studies to show the advantage of the DAME over the comparative approaches; and apply the DAME model to the United Nations voting network.

2 Dynamic Additive and Multiplicative Effects Model

2.1 Model Formulation

Our goal is to simultaneously model the sequence of $N \times N$ time-varying symmetric matrices $\mathbf{Y} = \{Y^1, ..., Y^T\}$, where the entry y_{ij}^t denote any relational data corresponding to the node i and node j, at timepoint t. Given the observed covariates $\mathbf{X} = \{X^1, ..., X^T\}$ where X^t consists of P number of covariates (i.e. $X^t = \{X_1^t, \ldots, X_P^t\}$), we define the model as

$$y_{ij}^t = \sum_{p=1}^{P} \beta_p^t X_{ijp}^t + z_{ij}^t, \tag{1}$$

independently for each $i = 2, ..., N$,$j = 1, ..., i-1$, and $t = 1, ..., T$, where X_{ijp}^t is the p^{th} edge covariates, β_p^t is the corresponding coefficient, and z_{ij}^t is the unobserved random effects. Following the latent factor model of Minhas et al. (2016a), we also model z_{ij}^t to account for potential higher-order dependencies in the additive and multiplicative form

$$z_{ij}^t = \theta_i^t + \theta_j^t + u_i^{t\prime} D^t u_j^t + \epsilon_{ij}^t, \tag{2}$$

where θ_i^t and θ_j^t are the additive nodal random effects of the node i and j, respectively, $u_i^{t\prime} D^t u_j^t$ is the multiplicative random effect with D^t denoting the $R \times R$ diagonal matrix (i.e. $D^t = \text{diag}(d_1^t, ..., d_R^t)$ and $u_i^t = [u_{i1}^t, ..., u_{iR}^t]'$ denoting the R-length vector of latent coordinates of node i, and ϵ_{ij}^t is the noise.

Following the prior specifications in Bhattacharya and Dunson (2011) and Durante and Dunson (2013), we assume independent Gaussian process (GP) priors for the parameters $\{\beta_p\}_{p=1}^P$, $\{\theta_i\}_{i=1}^N$, $\{d_r\}_{r=1}^R$ and $\{\{u_{ir}\}_{i=1}^N\}_{r=1}^R$. We also assign additional hierarchy to include a shrinkage parameters $\{\tau_p^\beta\}_{p=1}^P$, τ^θ, and $\{\tau_r^u\}_{r=1}^R$ using inverse-Gamma (IG) priors. Specifically, we let

1. $\beta_p \sim \text{GP}(0, \tau_p^\beta c_p^\beta)$ for $p = 1, ..., P$, where $\tau_p^\beta \sim \text{IG}(a_\beta, b_\beta)$ and $c_p^\beta(t, t') = f(\kappa_p^\beta, |t - t'|)$;
2. $\theta_i \sim \text{GP}(0, \tau^\theta c^\theta)$ for $i = 1, ..., N$, where $\tau^\theta \sim \text{IG}(a_\theta, b_\theta)$ and $c^\theta(t, t') = f(\kappa^\theta, |t - t'|)$;
3. $d_r \sim \text{GP}(0, \tau_r^d c_r^d)$ for $r = 1, ..., R$, where $\tau_r^d \sim \text{IG}(a_d, b_d)$ and $c_r^d(t, t') = f(\kappa_r^d, |t - t'|)$;
4. $u_{ir} \sim \text{GP}(0, I_T)$ for $i = 1, ..., N$ and $r = 1, ..., R$;
5. $\epsilon_{ij} \sim \text{N}(0, \sigma_e^2)$, where $\sigma_e^2 \sim \text{IG}(a_\sigma, b_\sigma)$,

where $\beta_p = [\beta_p^1, \ldots, \beta_p^T]'$ and the rest of vectors defined in the same manner, (c^β, c^θ, c^d) are the $T \times T$ dimensional GP covariance matrices corresponding to the respective parameters, and I_T is $T \times T$ identity matrix. Since the multiplicative random effect is measured by $u_i^{t\prime} D^t u_j^t$, we assign the temporal correlation structure and shirinkage parameter both to d_r^t to prevent from non-identifiability issue.

In the existing literatures (Bhattacharya and Dunson, 2011; Durante and Dunson, 2013, 2014b), Gaussian process (GP) has specified form of covariance functions, squared exponential correlation function, which is appropriate for modelling very smooth functions. Moreover, the parameter κ that characterize length-scale of the process is fixed as a input variable. However, smoothness is not necessarily required for discrete-time network modeling, and in practice, prior knowledge on how much the networks are correlated over time is hardly available. In that sense, choosing the right value of κ can be a big challenge, especially when the pre-specified κ determines the entire smoothness of functions. To better capture the network correlations over time as well as to allow large amount of flexibility, we instead estimate the value of κ using the standard Exponential covariance function:

$$f(\kappa, |t - t'|) = \exp\left(-\frac{|t - t'|}{\kappa}\right),$$

with the prior given as $\kappa \sim$ half-Cauchy$(0, 5)$—a relatively weakly-informative prior. Note that when timepoints are evenly spaced such as monthly or annually, the covariance matrix becomes symmetric Toeplitz matrix, a matrix in which each descending diagonal from left to right is constant. Following the convention, we compute the inverse of Toeplitz matrices using Trench and Durbin methods (Golub and Van Loan, 2012).

2.2 Posterior Computation

We take a Bayesian approach to make inference for the parameters of the DAME. Posterior computation is performed via simple Gibbs sampler to update the vector of time-varying regression coefficients and the vector of additive and multiplicative latent factors, while we use Metropolis-Hastings to sample the pair of variance and length parameters (τ, κ) in Gaussian process. Using the conjugate prior distributions, this section provides a Gibbs sampling scheme that approximates the joint posterior distribution $P(\beta, \theta, d, u, \tau^\beta, \tau^\theta, \tau^d, \kappa^\beta, \kappa^\theta, \kappa^d, \sigma_e^2 | \mathbf{Y}, \mathbf{X}, a_\beta, b_\beta, a_\theta, b_\theta, a_d, b_d, a_\sigma, b_\sigma)$.

