Bayesian Nonparametric Modeling of Dynamic International Relations

Modelli Bayesiani Nonparametrici per Relazioni Internazionali Dinamiche

Daniele Durante and David B. Dunson

Abstract Real world networks are often associated with a dynamic component and the development of flexible statistical methodologies to learn how the connectivity patterns are wired across time is a fundamental goal in many fields of application. News media reports currently provide a rich variety of complex dynamic networks along with novel applied questions. Our focus is on modeling dynamic networks of international relationships among countries during recent conflicts and crises. We accomplish this goal by leveraging a recent Bayesian nonparametric model which merges latent space representations of network data with Gaussian processes.
Abstract *Gli attuali dati di rete sono spesso associati ad una componente dinamica e lo sviluppo di metodologie statistiche per studiare come le connessioni evolvono nel tempo è un obiettivo importante in molte discipline. I mezzi d'informazione offrono una grande varietà di reti complesse così come nuovi interrogativi applicativi. Questo lavoro ha lo scopo di modellare reti dinamiche di relazioni tra stati durante i recenti conflitti. Per fare questo, si utilizzeranno recenti metodologie statistiche che fondono rappresentazioni a spazi latenti per dati di rete con processi Gaussiani.*

Key words: Dynamic network, Gaussian process, GDELT data, Latent space.

1 Introduction

The last two decades have abounded with key financial and conflict events strongly affecting the world geo-political system. Notable examples include the 1997–2000 dot-com bubble [34], the 2004-2007 United States housing bubble [3] and the sub-

Daniele Durante
Department of Statistical Sciences, University of Padua, Padua, Italy. e-mail: durante@stat.unipd.it

David B. Dunson
Department of Statistical Science, Duke University, Durham, NC, USA. e-mail: dunson@duke.edu

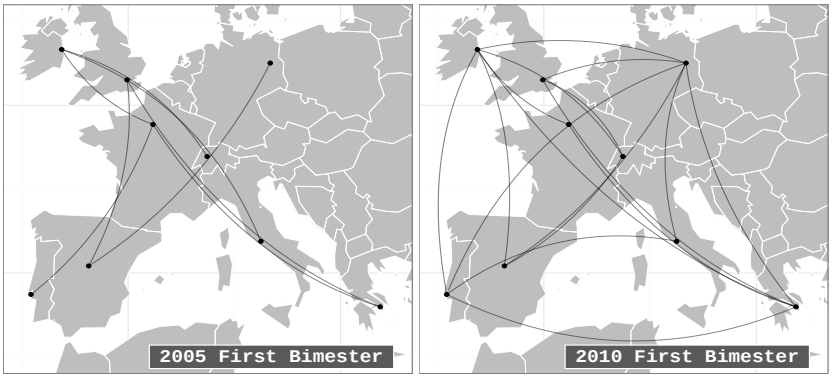

Fig. 1 For some European countries in our data, dynamic relationships for selected times.

sequent 2007–2009 global financial crisis [6] along with the 2008–2012 global recession and the 2010–2012 European sovereign-debt crisis [2]. Beside perturbing financial events, recent years have been additionally characterized by several conflicts including the 2008–2009 Russian-Ukrainian gas crisis [36] and the 2003–2011 Iraq war [12] – among others. These events potentially have a major impact on the dynamic evolution of international relationships among pairs of countries and predicting future patterns is a key to anticipate possible crises. Motivated by the importance of this endeavor and by the current availability of sophisticated algorithms monitoring world news media, we aim to exploit media reports to learn dynamic geo-political patterns in the last two decades.

There is growing interest in mining massive daily news media and web data to learn and predict social and political patterns; refer to [21], [41] and the references cited therein for examples. Although media reports do not necessarily provide an unbiased view on world events, they provide useful data regarding the overall tone of public opinions [37], including on relationships between countries. We focus on dynamic relationship networks among international countries based on the Global Database of Events, Language and Tone (GDELT) project.

GDELT is an open access database containing a comprehensive and high resolution catalog of geo-referenced sociopolitical events from 1979 to the present. Combining Conflict and Mediation Event Observations (CAMEO) taxonomy for political events and actors [29] and Textual Analysis by Augmented Replacement Instructions (TABARI) open software for the machine coding of text data [30], GDELT provides a platform that daily monitors the world's news media reports and translates them into relational data. Specifically each row in the data set corresponds to a specific event record for which a variety of spatio-temporal and contextual information are available including – among others – the two agents interacting, their country of affiliation, the type of relationship recorded and the calendar date in which it was first reported; see [22] for a more detailed overview. This project is attracting increasing interest in the machine learning community [13, 28] and has been successfully utilized in several applied settings, covering domestic protests

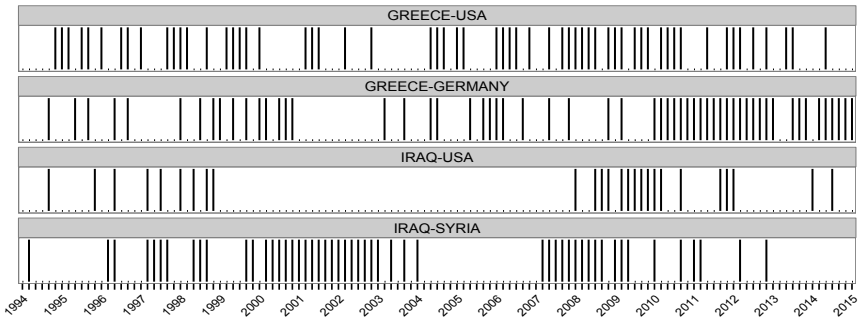

Fig. 2 For select pairs of countries, barcode plot of their edges in time. Bar at t_i means $A_{t_i[vu]} = 1$.

