An alternative class of models to position social network groups in latent spaces

Izabel Nolau\textsuperscript{a}, Gustavo S Ferreira\textsuperscript{b}

\textsuperscript{a}Federal University of Rio de Janeiro  
\textsuperscript{b}National School of Statistical Sciences

Abstract. Identifying key nodes, estimating the probability of connection between them, and distinguishing latent groups are some of the main objectives of social network analysis. In this paper, we propose a class of blockmodels to model stochastic equivalence and visualize groups in an unobservable space. In this setting, the proposed method is based on two approaches: latent distances and latent dissimilarities at the group level. The projection proposed in the paper is performed without needing to project individuals, unlike the main approaches in the literature. Our approach can be used in undirected or directed graphs and is flexible enough to cluster and quantify between and within-group tie probabilities in social networks. The effectiveness of the methodology in representing groups in latent spaces was analyzed under artificial datasets and in two case studies.

1 Introduction

In social network analysis, it is usual to examine the association of $n$ individuals through a matrix $Y_{n \times n}$, whose elements $y_{ij}$ describe the connection between the $i$-th and $j$-th components of the network. These elements can be represented as a graph, where each node (or vertex, or point) represents an individual, and the edges (or links, or ties) represent relationships among them. Identifying key nodes, estimating the probability of connection between them, and distinguishing latent groups are some of the main objectives of social network analysis. Several methods have been proposed, including deterministic techniques for graph analysis and, more recently, sophisticated statistical models using latent effects.

Among the classical approaches to model social networks data, Holland and Leinhardt (1981) described an exponential family of distributions that can be used to model directed graph or digraph (the $p_1$ model); Fienberg and Wasserman (1981) extended the $p_1$ model to model data with variables measuring nodal attributes or groups; Wang and Wong (1987) extended the $p_1$ model to include block structures; and Hoff, Raftery and Handcock (2002) developed a class of models where the probability of a relation between actors depends on their positions in a latent social space.

Some advances in social network studies focused on clustering problems, where one is interested in partitioning the elements of the social network in observable or unobservable groups (or blocks). The concept of structure equivalence, which is the basis of this approach, was firstly discussed in Lorrain and White (1971). In simple words, two individuals belonging to a specific group in a social network are structurally equivalent if they relate to individuals from other groups in the same way. Based on this, several procedures have been proposed to model relational data assuming prespecified groups (see Breiger, Boorman and Arabie, 1975; White, Boorman and Breiger, 1976; Doreian, Batagelj and Ferligoj, 2005, among others).

Fienberg and Wasserman (1981) and Holland, Laskey and Leinhardt (1983) extended the concept of structural equivalence to stochastic versions of blockmodeling. A stochastic blockmodel can be defined as a probability distribution for graphs of which the vertex set...
A network is partitioned into subsets called blocks, where this probability distribution is invariant under permutations of vertices within blocks (Snijders and Nowicki, 1997). In this setting, two elements classified in the same group are said to be stochastically equivalent, in the sense that the probabilities of ties with all other elements are the same for all elements in the same group.

In the context of the p1 model, Wasserman and Anderson (1987) extended the block-modeling to situations where the groups are defined without using exogenous information about network elements. This model, known as the posterior blockmodel, simultaneously partitions actors and performs relational data analyses using statistical methods that yield stochastic posterior groups.

Snijders and Nowicki (1997) proposed a stochastic posterior blockmodel to undirected graphs that considers two groups, where the block structure was not directly observed. In this model, the probability of a link between two elements depends only on the groups to which they belong, and it is modeled according to a Beta distribution with order constraints to avoid identifiability problems. Later, Nowicki and Snijders (2001) extended this approach, considering that relations could be directed and could have an arbitrary set of possible values, and the number of classes was arbitrary. These models are known as latent class models. Extending these ideas, Airoldi et al. (2008) proposed a mixed membership blockmodel, relaxing the assumption that each individual can only belong to one cluster, and Schweinberger and Handcock (2015) presented an alternative approach, based on random graph models with local dependence, which also produces groups of nodes.

Both models proposed by Snijders and Nowicki (1997) and Nowicki and Snijders (2001) can not describe transitivity — when individuals that have ties to a third individual are more likely to be tied than individuals that do not — and homophily on attributes — when individuals with similar attributes are more likely to tie to each other than dissimilar ones — within groups. Then, Tallberg (2004) extended this approach to include covariates on the individual level to represent homophily on attributes.

Representing social networks in latent spaces was initially introduced via Multidimensional Scaling (e.g. Breiger, Boorman and Arabie, 1975), where a set of dissimilarities between individuals in a social network is represented as a set of positions in an unobservable space. The larger the distance between two individuals, the higher their mutual dissimilarity. The set of dissimilarities can be obtained directly from the sociomatrix $Y$ or can be based on a measure of this matrix (e.g., correlation matrix). Since different dissimilarity measures are possible and this measure must be supplied to the algorithm, the results depend on this choice (Handcock, Raftery and Tantrum, 2007). Besides this, results obtained from classical Multidimensional Scaling do not take into account uncertainties about group dissimilarities. For interpretations of different distance measures to represent structural equivalence, see Faust and Romney (1985).

Hoff, Raftery and Handcock (2002) proposed a latent space model in which each individual has a latent position in an Euclidean space, and these positions are estimated through ML or Bayesian inference. In this case, there is no need to provide an arbitrary dissimilarity measure to the model. Then, Handcock, Raftery and Tantrum (2007) extended this approach to take account of transitivity, homophily on attributes, and clustering simultaneously. The resulting model can be viewed as a stochastic blockmodel with transitivity within blocks and homophily on attributes. Including random effects, Krivitsky et al. (2009) extended this model to represent all four features associated with a network: transitivity, homophily on attributes, clustering, and heterogeneity of individual degrees. Using observed dissimilarities as the response variable, Oh and Raftery (2007) proposed a model-based clustering method, where the latent positions in Euclidian space are generated from a mixture of multivariate normal distributions, each one corresponding to a cluster.
Alternative models to position social network groups

Alternative models to position social network groups

Other extensions of stochastic blockmodels were proposed in the literature, most of them aiming jointly to represent latent clustering, transitivity, and high-dimensional data (see Rohe et al., 2011; Choi, Wolfe and Airoldi, 2012; Ng et al., 2018; Fosdick et al., 2019, among others).

Alternative approaches focused on community detection and based on optimization and classical estimation methods are also found in the literature. Examples of these approaches are Cavallari et al. (2017) and Cavallari et al. (2019), which propose models for graph embedding based on embedding communities, and Wang et al. (2017), which presents models for network embedding that preserve both microscopic (in node level) and the mesoscopic structure (in group level), among others. This type of approach holds a high capacity to identify and represent communities and nodes in space.

Visual display of clusters in low-dimensional spaces is often desired, since it may provide useful information about group dissimilarities (Oh and Raftery, 2007; Hedenfalk et al., 2002; Nikkilä et al., 2002). Despite that, most of the recent advances focus on positioning individuals in latent spaces, instead of their groups. Even if these positions can be derived — e.g., placing the mean vector of each latent mixture component using the approach of Handcock, Raftery and Tantrum (2007) —, it is also needed to place each individual in the latent space. On the other hand, positioning individuals in latent spaces via Multidimensional Scaling to derive group positions may not be a suitable approach since one needs previously specify the set of dissimilarities that will be used as input to this method.

In this paper, we propose a simple latent class model to represent groups in an unobservable space. In this setting, we propose an alternative class of models to position groups based on two approaches: latent distances and latent dissimilarities at the group level. The novelty of our approach is to propose a class of simple blockmodels that allows to position groups in latent spaces aiming:

• to model the relationship between groups as in Hoff, Raftery and Handcock (2002) without needing to represent individuals in the latent space;
• to estimate within and between-groups probabilities of ties to properly represent stochastic equivalence, as in the traditional latent class models; and
• to use a random version of multidimensional scaling based on samples from the posterior distribution of a set of unknown dissimilarities provided by data.

The paper is organized as follows. In Section 2, an alternative class of models to positioning groups in latent spaces is introduced, and aspects of inference are discussed. Section 3 and 4 evaluate the effectiveness of the proposed models according to artificial datasets and real data, respectively. Finally, a discussion about the results and suggestions for further research are presented in Section 5.

1.1 Related models

Our general model, presented in Section 2, brings new features about blockmodelling and has similarities with alternative models in the literature. The Latent Dissimilarities Model - LDM - and the Latent Positions Model - LPM -, respectively described in Sections 2.1 and 2.2, extend the ideas from Nowicki and Snijders (2001) and Airoldi et al. (2008) to spatially represent network’s groups in unobserved spaces.

These representations are performed assuming that the latent positions are parameters, in the LPM model, and via Multidimensional Scaling, in the LDM model. Both models allow the visualization of the latent structure between groups and also give an intuitive interpretation for between-group probabilities.

Instead of representing individuals in latent spaces, as the distance model proposed by Hoff, Raftery and Handcock (2002), the LPM model places groups. On the other hand,
the LDM model can be seen as a reparametrization of a blockmodel by decomposing the between-group tie probability into two components: an intercept and a distance between groups, which are estimated from data. This set of distances is then used to represent latent positions in a second stage via Multidimensional Scaling. Thus, the novelty of the LDM model is to provide a set of distances for the Multidimensional Scaling directly estimated from data by using a simple and intuitive model structure.