Let E^t denote the $N \times N$ matrix of random noises, where the $(i,j)^{th}$ entry is defined as $E_{ij}^t = y_{ij}^t - \big(\sum_{p=1}^{P} \beta_p^t X_{ijp}^t + \theta_i^t + \theta_j^t + u_i^{t\prime} D^t u_j^t\big)$. Given that the distribution of the observed network $\mathbf{Y} = \{Y^1, ..., Y^T\}$ conditional on all the parameters can be written as

$$P(\boldsymbol{Y}|\boldsymbol{X}, \beta, \theta, d, u, \sigma_e^2) \propto \prod_{t=1}^{T} \prod_{i>j} (\sigma_e^2)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma_e^2} ||E_{ij}^t||^2\}, \tag{3}$$

we sequentially update each parameter (σ_e^2, τ^β, τ^θ, τ^d, κ^β, κ^θ, κ^d, β, θ, d, u) from its full conditional distribution in the following sampling scheme:

1. Sample $\sigma_e^2 \sim \text{IG}\big(\frac{T \cdot N(N-1)}{4} + a_\sigma, \frac{1}{2} \sum_{t=1}^{T} \sum_{i>j} (E_{ij}^t)^2 + b_\sigma\big)$;

2. For each $p \in \{1, ..., P\}$ in random order, sample β_p as follows:

   (a) Sample $(\tau_p^\beta, \kappa_p^\beta)$ using bivariate Metropolis-Hastings algorithm

   (b) Sample $\beta_p \sim \text{MVN}_T\big(\tilde{\mu}_{\beta_p}, \tilde{\Sigma}_{\beta_p}\big)$ with

   $$\tilde{\Sigma}_{\beta_p} = \Big((\tau_p^\beta c_p^\beta)^{-1} + \frac{\text{diag}\big(\{\sum_{i>j} X_{ijp}^{t2}\}_{t=1}^{t=T}\big)}{\sigma_e^2}\Big)^{-1} \text{ and } \tilde{\mu}_{\beta_p} = \Big(\frac{\{\sum_{i>j} (E_{ij[-p]}^t X_{ijp}^t)\}_{t=1}^{t=T}}{\sigma_e^2}\Big)\tilde{\Sigma}_{\beta_p},$$

   where $E_{ij[-p]}^t = E_{ij}^t + \beta_p^t X_{ijp}^t$.

3. Sample θ_i as follows:

   (a) Sample $(\tau^\theta, \kappa^\theta)$ using bivariate Metropolis-Hastings algorithm

(b) For each $i \in \{1, ..., N\}$ in random order, sample $\theta_i \sim \text{MVN}_T\big(\tilde{\mu}_{\theta_i}, \tilde{\Sigma}_{\theta_i}\big)$ wiwth

$$\tilde{\Sigma}_{\theta_i} = \Big((\tau^\theta c^\theta)^{-1} + \frac{(N-1)I_T}{\sigma_e^2}\Big)^{-1} \text{ and } \tilde{\mu}_{\theta_i} = \Big(\frac{\{\sum_{i=i,j\neq i} E^t_{ij[-i]}\}_{t=1}^{t=T}}{\sigma_e^2}\Big)\tilde{\Sigma}_{\theta_i},$$

where $E^t_{ij[-i]} = E^t_{ij} + \theta^t_i$.

4. For each $r \in \{1, ..., R\}$ in random order, sample d_r as follows:

   (a) Sample (τ^d_r, κ^d_r) using univariate Metropolis-Hastings algorithm

   (b) Sample $d_r \sim \text{MVN}_T\big(\tilde{\mu}_{d_r}, \tilde{\Sigma}_{d_r}\big)$ with

$$\tilde{\Sigma}_{d_r} = \Big((\tau^d_r c^d_r)^{-1} + \frac{\text{diag}\big(\{\sum_{i>j}(u^t_{ir}u^t_{jr})^2\}_{t=1}^{t=T}\big)}{\sigma_e^2}\Big)^{-1} \text{ and } \tilde{\mu}_{d_r} = \Big(\frac{\{\sum_{i>j}(E^t_{ij[-r]}u^t_{ir}u^t_{jr})\}_{t=1}^{t=T}}{\sigma_e^2}\Big)\tilde{\Sigma}_{d_r}),$$

   where $E^t_{ij[-r]} = E^t_{ij} + {u^t_{ir}}' d^t_r u^t_{jr}$.

5. For each $t \in \{1, ..., T\}$ and $i \in \{1, ..., N\}$ in random order, sample u^t_i as follows:

   (a) Sample $u^t_i \sim \text{MVN}_R\big(\tilde{\mu}_{u^t_i}, \tilde{\Sigma}_{u^t_i}\big)$ with

$$\tilde{\Sigma}_{u^t_i} = \Big(1 + \frac{\sum_{j\neq i} D^t u^t_j {u^t_j}' D^t}{\sigma_e^2}\Big)^{-1} \text{ and } \tilde{\mu}_{u^t_i} = \Big(\frac{\{\sum_{i=i,j\neq i}(E^t_{ij[-u]}{u^t_j}' D^t)'\}}{\sigma_e^2}\Big)\tilde{\Sigma}_{u^t_i},$$

   where $E^t_{ij[-u]} = E^t_{ij} + {u^t_i}' D^t u^t_j$,

where IG denotes inverse-Gamma distribution and MVN denotes multivariate normal distribution. Note that between every steps 1–5, $\mathbf{E} = \{E^1, \ldots, E^T\}$ has to be calculated again using the previously updated values, so that the next update is conditioned on the current values of all the other parameters. For the parameters which we know the full conditional distributions, we achieve easy implementation and fast computation via Markov Chain Monte Carlo (MCMC) with Gibbs samping. For joint sampling of κ and τ, we use the simplest case of Metropolis-Hastings algorithm using Normal proposals, where the detailed derivations can be found in the supplementary material of the paper.

2.3 Varying Number of Nodes

In many longitudinal networks, new node can join the network or existing node can disappear at any timepoint, and modeling the composition of nodes itself is a critical feature of dynamic network models. For example, continuous-time network models which are built upon survival analysis (Snijders et al., 2010; Butts, 2008; Vu et al., 2011; Hunter et al., 2011; Perry and Wolfe, 2013) naturally allow the number of nodes vary over time from the model formulation itself, while the stochastic actor oriented model (SAOM) (Snijders et al., 2010) manually added some data specification options to allow 'joiners' and 'leavers', by specifying structural zeros for all elements of the variables from actors (or nodes) who are absent at a given timepoint. Nevertheless, statistical methods to allow varying size of nodes have yet to be adapted to the most of exisiting dynamic latent distance models (Sarkar and Moore, 2005; Sewell and Chen, 2015; Friel et al., 2016), possibly due to their reliance on Markovian dependence to model the temporal evolution of networks. Although not assuming any Markovian properties, none of dynamic latent factor models (Durante and Dunson, 2013; Hoff, 2008; Minhas et al., 2016a) has been applied to longitudinal networks where the number of nodes change over time. The common practice is to model the dataset assuming fixed size of nodes over time and rely on imputations for missing values.

In our paper, we propose a straightforward method to allow nodes to join or leave the network, similar to the one used in Snijders et al. (2010). We define structural missing values where missing is not at random as 'structural zeros'. For example, when a country or company does not exist or disappears until/after specific timepoint, no single covariates is observed before or after then, thus

becoming structural zeros. In such cases, imputation could be either impossible or problematic and estimating the nodal additive or multiplicative effects for the non-existing period may be of no use. Even when imputation is possible, including the imputed values to estimate other (fixed) parameters could also reduce bias. To avoid those issues, we allow the model to differentiate two types of missingness, random missing and structural zeros, and handle the latter separately from the treatment of missing data.

