A latent space model for cognitive social structures data

Juan Sosa [a],[*], Abel Rodríguez [b]

[a] *Universidad Nacional de Colombia, Carrera 45 # 26-85 - Edificio Uriel Gutiérrez, Bogotá, Colombia*
[b] *University of Washington, Padelford Hall, Room B-313, Box 354322, Seattle, WA, 98195, United States*

A B S T R A C T

This paper introduces a novel approach for modeling a set of directed, binary networks in the context of cognitive social structures (CSSs) data. We adopt a relativist approach in which no assumption is made about the existence of an underlying true network. More specifically, we rely on a generalized linear model that incorporates a bilinear structure to model transitivity effects within networks, and a hierarchical specification on the bilinear effects to borrow information across networks. This is a spatial model, in which the perception of each individual about the strength of the relationships can be explained by the perceived position of the actors (themselves and others) on a latent social space. A key goal of the model is to provide a mechanism to formally assess the agreement between each actors' perception of their own social roles with that of the rest of the group. Our experiments with both real and simulated data show that the capabilities of our model are comparable with or, even superior to, other models for CSS data reported in the literature.

1. Introduction

Cognitive social structures (CSSs), also called triadic data in the social networks literature, naturally arise in diverse disciplines such as sociology, psychology, and organizational economics. Roughly speaking, a CSS is defined by a set of *cognitive* judgments that subjects form about the relationships among actors (themselves as well as others) who are embedded in a common environment. Hence, each subject reports a full description of the social network structure, resulting in a set of I sociomatrices $\mathbf{Y}_1, \ldots, \mathbf{Y}_I$, each with I actors, where $\mathbf{Y}_j = [y_{i,i'j}]$. In this notation, $i = 1, \ldots, I$ identifies the "sender" of the relation, $i' = 1, \ldots, I$ identifies the "receiver" of the relation, and $j = 1, \ldots, I$ identifies the "reporter" of the relation from i to i' (notice that $y_{i,i'j} = y_{i',ij}$ for undirected relations).

CSSs provide rich data that allow researchers to understand the patterns of social interactions as cognitively represented by each actor in the system. For instance, CSSs allow researchers to investigate the ability of social actors to precisely recognize the social network in which they are embedded, as well as the impact of such an ability on their success (Brands, 2013). It is widely assumed that actors who are skillful at understanding the relationships around them are better prepared to, for example, adjust their behavior according to the demands of a particular situation. The focus of this work is on developing statistical methods that allow us to identify such "highly adept" individuals.

The use of cognitive reports in the context of social network research goes back to Newcomb (1961) and Sampson (1968), but it was Krackhardt (1987) who formalized the study of CSSs and outlined key empirical methods based on aggregations. One of the primary aggregations discussed by Krackhardt is obtained by "collapsing" all the information onto a consensus network $\widetilde{Y} = [\widetilde{y}_{i,i'}]$ defined as:

$$\widetilde{y}_{i,i'} = \begin{cases} 1, & \text{if } \dfrac{1}{I}\sum_{j=1}^{I} y_{i,i'j} > \delta_0; \\ 0, & \text{otherwise}, \end{cases} \qquad (1)$$

where δ_0 is a fixed (but arbitrary!) threshold taking values from 0 to 1. This reduction is typically inappropriate because it usually results in a considerable loss of useful information.

Krackhardt's seminal work has been extended by numerous authors. Kumbasar et al. (1994) and Kumbasar (1996) went beyond Krackhardt's aggregations and employ multi-dimensional scaling and correspondence analysis to study the structure of triadic data. Subsequently, once again along the lines of data reduction schemes, Batchelder et al. (1997) proposed a model for aggregating separate reports into a single consensus network, allowing estimates of actor accuracy to be obtained along the way. Unfortunately, this model is quite restrictive and inference procedures are not straightforward. Later, Bond et al. (1997, 2000)

* Corresponding author.
  E-mail addresses: jcsosam@unal.edu.co (J. Sosa), abelrod@uw.edu (A. Rodríguez).
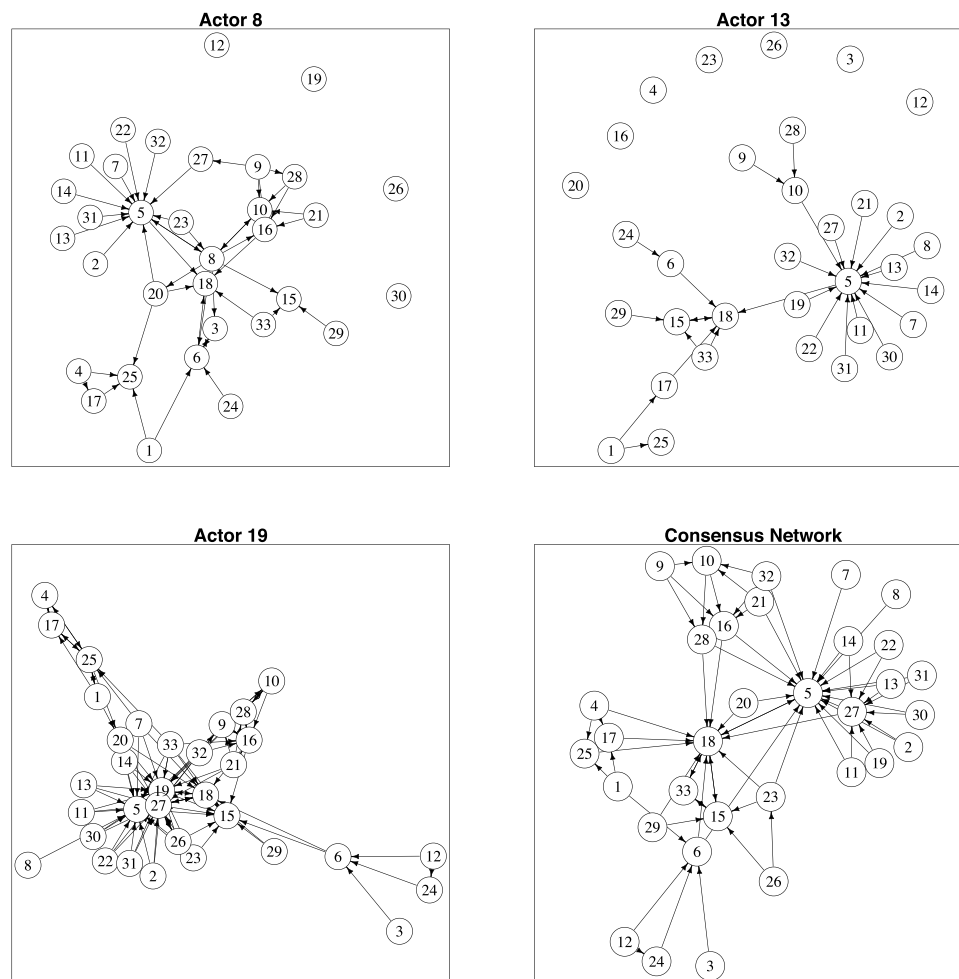
Fig. 1. Visualization of some networks in the CSS corresponding to Actors 8, 13, and 19. The bottom right panel shows the consensus network using $\delta_0 = 0.5$ as a threshold.

extended the social relations model of Kenny (1994) to analyze multivariate triadic relations; the later extension was the first attempt to incorporate covariances between measurements on different types of relations on the same pair of actors. Butts (2000), later published as Butts (2003), proposed a family of hierarchical Bayesian models which allows for simultaneous inference of informant accuracy and the underlying social structure in the presence of measurement error. In parallel, Koskinen (2002b) extended Batchelder's model in a Bayesian context treating the underling "true" network as a latent variable (and therefore part of the parameter space). Then, Koskinen (2002a) implemented a modified version of the previous model aiming to correlate bias on cognitive judgements with exogenous attributes of the reporters. Finally, Koskinen (2004) proposed an inference scheme where reference priors were used in order to allow some degree of automation in the model selection. More recently, Rodriguez (2015) developed a novel class of stochastic block models for CSS data by constructing a joint prior on the community structure of all networks using fragmentation and coagulation processes. Concurrently, Swartz et al. (2015), extending Bond's model, proposed a fully Bayesian logistic ANOVA model for triadic data. Their strategy uses a convenient parametrization that makes assessments of cognitive agreement among actors possible. Pattison (1994) and Brands (2013) provide excellent reviews about the early and modern literature on CSSs, respectively.

In this manuscript we introduce a random-effects model for triadic data that builds upon the latent space approach of Hoff (2005) and Hoff (2009). The model assumes that there is a true social space (a *K*-dimensional vector space in which each individual occupies a fixed position), that each individual has a perception of the true space (reflected as an assignment of coordinates to individuals within the same space that differ to a lesser or greater extent from the true positions), and that, when asked to provide CSS reports, each individual consults their individual spatial model, and then generates tie presence/absence reports independently. The positions in latent space can be interpreted as operationalizing the notion of (perceived) "social roles" (e.g., see Wasserman et al., 1994, Chapter 12). A consequence of this framework is that there is no "true" network: networks are purely a consequence of the behavior of individuals when we force them to express their (spatial) understanding in terms of a network instrument. The bilinear structure of our model is reminiscent of (multivariate) item response theory (IRT) models (e.g., see von Davier and Carstensen, 2007) and spatial voting models (e.g., see Clinton et al., 2004), which are widely used in various social sciences disciplines to map discrete responses into (latent) continuous scales.

Our main goal in building the model is to develop a test that would allow us to formally assess the level of agreement between an actor's self perception of their own position in a social environment and that of other actors embedded in the same social environment. This goal is accomplished though a hierarchical Bayesian model that incorporates a mixture prior. Such a prior enables the computation of posterior probabilities that assess whether the perception of an individual about its own position in social space agrees with the judgments of other actors. As a byproduct of this approach, we are able to estimate a *weighted* consensus network that summarizes the cognitive assessments of all reporters in the network by differentially weighting the information

according to the level of cognitive agreement among actors.