To compare our approach to other models that represent individuals in latent spaces, we will point out below the main differences and similarities between the proposed models, the Latent Position Cluster Model - LPCM -(Handcock, Raftery and Tantrum, 2007) and the Latent Space Stochastic Blockmodel - LSSBM (Fosdick et al., 2019).

The LPCM models within and between-group tie probabilities in the network, projecting all nodes in a latent space. In this setting, each latent position is modeled as Gaussian mixtures and the latent distances between all nodes are used to model tie probabilities, regardless of whether individuals are in the same group or not. Thus, while projecting all nodes in the latent space, the LPCM model makes no distinction between tie probabilities for nodes in the same or different groups and also does not provide the cluster’s projection in the latent space. Despite the fact that the vector of centers from the Gaussian mixture in the LPCM model can be indirectly assumed as a set of latent group’s positions — since the position of a node is affected by the weighted average of the positions of all centers and not only by the center of its group —, obtaining their estimates requires modeling the latent positions of all individuals in the network.

In turn, the LSSBM model decomposes network structure into two components: one that describes between-community relations, and another describing within-community relations. This approach also uses the concept of latent distances (Hoff, Raftery and Handcock, 2002), but only to model the within-group tie probabilities. Thus, unlike our approach, the LSSBM model provides latent representations only for nodes inside each group — in a multiresolution perspective — but does not provide groups’ latent positions. Furthermore, the LSSBM model also requires modeling the latent positions of all individuals in the network and focuses on undirected relations. On the other hand, LDM and LPM models do not require nodes’ latent positions to obtain the groups’ latent positions and are suitable to model undirected or directed relational data.

2 A new class of models to position latent groups

In general, social network data with \( n \) individuals produces a matrix \( Y_{n \times n} \), whose elements \( y_{ij} \) indicate the existence or non-existence of a connection — or simply the number of connections — between the \( i \)-th and \( j \)-th elements of the network. These matrices can be symmetric (undirected networks) or not (directed networks). In this paper, we consider the asymmetric case, where \( y_{ij} \neq y_{ji} \), but all results can be easily extended to undirected graphs by adjusting indexes (from \( i \neq j \) to \( i < j \)). Let \( \mu_{ij} \) be the probability that two individuals \( i \) and \( j \) share a connection — or the expected number of connections between them. Now, we propose a novel class of models to position groups in a latent space, which general formulation is presented below:

\[
Y_{ij} \sim F(\mu_{ij}), \quad i \neq j
\]

\[
g(\mu_{ij}) = \alpha_0 + \beta(c_i, c_j),
\]

\[
C_i \sim \text{Multinom}(1, p_i), \quad i = 1, \ldots, n
\]

\[
\alpha_0 \sim N(0, \sigma^2_\alpha),
\]

where
Alternative models to position social network groups

- $F$ represents a probability distribution with mean $\mu_{ij}$;
- $g(\cdot)$ is a link function;
- $\alpha_0$ is a common level for all individuals in the social network;
- $C_i = (C_{i1}, \ldots, C_{ik})'$ is a vector with $K - 1$ zeros and one number one, that indicates the group of the $i$–th individual and $C = (C_1, \ldots, C_n)'$;
- $c_i$ is the class of the $i$–th individual, i.e., $c_i = \{k \mid C_{ik} = 1\}$, for $i = 1, \ldots, n$;
- $p_i = (p_{i1}, \ldots, p_{ik})'$ is a vector indicating the prior probabilities of belonging to each group for the $i$–th individual;
- $K$ is the number of latent groups.

Here, the $\beta(c_i, c_j)$ parameter can be defined in many ways to accommodate, or not, between-groups’ and within-groups’ effects. The prior distribution of this quantity will be initially denoted by $p(\beta(c_i, c_j))$, for all $i \neq j$. On the other hand, the $\alpha_0$ parameter assumes the role of a sparsity parameter, controlling the average number of connections between individuals observed in the sociomatrix $Y$. Finally, the vector $p_i = (p_{i1}, \ldots, p_{ik})'$ can be fixed — e.g., assigning prior probabilities equal to $1/K$ — or modeled according to a Dirichlet distribution.

Under this formulation, the posterior distribution is given as follows:

$$p(\alpha_0, \beta, C \mid y) \propto \prod_{i=1}^{n} \prod_{j=1 \atop j \neq i}^{n} \mu_{ij}^{y_{ij}} (1 - \mu_{ij})^{1 - y_{ij}} \times p(\beta(c_i, c_j)) \times \prod_{i=1}^{n} \prod_{k=1}^{K} C_{ik}^{\frac{y_{ik}}{\alpha_0^2}} \times \exp \left( - \frac{\alpha_0^2}{2\sigma^2} \right).$$

The complexity level of the general model depends on the structure specified for $\beta(c_i, c_j)$. Henceforward, without loss of generality, it will be assumed that each observation $y_{ij}$ assumes only 0 or 1. In this way, set $F$ as the Bernoulli distribution, $\mu_{ij}$ as the probability that two individuals $i$ and $j$ share a connection, and $g$ as the logistic link function.

In Subsections 2.1 and 2.2, four possible formulations of the model (1) are presented. These different approaches are subdivided into two types: models based on latent dissimilarities and models based on latent positions.

### 2.1 Latent dissimilarities models

In this subsection, the proposed models aim to place groups in space based on the concept of latent dissimilarities. In this formulation, it is considered that the tie probability depends on a set of dissimilarities between groups. This approach allows a posterior spatial representation of groups through multidimensional scaling, without the need to previously specify the set of dissimilarities, denoted by $\delta$. Then, one can use samples drawn from the posterior distribution of the dissimilarities to place groups in the latent space, through Multidimensional Scaling.

#### 2.1.1 Latent dissimilarities model with no within-group variation

The latent dissimilarities model with no within-group variation (LDM) is the simplest proposed formulation of (1). The structure of this model depends on a set of $\binom{K}{2}$ latent dissimilarities (for each group pair, there is an associated dissimilarity) that account for estimating the between-group tie probabilities. This model’s construction considers that individuals belonging to the same group have a constant connection probability, regardless of the group that they are part of,
denying variation in within-group tie probabilities. In this case, the structure specified for \( \beta(c_i, c_j) \) is given by:

\[
\beta(c_i, c_j) = \begin{cases} 
0, & \text{if } c_i = c_j \\
-\delta_{c_ic_j}, & \text{if } c_i \neq c_j
\end{cases}, \quad \text{for } i, j = 1, \ldots, n, \ i \neq j,
\]

(2)

where \( \delta_{kl} \) represents a symmetric dissimilarity measure between groups \( k \) and \( l \). Note that, since the dissimilarity among groups \( k \) and \( l \) and between groups \( l \) and \( k \) are the same, then \( \delta_{kl} = \delta_{lk} \). Thus, to unify the notation, the groups that compose the subindex \( kl \) will be displayed in the ascending order. Moreover, if \( c_i = c_j \), the probability \( \mu_{ij} = \logit^{-1}(\alpha_0 + \beta(c_i, c_j)) \) will only depend on \( \alpha_0 \). Then, \( \logit^{-1}(\alpha_0) \) can be interpreted as the probability of a tie between two elements belonging to the same group. Moreover, the higher the dissimilarity between two groups, the smaller the probability of a tie between two elements from different groups.

Following the Bayesian paradigm, it is assumed that \( \delta_{kl} \sim \text{Gamma}(a_\delta, b_\delta) \), for \( l = 1, \ldots, K \) and \( k < l \), with \( a_\delta = 0.01 \) and \( b_\delta = 0.01 \) to ensure a non-informative prior. Under this formulation, the posterior distribution is given by:

\[
p(\alpha_0, \delta, \mathbf{C} \mid \mathbf{y}) \propto \prod_{j=1}^{n} \prod_{j \neq j}^{n} \left[ \frac{\exp(y_{ij}\alpha_0)}{\exp(\alpha_0) + 1} \right]^{1(y_{ij} = c_j)} \left[ \frac{\exp(y_{ij}(\alpha_0 - \delta_{c_i c_j}))}{\exp(\alpha_0 - \delta_{c_i c_j}) + 1} \right]^{1(y_{ij} \neq c_j)} \times \prod_{l=1}^{K} \prod_{k=1}^{K} C_{ik} \times \prod_{l=1}^{K} \prod_{k<l}^{K} \delta_{kl}^{a_\delta - 1} \exp(-b_\delta \delta_{kl}) \times \exp \left( -\frac{1}{2\sigma_\alpha^2} \right) \alpha_0^2,
\]

where \( \mathbb{I}(\cdot) \) represents the indicator function, which assumes 1, if the condition \( \cdot \) is satisfied, and 0 otherwise. For more information about the methods adopted to perform inference about LDM’s parameters, see Subsection 2.3.

2.1.2 Latent dissimilarities model with within-group variation

The latent dissimilarities model with within-group variation (GLDM) is an extension of the LDM, to represent variation among within-group tie probabilities, accommodating different levels of homogeneity inside each group. The structure of this model depends on a set of \( \binom{K}{2} \) latent dissimilarities that account for estimating the between-group tie probabilities, and on a set of \( K \) parameters, given in the vector \( \alpha = (\alpha_1, \ldots, \alpha_K) \), that account for estimating the within-group tie probabilities. It is required to include a sum-to-zero constraint in the vector \( \alpha \) to make its components identifiable, as described in Subsection 2.3. The effect of belonging to group \( k \) is represented by \( \alpha_k \), for \( k = 1, \ldots, K \). In this case, the structure specified for \( \beta(c_i, c_j) \) is given by:

\[
\beta(c_i, c_j) = \begin{cases} 
\alpha_{c_i}, & \text{if } c_i = c_j \\
-\delta_{c_ic_j}, & \text{if } c_i \neq c_j
\end{cases}, \quad \text{for } i, j = 1, \ldots, n, \ i \neq j,
\]

(3)

where \( \delta_{kl} \) represents a symmetric dissimilarity measure between groups \( k \) and \( l \).