[18], international conflicts [5], political instabilities [10] and global disasters [20]. We are specifically interested in modeling of dynamic variations in the relationships among countries across the last twenty years as reported by the news media.

Our data consist of a sequence of $V \times V$ dynamic symmetric adjacency matrices $A_{t_1}, \ldots, A_{t_n}$ having entries $A_{t_i[vu]} = A_{t_i[uv]} = 1$ if there is a positive overall cooperative relationship among countries $v = 2, \ldots, V$ and $u = 1, \ldots, v-1$ at time t_i and $A_{t_i[vu]} = A_{t_i[uv]} = 0$, otherwise. For interpretability, we consider bimonthly relationships among the $V = 25$ countries heavily involved in financial crises and international conflicts in the last twenty years. Refer to Figure 1 for an illustration. Dynamic networks $A_{t_1}, \ldots, A_{t_{127}}$ are constructed exploiting variable `QuadClass` in the GDELT data set to first obtain matrices $A^{\text{hel}}_{t_1}, \ldots, A^{\text{hel}}_{t_{127}}$ and $A^{\text{conf}}_{t_1}, \ldots, A^{\text{conf}}_{t_{127}}$. These dynamic matrices have entries $A^{\text{hel}}_{t_i[vu]} = A^{\text{hel}}_{t_i[uv]}$ and $A^{\text{conf}}_{t_i[vu]} = A^{\text{conf}}_{t_i[uv]}$ encoding the total number of unique events among pairs of agents affiliated with countries $v = 2, \ldots, V$ and $u = 1, \ldots, v-1$, respectively, at time interval t_i. Matrix $A^{\text{hel}}_{t_i}$ counts material help events, while $A^{\text{conf}}_{t_i}$ counts material conflict events. The difference $\Delta_{t_i} = A^{\text{hel}}_{t_i} - A^{\text{conf}}_{t_i}$ provides an aggregated measure of the strength of positive association between each pair of countries, with $A_{t_i[vu]} = A_{t_i[uv]} = 1(\Delta_{t_i[vu]} = \Delta_{t_i[uv]} \geq 0)$ indicating an overall positive cooperative relationship between countries v and u at time t_i. In the previous notation $1(\cdot)$ represents the indicator function. Examples of positive events include sharing intelligence and economic aid, while examples of negative events include imposing embargo and stopping military assistance. We avoid joint modeling the dynamic count matrices $(A^{\text{hel}}_{t_i}, A^{\text{conf}}_{t_i})$ directly, to improve robustness, limiting sensitivity to missed and duplicate events.

As shown in Figure 2, the edge trajectories cycle with varying patterns of duration and inter-dependence across time. Capturing such behavior is important in assessing how the dynamic relationships relate to key conflict and financial events. Occurrence times of such events can be included as time-varying predictors of the dynamic network. However, for simplicity and robustness, we focus on a dynamic network model, which is sufficiently general to account for dynamic changes in the network structure without requiring known events to be driving this variation.

1.1 Relevant Literature

There is a growing literature on dynamic networks, generalizing exponential random graph models (ERGMs) [16, 9], stochastic block models [23], mixed membership stochastic block models [1] and latent space models [15, 14] from the static setting to the time-varying one. Dynamic ERGMs are available via discrete time Markov models [26, 11] and continuous time specifications [33, 32]. Although dynamic ERGMs can leverage the several techniques available for the static case and explicitly account for suitable dependence structures among edges, these procedures rely on restrictive homogeneity assumptions forcing the model parameters to be shared among nodes and typically constant across time. Accommodating dynamic heterogenous behaviors is a key in our motivating application.

The importance of this endeavor has motivated an intense research aimed at finding alternative specifications to ERGM type models. Time-varying stochastic block models [40, 39] allow dynamic clustering of nodes according to their connection patterns. This framework can allow flexible changes over time, but imposes a community structure. More flexible models include dynamic mixed membership block models [38] and dynamic latent space models [27, 31]. Information is propagated across time via state space models, Markov processes and random walk trajectories, with several layers of approximation required for tractable inference.

Previous proposals lack computational tractability and few results on the flexibility of the models are available. In dynamic modeling of international relationships data we require procedures that can characterize heterogeneous dynamic patterns in the interconnection among countries, while efficiently borrowing information within the network and across time to reduce dimensionality. In accomplishing this goal [7] define the edge probabilities as a function of pairwise distances among nodes in a latent space and enforce flexible dynamic changes in these probabilities by allowing the nodes coordinates to evolve in continuous time via Gaussian processes (GP).

We evaluate their procedure in our motivating application. In particular, Section 2 summarizes the procedure developed by [7]. Key steps for posterior computation are outlined in Section 3, while a careful analysis of the results in our motivating application is presented in Section 4.