Historically, there have been two leading approaches to the analysis of CSS data, namely, the essentialist ("classical") and the relativist ("cognitivist") perspectives. Essentialists conveniently define a "real" network structure (known as criterion network, and usually constructed based on behaviors described by an external observer), and then study the discrepancies between actors' reports and this "truth" (Bond et al., 1997; Koskinen, 2002a,b, 2004; Butts, 2003; Kilduff et al., 2008). On the other hand, in the absence of an external truth, relativists compare actors' cognitive reconstructions of an underlying network with each other. Under the relativist approach, ties that actors perceive are considered to be more informative than ties reported by a third-party, external observer because people act upon what they consider real in their minds. Thus, differences between the perceptions of one individual and the perceptions of other actors represent the theoretical and empirical question of interest (Krackhardt, 1987). Our approach is built from a relativist perspective, as it relies exclusively in cognitive reports from actors embedded in the social environment, it does not assume the existence of a "true" network, and focuses on assessing "agreement" as opposed to "accuracy" (which requires the definition of an external "gold standard"). And while our approach can generate a weighted consensus network (which can be interpreted in terms of perceived affinities or social distances), we make no claim about whether this consensus provides an accurate representation of a "true" social structure.

The model we introduce in this paper is part of a growing literature on latent space models for tensor and multi-network data. Other examples of this literature include Salter-Townshend and McCormick (2017), Durante et al. (2017) and Wang et al. (2019) for replicated networks, Hoff (2015) for tensor data, and Sarkar and Moore (2006), Westveld et al. (2011), Durante and Dunson (2014) and Sewell and Chen (2015) for longitudinal network data. A key contribution of our approach vis-a-vis this existing literature is the hierarchical mixture structure in our prior that allows us to test perceptual agreement in the specific context of CSS data.

The remainder of the paper is organized as follows: Section 2 motivates our methodological developments using a CSS associated with 33 employees of a firm involved in maintenance of information systems originally presented in Krackhardt (1990). Section 3 introduces our model, discusses some of its properties, and describe our prior elicitation process. Section 4 discusses our approach to computation, including our approach to addressing identifiability issues and the selection of the dimension of the social space. Section 5 revisits the Krackhardt's dataset introduced in Section 2 and compares the results of applying our modelling strategy as well as that of Swartz et al. (2015). Section 6 further explores the performance of the model through a cross-validation exercise carried out on several datasets additional datasets. Finally, some concluding remarks and directions for future work are provided in Section 7.
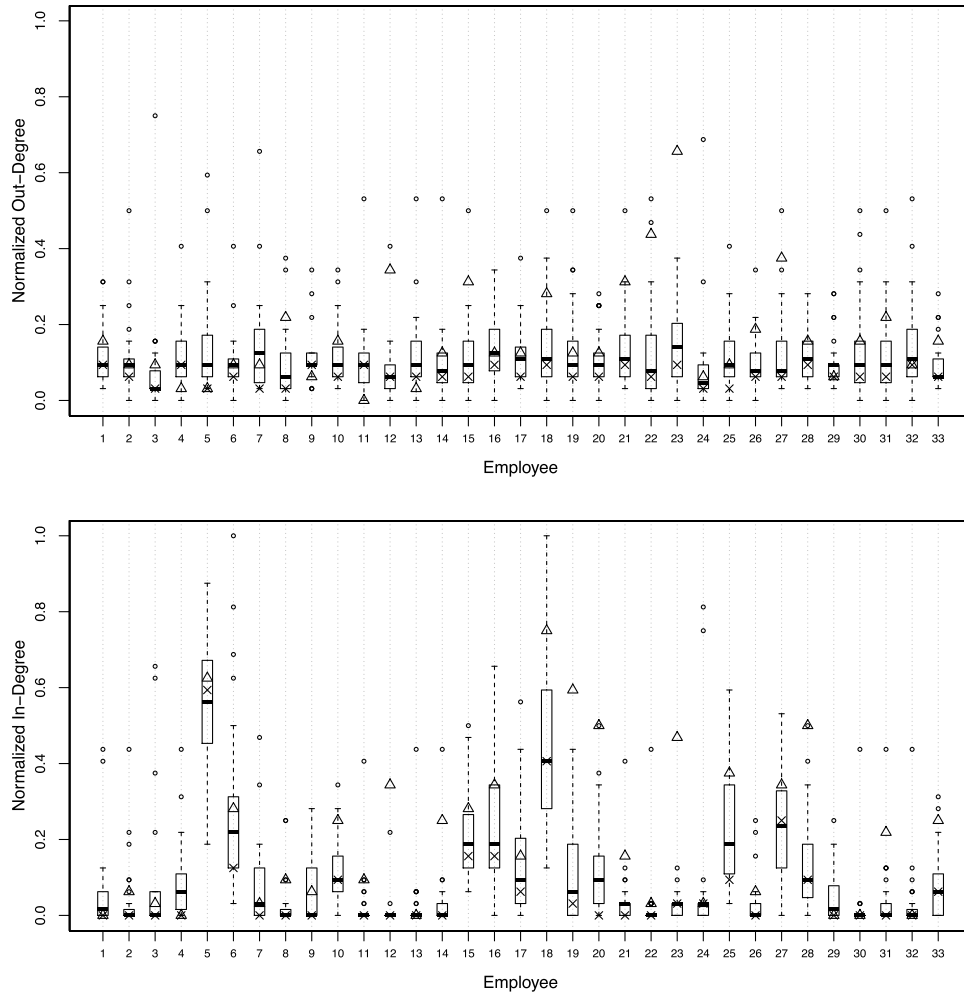

Fig. 2. Out-degree distribution d_{ij}^{out} (top panel) and in-degree distribution d_{ij}^{in} (bottom panel) across networks. The ith boxplot summarizes the distribution of the degree for all reporters except i, while the self-perceived degree is represented by a triangle ($\triangle$) and the degree in the consensus network by a cross ($\times$).

2. Krackhardt's 1990 dataset

To motivate our modeling approach, we consider the "advice" CSS reported by Krackhardt (1990). This CSS considers 33 employees in a small firm involved in maintenance of information systems. Using Krackhardt's (1987) methodology, each actor in the network was asked to fill out a questionnaire answering the following question: "Who would this person go to for help or advice at work?" That is, if a specific person had a question or ran into a problem at work, who would she/he likely go to ask for advice or help? All 33 respondents completed the questionnaire, each answering the same question about himself/herself and all the others employees. Hence, this advice dataset constitutes a complete CSS composed of $I = 33$ directed, binary networks $\mathbf{Y}_1, \ldots, \mathbf{Y}_{33}$, each one of size 33×33, in which $y_{i,i'\,j} = 1$ if Actor j informs that Actor i is likely to go for advice to Actor i', and $y_{i,i'\,j} = 0$ otherwise. The diagonal elements of each $\mathbf{Y}_j$ are treated as structural zeros. No nodal or link covariates are available for this dataset.

Fig. 1 presents the ties reported by three employees (which in our notation correspond to $\mathbf{Y}_8$, $\mathbf{Y}_{13}$ and $\mathbf{Y}_{19}$), along with the consensus network obtained by using (1) with $\delta_0 = 0.5$ (i.e., a link is retained if only if at least 50% of the actors agree on it). There is wide variability among the reporters, and homophily effects are clear. The number of links ranges from 27 (Actor 13) to 430 (Actor 12); while Actors 4, 8, 13, and 29 at most report 46 links each, 18 employees report more than 100 links each. Interestingly, Actor 11, who only perceives 94 links, is the only employee that reports never seeking advice from others. Similarly, Actors 1, 4, 13, 29, 30, and 32 report that no co-worker goes to them for guidance.

Fig. 2 illustrates the differences between the self-perception of each actor and the perception that others have of him/her by comparing the normalized out-degree $d_{i,i}^{\text{out}} = \sum_{i' \neq i} y_{i',i,i}/(I-1)$ (top panel) and in-degree $d_{i,i}^{\text{in}} = \sum_{i' \neq i} y_{i,i',i}/(I-1)$ (bottom panel) of a given actor when she/he is the reporter, against the distribution of the normalized in- and out-degrees when the reporter is any other subject in the network $d_{i,j}^{\text{in}} = \sum_{i' \neq i} y_{i,i'\,j}/(I-1)$ and $d_{i,j}^{\text{out}} = \sum_{i' \neq i} y_{i',i,j}/(I-1)$, $j \neq i$, as well as those obtained from the consensus network, $\widetilde{d}_i^{\text{in}} = \sum_{i' \neq i} \widetilde{y}_{i,i'}/(I-1)$ and $\widetilde{d}_i^{\text{out}} = \sum_{i' \neq i} \widetilde{y}_{i',i}/(I-1)$. Note that for many executives, both the in- and out-degree they perceive for themselves is smaller than the corresponding median computed from the networks perceived by the other actors, as well as that for the consensus network. This effect is particularly extreme for Actors 8, 12, 21, 22, and 23. Similar discrepancies can be observed for various Actors based on other descriptive measures such as clustering coefficient or assortativity index (plots not shown). We emphasize the fact that the perceptual discrepancies highlighted above are not established in relation to a "true" underlying network structure; instead, such discrepancies are cognitive disagreements with respect to others' reports.

3. A latent space model for CSSs data

Hoff (2005, 2009) proposes a bilinear latent space model for binary, directed networks $\mathbf{Y} = [y_{i,i'}]$. The model assumes that observations are conditionally independent Bernoulli draws, $y_{i,i'}|\theta_{i,i'} \sim \text{Ber}(\theta_{i,i'})$, $i, i' = 1, \ldots, I, i' \neq i$, and then proceeds to structure the probabilities as $\theta_{i,i'} = \Phi(\mathbf{x}_{i,i'}^T \boldsymbol{\beta} + \mathbf{u}_i^T \mathbf{v}_{i'})$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard Gaussian distribution. In addition to a global intercept, the p-dimensional vector of predictors $\mathbf{x}_{i,i'}$ incorporates known attributes associated with each pair of actors that might explain transitivity effects in the network. On the other hand, the bilinear term $\mathbf{u}_i^T \mathbf{v}_{i'}$ is used to account for residual transitivity effects not explained by the known attributes. The unknown vectors $\mathbf{u}_1, \ldots, \mathbf{u}_I$ and $\mathbf{v}_1, \ldots, \mathbf{v}_I$ can be interpreted as the positions of each actor in "sender" and "receiver" K-dimensional social spaces, respectively. The latent dimension K is treated as a known

constant, which needs to specified beforehand, or in our case, selected using a two-step approach (see Section 4.2 for details). The model is completed by setting (usually mutually independent) priors on the unknown vector of regression coefficients $\boldsymbol{\beta}$ and the unknown latent features $\mathbf{u}_1, \ldots, \mathbf{u}_I$ and $\mathbf{v}_1, \ldots, \mathbf{v}_I$.