The dynamic additive and multiplicative effects model takes the $N \times T$ matrix of availability A as an input, where $(n,t)^{th}$ element denotes the the availability of node n at time t. This matrix must contain all actors N who are part of the network at any observation time $t \in \{1, ..., T\}$. The indicator variable A_{nt} is then defined as below.

$$A_{nt} = \begin{cases} 1, & \text{node } n \text{ is available at timepoint } t \\ 0, & \text{node } n \text{ is not available at timepoint } t. \end{cases}$$

While random missing values (i.e. $A_{nt} = 1$) are imputed with posterior estimates and used in every steps of parameter estimation, those with structural zeros (i.e. $A_{nt} = 0$) are not imputed and left as missing, and the corresponding additive effect estimates θ_n^t and latent position estimates u_n^t are specified as 'not available (NA)'. Furthermore, other parameters including the fixed effects $\{\beta_p\}_{p=1}^P$ are estimated excluding the entries of structural zeros. Through this simple and straightforward approach, we can effectively model dynamic networks of varying number of nodes over time, which provides flexibility in fitting the model to large networks.

3 Simulation Study

We provide a simulation study with the aim to evaluate the performance of the proposed model on the ability to capture some important properties of the true data and correctly reconstruct the true underlying processes from the model estimates. There are two objectives in this simulation study: 1) show that understanding the correct covariance structure plays a key role in the model performance, in case of network that is highly correlated across time, and 2) prove that the eigenvalue formulation of the multiplicative random effects (i.e. $u'Du$) has the benefit of revealing negative transitivity effects.

3.1 Estimating Strong Correlations

We generated a set of relational data $\mathbf{Y}$ for $N = 20$ and $T = 10$ according to the generative process in Section 2.1, with $P = 1$ (intercept only), $R = 2$, all inverse-Gamma hyperparameters $(a, b) = (2, 1)$, and $(\kappa^\beta, \kappa^\theta, \kappa^d) = (10, 10, 10)$ such that the resulting dynamic network is highly correlated across time. We ran 6,000 MCMC iterations which proved to be enough for reaching convergence and discarded the first 1,000 samples, with thinning = 10.

To capture the overall correlation of the dataset, we define the new measure and refer to "lagged degree correlations" (dc). For the lag $l = 1, \ldots, T-1$,

$$dc_l = \rho(\text{degrees}^{1,\ldots(T-l)}, \text{degrees}^{(1+l),\ldots T}), \tag{4}$$

where $\rho(\cdot)$ is the Spearman correlation between two vectors, and $\text{degrees}^{1,\ldots(T-l)}$ and $\text{degrees}^{(1+l),\ldots T}$ indicate the vector of all degrees from timepoint $t = 1$ to $t = (T - l)$ and $t = (1 + l)$ to $t = T$, respectively, which result in both to be $N \times (T - l)$-length vector.

Figure 1 compares the lagged degree correlations of our model and the independence model —without estimating Gaussian process covariance parameters (and fix all $\kappa = 0$), which are constructed from their respective posterior estimates. This comparison highlights the good performance of the DAME, in correctly estimating the true temporal correlation across the timepoints. These can be noticed by comparing with the true degree correlations, where the independence model always exhibit lower posterior estimates than the true correlations.

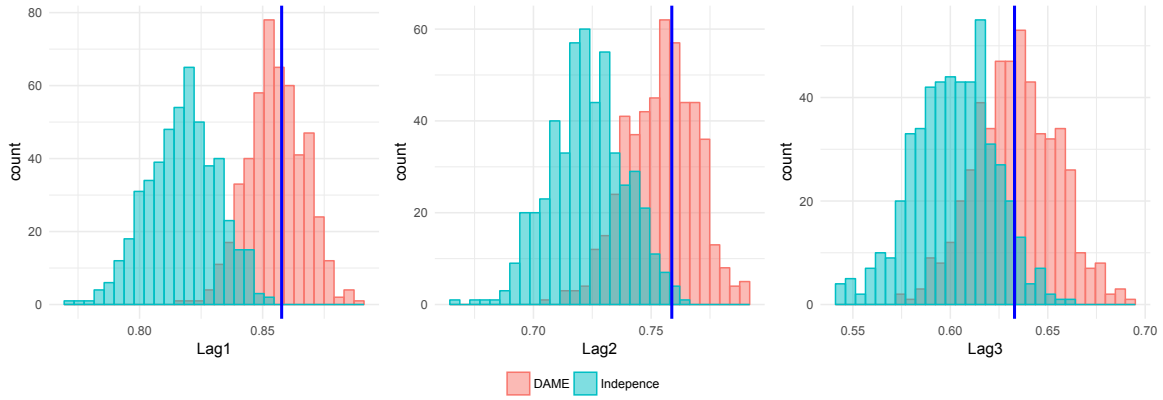

Figure 1: Histogram of posterior lagged degree correlations (from lag 1 to lag 3): the DAME (red) and independence model (blue), with the observed lagged degree correlations (blue lines)

3.2 Capturing Transitivity

Similar to Section 3.1, we generated $\mathbf{Y}$ with $N = 20$, $T = 10$, $P = 1$ (intercept only), $R = 2$, and $(a, b) = (2, 1)$ for all inverse-Gamma hyperparameters. On the other hand, considering that our new goal is not to estimate temporal correlations but to represent positive or negative transitivity, this time we modify two settings: 1) set $(\kappa^\beta, \kappa^\theta, \kappa^d) = (0.001, 0.001, 0.001)$ such that all paramters and the resulting dynamic network is nearly independent over time, and 2) force $d_r^t > 0$ or $d_r^t < 0$ for all $r = 1, 2$ and $t = 1, ..., 10$ such that the generated network exibits positive or negative transitive feature, respectively. Again, we ran 6,000 MCMC iterations and discarded the first 1,000, with thinning = 10, however, we fixed κ's at the true values and did not estimate the covariance paramters such that the difference in performance only comes from the multiplicative effects formulations, $u'Du$ and $u'u$. If we estimated κ's for this comparison, the difference in results may not only arise from the formulation of multiplicative effects, but also possibly come from lack of correlation structure in $u'u$ because our modeling framework do not assign any correlation on u's.

Figure 2 shows a graphical comparison between the two formulations of the multiplicative random effects with respect to the second and third moments of degree (i.e. degree($\hat{Y}^2$) and degree($\hat{Y}^3$), respectively), for a randomly chosen node (node 2). For the case of positive transitivity ($d > 0$), our model and its alternative (without D term) do not show significant differences; both formulations achieve great performance in replicating the first, second, and third moments of node 2's degree. On the contrary, when we fit the network with negative transitivity ($d < 0$), the two formulations reveal noticeable differences. While the DAME can still recover the true degree in all three moments, the model without D term shows inaccuracy in simulating a network that is close to the true data, especially in higher moments. Not only the $u'u$ model introduces bias, but also it shows lower precision than our model. This notable difference strongly supports our argument that $u'Du$ should be preferred over $u'u$, as a standard form of multiplicative random effects.