2 Time-varying Latent Space Formulation

Let A_{t_i} denote the adjacency matrix characterizing the undirected network with no self-relations at time $t_i \in \Re^+$. As self-relations are not of interest and A_{t_i} is symmetric, we model $A_{t_1},\ldots,A_{t_n}$ by defining a stochastic process for $\mathscr{L}(A_{t_1}),\ldots,\mathscr{L}(A_{t_n})$, with $\mathscr{L}(A_{t_i}) = (A_{t_i[21]}, A_{t_i[31]},\ldots,A_{t_i[V1]}, A_{t_i[32]},\ldots,A_{t_i[V2]},\ldots,A_{t_i[V(V-1)]})^{\mathrm{T}}$ the vector encoding the lower triangular elements of A_{t_i}, which uniquely characterize the network as $A_{t_i[vu]} = A_{t_i[uv]}$ for every $v = 2,\ldots,V$, $u = 1,\ldots,v-1$ and $t_i = t_1,\ldots,t_n$. As a result, $\mathscr{L}(A_{t_i})$ is a vector of binary elements $\mathscr{L}(A_{t_i})_l \in \{0,1\}$, $l = 1,\ldots,V(V-1)/2$, encoding the presence or absence of an edge for each pair of nodes l at time t_i.

Based on previous notation, developing a probabilistic representation for a sequence of time-varying undirected networks, translates into statistical modeling of a multivariate time series $\mathscr{L}(A_{t_1}),\ldots,\mathscr{L}(A_{t_n})$ arising from dynamic monitoring of $V(V-1)/2$ binary variables for n times. However, in accomplishing this goal it is important to explicitly account for the special structure of our data. Specifically, the key difference between a general unstructured multivariate time series and our dynamic vectors of edges is that the observed networks are potentially characterized by specific underlying patterns – such as community structures and transitive relations – which induce dependence among edges at each time t_i. As a result, by carefully accommodating the network structure in dynamic modeling of $\mathscr{L}(A_{t_1}),\ldots,\mathscr{L}(A_{t_n})$, one might borrow information within each $\mathscr{L}(A_{t_i})$ and across time, while reducing dimensionality and infer network properties along with their dynamic changes.

Consistently with the previous discussion, [7] assume data $\mathscr{L}(A_{t_1}),\ldots,\mathscr{L}(A_{t_n})$ as n snapshots of a continuous latent process $\{\mathscr{L}(\mathscr{A}_t) : t \in \mathbb{T} \subset \Re^+\}$ over a possibly unequally spaced time grid $t_1,\ldots,t_n$. In particular they let

$$\mathscr{L}(\mathscr{A}_t)_l \mid \pi_l(t) \sim \text{Bern}\{\pi_l(t)\}, \tag{1}$$

independently for each pair of nodes $l = 1,\ldots,V(V-1)/2$ and $t \in \mathbb{T}$, and focus on developing a careful representation for the edge probabilities $\pi_l(t) \in (0,1)$ – comprising the vector $\pi(t) = \{\pi_1(t),\ldots,\pi_{V(V-1)/2}(t)\}^{\mathrm{T}}$ – in order to reduce the dimensionality and account for the network structure. This is accomplished by expressing the log-odds $S_l(t) = \text{logit}\{\pi_l(t)\}$, $l = 1,\ldots,V(V-1)/2$ as a quadratic combination of a set of node-specific coordinates in a latent space. Specifically, focusing on the lth pair, corresponding to nodes v and u, $v > u$, they let

$$S_l(t) = \mu(t) + X_v(t)^{\mathrm{T}} X_u(t) = \mu(t) + \sum_{r=1}^{R} X_{vr}(t) X_{ur}(t), \tag{2}$$

for every $l = 1,\ldots,V(V-1)/2$ and $t \in \mathbb{T}$, where $X_v(t) = \{X_{v1}(t),\ldots,X_{vR}(t)\}^{\mathrm{T}} \in \Re^R$ and $X_u(t) = \{X_{u1}(t),\ldots,X_{uR}(t)\}^{\mathrm{T}} \in \Re^R$ are the vectors of latent coordinates for nodes v and u at time t, respectively, while $\mu(t) \in \Re$ is a baseline trajectory centering the latent similarity process. According to (1)–(2), nodes with latent coordinates in the same direction will be more similar and hence will have an higher edge probability. Recalling our motivating application on international relationships data, this construction has an appealing interpretation. In particular, each country is assigned a multifaceted latent position, which can be seen as representing its view on different debated topics or international policies. Countries having similar positions in the different attributes, both positive or negative, will be more likely to cooperate than countries with opposite positions. The similarity – or dissimilarity – will be higher the stronger the positions in the same direction – or opposite direction.

Moreover factorization (2) reduces dimensionality from $V(V-1)/2$ stochastic processes on the edge probabilities to $V \times R$ latent trajectories – typically $R \ll V$ – and one baseline process, with the shared dependence on a common set of node-specific latent coordinates inducing rich dependence structures and accommodating

recurring network properties; see for example [15, 14, 19] and [17]. Note also that the latent coordinates in (2) are not identified. However, since we focus inference on the edge probability trajectories $\pi_l(t)$, $l = 1,\ldots,V(V-1)/2$, we avoid identifiability constraints, as they are not necessary to ensure identifiability of $\pi(t)$.