The main difference between LDM and GLDM occurs when two individuals belong to the same group. Now, when \( i \) and \( j \) belong to the same group \( k \), the probability \( \mu_{ij} = \logit^{-1}(\alpha_0 + \beta(c_i, c_j)) \) will depend on \( \alpha_0 + \alpha_k \). Then, \( \logit^{-1}(\alpha_0 + \alpha_k) \) represent the tie probability between two elements belonging to the same group \( k \), for \( k = 1, \ldots, K \). Thus, \( \alpha_k \) is responsible for increase, or decrease, the probability that two individuals \( i \) and \( j \) belonging to the same group share a connection.

Following the Bayesian paradigm, it is assumed that \( \delta_{kl} \sim \text{Gamma}(a_\delta, b_\delta) \), for \( l = 1, \ldots, K \) and \( k < l \), with \( a_\delta = 0.01 \) and \( b_\delta = 0.01 \) to ensure a non-informative prior. For \( \alpha_k \) it is assumed a \( N(0, \sigma_\alpha^2) \) distribution, for \( k = 1, \ldots, K \), with \( \sigma_\alpha^2 = 9 \), that is a low-informative
prior due to the logit link function structure. Under this formulation, the posterior distribution
is given by:

\[
p(\alpha_0, \alpha, \delta, C \mid y) \propto \prod_{j=1}^{n} \prod_{i \neq j} \left( \frac{\exp(y_{ij}(\alpha_0 + \alpha_{c_i}))}{\exp(\alpha_0 + \alpha_{c_j}) + 1} \right)^{I(c_i = c_j)} \left( \frac{\exp(y_{ij}(\alpha_0 - \delta_{c_i,c_j}))}{\exp(\alpha_0 - \delta_{c_i,c_j}) + 1} \right)^{I(c_i \neq c_j)} \times \prod_{k=1}^{K} C_{ik} \times \prod_{l=1}^{K} \prod_{k<l} \delta_{kl}^{-1} \exp(-\delta_{kl}) \times \prod_{k=1}^{K} \exp\left(-\frac{1}{2\sigma^2_k}\right),
\]

where \( I(\cdot) \) represents the indicator function, which assumes 1, if the condition \( \cdot \) is satisfied,
and 0 otherwise. For more information about the methods adopted to infer about GLDM’s
parameters, see Subsection 2.3.

2.2 Latent positions models

In this subsection, the proposed models aim to place groups in space based on the concept of
latent distances (Hoff, Raftery and Handcock, 2002). In this formulation, it is considered that
the tie probability depends on the latent distance between groups’ positions. This approach
provides a spatial representation of groups through the set of positions in the latent space,
denoted by \( \mathbf{a} \). The LDM and GLDM depend on dissimilarities \( \delta_{kl} \), which can be interpreted
as the distance between groups \( k \) and \( l \) since dissimilarities are positive numbers, for \( l = 1, \ldots, K \) and \( k < l \). Thus, in this context, LDM and GLDM can be viewed as simplified
versions of the models based on latent positions, that will be presented in Subsections 2.2.1
and 2.2.2.

2.2.1 Latent positions model with no within-group variation

The latent positions model with no within-group variation (LPM) can be considered as an extension of the LDM to spatially place groups through a set of latent positions. The structure of this model depends on a set of \( K \times d \) latent positions (the position of each group is a point in the \( d \)-dimensional latent space) that account for estimating the between-group tie probabilities. This model’s construction considers that individuals belonging to the same group have a constant connection probability, regardless of the group that they are part of, denying variation in within-group tie probabilities. In this case, the structure specified for \( \beta(c_i, c_j) \) is given by:

\[
\beta(c_i, c_j) = -|\mathbf{a}_{c_i} - \mathbf{a}_{c_j}|, \quad \text{for } i, j = 1, \ldots, n, \ i \neq j,
\]

where \( |\cdot| \) is a distance measure satisfying the triangular inequality, \( \mathbf{a}_k = (a_{k1}, \ldots, a_{kd})' \) is the vector containing the position of class \( k \), for \( k = 1, \ldots, K \), in the latent space \( D \subset \mathbb{R}^d \) and \( \mathbf{a} = (\mathbf{a}_1, \ldots, \mathbf{a}_K)' \). Note that, if \( c_i = c_j \), the probability \( \mu_{ij} = \text{logit}^{-1}(\alpha_0 + \beta(c_i, c_j)) \) will only depend on \( \alpha_0 \) since \( |\mathbf{a}_{c_i} - \mathbf{a}_{c_j}| = 0 \). Then, \( \text{logit}^{-1}(\alpha_0) \) can be interpreted as the probability of a tie between two elements belonging to the same group. Moreover, the greater the distance between two groups’ position, the smaller the probability of a tie between two elements from different groups.

Following the Bayesian paradigm, it is assumed that \( a_{ki} \sim N(0, \sigma^2_a) \), for \( k = 1, \ldots, K \) and \( l = 1, \ldots, d \), with \( \sigma^2_a = 25 \), that is a low-informative prior due to the logit link function structure. Under this formulation, the posterior distribution is given by:

\[
p(\alpha_0, \mathbf{a}, C \mid y) \propto \left[ \prod_{j=1}^{n} \prod_{i \neq j} \frac{\exp(y_{ij}(\alpha_0 - |\mathbf{a}_{c_i} - \mathbf{a}_{c_j}|))}{\exp(\alpha_0 - |\mathbf{a}_{c_i} - \mathbf{a}_{c_j}|) + 1} \right] \times \left[ \prod_{k=1}^{K} C_{ik} \right] \times \prod_{k=1}^{K} \exp\left(-\frac{1}{2\sigma^2_a} \mathbf{a}_k' \mathbf{a}_k \right) \times \exp\left(-\frac{1}{2\sigma^2_{\alpha_0}} \alpha^2_0 \right).
\]
For more information about the methods adopted to infer about LPM’s parameters, see Subsection 2.3.

2.2.2 Latent positions model with within-group variation The latent positions model with within-group variation (GLPM) is an extension of the LPM, to represent variation among within-group tie probabilities, accommodating different levels of homogeneity inside each group; and can be considered as an extension of the GLDM to spatially place groups through a set of latent positions. The structure of this model depends on a set of \( K \times d \) latent positions that account for estimating the between-group tie probabilities, and on a set of \( K \) parameters, given in the vector \( \alpha = (\alpha_1, \ldots, \alpha_K) \), that account for estimating the within-group tie probabilities. It is required to include a sum-to-zero constraint in the vector \( \alpha \) to make its components identifiable, as described in Subsection 2.3. The effect of belonging to group \( k \) is represented by \( \alpha_k \), for \( k = 1, \ldots, K \). In this case, the structure specified for \( \beta(c_i, c_j) \) is given by:

\[
\beta(c_i, c_j) = \begin{cases} 
\alpha_{c_i}, & \text{if } c_i = c_j \\
-|a_{c_i} - a_{c_j}|, & \text{if } c_i \neq c_j
\end{cases}, \quad \text{for } i, j = 1, \ldots, n, \ i \neq j,
\]

where \(| \cdot |\) is a distance measure satisfying the triangular inequality, \( a_k = (a_{k1}, \ldots, a_{kd})' \) is the vector containing the position of class \( k \), for \( k = 1, \ldots, K \), in the latent space \( D \subset \mathbb{R}^d \) and \( a = (a_1, \ldots, a_K)' \).

Similarly to LDM and GLDM models, the main difference between LPM and GLPM occurs when two individuals belong to the same group. Thus, the interpretation of the parameters is similar to that of models based on latent dissimilarities.

Following the Bayesian paradigm, it is assumed that \( a_{kl} \sim N(0, \sigma^2_\alpha) \), for \( k = 1, \ldots, K \) and \( l = 1, \ldots, d \), with \( \sigma^2_\alpha = 25 \). For \( \alpha_k \) it is assumed a \( N(0, \sigma^2_\alpha) \) distribution, for \( k = 1, \ldots, K \), with \( \sigma^2_\alpha = 9 \). Both distributions can be considered as low-informative priors due to the logit link function structure. Under this formulation, the posterior distribution is given by:

\[
p(\alpha_0, \alpha, a, C | y) \propto \left[ \prod_{j=1}^{n} \left( \frac{\exp(y_{ij}(\alpha_0 + \alpha_{c_i}))}{\exp(\alpha_0 + \alpha_{c_i}) + 1} \right)^{I(c_i = c_j)} \left( \frac{\exp(y_{ij}(\alpha_0 - |a_{c_i} - a_{c_j}|))}{\exp(\alpha_0 - |a_{c_i} - a_{c_j}|) + 1} \right)^{I(c_i \neq c_j)} \right] \times \left[ \prod_{k=1}^{K} \prod_{l=1}^{d} p_{ik} \right] \times \left[ \prod_{k=1}^{K} \exp \left(-\frac{1}{2\sigma^2_\alpha} a_k' a_k \right) \right] \times \left[ \prod_{k=0}^{K} \exp \left(-\frac{1}{2\sigma^2_\alpha} \alpha_k' \alpha_k \right) \right],
\]

where \( I(\cdot) \) represents the indicator function, which assumes 1, if the condition \( \cdot \) is satisfied, and 0 otherwise. For more information about the methods adopted to infer about GLPM’s parameters, see Subsection 2.3.