We are interested in extending this model to accommodates the particular features associated with CSS data. As in Hoff (2005, 2009), we still assume that observations are conditionally independent, $y_{i,i'\,j}|\theta_{i,i'\,j} \sim \text{Ber}(\theta_{i,i'\,j})$, and construct a hierarchical prior for the array of probabilities $[\theta_{i,i'\,j}]$. To do so, we let

$$\theta_{i,i'\,j} = \Phi\left(\mathbf{x}_{i,i'}^T \boldsymbol{\beta}_j + \mathbf{u}_{i,j}^T \mathbf{v}_{i'\,j}\right), \tag{2}$$

where the additional index j makes explicit the reference to reporter j. If mutually independent priors were assigned to each reporter j, then this formulation would be equivalent to fitting the one-network model we described above independently for each reporter. Instead, we consider a hierarchical prior for the latent space positions that distinguishes between two cases. For values of $j \neq i$ (i.e., for the latent positions of actor i as perceived by all actors except him/herself) we assign conditionally independent Gaussian priors

$$\mathbf{u}_{i,j}|\boldsymbol{\eta}_i, \sigma_u^2 \overset{\text{ind}}{\sim} N(\boldsymbol{\eta}_i, \sigma_u^2 \mathbf{I}_K), \quad \mathbf{v}_{i,j}|\boldsymbol{\zeta}_i, \sigma_v^2 \overset{\text{ind}}{\sim} N(\boldsymbol{\zeta}_i, \sigma_v^2 \mathbf{I}_K), \tag{3}$$

where $\mathbf{I}_K$ denotes the $K \times K$ identity matrix. The means $\boldsymbol{\eta}_i$ and $\boldsymbol{\zeta}_i$ can be interpreted as the consensus positions in the sender and receiver spaces for each actor in the network. By placing priors on these consensus means we can capture similarities among the observed networks and borrow information across them. On the other hand, for $j = i$ we model the latent positions using two-component mixtures of the form

$$\mathbf{u}_{i,i}|\boldsymbol{\eta}_i, \sigma_u^2, \tau_u^2, \gamma_i \overset{\text{ind}}{\sim} \begin{cases} N(\boldsymbol{\eta}_i, \sigma_u^2 \mathbf{I}_K), & \gamma_i = 1, \\ N(\mathbf{0}, \tau_u^2 \mathbf{I}_K), & \gamma_i = 0, \end{cases} \tag{4}$$

$$\mathbf{v}_{i,i}|\boldsymbol{\zeta}_i, \sigma_v^2, \tau_v^2, \xi_i \overset{\text{ind}}{\sim} \begin{cases} N(\boldsymbol{\zeta}_i, \sigma_v^2 \mathbf{I}_K), & \xi_i = 1, \\ N(\mathbf{0}, \tau_v^2 \mathbf{I}_K), & \xi_i = 0, \end{cases} \tag{5}$$

where $\sigma_u^2 \sim \text{IGam}(a_\sigma, b_\sigma)$ and $\sigma_v^2 \sim \text{IGam}(a_\sigma, b_\sigma)$, while $\tau_u^2 \sim \text{IGam}(a_\tau, b_\tau)$ and $\tau_v^2 \sim \text{IGam}(a_\tau, b_\tau)$.

Note that if $\gamma_i = 1$ ($\xi_i = 1$) then the self-perception of Actor i's position in the sender (receiver) social space is drawn from the same general distribution as the perception of other actors about him/her. On the other hand, if $\gamma_i = 0$ ($\xi_i = 0$) then the perception that Actor i has its own position in sender (receiver) social space, which differs from that of the other actors. We treat $\gamma_1, \ldots, \gamma_I$ and $\xi_1, \ldots, \xi_I$ as unknown quantities and assign them common priors $\gamma_i|\psi \sim Ber(\psi)$ and $\xi_i|\psi \sim Ber(\psi)$, with $\psi \sim Beta(c, d)$. The choice of a common value of ψ for both the γ_is and the ξ_is is a (very mild) mechanism to share information between the sender and receiver spaces of the model, which are otherwise treated as independent a priori. This is most easily seen if ψ is integrated out of the model,

$$p(\gamma_1, \ldots, \gamma_I, \xi_1, \ldots, \xi_I) = \frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)} \frac{\Gamma(c + \gamma_\cdot + \xi_\cdot)\Gamma(d + 2I - \gamma_\cdot - \xi_\cdot)}{\Gamma(c + d + 2I)},$$

where $\gamma_\cdot = \sum_{i=1}^I \gamma_i$, $\xi_\cdot = \sum_{i=1}^I \xi_i$ and $\Gamma(\cdot)$ denotes the well-known Gamma function. The use of a common ψ simply reflects an assumption that the average "sparsity" in both sender and receiver spaces is identical. We believe that this is a natural assumption to make a priori, but if the data does not support it, the model is still free to deviate from it a posteriori.

Inferences on $\gamma_1, \ldots, \gamma_I$ and $\xi_1, \ldots, \xi_I$ allow us to make statements about the level of cognitive agreement between actor's self-perception and that of the rest of the actors. Furthermore, by learning $\gamma_1, \ldots, \gamma_I$ and $\xi_1, \ldots, \xi_I$ jointly from the data, the estimates of the consensus positions $\boldsymbol{\eta}_i$ and $\boldsymbol{\zeta}_i$ differentially weight an actor's self-perception and the
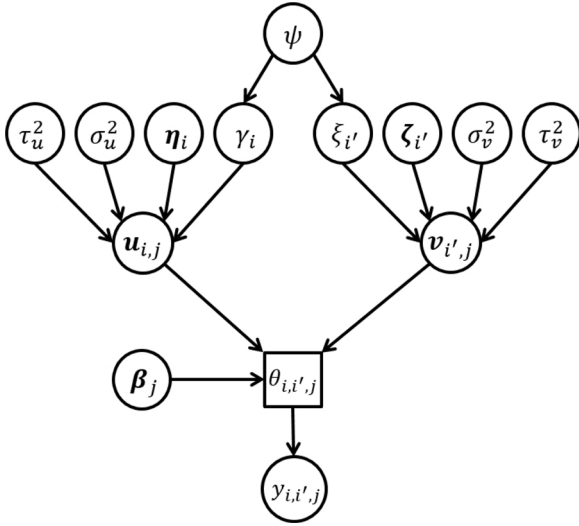
Fig. 3. Plate diagram summarizing the structure of our model.

$$\boldsymbol{\beta}_j | \boldsymbol{\nu}, \varsigma^2 \overset{\text{iid}}{\sim} N(\boldsymbol{\nu}, \varsigma^2 \, \mathbf{I}_p), \boldsymbol{\nu} \sim N(\mathbf{0}, \omega^2 \, \mathbf{I}_p), \varsigma^2 \sim \text{IGam}(a_\varsigma, b_\varsigma)$$

Fig. 3 presents a plate diagram that summarizes the structure of the model.

The construction we propose in this section is such that, in the case where no covariates are present, the resulting joint marginal distribution of the data is fully jointly exchangeable (i.e., that the distribution of the collection $\left\{y_{i,i'\,j}\right\}_{i,i'\,j=1}^{I}$ is the same as the distribution of $\left\{y_{\pi_1(i),\pi_2(i'),\pi_3(j)}\right\}_{i,i'\,j=1}^{I}$ only if the permutations π_1, π_2 and π_3 satisfy $\pi_1 = \pi_2 = \pi_3$, see Aldous, 1985). Full joint exchangeability (rather than a weaker form of exchangeability) is particularly attractive in this setting because all indexes i, i' and j refer to the same set of actors. This in particular implies that the marginal distribution of $y_{i,i'\,j}$ is not the same as that of $y_{i,i'\,,i}$, which should be clear from (3)–(5).

3.1. Hyperparameter elicitation

Careful elicitation of the hyperparameters c, d, ω^2, a_ς, b_ς, a_σ, b_σ, a_τ, b_τ, b_κ, and κ is key to ensure appropriate model performance. As is customary in the model selection literature (e.g., see Scott and Berger, 2010), we set $c = d = 1$, which implies a uniform prior on the number of actors that exhibit idiosyncratic self-perceptions. On the other hand, for the priors on the variance parameters, we set $a_\sigma = a_\tau = b_\kappa = 2$, which leads to a proper prior with finite mean but infinite variance. Furthermore, because of the symmetry of the model, it appears reasonable to set $b_\sigma = b_\tau = b$. We then jointly choose values of b, ω^2 and b_ς in a careful manner to ensure that $\mathbb{V}\text{ar}[\theta_{i,i'\,j}]$ is (roughly) constant as a function of the dimension of the latent spaces. We impose this constraint so that we can appropriately contrast models constructed with different values of K and avoid Barlett's paradox (Bartlett, 1957).
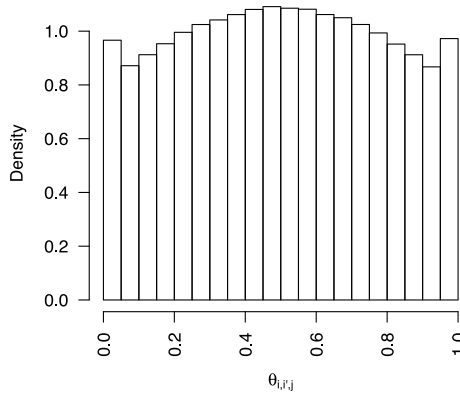
To accomplish this goal, assume that the covariates have been standardized and that we are looking at a (hypothetical) pair of actors for which the level of all covariates (except that associated with the intercept) is zero. Note that, for such an actor and moderate to large values of K, the linear predictor in (2) is approximately distributed a priori as a Gaussian distribution with (marginal) mean

$$\mathbb{E}\left[\beta_{j,0} + \boldsymbol{u}_{i,j}^T \boldsymbol{v}_{i'\,j}\right] = 0,$$
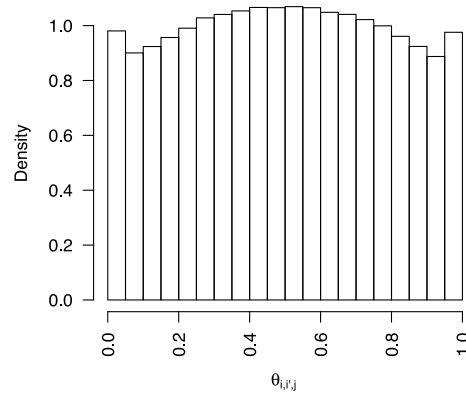
and (marginal) variance

$$\mathbb{V}\text{ar}\left[\beta_{j,0} + \boldsymbol{u}_{i,j}^T \boldsymbol{v}_{i'\,j}\right] = \{\omega^2 + b_\varsigma\} + K\{\kappa^2 + b\}^2. \tag{7}$$

Hence, setting $\{\omega^2 + b_\varsigma\} + K\{\kappa^2 + b\}^2 = 1$ in our probit model leads

opinions of other actors in the network. For example, for actors for which $\Pr(\gamma_i = 0 \mid \mathbf{Y})$ is very low, the model will heavily down-weight an actor's self perception in the estimates of $\boldsymbol{\eta}_i$. Finally, notice that our perceptual test is not based on the existence of an underlying "true" network structure; which implies that our inferences on $\gamma_1, \ldots, \gamma_I$ and $\xi_1, \ldots, \xi_I$ are limited to either agreement or disagreement regarding the others' cognitive perception, as opposed to accuracy or any other measure that involves any sort of an external gold standard.