To conclude the Bayesian specification, [7] choose independent priors for $\{\mu(t) : t \in \mathbb{T}\}$ and $\{X_{vr}(t) : t \in \mathbb{T}\}$, $v = 1,\ldots,V$ and $r = 1,\ldots,R$, to induce a prior on $\{\pi(t) : t \in \mathbb{T}\}$ which is able to borrow information across time, allow flexible trajectories in the edge probability processes and shrink towards lower-dimensional representations. This is accomplished by letting

$$X_{vr}(\cdot) \sim \text{GP}(0, \lambda_r c_X), \tag{3}$$

independently for $v = 1,\ldots,V$ and $r = 1,\ldots,R$, with $c_X(t_i,t_j) = \exp\{-\kappa_X(t_i-t_j)^2\}$, $\kappa_X > 0$. The squared exponential covariance function is appealing in our applications to enforce smoothness in analyzing cooperation data. Recalling [25], assumption (3) implies the following joint prior for the node-specific latent coordinates at the time grid $t_1,\ldots,t_n$ on which networks $\mathscr{L}(A_{t_1}),\ldots,\mathscr{L}(A_{t_n})$ are observed

$$\{X_{vr}(t_1),\ldots,X_{vr}(t_n)\}^{\text{T}} \sim \text{N}_n(0, \lambda_r K_X),$$

independently for $v = 1,\ldots,V$ and $r = 1,\ldots,R$, where the covariance matrix K_X has elements $K_{X[ij]} = \exp\{-\kappa_X(t_i-t_j)^2\}$ and $\lambda_r \geq 0$ represents a scaling effect so that when $\lambda_r \approx 0$ the latent coordinates trajectories for dimension r collapse around the zero mean function. Hence to favor adaptive shrinkage, [7] assign a hyperprior for the vector of scaling parameters $\lambda = (\lambda_1,\ldots,\lambda_R)^{\text{T}}$ that adaptively deletes redundant latent space dimensions which are not required to characterize the dynamic edge probability vectors according to the observed data. Adapting [4], they let

$$\lambda_r = \prod_{m=1}^{r} \frac{1}{\vartheta_m}, \quad \vartheta_1 \sim \text{Ga}(a_1, 1),\ \vartheta_{m>1} \sim \text{Ga}(a_2, 1), \quad r = 1,\ldots,R. \tag{4}$$

Prior (4) adaptively penalizes overparameterized representations favoring elements λ_r to be increasingly concentrated towards 0 as r increases for appropriate choice of a_2. See [4] for further discussion and theoretical properties. To conclude the prior specification, we let $\mu(\cdot) \sim \text{GP}(0, c_\mu)$, with $c_\mu(t_i,t_j) = \exp\{-\kappa_\mu(t_i-t_j)^2\}$, $\kappa_\mu > 0$.

3 Posterior Computation

Posterior computation exploits a recently developed Pólya-gamma data augmentation scheme for Bayesian logistic regression [24], which allows implementation of a simple Gibbs sampler. We outline in the following the key steps for implementation.

Algorithm 1 Gibbs sampler for the dynamic latent space model

[1] Sample Pólya-gamma augmented data

for each $l = 1,\dots,V(V-1)/2$ and $t_i = t_1,\dots,t_n$ **do**
Update each augmented data $\omega_l(t_i)$ from the full conditional Pólya-gamma

$$\omega_l(t_i) \mid - \sim \text{PG}\left\{1, \mu(t_i) + \mathscr{L}(X(t_i)X(t_i)^{\text{T}})_l\right\},$$

with $X(t_i)$ the $V \times R$ matrix containing the R coordinates of the V nodes at t_i.
end for

---

[2] Sample the baseline trajectory $\mu = \{\mu(t_1),\dots,\mu(t_n)\}^{\text{T}}$ from

$$\mu \mid - \sim \text{N}_n\left(\Sigma_\mu \begin{bmatrix} \sum_{l=1}^{V(V-1)/2}\{\mathscr{L}(A_{t_1})_l - 1/2 - \omega_l(t_1)\mathscr{L}(X(t_1)X(t_1)^{\text{T}})_l\} \\ \vdots \\ \sum_{l=1}^{V(V-1)/2}\{\mathscr{L}(A_{t_n})_l - 1/2 - \omega_l(t_n)\mathscr{L}(X(t_n)X(t_n)^{\text{T}})_l\} \end{bmatrix}, \Sigma_\mu\right),$$

with $\Sigma_\mu = \left\{\text{diag}\left(\sum_{l=1}^{V(V-1)/2}\omega_l(t_1),\dots,\sum_{l=1}^{V(V-1)/2}\omega_l(t_n)\right) + K_\mu^{-1}\right\}^{-1}$, and K_μ the Gaussian process covariance matrix with entries $K_{\mu[ij]} = \exp\{-\kappa_\mu(t_i - t_j)^2\}$.

---

[3] Sample the matrix of coordinates trajectories $X(t_1),\dots,X(t_n)$
for $v = 1,\dots,V$ **do**
Block-sample $\{X_v(t_1),\dots,X_v(t_n)\}$ given $X_{(-v)} = \{X_u(t_i) : u \neq v, t_i = t_1,\dots t_n\}$.

1. Define $X_{(v)} = \{X_{v1}(t_1),\dots,X_{v1}(t_n),\dots,X_{vR}(t_1),\dots,X_{vR}(t_n)\}^{\text{T}}$
2. Define a Bayesian logistic regression with $X_{(v)}$ acting as coefficient vector and having prior, according to GP, $X_{(v)} \sim \text{N}_{n\times R}\{0, \text{diag}(\lambda_1,\dots,\lambda_R) \otimes K_X\}$
3. The Bayesian logistic regression to update $X_{(v)}$ is a follow

$$\mathscr{L}(A)_{(v)} \sim \text{Bern}(\pi_{(v)}) \quad \text{logit}(\pi_{(v)}) = 1_{V-1} \otimes \mu + \tilde{X}_{(-v)}X_{(v)},$$

with $\mathscr{L}(A)_{(v)}$ obtained by stacking vectors $\{\mathscr{L}(A_{t_1})_l,\dots,\mathscr{L}(A_{t_n})_l\}^{\text{T}}$ for all pairs l having v as a one of the two nodes, and $\pi_{(v)}$ is the corresponding vector of edge probabilities. Finally, $\tilde{X}_{(-v)}$ is the matrix of regressors with entries suitably chosen from $X_{(-v)}$, to reproduce (2) for the sub-sample considered.