2.3 Model inference

We perform inference via MCMC to obtain samples from the resulting posterior distribution of each proposed model. The full conditional posterior distributions and the methods adopted to sample from each of them are detailed in Appendix A. To describe the inference procedure for the latent dissimilarity models presented in Section 2.1, we show the estimation procedure for the GLDM model, which consists of the following steps:

1. Initialize the counter \( j = 2 \) and set initial values for the parameters of the model: \( \alpha_0 \), \( \alpha \), \( \delta \) and \( C \);

2. Update the model parameters \( \alpha_0, \alpha_1, \ldots, \alpha_K \) from their conditional distributions;

3. Set \( \alpha_1 = -\sum_{k=2}^{K} \alpha_k \), according to the identification procedure described ahead;
(4) Update the model parameters $\delta$ and $C$ from their conditional distributions;

(5) Increment the counter $j$ to $j + 1$ and iterate from (2).

For the LDM model, the estimation procedure’s step (2) does not update $\alpha_k$ parameters, for $k = 1, \ldots, K$, since this model does not consider these parameters. The inference procedure for the latent positions models, presented in Section 2.2, follows the same steps of the GLDM estimation procedure, updating $a$ instead of $\delta$.

Some parameters of the model demand specific inference strategies to become identifiable. To identify the $K$ parameters responsible for estimating the within-group tie probabilities, $\alpha_1, \ldots, \alpha_K$, it is required to include a sum-to-zero constraint in the vector $\alpha$. This restriction is added to the GLDM and GLPM models. To estimate the set of positions in the latent space $a$, which provide a spatial representation of groups in the LPM and GLPM models, it is necessary to eliminate translation, rotation, and reflection effects in the configuration of the latent positions $a$ via procrustes transformation as in Hoff, Raftery and Handcock (2002). Henceforward, without loss of generality, it will be assumed that the distance measure is the Euclidean distance.

A usual problem in social network blockmodels is identifying the group labels for each partition of the $n$ individuals obtained while performing inference. More specifically, given a partition obtained, one can not directly determine which label is assigned to each group, since all $K!$ label permutations produce the same likelihood function value. To deal with this identification problem, known as label switching, we use a deterministic online procedure of classification, performed during the MCMC method, based on a classification method described in Celeux, Hurn and Robert (2000).

More specifically, the MCMC method is divided into two stages. In the first stage, we obtain reference centers and dispersion measures for a set of parameters $\xi$ from the first $m_1$ posterior distribution samples. From $j = m_1 + 1$ onwards, the permutation $p$ of labels of $C^{(j)}$ that produces a set of parameters $\xi^{(j)}_p$ closest to $\xi$ is chosen, and $C^{(j)}$ is switched. After identifying the optimal label configuration, the reference measures associated with $\xi$ are then updated, and this procedure is repeated in each of the next $m_2$ iterations.

This method is sensitive to the choice of $m_1$ since high values can be affected by label switching, and low values may not be enough to ensure a good estimate of $\xi$. To improve the label switching detection, we used the median and the median absolute deviation as robust measures of centrality and dispersion, respectively, and we also set $m_1 = 50$ after discarding the first 50 samples in MCMC.

Note that, since we have only $K!$ label permutations, this procedure has a low computational cost for a small number of groups, and is considered a reasonable and simple solution to the label switching problem in the simulations performed. Other methods based in order constraints or more sophisticated methods of classification are also available to deal with the switching label problem (see Snijders and Nowicki, 1997; Celeux, 1998; Rodriguez and Walker, 2014; Papastamoulis and Iliopoulos, 2010, 2013, among others), although they were not considered in this paper.

3 Simulated studies

In order to assess the effectiveness of the class of models defined in (1) under different scenarios, and to highlight the proposed models’ main features, artificial data was generated considering distinct configurations for each one of the four scenarios considered, which were created by varying $K$, $n$, $\alpha_0$, $\alpha$ and $\delta$ values. In particular, this study aimed to assess the goodness-of-fit of each model considering two network sizes, three number of latent classes,
Table 1 Different configurations of \( K, n, \alpha_0, \alpha \) and \( \delta \) parameters considered to generate artificial data.

<table>
<thead>
<tr>
<th>( K )</th>
<th>( n )</th>
<th>( n_1 )</th>
<th>( n_2 )</th>
<th>( n_3 )</th>
<th>( n_4 )</th>
<th>( \alpha_0 )</th>
<th>( \alpha_1 )</th>
<th>( \alpha_2 )</th>
<th>( \alpha_3 )</th>
<th>( \alpha_4 )</th>
<th>( \delta_{12} )</th>
<th>( \delta_{13} )</th>
<th>( \delta_{23} )</th>
<th>( \delta_{14} )</th>
<th>( \delta_{24} )</th>
<th>( \delta_{34} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>36</td>
<td>18</td>
<td>18</td>
<td>-</td>
<td>-</td>
<td>0.50</td>
<td>-2.00</td>
<td>2.00</td>
<td>-</td>
<td>-</td>
<td>1.50</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>-</td>
<td>0.50</td>
<td>-2.00</td>
<td>2.00</td>
<td>0.00</td>
<td>2.00</td>
<td>1.50</td>
<td>2.50</td>
<td>3.50</td>
<td>3.50</td>
<td>3.50</td>
<td>3.50</td>
</tr>
<tr>
<td>3</td>
<td>120</td>
<td>60</td>
<td>40</td>
<td>20</td>
<td>-</td>
<td>0.50</td>
<td>-0.25</td>
<td>0.00</td>
<td>0.25</td>
<td>0.25</td>
<td>1.00</td>
<td>1.50</td>
<td>2.00</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>120</td>
<td>45</td>
<td>35</td>
<td>25</td>
<td>15</td>
<td>0.50</td>
<td>-0.50</td>
<td>-0.25</td>
<td>0.25</td>
<td>0.50</td>
<td>1.00</td>
<td>1.25</td>
<td>1.50</td>
<td>2.00</td>
<td>2.25</td>
<td>3.00</td>
</tr>
</tbody>
</table>

Table 2 Different configurations of \( a \) considered to generate artificial data according to \( K, n \) and \( d \).

<table>
<thead>
<tr>
<th>( K )</th>
<th>( d )</th>
<th>( n )</th>
<th>( a_{11} )</th>
<th>( a_{21} )</th>
<th>( a_{31} )</th>
<th>( a_{41} )</th>
<th>( a_{12} )</th>
<th>( a_{22} )</th>
<th>( a_{32} )</th>
<th>( a_{42} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>36</td>
<td>-0.75</td>
<td>0.75</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>36</td>
<td>-0.42</td>
<td>-1.54</td>
<td>1.96</td>
<td>-</td>
<td>0.59</td>
<td>-0.40</td>
<td>-0.19</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>120</td>
<td>-0.25</td>
<td>-0.87</td>
<td>1.12</td>
<td>-</td>
<td>0.46</td>
<td>-0.32</td>
<td>-0.14</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>120</td>
<td>0.03</td>
<td>0.16</td>
<td>1.40</td>
<td>-1.59</td>
<td>0.13</td>
<td>-1.04</td>
<td>0.54</td>
<td>0.37</td>
</tr>
</tbody>
</table>

and two different configurations of within and between-group probabilities — a scenario with more distinguishable parameters and a scenario with closer parameters. Table 1 presents the different configurations of \( K, n, \alpha_0, \alpha \) and \( \delta \) considered in this study.

According to Table 1, the scenario with more distinguishable parameters is associated with the larger population, with \( n = 120 \) individuals, and the scenario with closer parameters, with \( n = 36 \) individuals. It is known that the closer the values of the parameters, the more difficult they are to be estimated. Thus, we decided to evaluate the estimation of the closest parameters in populations with a larger number of individuals. Note that, the latent positions \( a \) were not directly presented in Table 1. Despite that, to make the models more comparable, the dissimilarities \( \delta \) of the LDM and GLDM models are associated with the distances between the latent positions \( a \) of the LPM and GLPM models, i.e., we set latent positions \( a \) whose Euclidean distances among them were compatible with the dissimilarities \( \delta \). Table 2 presents the different configurations of \( a \) considered to generate artificial data in the LPM and GLPM models.

Remark that the structure of LPM and GLPM models depends on a set of \( K \times d \) latent positions. Thus, Table 2 also supplies the dimension \( d \) of the latent space. In cases where the number of latent groups was equal to 2, we set \( d = 1 \), i.e., the groups’ positions were projected in a one-dimensional latent space; and for \( K > 2 \), we set \( d = 2 \), i.e., the groups’ positions were projected in a two-dimensional latent space. Both LPM and GLPM models were fitted to the resulting datasets in order to evaluate the models’ effectiveness to recover the true latent positions \( a \) through inference. In contrast, the LDM and GLDM models were fitted to these data in order to evaluate the models’ ability to recover the respective resulting dissimilarities \( \delta \).

The MCMC algorithm was implemented in the \( R \) programming language, v. 3.6.1 (Team, 2019), as presented in Subsection 2.3. For each scenario and fitted model, we let each chain run for 70,000 iterations, discarded the first 20,000 as burn-in, and stored every 5th iteration to obtain 10,000 independent samples. The chains’ convergence was checked through the time-series plot evaluation for each model and parameters considered in this study. As a result, Table 3 displays the posterior mean and the respective 95% credibility interval for the models’ parameters, obtained by fitting the proposed models to artificial data generated under each scenario.