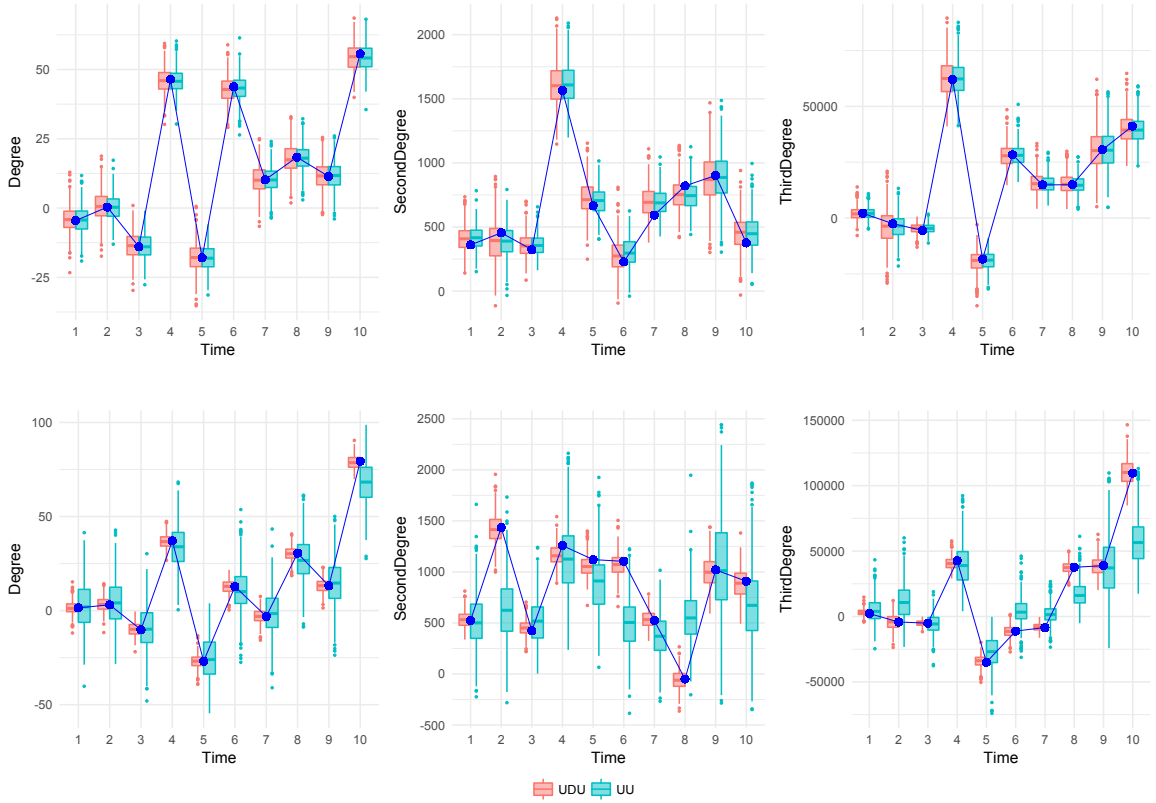

Figure 2: Posterior distribution of the first (left), second (middle), and third (right) moment degrees for node 2 using positive transitive network (upper) and negative transitive network (lower): the DAME (red) and without D model (green), with the observed moment degrees (blue dots)

4 Analysis of the United Nations Voting Network

4.1 Data

Votes in the United Nations General Assembly (UNGA) have been analyzed in many political science papers (Voeten, 2000, 2004; Bearce and Bondanella, 2007; Mattes et al., 2015; Bailey et al., 2017), and became one of the standard data sources to measure the state preferences. Unfortunately, to our knowledge, the existing studies ignored two important features of the dataset. First, votes are highly correlated across timepoints, because they are the reflections from history. Bailey et al. (2017) proposed a dynamic ordinal spatial IRT (item response theory) model and allowed for better inter-temporal comparisons, however, their model limited the temporal dependence to be lag 1 (i.e. Markovian assumption — votes at time $(t+1)$ is only dependent on votes at time t). Second, although the researchers have viewed "voting" as dyadic behavior and thus used dyadic similarity indicators such as affinity or S scores (Gartzke, 1998; Signorino and Ritter, 1999), to our knowledge, the United Nations voting data has never been analyzed using the dynamic network models.

To overcome the limitations and provide the new insight, we apply the DAME to the voting behavior of 74 countries in the United Nations (UN) General Assembly roll-call votes over the 32 year period, from 1983 through 2014. Full list of the 74 countries included in this analysis and their abbreiviationsare listed in Table 5 in Appendix A. These data were obtained from Voeten et al. (2016), where we use the subset of the votes called 'important votes', the votes identified as important by U.S. State Department report Voting Practices in the United Nations. Throughout 1983 to 2014, the number of important votes on average is 12 per year, ranges from 6 to 28. Specifically, we construct the response $\mathbf{Y} = \{Y^1, \ldots, Y^{32}\}$, where each Y^t is 74×74 matrix, using the variable

'agree3unimportant' in the original dataset, which is a voting similarity index (0–1) computed using 3 category vote data (Y = "yes" or approval for an issue; A = abstain; N = "no" or disapproval for an issue)[1]. Note that abstention is counted as half-agreement with a yes or no vote, while two abstention is treated as full-agreement. For basic summary of Unite Nations voting data including the average number of common votes and the averge proportion of agreement or voting similarity index per year, see Table 3 in Appendix B.

As an exploratory data analysis, Table 1 illustrates the lagged degree correlation (defined in Equation (4)) of the dataset, which is used in Section 3.1, to measure how much Unite Nations voting data is correlated over time. There exists strong positive correlation in how countries vote in the United Nations General Assembly over time, and as the distance between two timepoints become larger the correlation gets weaker. This provides solid evidence that our model with Gaussian process assumption is one of the appropriate appraoches to effectively analyze this dataset.

Lag	1	2	3	4	5	6	7	8	9	10
Correlation	0.800	0.745	0.669	0.589	0.541	0.430	0.365	0.246	0.218	0.203

Table 1: Lagged degree correlation of United Nations Voting Data

Next, to dynamically model the voting network in relation to other variables reflecting international relations, we combine $P = 5$ different dyadic variables from the Correlates of War (COW) data (Gibler, 2008), Polity IV data (Marshall et al., 2014) and the International Monetary Fund (IMF)'s Direction of Trade Statistics (DOTS) and International Financial Statistics (IFS) data, and construct the observed edge covariates $\mathbf{X}$. For every $p = 1, ..., 5$, we set the explanatory variable X^t_{ijp} as below:

1. Intercept$^t_{ij}$: constant 1,
2. log(distance)$^t_{ij}$: log of the geographic distance between country i and country j,
3. Alliance$^t_{ij}$: 1 if country i and country j have alliance, and 0 otherwise,
4. Polity difference$^t_{ij}$: absolute difference in polity IV number between country i and country j,
5. Lower trade-to-GDP ratio$^t_{ij}$: index of economic dependence using bilateral trade weighted by each country's domestic product (GDP), defined in Gartzke (2000)

$$min\Big(\frac{\text{Trade}_{ijt}}{\text{GDP}_{it}}, \frac{\text{Trade}_{ijt}}{\text{GDP}_{jt}}\Big),$$

6. Common language$^t_{ij}$: indicator of whether country i and country j share the same language,

where all covariates are symmetric $X^t_{ijp} = X^t_{jip}$. Note that intercept is included to account for the baseline degree of agreement at each timepoint. Two variables, log(distance) and common language are time-invariant covariates, although their coefficients may vary over time.