4. Hence the Pólya-gamma sampling provides

$$X_{(v)} \mid - \sim \text{N}_{n\times R}\left(\mu_{x_{(v)}}, \Sigma_{x_{(v)}}\right),$$

with $\Sigma_{x^{(v)}} = \left\{\tilde{X}_{(-v)}^{\text{T}}\Omega_{(v)}\tilde{X}_{(-v)} + \text{diag}(\lambda_1^{-1},\dots,\lambda_R^{-1}) \otimes K_X^{-1}\right\}^{-1}$, $\Omega_{(v)}$ a diagonal matrix with the corresponding Pólya-gamma augmented data and mean vector $\mu_{x_{(v)}} = \Sigma_{x_{(v)}}\left[\tilde{X}_{(-v)}^{\text{T}}\left\{\mathscr{L}(A)_{(v)} - 1_{V-1} \otimes 1_n 0.5 - \Omega_{(v)}(1_{V-1} \otimes \mu)\right\}\right]$.

end for

---

[4] Sample the gamma quantities defining the shrinkage parameters $\lambda_1,\ldots,\lambda_R$

$$\vartheta_1 \mid - \sim \text{Ga}\left\{a_1+\frac{V\times n\times R}{2}, 1+\frac{1}{2}\sum_{m=1}^{R}\theta_m^{(-1)}\sum_{v=1}^{V}X_{vm}^{\text{T}}K_X^{-1}X_{vm}\right\},$$

$$\vartheta_r \mid - \sim \text{Ga}\left\{a_2+\frac{V\times n\times (R-r+1)}{2}, 1+\frac{1}{2}\sum_{m=r}^{R}\theta_m^{(-r)}\sum_{v=1}^{V}X_{vm}^{\text{T}}K_X^{-1}X_{vm}\right\},$$

where $\theta_m^{(-r)}=\prod_{t=1,t\neq r}^{m}\vartheta_t$ for $r=1,\ldots,R$ and $X_{vm}=\{X_{vm}(t_1),\ldots,X_{vm}(t_n)\}^{\text{T}}$.

---

[5] Compute the dynamic edge probabilities $\pi_l(t_1),\ldots,\pi_l(t_n)$ as

$$\pi_l(t_i)=\left[1+\exp\left\{-\mu(t_i)-\sum_{r=1}^{R}X_{vr}(t_i)X_{ur}(t_i)\right\}\right]^{-1}$$

for every $t_i=t_1,\ldots,t_n$ and $l=1,\ldots,V(V-1)/2$, with v and u, $v>u$, the nodes characterizing the lth pair.

In performing posterior computation, R is fixed at a conservative upper bound, allowing unnecessary extra dimensions to be effectively removed through posterior distributions for λ_r that are concentrated near zero. The results are not sensitive to R unless R is chosen to be too small, in which case λ_R not concentrated near zero provides evidence that R should be increased.

4 Application to International Relationships Networks

We apply the dynamic network model outlined in Section 2 to the GDELT relationships data $A_{t_1},\ldots,A_{t_{127}}$ described in Section 1. For inference we choose $R=10$ and length scales $\kappa_\mu=\kappa_X=0.05$. The hyperparameters in (4) are set equal to $a_1=2.5$ and $a_2=3.5$ in order to allow the priors for λ_r to increasingly concentrate mass around 0 as r increases, facilitating shrinkage. We consider 5,000 Gibbs iterations, and discard the first 1,000. Mixing has been assessed via effective sample sizes for the quantities of interest, represented by $\pi_l(t_i)$, for $l=1,\ldots,V(V-1)/2$ and $t_i=t_1,\ldots,t_{127}$ after burn-in. Most of these values were around 1,500 out of 4,000, suggesting good mixing.

The trajectory of the posterior mean for the expected network density in Figure 3 provides an appealing overview of the overall dynamic connectivity behavior in relation to key financial and economic international events. Note that the posterior distribution of this quantity – at every time t_i – can be easily derived as a function of the posterior samples for the edge probabilities, as $\text{E}[\sum_{l=1}^{V(V-1)/2}\mathscr{L}(\mathscr{A}_{t_i})_l/\{V(V-1)/2\}]=\sum_{l=1}^{V(V-1)/2}\text{E}\{\mathscr{L}(\mathscr{A}_{t_i})_l\}/\{V(V-1)/2\}=\sum_{l=1}^{V(V-1)/2}\pi_l(t_i)/\{V(V-1)/2\}$. It is first interesting to notice how the posterior mean for the time-varying expected