The results shown in Table 3 point that all models were able to accurately recover the real parameters’ value since the posterior mean of the considered parameters is closer to true values, which are contemplated by its respective 95% credibility intervals. In LPM and GLPM
models, the dissimilarities $\delta$ were set as the distances between the latent positions $a$. Despite the variability of posterior distributions — reflected in the range of the credibility intervals — point estimates are distinguishable and become more accurate when the size of the network increases. Moreover, as expected, scenarios with more distinguishable parameters provided better results than scenarios with closer parameters. Also, there are no significant differences between results obtained by models based on latent distances (LPM and GLPM) and models based on dissimilarities (LDM and GLDM).

Table 4 presents the posterior mean and the respective 95% credibility interval for the latent positions $a$. In this aspect, all posterior means and respective 95% credibility intervals contemplate the real values of $a$. Moreover, these intervals’ respective ranges seem smaller in the scenario with closer parameters. Also, there are no significant differences between results obtained by models based on latent distances (LPM and GLPM) and models based on dissimilarities (LDM and GLDM).

Table 4 Posterior mean and respective 95% credibility interval for $a_0$, $a$ and $\delta$, obtained for the fits of the models, considering each scenario.

<table>
<thead>
<tr>
<th>$K$</th>
<th>$n$</th>
<th>Parameter</th>
<th>LDM</th>
<th>LPM</th>
<th>GLDM</th>
<th>GLPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>36</td>
<td>$a_0 = 0.50$</td>
<td>0.48(0.31, 0.65)</td>
<td>0.49(0.33, 0.64)</td>
<td>0.55(0.27, 0.86)</td>
<td>0.55(0.29, 0.83)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_1 = -2.00$</td>
<td>-</td>
<td>-</td>
<td>-1.88(-2.29, -1.48)</td>
<td>-1.88(-2.28, -1.49)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_2 = 2.00$</td>
<td>-</td>
<td>-</td>
<td>1.88(1.48, 2.29)</td>
<td>1.88(1.49, 2.28)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{12} = 1.50$</td>
<td>1.55(1.30, 1.80)</td>
<td>1.56(1.33, 1.80)</td>
<td>1.62(1.29, 1.97)</td>
<td>1.62(1.31, 1.95)</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>$a_0 = 0.50$</td>
<td>0.52(0.31, 0.72)</td>
<td>0.53(0.33, 0.73)</td>
<td>0.65(0.33, 1.05)</td>
<td>0.68(0.35, 1.08)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_1 = -2.00$</td>
<td>-</td>
<td>-</td>
<td>-1.76(-2.30, -1.24)</td>
<td>-1.80(-2.34, -1.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_2 = 2.00$</td>
<td>-</td>
<td>-</td>
<td>-0.08(-0.59, 0.43)</td>
<td>-0.12(-0.67, 0.43)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{12} = 1.50$</td>
<td>1.60(1.28, 1.81)</td>
<td>1.61(1.28, 1.81)</td>
<td>1.71(1.30, 2.18)</td>
<td>1.74(1.31, 2.18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{13} = 2.50$</td>
<td>2.53(1.24, 3.22)</td>
<td>2.54(1.25, 3.22)</td>
<td>2.68(2.21, 3.22)</td>
<td>2.70(2.21, 3.22)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{23} = 3.50$</td>
<td>3.55(3.01, 4.37)</td>
<td>3.66(3.08, 4.37)</td>
<td>3.67(3.04, 4.37)</td>
<td>3.82(3.14, 4.37)</td>
</tr>
<tr>
<td>3</td>
<td>120</td>
<td>$a_0 = 0.50$</td>
<td>0.51(0.45, 0.56)</td>
<td>0.51(0.45, 0.56)</td>
<td>0.53(0.42, 0.66)</td>
<td>0.52(0.38, 0.66)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_1 = -0.25$</td>
<td>-</td>
<td>-</td>
<td>-0.23(-0.38, -0.10)</td>
<td>-0.23(-0.38, -0.09)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_2 = 0.00$</td>
<td>-</td>
<td>-</td>
<td>-0.07(-0.23, 0.06)</td>
<td>-0.07(-0.23, 0.06)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_3 = 0.25$</td>
<td>-</td>
<td>-</td>
<td>0.29(0.08, 0.58)</td>
<td>0.29(0.07, 0.58)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{12} = 1.00$</td>
<td>0.98(0.90, 1.23)</td>
<td>0.98(0.90, 1.23)</td>
<td>1.01(0.89, 1.23)</td>
<td>1.01(0.86, 1.23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{13} = 1.50$</td>
<td>1.52(1.42, 1.72)</td>
<td>1.53(1.42, 1.72)</td>
<td>1.53(1.24, 1.72)</td>
<td>1.52(1.11, 1.72)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{23} = 2.00$</td>
<td>2.01(1.87, 2.20)</td>
<td>2.02(1.88, 2.20)</td>
<td>2.02(1.79, 2.20)</td>
<td>1.99(1.41, 2.20)</td>
</tr>
<tr>
<td>4</td>
<td>120</td>
<td>$a_0 = 0.50$</td>
<td>0.48(0.41, 0.54)</td>
<td>0.49(0.43, 0.55)</td>
<td>0.43(0.29, 0.57)</td>
<td>0.48(0.34, 0.63)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_1 = -0.50$</td>
<td>-</td>
<td>-</td>
<td>-0.55(-0.72, -0.39)</td>
<td>-0.61(-0.78, -0.44)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_2 = -0.25$</td>
<td>-</td>
<td>-</td>
<td>-0.33(-0.53, 0.06)</td>
<td>-0.37(-0.56, 0.06)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_3 = 0.25$</td>
<td>-</td>
<td>-</td>
<td>0.14(-0.09, 0.33)</td>
<td>0.09(-0.13, 0.30)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_4 = 0.50$</td>
<td>-</td>
<td>-</td>
<td>0.74(0.34, 1.24)</td>
<td>0.89(0.43, 1.38)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{12} = 1.00$</td>
<td>0.98(0.88, 1.07)</td>
<td>1.14(1.05, 1.24)</td>
<td>0.92(0.77, 1.09)</td>
<td>1.13(0.98, 1.29)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{13} = 1.25$</td>
<td>1.21(1.11, 1.32)</td>
<td>1.44(1.33, 1.54)</td>
<td>1.18(1.01, 1.37)</td>
<td>1.43(1.28, 1.58)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{23} = 1.50$</td>
<td>1.54(1.41, 1.68)</td>
<td>1.68(1.56, 1.80)</td>
<td>1.48(1.19, 1.69)</td>
<td>1.67(1.50, 1.84)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{14} = 2.00$</td>
<td>2.00(1.86, 2.13)</td>
<td>2.01(1.87, 2.15)</td>
<td>1.94(1.74, 2.14)</td>
<td>2.00(1.80, 2.20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{24} = 2.25$</td>
<td>2.24(2.05, 2.42)</td>
<td>2.24(2.07, 2.43)</td>
<td>2.18(1.90, 2.41)</td>
<td>2.23(1.98, 2.48)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_{34} = 3.00$</td>
<td>3.12(2.83, 3.42)</td>
<td>3.03(2.83, 3.24)</td>
<td>3.06(2.74, 3.39)</td>
<td>3.01(2.69, 3.34)</td>
</tr>
</tbody>
</table>
edge about the number of groups and the presence of variation in within-group tie probabilities. However, in practical situations, this previous information may not be available, and an incorrect model specification can lead to poor results.

Finally, let \( \mu \) be the set of probabilities that two individuals \( i \) and \( j \) share a connection, for \( i, j = 1, \ldots, n, i \neq j \). The mean squared error associated with the estimation of the set of probabilities \( \mu \) was smaller than 0.002 in all simulations, evidencing the model’s capability to recover the tie probabilities between any two elements from the network. Differences in the mean square error between models based on latent positions and models based on latent dissimilarities were negligible in the simulations performed.

### 4 Case studies

In this section, two case studies are presented for evaluating the performance of the proposed models. In the first study, a detailed analysis shows the ability of models to both cluster and project groups in latent spaces. The second dataset presents an example of a small network, where the groups’ cohesion probabilities are obtained, showing that the proposed models have desirable features of blockmodels in social networks.

#### 4.1 Case study I

In this section, we analyze the performance of each of the four proposed models using a real dataset. The data contains information about the relationship of trust among eighteen monks in an American monastery. Sampson (1968), which introduced this case research, suggested the existence of four factions in the monastery: Loyal Opposition (LO), Young Turks (YT), Outcasts (O) and Waverers (W). The Loyal Opposition represents the oldest members of the
monastery, while the Young Turks are the newest ones. The Outcasts are the members of the monastery that were not accepted in any of the previous groups, and the Waverers are the members who did not take sides. These monks are usually labeled according to their names, or according to a sequence of numbers from 1 to 18.

The sociomatrix $Y$ with the relationship of trust between the eighteen monks was obtained from \textit{latentnet} (Krivitsky and Handcock, 2008), an \texttt{R} package utilized to fit and evaluate the statistical latent position and cluster models for networks. Sampson (1968) performed three studies over time, which resulted in a dataset of a time-aggregated network. In this dataset, a tie from monk A to monk B exists if A nominated B as one of his three (or four, in case of a draw) best friends at any of the three-time points. Based on Figure 1, which presents the sociomatrix of monks data (a black pixel represents a link between two monks), it is possible to distinguish block structures of the social relationship between each monk faction. Note that, analyzing the communication within-blocks (black pixels), the Loyal Opposition group presents fewer connections between its members than other groups.