The model is completed by specifying priors on the remaining model parameters. The consensus positions are assigned independent priors

$$\boldsymbol{\eta}_i | \kappa_\eta^2 \overset{\text{iid}}{\sim} N(\mathbf{0}, \kappa_\eta^2 \, \mathbf{I}_K), \quad \boldsymbol{\zeta}_i | \kappa_\zeta^2 \overset{\text{iid}}{\sim} N(\mathbf{0}, \kappa_\zeta^2 \, \mathbf{I}_K), \tag{6}$$

where $\kappa_\eta^2 = \kappa_\zeta^2 = \kappa^2$ are fixed. Note that this resembles the prior specification we used in (4) for the case $\gamma_i = 0$ and in (5) for the case $\xi_i = 0$. Centering the latent positions in bilinear models around zero is a common practice in the literature, in part because it implies that the model is, roughly speaking, centered around an Erdos-Renyi model (Erdös and Rényi, 1959). Furthermore, the choice of a zero mean and a covariance matrix that is proportional to the identity ensures that the priors, as well as the likelihood, are invariant to rotations of the latent space. This property will be important in our discussion of identifiability (see Section 4.1 below).

Finally, when covariates are available, the corresponding regression coefficients are assigned a hierarchical prior,

(a) K = 3

(b) K = 6

Fig. 4. Histogram of the marginal prior distribution on $\theta_{i,i'\,j}$, $j \neq i$, i', for $K = 3$ and $K = 6$.

to a prior for $\theta_{i,i'j}$ that is close to the uniform. To select the values of the hyperparameters, we split the prior variance of the linear predictor equally among all terms, leading to $\omega^2 = b_\zeta = 1/4$ and $\kappa^2 = b = 1/\sqrt{8K}$. Fig. 4 presents histograms of 100,000 independent realizations from the induced marginal prior on $\theta_{i,i'j}$ for $K = 3$ and $K = 6$. Both are very similar (both are somewhat "trimodal", with modes at $\theta_{i,i'j} = 0$, $\theta_{i,i'j} = 1/2$ and $\theta_{i,i'j} = 1$), but for $K = 6$ the distribution is, as we would expect, slightly less peaked.

We note that, while keeping the variance of the linear predictor constant is, in our experience, critical for good dimension selection performance, the specific value of 1 that we have chosen to use as a default is not. In particular, the mean and the variance of $\theta_{i,i'j}$ could be tuned to reflect prior information by slightly modifying the elicitation process we just described.

4. Computation

For a given K the posterior distribution of the parameters can be explored using Markov chain Monte Carlo (MCMC) algorithms in which the posterior distribution is approximated using dependent but approximately identically distributed samples $\Upsilon^{(1)}, \ldots, \Upsilon^{(S)}$, where

$$\Theta^{(s)} = \left(\beta_1^{(s)}, \ldots, \beta_I^{(s)}, u_{1,1}^{(s)}, \ldots, u_{I,I}^{(s)}, v_{1,1}^{(s)}, \ldots, v_{I,I}^{(s)}, \eta_1^{(s)}, \ldots, \eta_I^{(s)}, \zeta_1^{(s)}, \ldots, \zeta_I^{(s)}, \gamma_1^{(s)}, \ldots, \gamma_I^{(s)}, \xi_1^{(s)}, \ldots, \xi_I^{(s)}, \sigma_u^{(s)}, \tau_u^{(s)}, \sigma_v^{(s)}, \tau_v^{(s)}, \nu^{(s)}, \varsigma^{(s)}, \psi^{(s)} \right)$$

Point and interval estimates can be approximated from the empirical distributions. To facilitate computation we follow Albert and Chib (1993) and introduce independent auxiliary variables $z_{i,i'j}|\beta_j, u_{i,j}$, $v_{i'j} \overset{\text{ind}}{\sim} N\left(x_{i,i'}^T \beta_j + u_{i,j}^T v_{i'j}, 1 \right)$ and let

$$y_{i,i'j}|z_{i,i'j} = \begin{cases} 1, & z_{i,i'j} \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Marginalizing over the $z_{i,i'j}$ leads to our original Bernoulli likelihood with success probability given by (2). After introducing these latent variables, all full conditional distributions reduce to standard families, greatly simplifying computation. Details of the MCMC algorithm can be seen in the Supplementary Materials.

4.1. Identifiability

Bilinear models are invariant to rotations and reflections of the social space. Indeed, for any $K \times K$ orthogonal matrix $\mathbf{Q}$, the likelihood associated with the reparametrization $\tilde{u}_{ij} = \mathbf{Q}u_{ij}$, $\tilde{v}_{ij} = \mathbf{Q}v_{ij}$, $\tilde{\eta}_i = \mathbf{Q}\eta_i$ and $\tilde{\zeta}_i = \mathbf{Q}\zeta_i$ is independent of $\mathbf{Q}$. This lack of identifiability does not affect our ability to make inferences on the $\theta_{i,i'j}$s (which are identifiable), or on the indicators $\gamma_1, \ldots, \gamma_I, \xi_1, \ldots, \xi_I$, whose posterior distribution is a function of the $u_{i,j}$s, $v_{i,j}$s, η_is and ζ_is only through quadratic or bilinear functions of them (and are therefore identifiable as well). However, it does hinder us when trying to provide posterior estimates of the latent positions.

We address this invariance issue using a parameter expansion approach similar to that described in Hoff (2005). In particular, we address the identifiability issues through a post-processing step in which posterior samples are rotated/reflected to a shared coordinate system. For each sample $\Upsilon^{(s)}$, an orthogonal transformation matrix $\mathbf{Q}^{(s)}$ is obtained by minimizing the Procrustes distance,

$$\tilde{\mathbf{Q}}^{(s)} = argmin_{\mathbf{Q} \in \mathcal{S}^K} \text{tr}\left\{ \left(\mathbf{W}^{(1)} - \mathbf{W}^{(s)}\mathbf{Q} \right)^T \left(\mathbf{W}^{(1)} - \mathbf{W}^{(s)}\mathbf{Q} \right) \right\} \tag{8}$$

where $\mathcal{S}^K$ denotes the set of $K \times K$ orthogonal matrices and $\mathbf{W}^{(s)}$ is the $2I \times K$ matrix whose first I rows correspond to the transposes of $\eta_1^{(s)}, \ldots, \eta_I^{(s)}$ and the rest correspond to the transposes of $\zeta_1^{(s)}, \ldots, \zeta_I^{(s)}$. The minimization problem in (8) can be easily solved using singular value decompositions (e.g., see Borg and Groenen, 2005, Section 20.2). Once the matrices $\tilde{\mathbf{Q}}^{(1)}, \ldots, \tilde{\mathbf{Q}}^{(S)}$ have been obtained, posterior inference for the latent positions are based on the transformed coordinates $\tilde{u}_{ij}^{(s)} = \tilde{\mathbf{Q}}^{(s)} u_{ij}^{(s)}$, $\tilde{v}_{ij}^{(s)} = \tilde{\mathbf{Q}}^{(s)} v_{ij}^{(s)}$, $\tilde{\eta}_i^{(s)} = \tilde{\mathbf{Q}}^{(s)} \eta_i^{(s)}$ and $\tilde{\zeta}_i^{(s)} = \tilde{\mathbf{Q}}^{(s)} \zeta_i^{(s)}$.

4.2. Selection of the latent dimension K

The choice $K = 2$ is popular in the network literature. Indeed, setting $K = 2$ simplifies visualization and interpretation, and is therefore particularly useful when the main goal of the analysis is to provide a description of the social relationships. However, our model focuses on testing structural hypotheses associated with the cognitive data, and the value of K can potentially play a critical role in the results. Hence, we investigate some methodologies for selecting the dimension of the social space.

The network literature has largely focused on the Bayesian Information Criteria (BIC) (e.g., see Hoff, 2005, Handcock et al., 2007 and Airoldi et al., 2009) as a tool for model selection. However, BIC is inappropriate for hierarchical models since the hierarchical structure implies that the effective number of parameters will typically be lower than the actual number of parameters in the likelihood. An alternative to BIC that addresses this issue is the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002, 2014; Gelman et al., 2014b),

$$\text{DIC}\left(K\right) = -2logp\left(\mathbf{Y}|\hat{\Theta}_K\right) + 2p_{\text{DIC}}, \tag{9}$$

where $\hat{\Theta}_K$ denotes the posterior mean of model parameters assuming that the dimension of the social space is K, and the penalty term p_{DIC} on the model complexity is given by

$$p_{\text{DIC}} = 2logp\left(\mathbf{Y}|\hat{\Theta}_K\right) - 2\,\mathbb{E}[logp(\mathbf{Y}|\Theta_K)]$$

An alternative to DIC is the Watanabe-Akaike Information Criterion (WAIC) (Watanabe, 2010, 2013; Gelman et al., 2014b),

$$\text{WAIC}\left(K\right) = -2\sum_{j,i<i} log\mathbb{E}[p(y_{i,i'j}|\Theta_K)] + 2p_{\text{WAIC}},$$

where the complexity penalty is given by

$$p_{\text{WAIC}} = 2\sum_{j,i<i} \left\{ log\mathbb{E}[p(y_{i,i'j}|\Theta_K)] - \mathbb{E}[logp(y_{i,i'j}|\Theta_K)] \right\}$$

Note that in the previous expressions all expectations, which are computed with respect to the posterior distribution, can be approximated by averaging over MCMC samples.