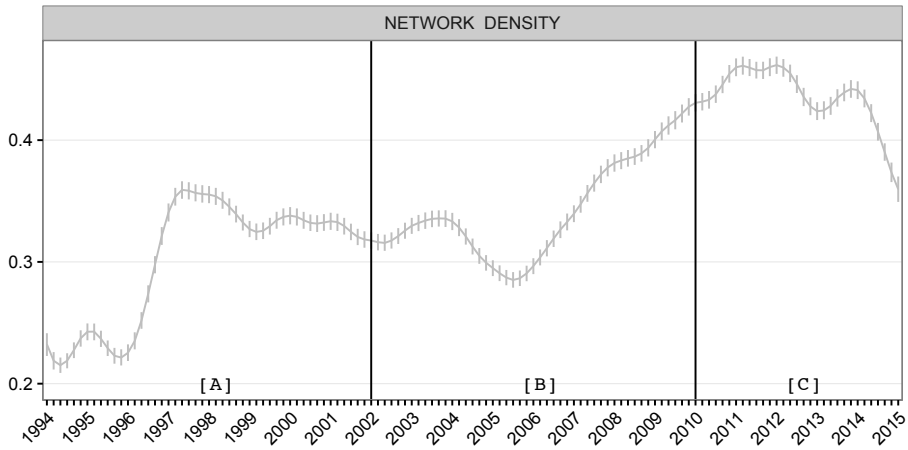

Fig. 3 Trajectory of the posterior mean for the expected network density (gray line) and pointwise posterior interquartile range (gray segments). [A] Mexican economic crisis (≈1995), Asian financial crisis (≈1997–1998), Russian financial crisis (≈1998), Turkish financial crisis (≈2000–2001), Argentinian financial crisis (≈1999–2001), raise and burst of Dot-com bubble (≈1997–2001); [B] Raise and burst of housing bubble (≈2002–2007) and global financial crisis (≈2007–2009); [C] European debt crisis (≈2010–2013), Russian financial crisis (≈2014).

network density evolves on a range between 0.2 and 0.5, meaning that there is not an overall strong tendency towards material cooperations compared to material conflicts in the time window considered. This can be partially explained by the fact that 7 nodes on a total of 25 represent Arab countries, which traversed long conflictual periods with the other nations, but may also reflect an overall general tendency of mass media towards negative news reports, such as material conflicts, rather than positive ones; see for example [35] and the references cited therein.

Although evolving on low values, the overall dynamic connectivity behavior is characterized by a positive trend, which traverses several changes and cyclical periods interestingly related to the key financial events occurring in the time window considered. We observe a rapid change in the expected network density at the burst of the Asian financial crisis in 1997, which then remains on similar high levels in the subsequent years, while displaying bumps in correspondence to the main crises occurred in [A] – i.e. the Mexican peso crisis in 1995, the 1997–1998 Asian financial crisis, the Russian flu in 1998, the 2000–2001 Turkish economic crisis and the Argentina great depression of 1999–2001. Refer to [8] and the references cited therein. Previous crises are generally accompanied by rescue packages and increased material cooperation relationships among international countries to organize bailout investments and avoid facing spread of contagion in case of financial collapse of the countries affected [8]. Our estimated increments for the overall propensity towards material cooperations in correspondence of previous crises confirm this behavior, with the persistent high levels after 1997 potentially related to the growth of the 1997–2001 Dot-com bubble, which facilitated worldwide investments.

The estimated expected network density remains approximately on the same level in later years until further increasing from 2006, with the burst of the United States housing bubble [34] and the subsequent 2007–2009 Global financial crisis [6] [B]. This behavior is interestingly consistent with our previous conclusions for time window [A] and key analyses from [3], [6] and [34]. In particular, similarly to the Dot-com bubble, the behavior prior to 2006 can be related to the growth of the United States housing bubble, which was stimulated by the unusually low interest rates decision of the Federal Reserve to mitigate the effects of the Dot-com bubble [34] and facilitated a wide network of material investments among countries under a "global saving glut" scenario [3]. The increase of the expected network density in later years is instead reasonably associated with need for international material cooperation to provide the many bailouts and bank rescue packages in order to avoid bankruptcy or spread of financial collapses. Refer to [6] for an overview of the interventions required on key financial institutions covering Northern Rock, Bear Stearns, JP Morgan, Lehman Brother and others. A similar scenario applies during the subsequent European debt crisis, which required important bailout investments by the European Stability Mechanism and the International Monetary Fund to face the most acute phases of the crises for Greece, Ireland, Portugal in 2010–2011 and Spain in 2012 [2]. Our dynamic network model captures also these events with high levels in 2010–2012, which are followed by a last increment in correspondence with the recent 2014 Russia ruble crisis.

We provide further insights to specific events by focusing on the estimated dynamic cooperation probabilities between selected pairs of countries, outlined in Figure 4. Results in the top panels confirm previous discussion on the European debt crisis with a specific focus to Greece. Consistently with the leading role of Germany in guaranteeing financial stability of the Eurozone, the estimated cooperation probability between Greece and Germany rapidly increases exactly at the burst of the Greek debt crisis and later stabilizes at very high levels. Conversely, the relationships between United States and Greece are instead characterized by a decreasing trend starting in 2010. This may be a result of efforts to reduce inter-connection with a country in crisis.