Several authors that analyzed this dataset point that there were three prominent latent groups in the monastery (see Breiger, Boorman and Arabie, 1975; White, Boorman and Breiger, 1976; Airoldi et al., 2008; Handcock, Raftery and Tantrum, 2007; Hoff, Raftery and Handcock, 2002; Schweinberger and Handcock, 2015, among others). Since the main objective of this analysis is to determine the latent social structure within the monastery, the LDM, LPM, GLDM and GLPM models were fitted to this data considering the existence of $K = 2$ and $K = 3$ latent classes. For each scenario and number of latent groups, we let each chain run for 60,000 iterations, discarded the first 10,000 as burn-in, and stored every 5th iteration to obtain 10,000 independent samples.

Table 5 displays the log-likelihood function evaluated in the posterior mean of $\mu$, i.e., $\log p(y|\mu)$, and the Expected Akaike Information Criterion (EAIC) as a measure of goodness-of-fit. Note that the greater the number of latent positions considered, the higher the log-likelihood value, indicating that the models considering $K = 3$ performed better than those with $K = 2$ latent groups. These results corroborate with the EAIC values since they are notoriously lower for $K = 3$. Among the models based on latent positions, the ones with within-group variation have a higher log-likelihood value. Furthermore, among the models based on dissimilarities, those with variation within the group also have a higher log-likelihood value. This conclusion is analogous to the models based on latent positions. Finally, there is evidence of a better fitting of the GLPM model with $K = 3$ latent classes.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{Monks sociomatrix $Y$ blocked according to the different factions (black pixels indicate the ties between monks).}
\end{figure}
Table 5  Log-likelihood function evaluated in the posterior mean of $\mu$ and EAIC, obtained for the fits of the models to monks’ dataset.

<table>
<thead>
<tr>
<th>K</th>
<th>LDM</th>
<th>LPM</th>
<th>GLDM</th>
<th>GLPM</th>
<th>LDM</th>
<th>LPM</th>
<th>GLDM</th>
<th>GLPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-156.58</td>
<td>-156.38</td>
<td>-154.19</td>
<td>-156.08</td>
<td>323.94</td>
<td>325.63</td>
<td>327.71</td>
<td>328.96</td>
</tr>
<tr>
<td>3</td>
<td>-136.75</td>
<td>-136.64</td>
<td>-134.02</td>
<td>-133.33</td>
<td>286.33</td>
<td>291.28</td>
<td>297.61</td>
<td>304.94</td>
</tr>
</tbody>
</table>

Figure 2  Posterior mean of the tie probabilities between monks obtained from LDM, LPM, GLDM, and GLPM models (bordered pixels represent true ties and the shades of gray represent the estimated tie probabilities).

Figure 2 presents the posterior mean of tie probabilities $\mu$ between monks obtained from each model. Note that models assuming the existence of $K = 3$ latent groups performed better than models assuming $K = 2$ since they were able to distinguish Outcasts and Waverers from Young Turks and Loyal Opposition groups. Moreover, based on the range of values (represented by the shades of gray), it is possible to point out that models considering variation in the within-group tie probabilities produce more extreme values, i.e., values closer to 0 and 1. These results highlight the ability of GLDM and GLPM models to properly predict ties in comparison to LDM and LPM models, considering this dataset. Finally, the results corroborate the previous evidence, given in Table 5, of a better fitting of the models with different within-group tie probabilities and $K = 3$ latent classes.

In the present study, the latent groups to which the monks belong to are not previously known, except for Sampson (1968)’s factions suggestion. Thus, it is possible to analyze how the proposed models classify the monastery’s monks into 2 and 3 latent groups in comparison with Sampson (1968)’ classification. Table 6 presents the clustering obtained from the posterior mode of $C$, considering each number of latent positions, for all proposed models. Note that all models were able to distinguish the Young Turks and the Loyal Opposition groups satisfactorily. Moreover, for $K = 2$ latent groups, the LDM and LPM models grouped the monks in the same way, as well as the GLDM and LPM models, and for $K = 3$, all four proposed models equally grouped monks.

According to Table 6, considering $K = 2$, LDM and LPM grouped the Young Turks, the Outcasts, and Amand (from Waverers) in the same class, and grouped the Loyal Opposition and the remaining Waverers in another class. On the other hand, GLDM and GLPM grouped all Outcasts, Waverers, and Loyal Opposition groups in the same class. In the $K = 3$ latent groups’ configuration, all models led to the same partition composed of Young Turks in a class, the Outcasts, and Amand (from Waverers) in another class, and the Loyal Opposition
Table 6  Monks grouping obtained from the posterior mode of $C$ for the fits of the models to monks’ dataset, considering 2 and 3 latent groups.

<table>
<thead>
<tr>
<th>Monks</th>
<th>Sampson (1968)</th>
<th>$K = 2$</th>
<th>$K = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LDM</td>
<td>LPM</td>
<td>GLDM</td>
</tr>
<tr>
<td>Albert</td>
<td>YT</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Boniface</td>
<td>YT</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Gregory</td>
<td>YT</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hugh</td>
<td>YT</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>John Bosco</td>
<td>YT</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mark</td>
<td>YT</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Winfrid</td>
<td>YT</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Amand</td>
<td>W</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Romuald</td>
<td>W</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Victor</td>
<td>W</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Basil</td>
<td>O</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Elias</td>
<td>O</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Simplicius</td>
<td>O</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ambrose</td>
<td>LO</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Berthold</td>
<td>LO</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Bonaventure</td>
<td>LO</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Louis</td>
<td>LO</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Peter</td>
<td>LO</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 7  Monks’ factions suggested by [0] Sampson (1968) and latent groups obtained according to [1] Krivitsky et al. (2009), White, Boorman and Breiger (1976), Handcock, Raftery and Tantrum (2007) and Schweinberger and Handcock (2015), [2] Airoldi et al. (2008) and [3] Breiger, Boorman and Arabie (1975) in comparison with the results obtained from the GLPM model, and posterior probabilities of belonging to each class.

<table>
<thead>
<tr>
<th>i</th>
<th>Monks</th>
<th>[0]</th>
<th>[1]</th>
<th>[2]</th>
<th>[3]</th>
<th>GLPM</th>
<th>$P(c_i = 1)$</th>
<th>$P(c_i = 2)$</th>
<th>$P(c_i = 3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Albert</td>
<td>YT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.723</td>
<td>0.146</td>
<td>0.131</td>
</tr>
<tr>
<td>2</td>
<td>Bonifaceman</td>
<td>YT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.730</td>
<td>0.155</td>
<td>0.115</td>
</tr>
<tr>
<td>3</td>
<td>Gregory</td>
<td>YT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.732</td>
<td>0.155</td>
<td>0.113</td>
</tr>
<tr>
<td>4</td>
<td>Hugh</td>
<td>YT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.729</td>
<td>0.154</td>
<td>0.117</td>
</tr>
<tr>
<td>5</td>
<td>John Bosco</td>
<td>YT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.731</td>
<td>0.153</td>
<td>0.117</td>
</tr>
<tr>
<td>6</td>
<td>Mark</td>
<td>YT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.731</td>
<td>0.152</td>
<td>0.117</td>
</tr>
<tr>
<td>7</td>
<td>Winfrid</td>
<td>YT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.731</td>
<td>0.156</td>
<td>0.113</td>
</tr>
<tr>
<td>8</td>
<td>Amand</td>
<td>W</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0.125</td>
<td>0.740</td>
<td>0.135</td>
</tr>
<tr>
<td>9</td>
<td>Romuald</td>
<td>W</td>
<td>3</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>0.146</td>
<td>0.083</td>
<td>0.771</td>
</tr>
<tr>
<td>10</td>
<td>Victor</td>
<td>W</td>
<td>3</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>0.145</td>
<td>0.079</td>
<td>0.776</td>
</tr>
<tr>
<td>11</td>
<td>Basil</td>
<td>O</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.126</td>
<td>0.756</td>
<td>0.118</td>
</tr>
<tr>
<td>12</td>
<td>Elias</td>
<td>O</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.125</td>
<td>0.759</td>
<td>0.116</td>
</tr>
<tr>
<td>13</td>
<td>Simplicius</td>
<td>O</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.125</td>
<td>0.759</td>
<td>0.116</td>
</tr>
<tr>
<td>14</td>
<td>Ambrose</td>
<td>LO</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.141</td>
<td>0.077</td>
<td>0.782</td>
</tr>
<tr>
<td>15</td>
<td>Berthold</td>
<td>LO</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.138</td>
<td>0.079</td>
<td>0.783</td>
</tr>
<tr>
<td>16</td>
<td>Bonaventure</td>
<td>LO</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.144</td>
<td>0.079</td>
<td>0.776</td>
</tr>
<tr>
<td>17</td>
<td>Louis</td>
<td>LO</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.143</td>
<td>0.076</td>
<td>0.781</td>
</tr>
<tr>
<td>18</td>
<td>Peter</td>
<td>LO</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.145</td>
<td>0.084</td>
<td>0.771</td>
</tr>
</tbody>
</table>

and the remaining Waverers in a third class. These results are in accordance with White, Boorman and Breiger (1976), Handcock, Raftery and Tantrum (2007), Krivitsky et al. (2009) and Schweinberger and Handcock (2015) as showed in Table 7, which also presents the posterior probabilities of belonging to each class for each monk, for the GLPM model.