A key advantage of the WAIC criteria is its invariance to reparameterizations, which makes it particularly helpful for models (such as ours) with hierarchical structures, for which the number of parameters increases with sample size (Gelman et al., 2014a; Spiegelhalter et al., 2014).

Table 1

Values of DIC and WAIC for selecting the dimension K of the latent space for LATENT using Krackhardt's (1990) data.

K	2	3	4	5	6	7	8
DIC	13,389.0	11,122.6	9470.2	8903.8	**8548.9**	8575.1	8576.7
WAIC	13,718.9	11,730.7	10,207.3	9707.4	**9384.7**	9404.3	9409.5

Bold values correspond to those values that minimize the criteria reported in the table.

5. Krackhardt's (1990) dataset revisited

In this section we analyze the Krackhardt's data introduced in Section 2 using our model from Section 3 with $\beta_j \equiv \beta_j$, for all j, $x_{i,i'} \equiv 1$ for all i and i', and $\nu \equiv \nu$ (referred to as LATENT for short), as well as the Bayesian models introduced in Swartz et al. (2015) (SWARTZ for short; see Section 1). SWARTZ is a logistic ANOVA model that includes all first and second order interactions among reporter, sender and receiver levels. This parameterization allows the practitioner to assess the level of agreement between self and group perception by contrasting the value of appropriate interaction parameters. However, it does not yield a straightforward mechanism to construct a consensus network. The results we report in this section are based on $S = 40,000$ samples obtained after thinning the original Markov chains every 25 observations. Execution times, at 2.169 h. for SWARTZ and between 2.186 ($K = 2$) and 4.720 ($K = 9$) h. for LATENT using a single core of an i7 Intel processor, are comparable for these two models. Convergence was monitored by tracking the variability of the joint distribution of data and parameters

using the multi-chain procedure discussed in Gelman and Rubin (1992).

5.1. Dimension of the latent space

Table 1 presents the values of the DIC and WAIC associated with versions of our model that differ in K, the number of dimensions of the latent social space. Note that both criteria favor a choice of $K = 6$, which is the value we use for all of our analyzes.

5.2. Self-perception assessment

As discussed in Sectionss 1 and 2, one important goal in the analysis of CSS data is to assess the agreement of the actors' self-perception with that of other members of the system. In the context of LATENT, the posterior probabilities $\Pr[\gamma_i = 1|Y]$ and $\Pr[\xi_i = 1|Y]$ provide the desired measures of Actor i's agreement in its role as sender and receiver, respectively. In particular, remember that posterior probabilities close to one correspond to high levels of agreement. In the case of SWARTZ, a
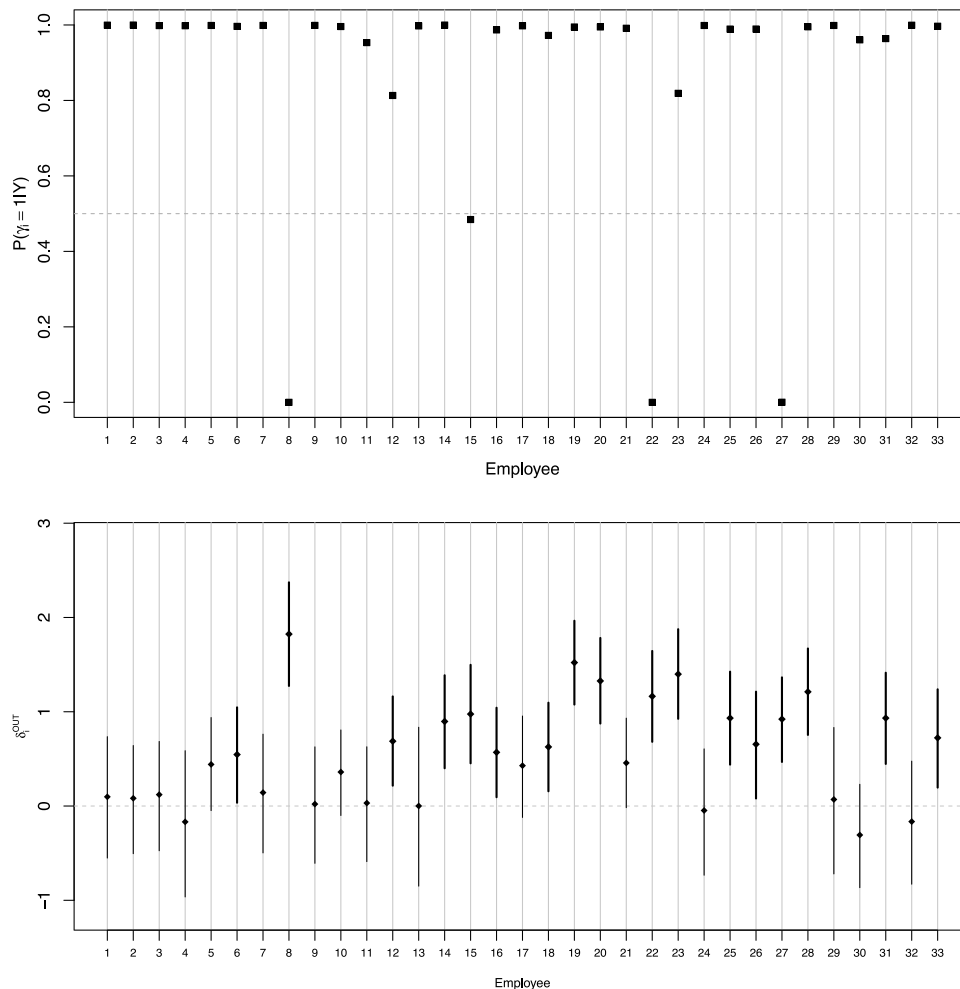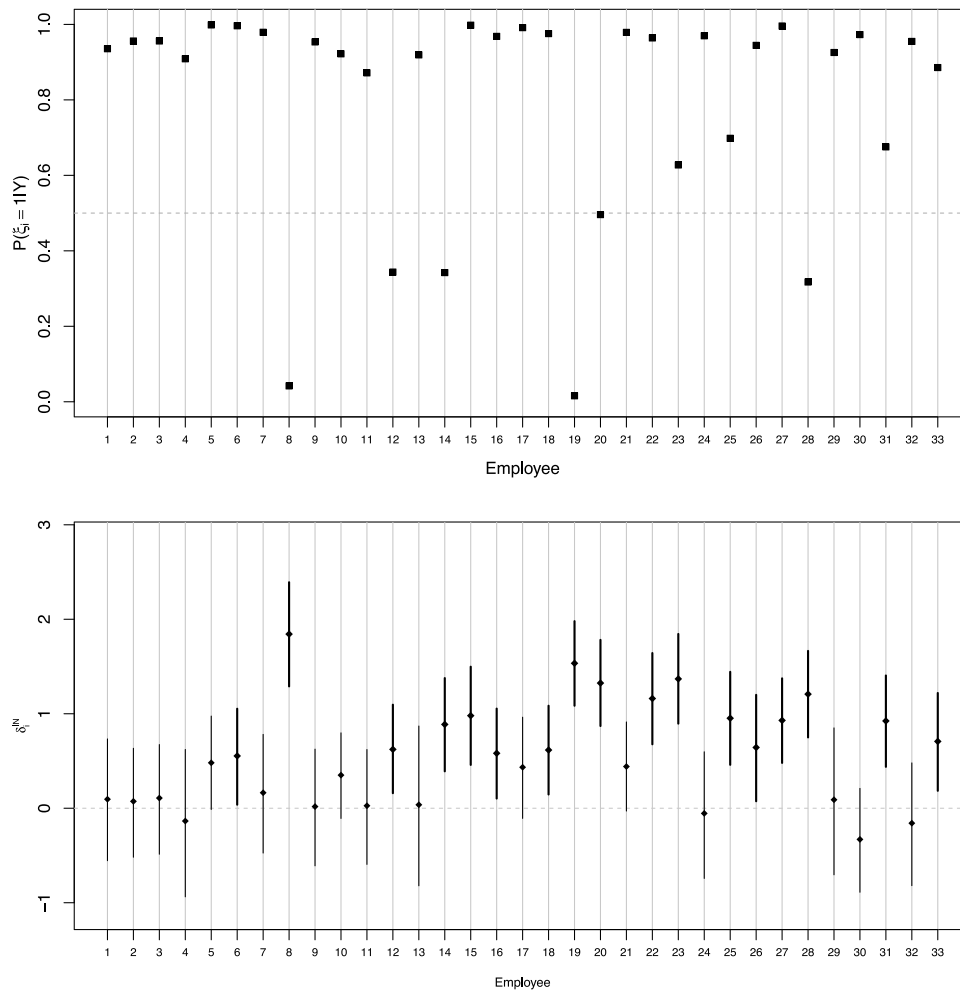

Fig. 5. Comparison of sender self perception assessments under LATENT and SWARTZ for Krackhardt's (1990) data. Top panel: For LATENT, posterior probabilities $\Pr[\gamma_i = 1|Y]$. Bottom panel: For SWARTZ, 95% credible intervals and posterior means for the distribution of the personal assessment parameters δ_i^{OUT}. Thicker lines correspond to credible intervals that do not contain zero.

Fig. 6. Comparison of receiver self perception assessment under LATENT and SWARTZ for Krackhardt's (1990) data. Top panel: For LATENT, posterior probabilities $\mathbb{Pr}[\xi_i = 1|Y]$. Bottom panel: For SWARTZ, 95% credible intervals and posterior means for the distribution of the personal assessment parameters δ_i^{IN}. Thicker lines correspond to credible intervals that do not contain zero.

sender (receiver) agreement measure δ_i^{OUT} (δ_i^{IN}) can be defined as the difference between the interaction term associated with reporter i and sender (receiver) i' and the average value of the same interaction term associated with all other reporters (see Swartz et al., 2015, Section 2 for details). For metrics δ_i^{OUT} and δ_i^{IN}, differences close to zero in absolute value correspond to actors whose self-perception strongly agrees with the consensus.