Moreover, we specified the matrix of availability A introduced in Section 2.3 to reflect six countries' nonparticipation in the United Nations Genearl Assembly:

1. North Korea (PRK) have structural zeros from $t = 1$ to $t = 8$:
   - North Korea did not vote until North Korea and South Korea have simultaneously admitted to United Nations in 1991.
2. South Korea (ROK) have structural zeros from $t = 1$ to $t = 8$:
   - South Korea did not vote until North Korea and South Korea have simultaneously admitted to United Nations in 1991.

[1] voting similarity index between the two countries i and j at year t is calculated as (Number of votes i and j agreed at year t) / (Number of votes i and j both participated at year t)

3. Ukraine (UKR) has structural zeros from $t = 1$ to $t = 9$:
   - Previously Ukrainian Soviet Socialist Republic, the country changed its name to Ukraine on 1991 and started voting with new name from 1992.

4. Russia (RUS) has structural zeros from $t = 1$ to $t = 9$:
   - Russia succeeded the Soviet Union's seat, including its permanent membership on the Security Council in the United Nations after the dissolution of the Soviet Union in 1991.

5. Georgia (GRG) has structural zeros from $t = 1$ to $t = 10$:
   - Georgia, one of the former Soviet Republics, was admitted to United Nations on 1992, thus no voting records until 1992.

6. Iraq (IRQ) has structural zeros from $t = 13$ to $t = 21$:
   - The country did not participate the UNGA roll-call votes during 1995 - 2003. Since 1990, Iraq was under severe sanctions from the international community including the United Nations.

Therefore, any missing values corresponding the country and missing period are treated as structural zeros. In Section 2.3, we mentioned that other non-structural missing values are treated as random missing (and thus imputed), however, this dataset does not include any random missings.

4.2 Reduced-Rank Structure of Voting Network

In this section, we demonstrate the special low-rank structure of the voting network, which can be generalized into any other similarly constructed agreement networks. First, for simplicity, we assume static network by fixing $T = 1$ and consider the $N \times N$ adjacency matrix V^1 representing the network consists of single vote. Since each entry of V^1 denotes a voting similarity index (0-1) computed using 3 categories ("yes", "abstain", and "no"), the $(i, j)^{th}$ element V^1_{ij} can only have three possible values,

$$V^1_{ij} = \begin{cases} 1, & \text{agreement} \\ 0.5, & \text{half-agreement (one abstain)} \\ 0, & \text{disagreement} \end{cases}$$

which implies perfect transitivity of the agreement network, i.e. if there is an agreement between i and j, and also between j and h, then there must be an agreement between i to h, and therefore any two edges sharing one node, such as V^1_{ij} and V^1_{jh}, automatically determines the third edge between the unshared nodes V^1_{ih}. This constraint makes the maximum rank of V^1 to be 3, thus we can use low rank factorization of the matrix V^1 using $R = 3$. If we were to apply the DAME model to this type of single vote network (without additive effects and explanatory variables), the maximum size of dimension for the latent factors we could fit is $R = 3$ and the estimated latent factors can be viewed as the distinct constructs behind the vote.

Although it is worth understanding the constraint behind single votes, this is not a practical issue in modeling United Nations voting network, where we aggregate multiple votes per year (minimum number of important votes per year is 6). When we add M number of matrices representing each vote, the aggregated matrix $V^1 + V^2 + ... + V^M$ gets closer to the full rank thus we no longer need to worry about the maximum dimension of the latent factors in fitting the DAME model. Moreover, including the observed covariates X and additive random effects θ also relax the reduced rank structure, since the multiplicative latent effect $u'Du$ is modeled after we subtract those effects from the response matrix or array (See Section 2). Accounting for large variability in the observed covariates and additive random effects, the residuals no longer maintains the perfect transivity structure.

4.3 Model Validation

To perform a model validation as well as to understand the role of different effects at the same time, we fitted four different dynamic AME models, one with additive and multiplicative effects, one with only multiplicative effects, one with only additive effects, and the last without any random effects, where all four models included 6 fixed effects in Section 4.1. We constructed 95% credible intervals

for degree statistics of the model-simulated response $\hat{\mathbf{Y}}$ at every 10^{th} sample after burn-in.

Figure 3 shows the posterior predictive plots for the country "Israel" comparing the four models. First of all, there is huge benefit of correcting the bias when we add additive effects, compared to the model with no random effects. Next, when we compare between AE (additive only) and ME (multiplicative only), it is apparent that that the multiplicative effects significantly reduced the interval of estimates. This happens since the additive effects have very small variances (≈ 0) while the multiplicative effects have larger variance estimates, which significantly lowered the error variance estimates $\hat{\sigma}_e^2$ that plays a big role in the precision of simulated data. Our model with both additive and multiplicative effects seems to perform similar to the ME model (in this aggregated plot), however, further analysis showed that the DAME still outperform the ME model in terms of both accuracy and precision. Overall, not only the DAME model showed the most accurate estimates over time, but it also had the narrowest 95% credible intervals among the four. This finding emphasizes the importance of including the additive and multiplicative term in the model, capturing some features not explicable by fixed effects and maximizing model performance. More results using the full DAME model are demonstrated in Section 4.4.

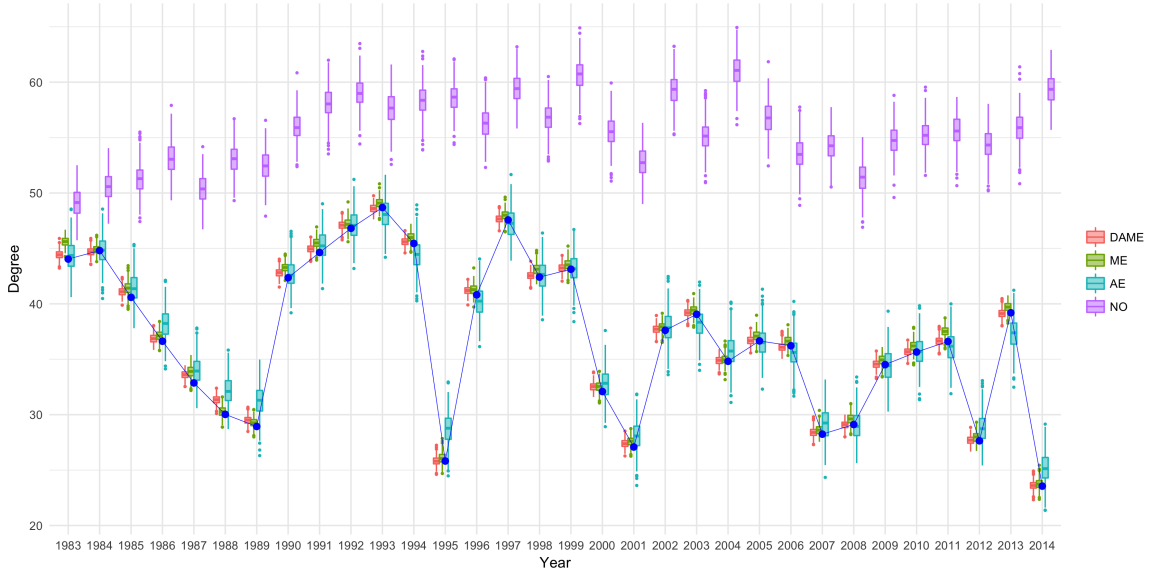

Figure 3: Posterior predictive boxplots based on 95% credible intervals for degree staitstics of Israel (ISR), comparing three models: the DAME (red), only multiplicative effects (green), only additive effects (blue), and no random effects (blue), with the blue dots represent the country's observed degree statistic over 32 years.