According to Table 7, the main difference between the partitions found in the literature involves the classification of monk Amand, probably due to his ambiguous positioning in the monastery (see Breiger, Boorman and Arabie, 1975; Airoldi et al., 2008, among others). Moreover, for all monks, the modal class presents the highest probability, indicating a small
Figure 3 Visual display of monks clustering in the latent space obtained from the GLPM model (circles radius are proportional to $1 - \text{within-group tie probability}$).

probability of belonging to any other cluster class. These high probabilities are also following several authors in literature (see Handcock, Raftery and Tantrum, 2007; Krivitsky et al., 2009, among others).

The visual display of clustering in the latent space for the GLPM model can be seen in Figure 3. This display presents both the different group cohesion levels — represented by distinct circles with a radius proportional to $1 - \text{within-group tie probability}$ — and a satisfactory notion of the distance between groups — represented by the Euclidean distance between estimated latent positions.

The estimated within and between-group tie probabilities for all models are presented in Table 8. As expected, the between-groups probabilities are considerably lower than within-groups probabilities, and the three groups presented different cohesion levels (see White, Boorman and Breiger, 1976; Handcock, Raftery and Tantrum, 2007; Airoldi et al., 2008, among others). As observed in the sociomatrix $Y$ (see Figure 1), the within-group tie probabilities of the class containing the Loyal Opposition members is lower than the one that holds the Young Turks members, for all four proposed models. In the models with $K = 3$ latent classes, all models presented higher within-group tie probabilities for the group containing the Outcasts members, as already pointed out by several authors in the literature (see Sampson, 1968; Breiger, Boorman and Arabie, 1975; Handcock, Raftery and Tantrum, 2007, among others).

To obtain the latent group positions for the LDM and GLDM models, the multidimensional scaling method was performed using the dissimilarity samples $\delta$ as input. The chosen approach follows the analysis of Mardia (1978) and was performed via `cmdscale` function in R software. The posterior distributions of the latent group positions for all four proposed models are presented in Figure 4.

The latent group positions seem similar for all four proposed models, despite the models based on latent positions present more variability than the models based on latent dissimilarities. The resulting latent positions from models that allow variation in the within-group tie probabilities are quite similar to those obtained from models that do not allow this variation. Note that all models were successful in distinguishing the latent positions from all groups.

4.2 Case study II

In this section, we analyze a famous dataset initially introduced by Zachary (1977) and extensively studied in the literature of fission in small groups. The objective here is to recover
Alternative models to position social network groups

Table 8  Posterior mean of average level (AP), between-group (BP) and within-group probabilities (WP) obtained from the fits of the LDM, LPM, GLDM, and GLPM models for $K = 2$ and $K = 3$. The within-group probabilities are ordered according to labels 1, 2, and 3, and the between-group probabilities are ordered according to pairs (1,2), (1,3) and (2,3).

<table>
<thead>
<tr>
<th>K</th>
<th>LDM</th>
<th>LPM</th>
<th>GLDM</th>
<th>GLPM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AP</td>
<td>BP</td>
<td>WP</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.465 (0.386, 0.548)</td>
<td>0.118 (0.068, 0.179)</td>
<td>0.678 (0.484, 0.859)</td>
<td>0.675 (0.383, 0.885)</td>
</tr>
<tr>
<td>3</td>
<td>0.646 (0.545, 0.744)</td>
<td>0.144 (0.083, 0.220)</td>
<td>0.718 (0.529, 0.880)</td>
<td>0.666 (0.272, 0.929)</td>
</tr>
</tbody>
</table>

Figure 4  Posterior distribution of the latent group positions for all four proposed models.

groups of students from a karate club, in which a factional division led to fission into two groups (Mr. Hi’s group and John’s group). This network can be represented by a nondirectional graph with its respective symmetric sociomatrix, whose edges represent the friendship between two students related to outside club activities.

The four proposed models were fitted to this dataset, aiming to identify the faction membership (before split) and looking for evidence of differences in within-group tie probabilities, which can represent factions’ cohesion levels.

Table 9 presents measures of goodness-of-fit obtained for the fits of the proposed models to karate’s dataset considering $K = 2$ factions. The GLDM model performed better than the other models on both criteria despite the subtle differences between them. Although omitted here, the partition found by the GLDM model agrees 100% with the ethnographic faction membership presented in Zachary (1977).

The estimated within and between-group tie probabilities, considering the GLDM model, are shown in Table 10. As pointed out in the literature, both factions present different cohesion levels. John’s faction within-group probability was lower than Mr. Hi’s (0.217 against
Table 9 Log-likelihood function evaluated in the posterior mean of $\mu$ and EAIC, obtained for the fits of the models to karate’s dataset considering $K = 2$.

<table>
<thead>
<tr>
<th></th>
<th>log $L(\hat{\mu})$</th>
<th></th>
<th>EAIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDM</td>
<td>-868.17</td>
<td>LDM</td>
<td>1726.74</td>
</tr>
<tr>
<td>LPM</td>
<td>-871.95</td>
<td>LPM</td>
<td>1736.44</td>
</tr>
<tr>
<td>GLDM</td>
<td>-857.23</td>
<td>GLDM</td>
<td>1726.50</td>
</tr>
<tr>
<td>GLPM</td>
<td>-883.57</td>
<td>GLPM</td>
<td>1828.05</td>
</tr>
</tbody>
</table>

0.248) due to weak supporters being part of this faction, e.g., the student labeled as number 9 in the original study (Cavallari et al., 2019).

Table 10 Posterior mean of average level (AP), between-group (BP) and within-group probabilities (WP) obtained from the fit of the GLDM model for $K = 2$.

<table>
<thead>
<tr>
<th></th>
<th>AP</th>
<th>BP</th>
<th>WP - Mr. Hi</th>
<th>WP - John</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.248 (0.212, 0.286)</td>
<td>0.037 (0.023, 0.055)</td>
<td>0.284 (0.228, 0.366)</td>
<td>0.217 (0.156, 0.269)</td>
</tr>
</tbody>
</table>

Unfortunately, the GLDM model results obtained for $K = 2$ groups do not allow one to represent groups in latent spaces via Multidimensional Scaling. On the other hand, the results showed that there are different levels of cohesion in factions. In addition, the GLDM model seems to have desirable features of blockmodels in social networks, despite its simplicity.

5 Discussion

In this paper, we proposed an alternative class of models for social networks to represent groups in an unobservable space. This class of models encompasses approaches based on latent dissimilarities — the LDM and GLDM models — and latent positions — the LPM and GLPM models —, allowing the researcher to visualize the latent groups of the social network; quantify tie probabilities for individuals belonging to the same or different groups, without representing individuals in the latent space; and use a random version of multidimensional scaling based on samples from the posterior distribution of a set of unknown dissimilarities provided by data. Both approaches can be used in undirected or directed graphs, i.e., in cases where the sociomatrix $Y$ is symmetric or non-symmetric, respectively. Remark that, if the sociomatrix is symmetrical, there will be less available data information to estimate the model parameters since $Y$ will be an upper triangular matrix.

All four models proposed in this paper are related and suitable for classification problems in social networks. More specifically, the models where $\beta$ are functions of dissimilarities $\delta$ — the LDM and GLDM models —, can be view as simplified versions of the models based on latent positions $a$ — the LPM and GLPM models. Regarding a scenario where the positions are represented in a two-dimensional latent space, the LDM and GLDM models can be more suitable when the number of groups is small. However, if $K > 5$, these models will contain more parameters associated with between-group probabilities than its associated models based on latent positions $a$ and, consequently, it may not be the most suitable choice (see Figure 5).

The LDM and GLDM models aim to represent groups in space based on the concept of latent dissimilarities. In the two-stage proposed methodology, the set of dissimilarities between groups is estimated in the first step, and the samples drawn from the posterior distribution of the dissimilarities are used as input in a multidimensional scaling technique, in the second stage. Once there is a sample of the posterior distribution of the set of dissimilarities available, it is possible to take into account the uncertainty associated with the multidimensional
Alternative models to position social network groups

scaling result. Although there are several methods for modeling data through Multidimensional Scaling, the main differences between existing approaches have not been addressed in this paper.

Under the GLDM and GLPM models, individuals belonging to different groups have the same probabilities of tie to each other. Despite that, all four proposed models presented in this work can not properly represent transitivity and homophily on attributes. To represent these features, it would be required to model relational data at the individual level, or to include individuals’ information. The individual-level modeling was not an aim of this paper, but rather to properly represent the relationships between groups in latent spaces.

The Multinomial distribution assigned for $C_i$ — responsible for indicating the group of the $i$–th individual — depends on the hyperparameters $p_i$, which represent prior probabilities of belonging to each group for the $i$–th individual, for $i = 1, \ldots, n$. For all cases analysed in this paper, we set $p_i = K^{-1}1_K$, for $i = 1, \ldots, n$, where $1_K$ represents the $K$-dimensional vector of ones. Alternative approaches include assigning a prior distribution to these quantities, e.g., a Dirichlet distribution. However, simulated examples modeling $p_i$ through the Dirichlet distribution, for $i = 1, \ldots, n$, have shown that inference about these parameters is quite sensitive to the choice of its hyperparameters.

