Mixed Membership Estimation in Two-Path Partial Correlation Network Model

Siao Xu*
Department of Economics, University of Mannheim

June 11, 2026

Abstract

We develop a two-path partial correlation network model (2PPCN) for panel of time-series data. The model allows the cross-sectional dependence structure to be driven jointly by mixed membership and node-specific sociability. These two latent structures are embedded in the covariance matrix through a joint distribution specified by a Bayesian network and enter the panel model through the innovation process. The resulting partial-correlation network separates sparsity structure from connection intensity through two distinct latent pathways. We propose an algorithm for estimating the mixed-membership structure, prove consistency of the estimator, and document its finite-sample performance in simulation studies. An empirical application to U.S. state-level employment growth rate illustrates how the method can uncover overlapping latent groups in economic panel data.

Keywords: Partial Correlation Network, concentration matrix, Mixed Membership Estimation, Probabilistic Graph, Bayesian Belief Network

*We are very grateful to Carsten Trenkler, Jean-Marie Dufour, Matthew O. Jackson, Mengshan Xu, Evan Munro, Christopher Rothe, Frank Schorfheide, Endong Wang, Amilcar Velez, Bin Peng, Juan Carlos Escanciano, for their very insightful suggestions and discussions. We thank all participants and organizers of 2025 $(EC)^2$ conference in Lugano, 2026 CSW Asia Meeting (Main) in Abu Dhabi, Mannheim Econometrics Seminar, the 12th HKMetrics Workshop in Mannheim, the 45th International Symposium on Forecasting in Peking, 2025 Harbin International Conference on Econometrics and Statistics in Harbin, Time Series AI in Macroeconomics and Finance in Bristol, ENTER Jamboree in Carlos 3, Madrid for their comments, suggestions and feedback. All faults are our own. *Graduate School of Economic and Social Sciences, University of Mannheim. E-mail: siao.xu@students.uni-mannheim.de*

1 Introduction

Latent group structure in panel data has received increasing attention in recent econometric research. One strand of the literature estimates grouped heterogeneity in panel models, allowing parameters to differ across groups while imposing homogeneity within each group, for example in fixed effects or slope coefficients (Bonhomme & Manresa 2015, Su et al. 2016). A second strand recovers group structure from the interdependence patterns of large multivariate systems by imposing community or block structure (Guðmundsson & Brownlees 2021, Guðmundsson 2026, Chen et al. 2023, Zhu et al. 2023). These approaches are powerful, but they typically rely on pure membership: each unit is assigned to exactly one latent group. This restriction may be inappropriate in many economic and social systems, where units often belong to multiple latent groups simultaneously. For example, firms may be linked to several ownership, supply-chain, or industry networks; voters may shift across political parties; individuals may have multiethnic backgrounds; and researchers often span multiple disciplines. We address this limitation by developing a method for estimating mixed-membership structure in large panels of time series through the partial-correlation network induced by the concentration matrix of the underlying multivariate system.

In many real-world networks, vertices exhibit group structure: link formation depends systematically on whether two vertices belong to the same latent group. This empirical regularity is commonly referred to as community structure or clustering. A

widely used model for such networks is the degree-corrected stochastic block model (DC-SBM) (Karrer & Newman 2011). The DC-SBM is a random graph model in which vertices are assigned to latent groups, and the probability of an edge depends on group membership as well as node-specific degree heterogeneity. In the canonical assortative case, vertices in the same group are more likely to be connected than vertices in different groups.

Recent work has introduced community structure into large panels of time series by assuming that the series are partitioned into latent groups and that dependence is stronger within groups than across groups (Brownlees et al. 2022). This is achieved by allowing the partial-correlation structure of the N series to be governed by a latent DC-SBM. While this approach provides a useful framework for modeling grouped dependence in high-dimensional panels, it has two limitations for our purpose. First, each series is restricted to belong exclusively to one group, thereby ruling out overlapping or mixed-membership patterns. Second, the DC-SBM primarily characterizes the sparsity pattern of the partial-correlation network, whereas it does not separately model the intensity of conditional dependence in the system.