Figs. 5 and 6 present estimates of cognitive agreement for Krackhardt's data under these two models. Overall, LATENT is the most conservative model, identifying only four individuals whose sender self-perception does not agree with the consensus (actors 8, 15, 22 and 27), and five individuals whose receiver self-perception disagrees with the consensus (8, 12, 14, 19 and 28). SWARTZ identifies a total of 17 actors whose sender self-perception disagrees with the consensus (the same four identified by LATENT plus 6, 15, 16, 18, 19, 20, 22, 23, 25, 26, 27, 31, and 33), and another 17 whose receiver self-perception disagrees with the consensus (again, the five identified by LATENT plus 6, 15, 16, 18, 20, 22, 23, 25, 26, 27, 31 and 33). It is particularly striking that both lists are identical for SWARTZ. The difference between LATENT and SWARTZ can be partially explained by the well-known tendency of Bayesian procedures to automatically adjust for multiple comparisons when appropriate hierarchical priors are used (e.g., see Scott and Berger, 2006 and Scott and Berger, 2010). Achieving the same adjustment using SWARTZ approach would require adjusting the credibility level of the

intervals up from 95% according to, for example, Bonferroni's correction. One challenge with making these adjustments is that, in this setting, the different hypotheses being testing here are likely to be highly correlated.

5.3. Consensus network

In LATENT, the average positions $\eta_1, \ldots, \eta_I$ and $\zeta_1, \ldots, \zeta_I$ can be used to generate a *weighted* network, $\vartheta_{i,i'} = \Phi(\nu + \eta_i^T \zeta_{i'})$, which can be interpreted as the consensus "affinity" between Actors i and i'. Fig. 7 presents the matrix of posterior means $\mathbb{E}[\vartheta_{i,i'}|Y]$ for Krackhardt's (1990) data, along with a heat-map of the proportion of actors reporting each link, $\frac{1}{I}\sum_{j=1}^{I} y_{i,i',j}$. We do not present a consensus network for SWARTZ because this model does not provide any straightforward mechanism to construct such a network. Note that the consensus network from LATENT tends to be sparse, probably due to shrinkage and the fact that our model discounts information from actors whose self-perception disagrees with the rest.

5.4. Projections in social space

The latent positions $u_{i,j}$ and $v_{i,j}$ provide a powerful tool to describe social interactions. To illustrate this, we show in Fig. 8 the coordinates
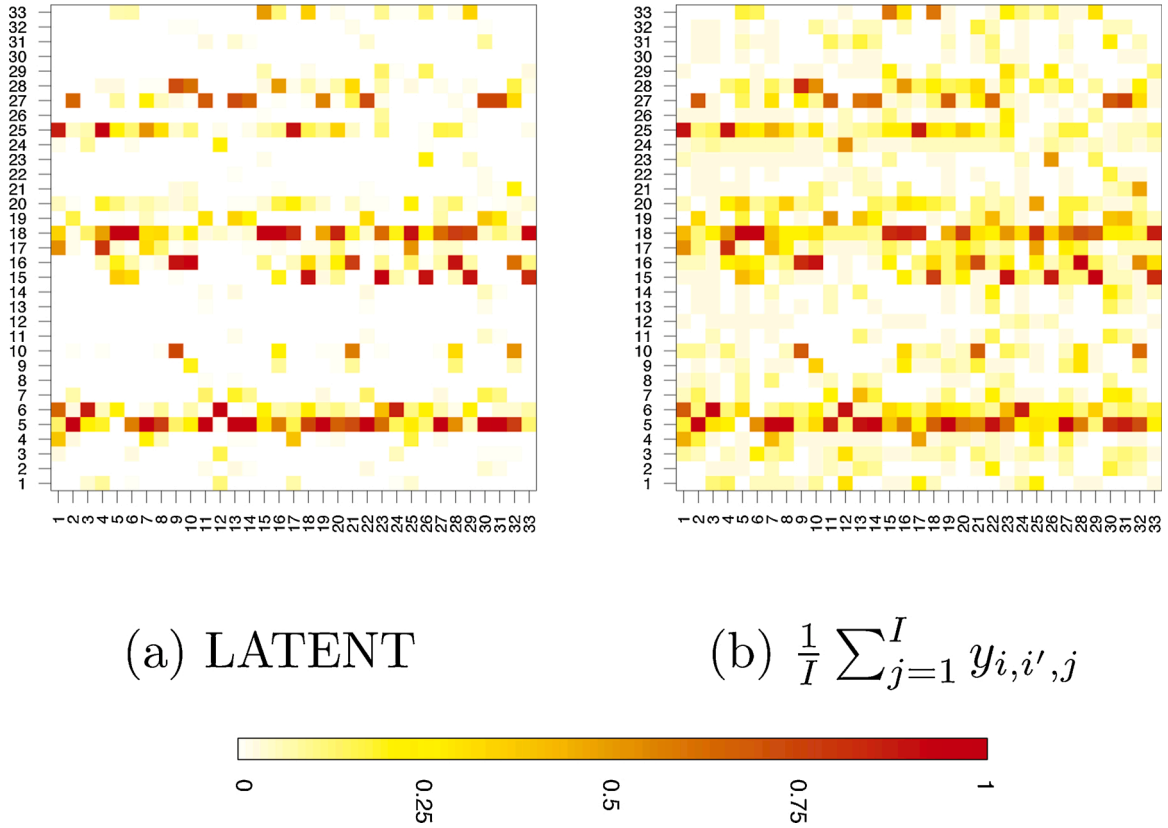
(a) LATENT

(b) $\frac{1}{I}\sum_{j=1}^{I} y_{i,i',j}$

Fig. 7. Estimates of the *weighted* consensus network for Krackhardt's (1990) data. The left panel provides the posterior mean under LATENT, and the right panel shows the proportion of actors that report perceiving each possible link, $\frac{1}{I}\sum_{j=1}^{I} y_{i,i'j}$.

along the two highest-variability dimensions of $\boldsymbol{u}_{i,j}$ (top row) and $\boldsymbol{v}_{i,j}$ (bottom row) for a two fixed values of i (in this case, $i = 1$ and $i = 8$) and every possible j (i.e., the position of these two actors as perceived by the different members of the social network, including themselves). These graphs are consistent with those from Figs. 5 and 6. In the first column of the figure, Actor 1's self-perceived position clusters with his/her position as perceived by other actors in both graphs, which is consistent with high values for both $\Pr[\gamma_1 = 1|Y]$ and $\Pr[\xi_1 = 1|Y]$. On the other hand, in the second column of Fig. 8 we see that, in both cases, Actor 8 is clearly isolated from the other actors. This is again consistent with the low values we reported for $\Pr[\gamma_8 = 1|Y]$ and $\Pr[\xi_8 = 1|Y]$.

It is interesting to contrast these results with those presented from Fig. 9, which instead presents the coordinates $\boldsymbol{u}_{i,j}$ (top row) and $\boldsymbol{v}_{i,j}$ (bottom row) for two fixed values of j and every possible i (i.e., the position of all actors as perceived by each of the two different observers). For example, from the second column we see that Actor 8 sees him/herself as occupying a somewhat isolated position in sender space (although certainly not as isolated as the one that other actors perceive him/her to be). However, in terms of the receiver space, Actor 8 perceives him/herself as being located right with the other actors. As we discussed above, this is the opposite of how the other actors perceive him/her.

5.5. Model fit

To asses the fit of our model, we first complement the results presented in Table 1 by computing the DIC and WAIC values for SWARTZ. For this model the DIC is 11,073.52, whereas the WAIC is 11,406.92. While the performance of SWARTZ is competitive with that of a low-dimensional LATENT models, it is clear that, for an optimally selected

latent-space dimension, LATENT outperforms SWARTZ under both criteria.

Next, following Gelman et al. (2014a, Chapter 6) and Kolaczyk and Csárdi (2014, Chapter 4), we replicate pseudo-data from both fitted models and calculate a battery of summary statistics (in our case, the average density, assortativity, clustering coefficient, and mean path length over the different networks) for each sample. This allows us to generate an estimate of the posterior predictive distribution of the summaries, which can then be compared against the value observed in the original sample (see Fig. 10). Our model clearly provides a better fit to the data, outperforming SWARTZ. This is most obvious for the assortativity, transitivity and mean path length indexes.

6. Cross-validation

As an additional goodness-of-fit assessment, we carry out cross-validation experiments on several CSS datasets exhibiting different kinds of actors, sizes, and relations, including two datasets collected as part of our research (see Table 2 for details). More specifically, for each combination of dataset and model, we performed an *L*-fold cross-validation (CV) in which *L* randomly selected subsets of roughly equal size in the CSS are treated as missing and then predicted using the rest of the data.