4.4 Parameter Estimation and Interpretation

In this section, we apply the model with the dimension of latent factors chosen as $R = 2$, based on some preliminary experiments where increasing the dimension of latent factors did not significantly improve the model fitting. For example, when $R = 3$, the last value of estimated eigenvalue d_3. is very close to zero for the most of the timepoints $t = 1, ..., 32$. Thus, we present the results from the same settings as Section 4.3, which are $6,000$ Gibbs iterations with a burn-in of $1,000$, where every 10^{th} sample was taken as a thinning. All inverse-Gamma hyperparameters $(a, b) = (2, 1)$, and the covariance parameters (κ, τ) were estimated, following Section 2.2. Figure 4 shows the posterior estimates of the fixed effect coefficients $\{\beta_p\}_{p=1}^6$ with their 95% credible intervals. Interpretation of fixed effect plot

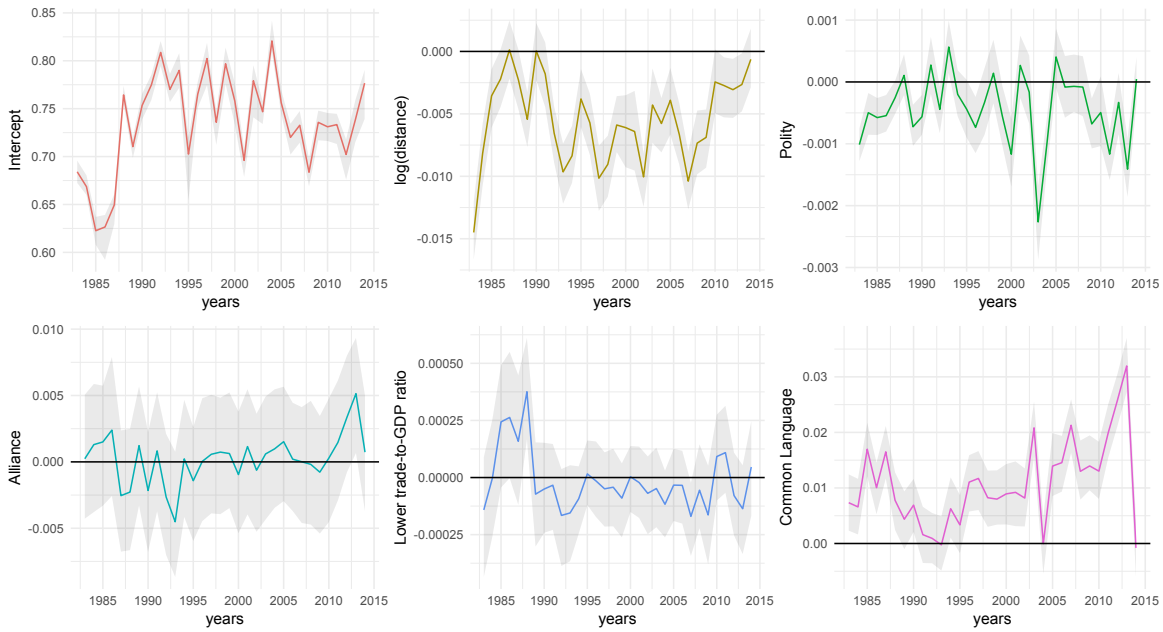

Figure 4: Posterior mean for the fixed effect coefficients $\{\beta_p\}_{p=1}^{6}$ (colored line): Intercept, log(distance), Alliance, Polity difference, Lower trade-to-GDP ratio, and Common Language, and their corresponding 95% credible intervals (grey areas).

After controlling for the observed covariates, we move on to the analysis of random effects, both additive and multiplicative ones. For clear visualization, we only present the results from 23 countries, where the countries are chosen based on the most active countries during the ten year period of 2004 – 2014 (Hoff, 2015). Here, the action types include negative material actions, positive material actions, negative verbal actions and positive verbal actions. The 23 countries are marked with $*$ in Appendix A. Figure 5 is the posterior mean estimates of each node or country's additive random effects which capture the node-specific and within-dyad dependence. Interpretation of additive random effects

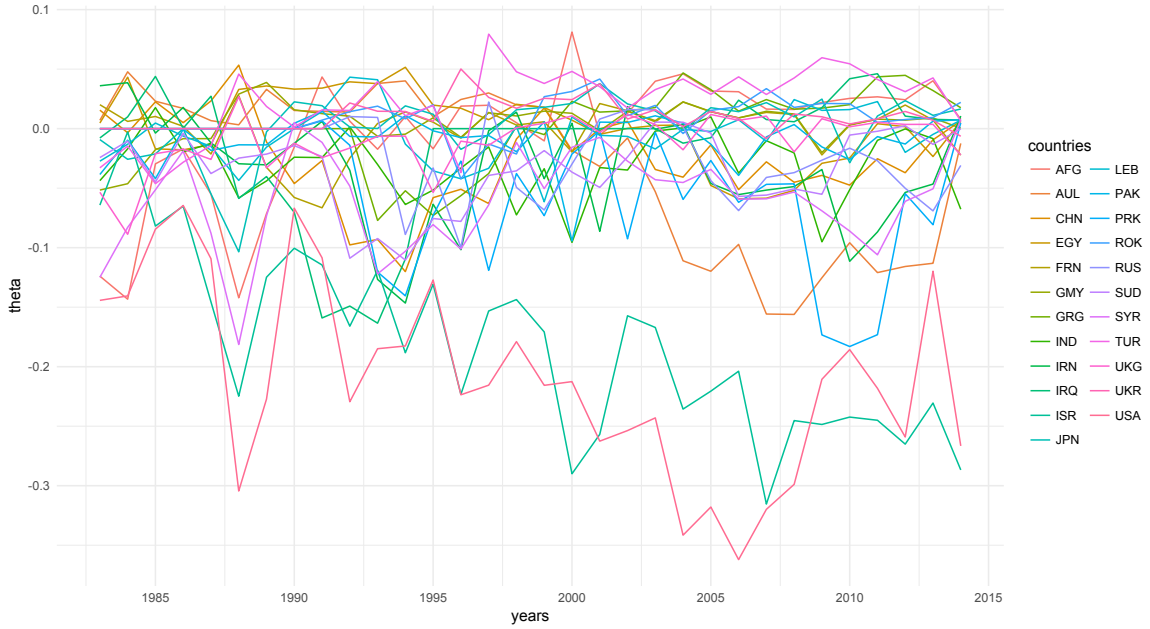

Figure 5: Posterior mean estimates for the additive random effect estimates θ.