Regarding the inference procedure, MCMC methods were used to obtain samples from the resulting posterior distribution from the proposed models. In this context, different distributions can be used to generate proposals in MCMC. Normal proposal distributions were assigned for $\alpha_0$, $\alpha$ and $\alpha$ parameters. For the set of dissimilarities $\delta$, both Gamma and Truncated Normal distributions were examined, and the posterior results were quite similar. As described in Appendix A, the posterior full conditional of $C$ has an analytical closed-form. Despite that, to obtain a more efficient sampling scheme, the Metropolis-Hastings step was used to update $C$’s chain instead of using Gibbs Sampling in MCMC (Besag et al., 1995; Gamerman and Lopes, 2006). Two approaches were considered to generate proposal values for $C$: Multinomial proposals based on prior probabilities $p$; and proposals based on mutations of some elements of the configuration $C$ drawn in the previous iteration. The main advantage of the latter approach is that it allows better control of the MCMC’s acceptance rate.

The performance of the proposed models was assessed through artificial studies and a case study. In both analyses, $Y_{ij}$ is modeled according to the Bernoulli distribution. Nevertheless, in practical situations, Binomial, Negative Binomial, and Poisson distributions — with their associated canonical link functions — are usually used to model relational data. In the proposed methodology, these other ways to model the relationship between individuals can be naturally employed.
Some parameters of the model required specific inference strategies to become identifiable, as well as the latent groups’ labels. In all proposed models, a deterministic classification procedure was performed during the MCMC method to overcome the label switching. Despite the satisfactory results obtained through this approach, other procedures can be successfully performed. As an alternative approach, one could previously estimate \( C \) based on a traditional cluster method, e.g., the K-means algorithm (MacQueen, 1967) — or using a crude estimate of \( C \) —, and then, use this estimate as a reference configuration to \( C \). To perform this, in each MCMC iteration, the configuration of labels \( C^{(j)} \) drawn in the \( j \)–th iteration would be switched to the configuration closest to the reference configuration of \( C \). The main limitation of this approach relies on the quality of the solution obtained to build the reference configuration, which can lead to unsatisfactory partitions of the \( n \) network individuals.

Since the labels associated with each latent group are arbitrary, there are \( K! \) ways to represent identical groupings of the \( n \) individuals in the network. Thus, in all performed studies, the labels associated with the latent groups were relabeled to become comparable to each other. In particular, since the groups’ labels are previously known when we are dealing with artificial data, the latent groups were relabeled according to the true partition.

Simulations with artificial data have shown that the proposed models can successfully distinguish latent groups, providing a 100% hit rate in all artificial simulations conducted. The performance of all proposed models improves as the number of individuals in the network increases and as the groups are balanced. Nevertheless, the computational cost also increases with the number of individuals and groups in the network. For example, the GLDM and GLPM models with \( n = 120 \) individuals and \( K = 3 \) latent groups take about 2 hours to run 60,000 iterations in MCMC on a computer with an Intel Core i7 processor, 2.4GHz and 16 GB of RAM.

In case study I, the performance of each of the four proposed models was analyzed using the Monks’ dataset. The models considering \( K = 3 \) performed better than the ones with \( K = 2 \) latent groups. In particular, there is evidence of a better fitting of the model based on latent positions with different within-group tie probabilities, the GLPM model. Considering \( K = 3 \) latent groups, all models led to the same partition, which follows several authors in literature.

Still in case study I, the performance of each of the four proposed models was analyzed using the Monks’ dataset. The models considering \( K = 3 \) performed better than the ones with \( K = 2 \) latent groups. In particular, there is evidence of a better fitting of the model based on latent positions with different within-group tie probabilities, the GLPM model. Considering \( K = 3 \) latent groups, all models led to the same partition, well distinguishing Young Turks, Loyal Opposition, and Outcasts groups. All four proposed models’ fit to this dataset showed that this class of models is flexible enough to properly clustering monks, and also to quantify between and within-group tie probabilities according to several authors in the literature.

In practical situations, the most common options for the dimension of the latent space are \( d = 1 \) or \( d = 2 \). The projection of networks or groups in latent spaces intends to allow better visualization of the network. Except in cases where interactive 3-D graphics are available, it is hard to achieve this aim considering \( d > 2 \). However, in cases where \( K = 2 \), the GLDM and LDM models do not provide projections of the groups in the latent space due to limitations of the multidimensional scaling technique.

The main findings of this work encourage an extension of the proposed class of models to consider the number of groups as a random variable, i.e., assigning a zero truncated binomial distribution for \( K \). In this case, it will also be necessary to specify a limit \( L \) for the number of classes \( K \) or to control its variation by choosing suitable hyperparameters in its prior distribution.
Finally, obtaining both the latent positions of nodes and groups is a challenging and promising future work. Since the groups have random latent positions, it would be needed to propose a latent set of nodes for each group at each MCMC iteration. In addition, the switching label problem in this situation would need advanced treatment. Sequential approaches could be performed in this case, e.g., methods based on latent configurations obtained from Hoff, Raftery and Handcock (2002)’ model followed by a post-processing technique to estimate latent group positions.

References


Appendix A: Full conditional posterior distributions of the parameters in the proposed models

In this section, we present the general posterior full conditional distributions of the components of the parameter vector $\Theta$ associated with the proposed models. We denote the posterior full conditional of a parameter $\theta$ in $\Theta$ by $p(\theta | \cdots)$. 

References


A.1 Full conditional posterior distribution of $\alpha_0$

The posterior full conditional of $\alpha_0$ is proportional to:

$$p(\alpha_0 \mid \cdots) \propto \prod_{j=1}^{n} \prod_{i \neq j} \mu_{ij} \times \exp \left( -\frac{1}{2\sigma_0^2} \alpha_0^2 \right),$$

which does not have an analytical closed-form. We use the Metropolis-Hastings algorithm with a Normal proposal, whose mean is the current value of the parameter and the variance is fixed at $\sigma_{\alpha_0}^2$, where the term $\sigma_{\alpha_0}^2$ controls the acceptance rate.

A.2 Full conditional posterior distribution of $\alpha_k$, for $k = 2, \ldots, K$

The posterior full conditional of $\alpha_k$, for $k = 2, \ldots, K$, is proportional to:

$$p(\alpha_k \mid \cdots) \propto \prod_{j=1}^{n} \prod_{i \neq j} \mu_{ij}^{\mathbb{I}(c_i = c_j) + \mathbb{I}(c_i = l) + \mathbb{I}(c_j = k)} \times \exp \left( -\frac{1}{2\sigma_k^2} \alpha_k^2 \right),$$

which does not have an analytical closed-form. We use the Metropolis-Hastings algorithm with a Normal proposal, whose mean is the current value of the parameter and the variance is fixed at $\sigma_{\alpha_k}^2$, where the term $\sigma_{\alpha_k}^2$ controls the acceptance rate, for $k = 2, \ldots, K$. After generating $\alpha_2, \ldots, \alpha_K$, we set $\alpha_1 = -\sum_{k=2}^{K} \alpha_k$, according to Subsection 2.3.

A.3 Full conditional posterior distribution of $\delta_{kl}$, for $l = 1, \ldots, K$

The posterior full conditional of $\delta_{kl}$, for $l = 1, \ldots, K$ and $k < l$, is proportional to:

$$p(\delta_{kl} \mid \cdots) \propto \prod_{j=1}^{n} \prod_{i \neq j} \mu_{ij}^{\mathbb{I}(c_i = l) + \mathbb{I}(c_j = k) + \mathbb{I}(c_i = c_j) + \mathbb{I}(c_j = k)} \times \delta_{kl}^{\alpha_{kl} - 1} \exp(-b_\delta \delta_{kl}),$$

which does not have an analytical closed-form. We use the Metropolis-Hastings algorithm with a zero-truncated Normal proposal, whose mean is the current value of the parameter and the variance is fixed at $\sigma_{\delta_{kl}}^2$, where the term $\sigma_{\delta_{kl}}^2$ controls the acceptance rate, for $l = 1, \ldots, K$ and $k < l$.

A.4 Full conditional posterior distribution of $a$

The posterior full conditional of $a$ is proportional to:

$$p(a \mid \cdots) \propto \prod_{j=1}^{n} \prod_{i \neq j} \mu_{ij}^{\mathbb{I}(c_i \neq c_j)} \times \exp \left( -\frac{1}{2\sigma_a^2} a^2 \right),$$

which does not have an analytical closed-form. We use the block Metropolis-Hastings algorithm considering a Normal proposal for each $a_{kd}$, whose mean is the current value of the parameter and the variance is fixed at $\sigma_{a_k}^2$, where the term $\sigma_{a_k}^2$ controls the acceptance rate. Translation, rotation and reflection effects of the proposed $a$ are eliminated via procrustes transformation, according to Subsection 2.3.
A.5 Full conditional posterior distribution of C

The posterior full conditional of C is proportional to:

\[ p(C | \cdots) \propto \left[ \prod_{j=1}^{n} \prod_{i \neq j} \mu_{ij} \right] \times \left[ \prod_{i=1}^{n} \prod_{k=1}^{K} p_{ik}^{C_{ik}} \right], \]

which has an analytical closed-form. Despite that, we use the Metropolis-Hastings algorithm with a proposal distribution that depends on the current value of $C$, $C^*$. We sample $q$ individuals to have their groups changed. Then, $q$ new groups for the sampled individuals are drawn from a Multinomial distribution, supposing the same probability for belonging to each group. The proposed $C$ is $C^*$ with $q$ individuals’ groups changed. The number of individuals to be sampled, $q$, controls the acceptance rate.