We summarize our findings in Table 3, where we report the average Area Under the ROC Curve (AUC) corresponding to the prediction of missing links for LATENT and SWARTZ, and also each dataset presented in Table 2. In this context, the AUC is a measure of how well a given model is capable of predicting missing links (the higher the AUC, the better the model is at predicting 0s as 0s and 1s as 1s). For LATENT, we report the AUC for the model with the optimal value of *K* according to the WAIC criteria.
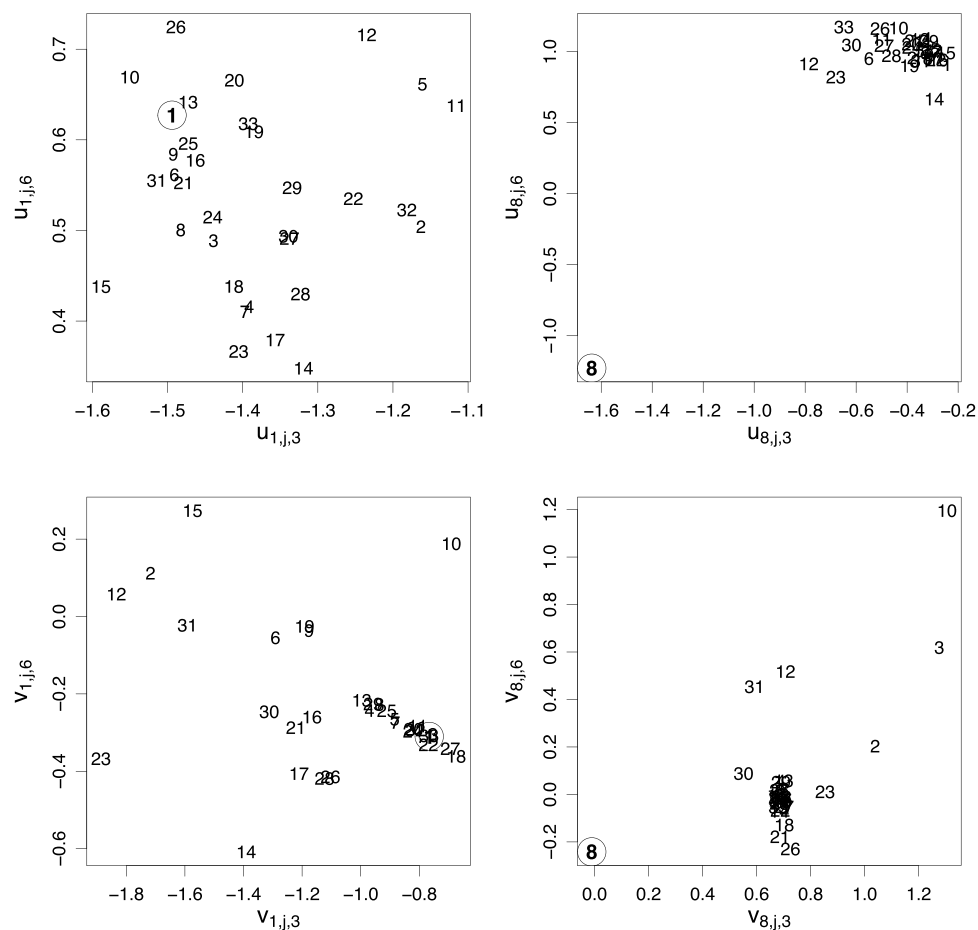
Fig. 8. Posterior means of the positions in social space along the two dimensions with highest variance for Actors 1 and 8 in Krackhardt's (1990) data, as perceived by all actors.

We can see from Table 3 that LATENT is the best model in terms of predicting missing links. Specifically, for the K21 and Six datasets, the predictive performance of LATENT and SWARTZ is practically the same; however, for all the other CSSs, LATENT has a better predictive performance than SWARTZ. These results strongly suggest that the predictive capabilities of LATENT are indeed comparable with or even better than those offered by its competitor.

7. Discussion

We have presented a novel approach for modeling CSS data with a focus on assessing cognitive agreement across actors in a social network. The spatial nature of the model (which operationalizes the notion of social roles) and its close relationship with latent factor models widely used for scaling categorical data in the social sciences make the model highly interpretable.

During the review process, one of the referees suggested that, if the spatial representation is the correct, then traditional CSS instruments are not appropriate, and that other approaches to data collection, such as introducing variants of social distance scales or attempting to directly elicit perceived positions in the underlying spatial representation, would be needed. While we agree that alternative instruments have the potential to provide higher quality data that would allow for more accurate estimates, we see no structural impediment to using traditional CSS instruments in the context of our model. Indeed, there is a rich tradition of using binary data to estimate spatial models in various social sciences disciplines. The most obvious example is the case of spatial voting models (Poole and Rosenthal, 1985; Clinton et al., 2004; Rodriguez and

Moser, 2015; Lofland et al., 2020). In fact, the model discussed in Clinton et al. (2004) is a bilinear model that is strongly related to the one we discuss in this manuscript. What these models do is map the binary responses into (social) distance scales without directly asking about them. In our case, the fact that the data is binary can be seen as measurement error introduced by the CSS instrument.

Our model can be easily adapted to account for undirected and/or weighted CSS data and to deal with missing values. Furthermore, a number of extension are possible. For example, we could incorporate explicit reporter-specific popularity terms for each actor (e.g., see Hoff, 2005), which can also be modeled hierarchically and tested for. Similarly, alternative models for transitivity (such as the eigenvalue model of Hoff, 2009) could be adapted to provide additional flexibility to the model. Also, the agreement probabilities of the indicators in Eqs. (4) and (5) could be replaced for subject-specific parameters in order to introduce more sensitivity into the model.

In this paper we adopted a two step procedure to selecting the dimension of the latent space using DIC and WAIC as model selection procedures. We have recently been successful at incorporating the selection of the dimension as part of the model estimation process in the context of dynamic networks (please see Guhaniyogi et al., 2020). In the future, we plan to investigate similar approaches to dimensionality selection in the context of CSS data.

One of the referees raised questions about the use of Gaussian distributions as priors for the latent positions. This choice, which is in line with most of the literature on the topic, is mostly one of convenience and is potentially restrictive. In the future, we plan to explore extensions of the model that would replace the Gaussian distributions in Eqs. (3), (4),
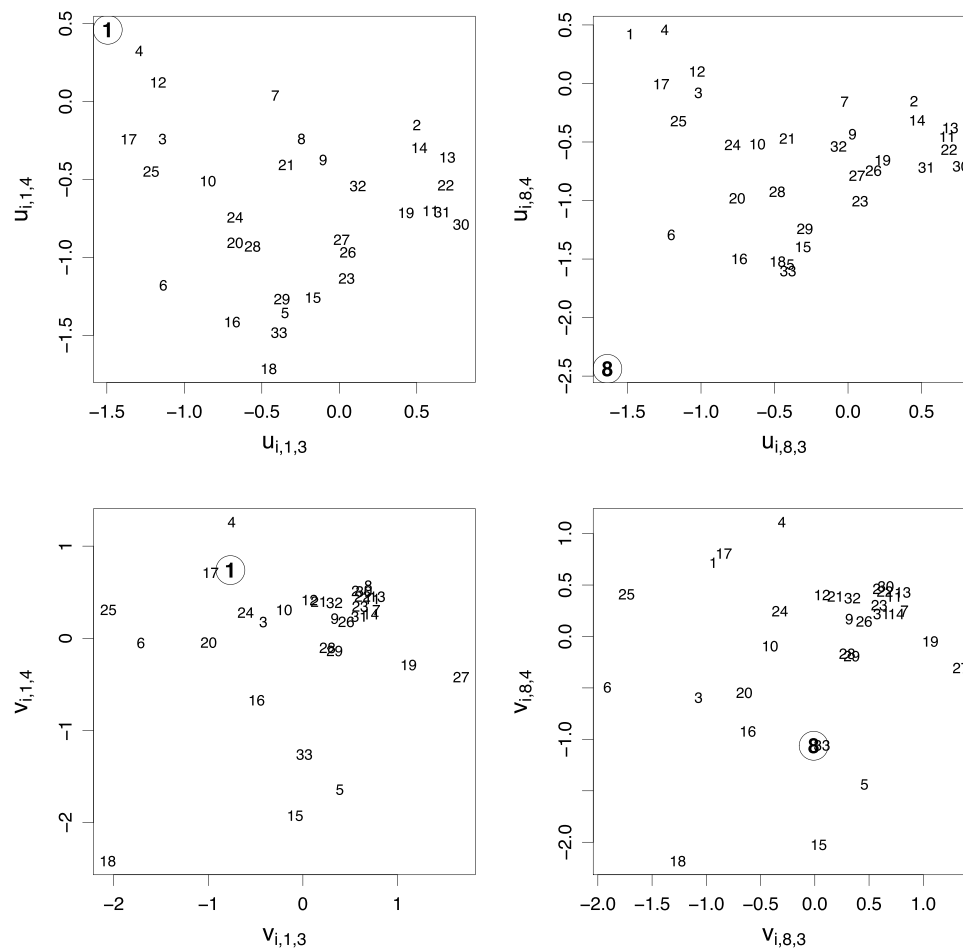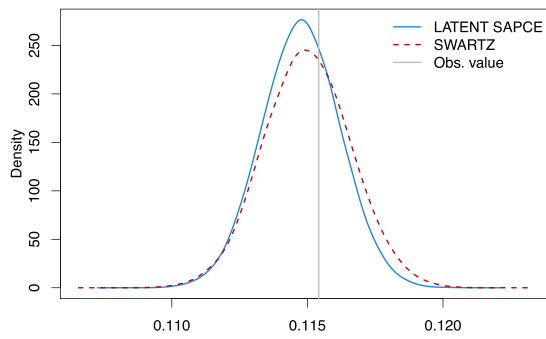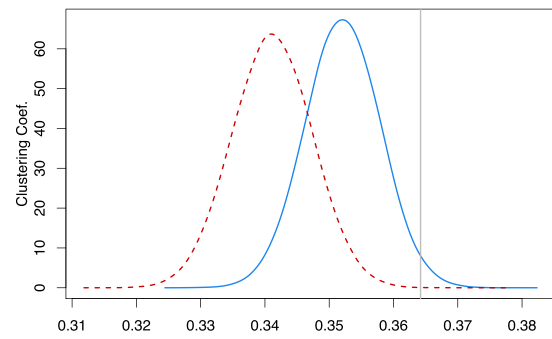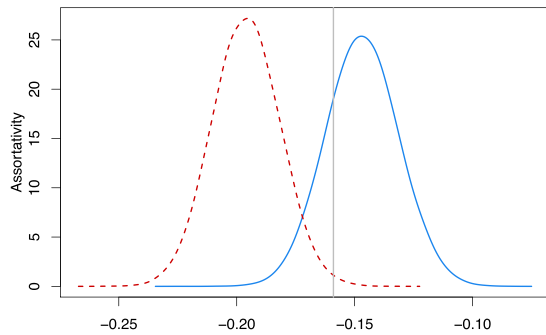

Fig. 9. Posterior means of the positions in social space along the two dimensions with highest variance for all actors as perceived by Actors 1 and 8 in Krackhardt's (1990) data.
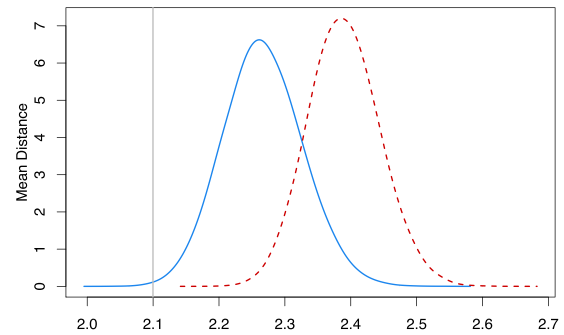
(a) Density

(b) Transitivity

(c) Assortativity

(d) Mean path length

Fig. 10. Kernel estimates corresponding to the empirical distribution of the test statistics for replicated data along with the observed value in Krackhardt's (1990) data.