The estimated latent factor coordinates in Figure 6, which can be viewed as the weighted latent positions of nodes for every 4 years between 1983 and 2014, demonstrate the remaining higher-order network dependencies in United Nations voting behaviors. To determine the posterior estimates of u and D without identifiability issue, we calculated posterior mean of the multiplicative random effect matrix $u'Du$ and applied eigen-decomposition to the posterior mean matrix. D is then the diagonal matrix with eigenvalues, and u becomes the eigenvectors. For this specific plot, we multiplied d_r and u_r for each dimension such that the plots reflect the magnitude of eigenvalues. As mentioned in Section 4.2, it turned out that none of the eigenvalues were negative, implying that there is no negative transitivity effect in United Nations voting behaviors. Interpretations — What shoud be said about this multiplicative effect plots? As illustrated, the model estimated additive and multiplicative latent effects and the movement of latent positions reveal that United Nations voting network reflects interesting and meaningful foreign policy positions and alliance of various countries, even after controlling for other covariates that are considered critical in the international relation studies.

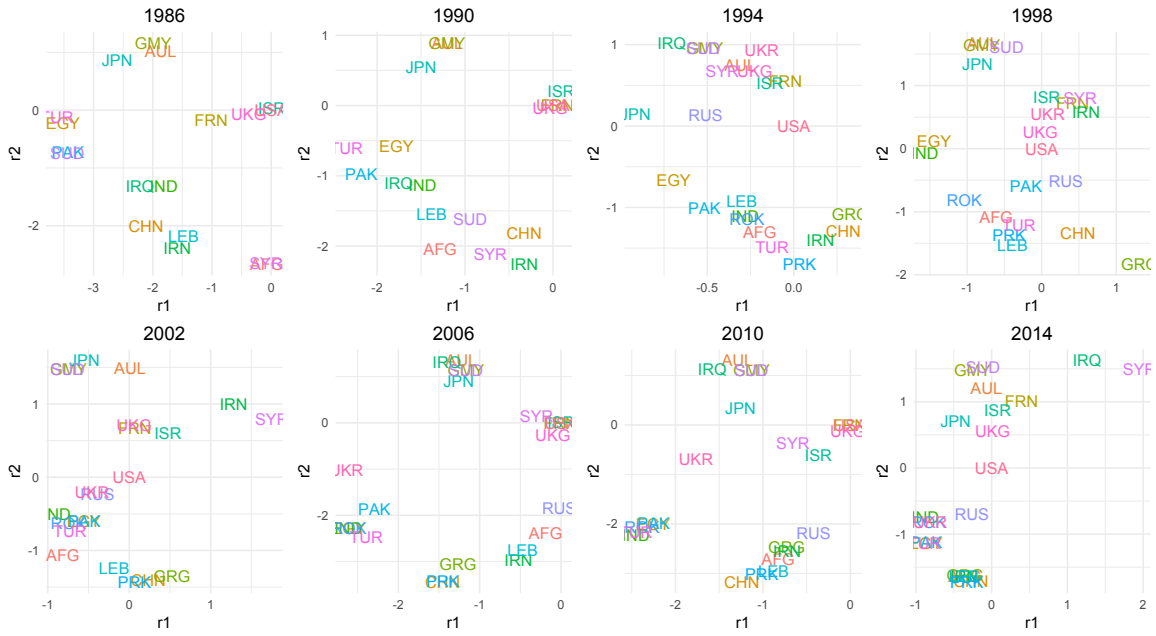

Figure 6: Latent coordinates of the 23 selected countries for every 4 years between 1983 and 2014, obtained from eigen-decomposition of posterior mean of the latent factor matrix $u'Du$. The X-coordinates correspond to $d_1^t \times u_{i1}^t$ and the Y-coordinates correspond to $d_2^t \times u_{i2}^t$.

5 Discussion

Using the correlation associated with network at each time point makes better use of dynamic networks than modeling them as a separate network snapshots. Capturing the temporal dependence from the data is highly informative and should lead to more accurate inference. Our algorithm also eliminates the need to make an arbitrary user-defined covariance structure for determining the prior on the parameters. As a time-varying coefficient extension of the additive and multiplicative effects (AME) model, the dynamic additive and multiplicative nework effects (DAME) model can flexibly learn time-varying network structures, while inferring the effects of latent variables. Further, the visualization of the model estimated time-varying parameters provides an effective temporal trend analysis of dynamic networks, as well as the descriptive visualization of higher-order dependencies from the latent positions over time.

We have demonstrated such effectiveness of our model by modeling the United Nations Voting networks. Although we illustrated the entire framework in the context of symmetric or undirected networks, our model can be easily extended to allow directed networks, following the specification of additive and multiplicative effects model for the directed network (Minhas et al., 2016a). Furthermore, the approach can be applied to binary and ordinal network data with appropriate link functions, while we currently only provide the application to continuous-valued networks. Finally, considering the recent explosion of network dataset with extremely large number of nodes or time-points, our model has a broad range of applicability, suggesting promising computational methods that can accommodate huge networks.

References

Bailey, M. A., Strezhnev, A., and Voeten, E. (2017). Estimating dynamic state preferences from united nations voting data. *Journal of Conflict Resolution*, 61(2):430–456.

Bearce, D. H. and Bondanella, S. (2007). Intergovernmental organizations, socialization, and member-state interest convergence. *International Organization*, 61(4):703–733.

Bhattacharya, A. and Dunson, D. B. (2011). Sparse bayesian infinite factor models. *Biometrika*, 98(2):291–306.

Butts, C. T. (2008). A relational event framework for social action. *Sociological Methodology*, 38(1):155–200.

Durante, D. and Dunson, D. (2014a). Bayesian logistic gaussian process models for dynamic networks. In *Artificial Intelligence and Statistics*, pages 194–201.

Durante, D. and Dunson, D. B. (2013). Nonparametric bayes dynamic modeling of relational data. *arXiv preprint arXiv:1311.4669*.

Durante, D. and Dunson, D. B. (2014b). Bayesian dynamic financial networks with time-varying predictors. *Statistics & Probability Letters*, 93:19–26.

Friel, N., Rastelli, R., Wyse, J., and Raftery, A. E. (2016). Interlocking directorates in irish companies using a latent space model for bipartite networks. *Proceedings of the National Academy of Sciences*, 113(24):6629–6634.

Gartzke, E. (1998). Kant we all just get along? opportunity, willingness, and the origins of the democratic peace. *American Journal of Political Science*, pages 1–27.

Gartzke, E. (2000). Preferences and the democratic peace. *International Studies Quarterly*, 44(2):191–212.

Gibler, D. M. (2008). *International military alliances, 1648-2008*. CQ Press.