The last two panels provide insights on the effect of recent conflicts on the dynamics of the estimated cooperation probabilities. In the three time windows of the middle panels we learn opposite behavior of United States and Russia in their cooperation relationships with Ukraine. In particular, window [A.2] refers to Viktor Yushchenko president (2005–2010) and Yulia Tymoshenko prime minister (2007–2010) period who deepened relations with United States after the Orange Revolution in 2004 and supported NATO membership for Ukraine while progressively increasing conflicts occasions with Russia, which culminated in the 2008–2009 Russian-Ukrainian gas crisis [36]. Consistently with the previous political background, we learn evident lower cooperation relationships between Russia and Ukraine compared to United States and Ukraine, with the latter evolving on very high levels after a bump in 2007 when the prime minister Viktor Yanukovych was succeeded by Yulia Tymoshenko. Differently from Yushchenko and Tymoshenko, Yanukovych improved relations with Russia since he was elected president in 2010, renouncing

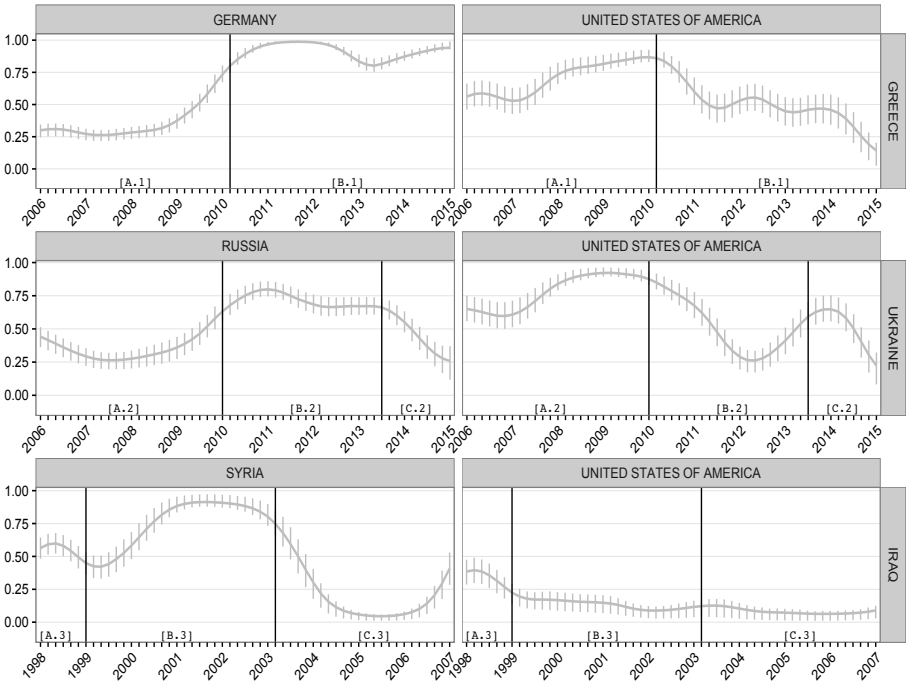

Fig. 4 For selected time windows. Upper panels: posterior mean (gray lines) and pointwise posterior interquartile range (gray segments) for the dynamic cooperation probabilities between Greece–Germany and Greece–USA. Middle panels: same quantities with respect to Ukraine–Russia and Ukraine–USA. Lower panels: same quantities with respect to Iraq–Syria and Iraq–USA.

any aspirations to join NATO and allowing Russia's Black Sea Fleet to stay in the Crimean port of Sevastopol [36]. This change of regime is evident in the two trajectories of the estimated edge probabilities in [B.2] which cross in correspondence of Yanukovych election to reach higher and lower levels for Russia and United States, respectively. As expected, a further variation is evident in [C.2] during the 2013–2015 Ukrainian–Russia crisis and the related ousting of Yanukovych, when the estimated relationships between Ukraine and United States returns to high levels while those with Russia sharply drop.

The lower panels focus instead on the Iraq war. As expected, relationships between United States and Iraq evolve on very low values, with an estimated decrement at the end of 1998 in correspondence of the 1998 Iraq Liberation Act and the subsequent operation Desert Fox in December 1998. Cooperation relationships between Syria and Iraq register instead an increment around the 2000 with the raise of Bashar al-Assad as president of Syria [B.3], and remains on high levels until the 2003–2011 Iraq war [C.3]. These results appear to be consistent with improved economic relations between Iraq and Syria under the Bashar al-Assad regime, mostly related to Iraq oil exports at subsidized prices, which was later shut down by the United States invasion of Iraq in 2003 [12].