Table 2

CSS datasets for which a series of cross-validation experiments are performed using LATENT and SWARTZ. Note that the K33 and K21 datasets are widely analyzed in Sections 2 and 2 from the supplementary material, respectively.

Acronym	Reference	Actors	Type	No of actors
K33	Krackhardt (1990)	Executives	Advise	33
It1	Casciaro (1998)	Researchers	Advise	24
Eco	Sosa and Rodriguez	College students	Advise	28
K21	Krackhardt (1987)	Executives	Friendship	21
It2	Casciaro (1998)	Researchers	Friendship	24
Six	Neal (2008)	6th graders	Friendship	15
Fin	Sosa and Rodriguez	College students	Friendship	19

Table 3

Average AUCs corresponding to the prediction of missing links in a series of CV experiments to assess the predictive performance of LATENT and SWARTZ, using each CSS provided in Table 2.

Dataset	LATENT	SWARTZ
K33	**0.962**	0.950
It1	**0.944**	0.928
Eco	**0.960**	0.923
K21	**0.917**	**0.913**
It2	**0.941**	0.930
Six	**0.928**	**0.926**
Fin	**0.931**	0.918

Bold values correspond to those values that minimize the criteria reported in the table.

and (5) with more general (possibly non-parametric) mixture priors. Such model structure, which is somewhat reminiscent of Handcock et al. (2007), is motivated by the results presented in Figs. 8 and 9. Indeed, from these figures it is apparent that the perception of the position of different actors in social space tend to form clusters, and that the use of Gaussian random effects might be inappropriate. Such extensions of the model will be discussed elsewhere.

Supplementary materials

The online supplementary materials include details of the Markov chain Monte Carlo algorithm used to fit our model, as well as an additional illustration focused on the dataset introduced in Krackhardt (1987), as well as detailed simulation studies that assess the accuracy of LATENT and SWARTZ to identify individuals with perceptual disagreement, the ability of our two-step approach to correctly select the dimension of the latent space, and an evaluation of the computational performance of our algorithm.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.socnet.2020.12.002.

References

Airoldi, E.M., Blei, D.M., Fienberg, S.E., Xing, E.P., 2009. Mixed membership stochastic blockmodels. Advances in Neural Information Processing Systems, pp. 33–40.

Albert, J.H., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. J. Am. Stat. Assoc. 88 (422), 669–679.

Aldous, D.J., 1985. Exchangeability and related topics. École d'Été de Probabilités de Saint-Flour XIII-1983 1–198.

Bartlett, M., 1957. A comment on d. v. lindley's statistical paradox. Biometrika 44 (3), 533–534.

Batchelder, W.H., Kumbasar, E., Boyd, J.P., 1997. Consensus analysis of three-way social network data. J. Math. Sociol. 22 (1), 29–58.

Bond, C.F., Horn, E.M., Kenny, D.A., 1997. A model for triadic relations. Psychol. Methods 2 (1), 79.

Bond, C.F., Kenny, D.A., Broome, E.H., Stokes-Zoota, J.J., Richard, F.D., 2000. Multivariate analysis of triadic relationst. Multivar. Behav. Res. 35 (3), 397–426.

Borg, I., Groenen, P.J., 2005. Modern Multidimensional Scaling: Theory and Applications. Springer Science & Business Media.

Brands, R.A., 2013. Cognitive social structures in social network research: a review. J. Organ. Behav. 34 (S1), S82–S103.

Butts, C.T., 2000. Network Inference, Error, and Informant (in)Accuracy: A Bayesian Approach. Carnegie Mellon University. Technical Report.

Butts, C.T., 2003. Network inference, error, and informant (in) accuracy: a Bayesian approach. Soc. Netw. 25 (2), 103–140.

Casciaro, T., 1998. Seeing things clearly: social structure, personality, and accuracy in social network perception. Soc. Netw. 20 (4), 331–351.

Clinton, J., Jackman, S., Rivers, D., 2004. The statistical analysis of roll call data. Am. Polit. Sci. Rev. 98 (2), 355–370.

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

Durante, D., Dunson, D.B., Vogelstein, J.T., 2017. Nonparametric bayes modeling of populations of networks. J. Am. Stat. Assoc. 112 (520), 1516–1530.

Erdős, P., Rényi, A., 1959. On random graphs. Publ. Math. 6 (290–297), 5.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2014a. Bayesian Data Analysis, vol. 2. Chapman & Hall/CRC, Boca Raton, FL, USA.

Gelman, A., Hwang, J., Vehtari, A., 2014b. Understanding predictive information criteria for Bayesian models. Stat. Comput. 24 (6), 997–1016.

Gelman, A., Rubin, D., 1992. Inferences from iterative simulation using multiple sequences. Stat. Sci. 7, 457–472.

Guhaniyogi, R., Rodriguez, A., et al., 2020. Joint modeling of longitudinal relational data and exogenous variables. Bayesian Anal.

Handcock, M.S., Raftery, A.E., Tantrum, J.M., 2007. Model-based clustering for social networks. J. R. Stat. Soc.: Ser. A (Stat. Soc.) 170 (2), 301–354.

Hoff, P.D., 2005. Bilinear mixed-effects models for dyadic data. J. Am. Stat. Assoc. 100 (469), 286–295.

Hoff, P.D., 2009. Multiplicative latent factor models for description and prediction of social networks. Comput. Math. Organ. Theory 15 (4), 261–272.

Hoff, P.D., 2015. Multilinear tensor regression for longitudinal relational data. Ann. Appl. Stat. 9 (3), 1169.

Kenny, D.A., 1994. Interpersonal Perception: A Social Relations Analysis. Guilford Press.

Kilduff, M., Crossland, C., Tsai, W., Krackhardt, D., 2008. Organizational network perceptions versus reality: a small world after all? Organ. Behav. Hum. Decis. Process. 107 (1), 15–28.

Kolaczyk, E.D., Csárdi, G., 2014. Statistical Analysis of Network Data With R, vol. 65. Springer.

Koskinen, J., 2002a. Bayesian Analysis of Cognitive Social Structures With Covariates. Department of Statistics, Stockholm University. Technical Report.

Koskinen, J., 2002b. Bayesian Analysis of Perceived Social Networks. Department of Statistics, Stockholm University. Technical Report.

Koskinen, J., 2004. Model Selection for Cognitive Social Structures. Department of Statistics, Stockholm University. Technical Report.

Krackhardt, D., 1987. Cognitive social structures. Soc. Netw. 9 (2), 109–134.

Krackhardt, D., 1990. Assessing the political landscape: structure, cognition, and power in organizations. Adm. Sci. Q. 342–369.

Kumbasar, E., 1996. Methods for analyzing three-way cognitive network data. J. Quant. Anthropol. 6 (1–2), 15–34.

Kumbasar, E., Rommey, A.K., Batchelder, W.H., 1994. Systematic biases in social perception. Am. J. Sociol. 477–505.

Lofland, C.L., Moser, S., Rodriguez, A., 2020. Multiple ideal points: revealed preferences in different domains. Polit. Anal.

Neal, J.W., 2008. "kracking" the missing data problem: applying krackhardt's cognitive social structures to school-based social networks. Sociol. Educ. 81 (2), 140–162.

Newcomb, T., 1961. The Acquaintance Process. Holt, Rinehart & Winston, New York.

Pattison, P., 1994. Social Cognition in Context. Advances in Social Network Analysis. Sage, Thousand Oaks, pp. 79–109.

Poole, K.T., Rosenthal, H., 1985. A spatial model for legislative roll call analysis. Am. J. Polit. Sci. 357–384.

Rodriguez, A., 2015. A Bayesian Nonparametric Model for Exchangeable Multinetwork Data Based on Fragmentation and Coagulation Processes. University of California, Santa Cruz. Technical Report.

Rodriguez, A., Moser, S., 2015. Measuring and accounting for strategic abstentions in the us senate, 1989–2012. J. R. Stat. Soc.: Ser. C (Appl. Stat.) 64 (5), 779–797.

Salter-Townshend, M., McCormick, T.H., 2017. Latent space models for multiview network data. Ann. Appl. Stat. 11 (3), 1217.

Sampson, S.F., 1968. A Novitiate in a Period of Change: An Experimental and Case Study of Social Relationships. PhD Thesis. Cornell University. September.

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

Scott, J.G., Berger, J.O., 2006. An exploration of aspects of Bayesian multiple testing. J. Stat. Plan. Inference 136, 2144–2162.

Scott, J.G., Berger, J.O., 2010. Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problems. Ann. Stat. 38, 2587–2619.

Sewell, D.K., Chen, Y., 2015. Latent space models for dynamic networks. J. Am. Stat. Assoc. 110 (512), 1646–1657.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A., 2014. The deviance information criterion: 12 years on. J. R. Stat. Soc.: Ser. B (Stat. Methodol.) 76 (3), 485–493.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc.: Ser. B (Stat. Methodol.) 64 (4), 583–639.

Swartz, T.B., Gill, P.S., Muthukumarana, S., 2015. A Bayesian approach for the analysis of triadic data in cognitive social structures. J. R. Stat. Soc.: Ser. C (Appl. Stat.) 64 (4), 593–610.

von Davier, M., Carstensen, C.H., 2007. Multivariate and Mixture Distribution Rasch Models: Extensions and Applications. Springer.

Wang, L., Zhang, Z., Dunson, D., et al., 2019. Common and individual structure of brain networks. Ann. Appl. Stat. 13 (1), 85–112.

Wasserman, S., Faust, K., et al., 1994. Social Network Analysis: Methods and Applications, vol. 8. Cambridge University Press.

Watanabe, S., 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. 11 (December), 3571–3594.

Watanabe, S., 2013. A widely applicable bayesian information criterion. J. Mach. Learn. Res. 14 (March), 867–897.

Westveld, A.H., Hoff, P.D., et al., 2011. A mixed effects model for longitudinal relational and network data, with applications to international trade and conflict. Ann. Appl. Stat. 5 (2A), 843–872.