Golub, G. H. and Van Loan, C. F. (2012). *Matrix computations*, volume 3. JHU Press.

Hanneke, S., Fu, W., Xing, E. P., et al. (2010). Discrete temporal models of social networks. *Electronic Journal of Statistics*, 4:585–605.

Hoff, P. (2008). Modeling homophily and stochastic equivalence in symmetric relational data. In *Advances in Neural Information Processing Systems*, pages 657–664.

Hoff, P., Fosdick, B., Volfovsky, A., and Stovel, K. (2014). amen: Additive and multiplicative effects modeling of networks and relational data. *R package version 0.999. URL: http://CRAN. R-project. org/package= amen*.

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):261–272.

Hoff, P. D. (2011). Hierarchical multilinear models for multiway data. *Computational Statistics & Data Analysis*, 55(1):530–543.

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

Hoff, P. D. et al. (2011). Separable covariance arrays via the tucker product, with applications to multivariate relational data. *Bayesian Analysis*, 6(2):179–196.

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

Hunter, D., Smyth, P., Vu, D. Q., and Asuncion, A. U. (2011). Dynamic egocentric models for citation networks. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pages 857–864.

Marshall, M. G., Jaggers, K., and Gurr, T. R. (2014). Polity iv annual time-series, 1800–2013. *Center for International Development and Conflict Management at the University of Maryland College Park*.

Mattes, M., Leeds, B. A., and Carroll, R. (2015). Leadership turnover and foreign policy change: Societal interests, domestic institutions, and voting in the united nations. *International Studies Quarterly*, 59(2):280–290.

Minhas, S., Hoff, P. D., and Ward, M. D. (2016a). Inferential approaches for network analyses: Amen for latent factor models. *arXiv preprint arXiv:1611.00460*.

Minhas, S., Hoff, P. D., and Ward, M. D. (2016b). A new approach to analyzing coevolving longitudinal networks in international relations. *Journal of Peace Research*, 53(3):491–505.

Perry, P. O. and Wolfe, P. J. (2013). Point process modelling for directed interaction networks. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(5):821–849.

Sarkar, P. and Moore, A. W. (2005). Dynamic social network analysis using latent space models. *ACM SIGKDD Explorations Newsletter*, 7(2):31–40.

Sarkar, P., Siddiqi, S. M., and Gordon, G. J. (2007). A latent space approach to dynamic embedding of co-occurrence data. In *AISTATS*, pages 420–427.

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

Sewell, D. K. and Chen, Y. (2016). Latent space models for dynamic networks with weighted edges. *Social Networks*, 44:105–116.

Signorino, C. S. and Ritter, J. M. (1999). Tau-b or not tau-b: measuring the similarity of foreign policy positions. *International Studies Quarterly*, 43(1):115–144.

Snijders, T. A., Van de Bunt, G. G., and Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. *Social networks*, 32(1):44–60.

Voeten, E. (2000). Clashes in the assembly. *International organization*, 54(2):185–215.

Voeten, E. (2004). Resisting the lonely superpower: Responses of states in the united nations to us dominance. *Journal of Politics*, 66(3):729–754.

Voeten, E., Strezhnev, A., and Bailey, M. (2016). United nations general assembly voting data.

Vu, D. Q., Hunter, D., Smyth, P., and Asuncion, A. U. (2011). Continuous-time regression models for longitudinal networks. In *Advances in neural information processing systems*, pages 2492–2500.

Ward, M. D., Ahlquist, J. S., Rozenas, A., et al. (2013). Gravity's rainbow: A dynamic latent space model for the world trade network. *Network Science*, 1(01):95–118.

Ward, M. D. and Hoff, P. D. (2007). Persistent patterns of international commerce. *Journal of Peace Research*, 44(2):157–175.

Xu, K. S. and Hero III, A. O. (2013). Dynamic stochastic blockmodels: Statistical models for time-evolving networks. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, pages 201–210. Springer.

Appendix

Appendix A: List of Countries in Voting Network

Abbreviation	Country or area name	Abbreviation	Country or area name
AFG*	Afghanistan	LEB*	Lebanon
ARG	Argentina	MAL	Malaysia
AUL*	Australia	MEX	Mexico
AUS	Austria	MOR	Morocco
BAH	Bahrain	NEW	New Zealand
BNG	Bangladesh	NIG	Nigeria
BOL	Bolivia	NOR	Norway
BRA	Brazil	NTH	Netherlands
BUL	Bulgaria	OMA	Oman
CAN	Canada	PAK*	Pakistan
CAO	Cameroon	PAN	Panama
CHL	Chile	PAR	Paraguay
CHN*	China	PER	Peru
COL	Colombia	PHI	Philippines
COS	Costa Rica	POR	Portugal
CYP	Cyprus	PRK*	North Korea
DEN	Denmark	QAT	Qatar
ECU	Ecuador	ROK*	South Korea
EGY*	Egypt	RUS*	Russia
FIN	Finland	SAU	Saudi Arabia
FRN*	France	SEN	Senegal
GMY*	Germany	SIN	Singapore
GRC	Greece	SPN	Spain
GRG*	Georgia	SRI	Sri Lanka
GUA	Guatemala	SUD*	Sudan
HON	Honduras	SWD	Sweden
IND*	India	SYR*	Syrian Arab Republic
INS	Indonesia	THI	Thailand
IRE	Ireland	TRI	Trinidad and Tobago
IRN*	Iran (Islamic Republic of)	TUN	Tunisia
IRQ*	Iraq	TUR*	Turkey
ISR*	Israel	UAE	United Arab Emirates
ITA	Italy	UKG*	United Kingdom
JOR	Jordan	UKR*	Ukraine
JPN*	Japan	URU	Uruguay
KEN	Kenya	USA*	United States of America
KUW	Kuwait	ZAM	Zambia

Table 2: List of 74 countries included in the analysis

Appendix B: Summary of UN Voting Network

Year	1983	1984	1985	1986	1987	1988	1989	1990	1991
Joint votes	7.447	7.501	8.371	9.424	8.532	5.273	13.213	7.557	9.039
Agreement	0.694	0.716	0.733	0.761	0.728	0.773	0.767	0.817	0.818
Year	1992	1993	1994	1995	1996	1997	1998	1999	2000
Joint votes	14.424	11.317	13.922	26.541	9.920	10.460	9.105	11.808	9.276
Agreement	0.804	0.766	0.785	0.814	0.763	0.803	0.770	0.833	0.747
Year	2001	2002	2003	2004	2005	2006	2007	2008	2009
Joint votes	9.568	13.511	12.538	8.812	9.546	11.467	11.362	11.652	11.602
Agreement	0.706	0.805	0.731	0.819	0.754	0.709	0.715	0.670	0.721
Year	2010	2011	2012	2013	2014				
Joint votes	12.678	9.679	7.892	10.758	12.626				
Agreement	0.722	0.735	0.714	0.739	0.810				

Table 3: Summary of United Nations Voting Data: Average number of common votes (upper) and averge proportion of agreement or voting similarity index (lower) per year