References

1. Airoldi, E.M., Blei, D.M., Fienberg, S.E., Xing, E.P.: Mixed Membership Stochastic Blockmodels. Journal of Machine Learning Research **9**, 1981–2014 (2008)
2. Belkin, P., Nelson R.M. Mix, D., Weiss, M.: The eurozone crisis: Overview and issues for congress. Tech. Rep. R42377, Congressional Research Service (2012)
3. Bernanke, B.S.: Global imbalances: recent developments and prospects. Speech 317, Board of Governors of the Federal Reserve System, U.S (2007)
4. Bhattacharya, A., Dunson, D.B.: Sparse bayesian infinite factor models. Biometrika **98**(2), 291–306 (2011)
5. Brandt, P.T., Freeman, J.R., Lin, T., Schrodt, P.A.: Forecasting conflict in the cross-straits: Long term and short term predictions. In: Annual Meeting of the American Political Science Association (2013)
6. Brunnermeier, M.K.: Deciphering the liquidity and credit crunch 2007–2008. Journal of Economic Perspectives **23**(1), 77–100 (2009)
7. Durante, D., Dunson, D.B.: Nonparametric bayes dynamic modelling of relational data. Biometrika **101**(4), 883–898 (2014)
8. Eun, C.S., Resnick, B.G.: International Financial Management. Tata McGraw-Hill Education (2010)
9. Frank, O., Strauss, D.: Markov graphs. Journal of the American Statistical Association **81**(395), 832–842 (1986)
10. Gao, J., Leetaru, K.H., Hu, J., Cioffi-Revilla, C., Schrodt, P.: Massive media event data analysis to assess world-wide political conflict and instability. In: Social Computing, Behavioral-Cultural Modeling and Prediction, pp. 284–292 (2013)
11. Hanneke, S., Fu, W., Xing, E.P., et al.: Discrete temporal models of social networks. Electronic Journal of Statistics **4**, 585–605 (2010)
12. Hinnebush, R.: Syrian foreign policy under bashar al-asad. Ortadoğu Etütleri **1**, 7–26 (2009)
13. Hoff, P.: Multilinear tensor regression for longitudinal relational data. Tech. Rep. 631, Department of Statistics, University of Washington (2014)
14. Hoff, P.D.: Modeling homophily and stochastic equivalence in symmetric relational data. In: J. Platt, D. Koller, Y. Singer, S. Roweis (eds.) Advances in Neural Information Processing Systems 20, pp. 657–664. Curran Associates, Inc. (2008)
15. Hoff, P.D., Raftery, A.E., Handcock, M.S.: Latent space approaches to social network analysis. Journal of the American Statistical Association **97**(460), 1090–1098 (2002)
16. Holland, P.W., Leinhardt, S.: An exponential family of probability distributions for directed graphs. Journal of the American Statistical Association **76**(373), 33–50 (1981)
17. Hunter, D.R., Krivitsky, P.N., Schweinberger, M.: Computational statistical methods for social network models. Journal of Computational and Graphical Statistics **21**(4), 856–882 (2012)
18. Keneshloo, Y., Cadena, J., Korkmaz, G., Ramakrishnan, N.: Detecting and forecasting domestic political crises. In: Proceedings of the 2014 ACM conference on Web science - WebSci'14 (2014)
19. Krivitsky, P.N., Handcock, M.S., Raftery, A.E., Hoff, P.D.: Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models. Social Networks **31**(3), 204–213 (2009)
20. Kwak, H., An, J.: A first look at global news coverage of disasters by using the GDELT dataset. In: Lecture Notes in Computer Science, pp. 300–308 (2014)
21. Leetaru, K.: Culturomics 2.0: Forecasting large-scale human behavior using global news media tone in time and space. First Monday **16**(9) (2011)
22. Leetaru, K., Schrodt, P.A.: Gdelt: Global data on events, location and tone, 1979-2012. In: International Studies Association Annual Conference (2013)
23. Nowicki, K., Snijders, T.A.B.: Estimation and prediction for stochastic blockstructures. Journal of the American Statistical Association **96**(455), 1077–1087 (2001)
24. Polson, N.G., Scott, J.G., Windle, J.: Bayesian inference for logistic models using pólya–gamma latent variables. Journal of the American Statistical Association **108**(504), 1339–1349 (2013)

25. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. The MIT Press (2006)
26. Robins, G., Pattison, P.: Random graph models for temporal processes in social networks. The Journal of Mathematical Sociology **25**(1), 5–41 (2001)
27. Sarkar, P., Moore, A.W.: Dynamic social network analysis using latent space models. SIGKDD Explor. Newsl. **7**(2), 31–40 (2005)
28. Schein, A., Paisley, J., Blei, D., Wallach, H.: Inferring polyadic events with poisson tensor factorization. In: NIPS 2014 Workshop on Networks (2014)
29. Schrodt, P.: Cameo conflict and mediation event observations event and actor codebook. http://data.gdeltproject.org/documentation/CAMEO.Manual.1.1b3.pdf (2012)
30. Schrodt, P.: Tabari textual analysis by augmented replacement instructions. http://eventdata.parusanalytics.com/tabari.dir/TABARI.0.8.4b3.manual.pdf (2014)
31. Sewell, D.K., Chen, Y.: Latent space models for dynamic networks. Journal of the American Statistical Association **to appear** (2015)
32. Snijders, T.A., van de Bunt, G.G., Steglich, C.E.: Introduction to stochastic actor-based models for network dynamics. Social Networks **32**(1), 44–60 (2010)
33. Snijders, T.A.B.: Models for longitudinal network data. In: Models and Methods in Social Network Analysis, pp. 215–247 (2005)
34. Taylor, J.B.: The financial crisis and the policy responses: An empirical analysis of what went wrong. NBER Working Papers 14631, National Bureau of Economic Research, Inc (2009)
35. Thapthiang, N.: An analysis of news reporting and its effects, using ibil model: Lee gardens plaza and cs pattani hotels cases. Procedia-Social and Behavioral Sciences **91**, 411–420 (2013)
36. Tsygankov, A.: Vladimir putin's last stand: the sources of russia's ukraine policy. Post-Soviet Affairs **31**(4), 279–303 (2015)
37. Wanta, W., Golan, G., Lee, C.: Agenda setting and international news: Media influence on public perceptions of foreign nations. Journalism & Mass Communication Quarterly **81**(2), 364–377 (2004)
38. Xing, E.P., Fu, W., Song, L.: A state-space mixed membership blockmodel for dynamic network tomography. Annal of Applied Statistics **4**(2), 535–566 (2010)
39. Xu, K.S.: Stochastic block transition models for dynamic networks. Journal of Machine Learning, Workshops & Proceedings **38**, 1079–1087 (2015)
40. Xu, K.S., Hero, A.O.: Dynamic stochastic blockmodels for time-evolving social networks. IEEE Journal of Selected Topics in Signal Processing **8**(4), 552–562 (2014)
41. Zaman, T., Fox, E.B., Bradlow, E.T.: A bayesian approach for predicting the popularity of tweets. Annal of Applied Statistics **8**(3), 1583–1611 (2014)