Motivated by recent developments in mixed-membership network models (Airoldi et al. 2008, Fan et al. 2022), we introduce a two-path partial correlation network model for large panels of time series. The model separates conditional dependence into two latent pathways. The first pathway determines the sparsity pattern of the partial-correlation network through a latent mixed-membership structure, while the

second pathway determines the magnitude of nonzero partial correlations through an independent latent sociability structure. Rather than imposing latent structure directly on a given partial-correlation network as in Brownlees et al. (2022), we jointly model the partial correlations, network structure, mixed membership, and sociability through a probabilistic graph specified by a Bayesian belief network. This construction embeds the latent structures into the innovation covariance matrix, thereby allowing mixed membership to affect the panel through the innovation process.

To estimate the latent mixed-membership structure, we develop PartialCorr-Mixed-SCORE, a spectral algorithm that extends SCORE (Jin 2015) and Mixed-SCORE (Jin et al. 2024) to high-dimensional panel time-series settings. Unlike the original SCORE framework, which recovers pure- or mixed-membership communities from binary adjacency matrices generated by degree-corrected stochastic block models or mixed-membership network models, our method operates on the sample concentration matrix estimated from large panels. The algorithm first applies eigen-ratio normalization to remove node-specific heterogeneity and sociability effects. It then exploits the simplex geometry of Mixed-SCORE to recover mixed-membership weights from the normalized spectral representation.

We conduct simulation studies to assess the finite-sample performance of the PartialCorr-Mixed-SCORE algorithm. We further establish consistency by adapting the theoretical framework of Rohe et al. (2011). The analysis proceeds in two

steps. First, we show that the algorithm recovers the mixed-membership structure at the population level under the model assumptions. Second, we derive bounds on the deviation between the sample and population mixed membership, thereby quantifying the effect of sampling noise on membership estimation.

We illustrate the empirical relevance of the proposed method using U.S. state-level employment growth data. The same dataset was analyzed by Hamilton & Owyang (2012) to infer posterior distributions of group memberships and by Brownlees et al. (2022) to detect pure-membership groups. We instead apply PartialCorr-Mixed-SCORE to estimate mixed memberships, allowing each state to partially affiliate with multiple latent groups. The resulting membership patterns share some features with earlier findings but also reveal important differences, suggesting that mixed membership provides a complementary view of the latent dependence structure underlying U.S. employment dynamics.

Our paper is related to several strands of the literature. First, it contributes to the body of work on identifying latent group structure in panel data, see e.g., Brownlees et al. (2022), Su et al. (2016), Bonhomme & Manresa (2015), Guðmundsson & Brownlees (2021) and Zhang et al. (2019). Second, it connects to the literature on mixed membership network and estimation, including, among others, Airoldi et al. (2008), Jin et al. (2024), and Zhang et al. (2020). Third, we build upon the probabilistic graphical model and Bayesian belief network, see e.g., Blei et al. (2003), Munro & Ng (2022) and Koller & Friedman (2009). Finally, the paper is also related

to social effect structure in panel data, see e.g., Chen et al. (2022) and Zhu et al. (2019).

The rest of the paper is organized as follows. Section 2 presents the methodology, including the main model, estimation algorithm, theoretical results and discussion. Section 3 reports the results of a simulation study. Section 4 is devoted to the empirical application. Finally, Section 5 concludes. Technical proofs, identification strategy and additional materials are provided in the supplementary materials.

2 Methodology

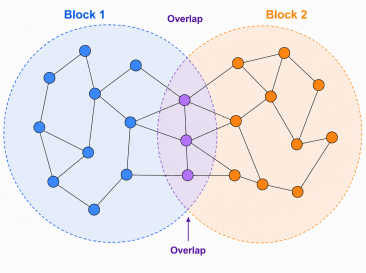
2.1 The Two-path Partial Correlation Network

Consider an (N)-dimensional panel time series $\{Y_t\}_{t=1}^{T}$, where the innovation process satisfies $\epsilon_t \sim \mathcal{N}\Big(0, \Sigma\Big)$. For simplicity, we first consider the case $Y_t = \epsilon_t$. Let $\Lambda = \Sigma^{-1}$ denote the concentration, or precision, matrix (Dempster 1972). Whereas the covariance matrix captures marginal dependence, the concentration matrix characterizes conditional dependence after controlling for the remaining variables. Under Gaussianity, the partial correlation between units i and j is given by $\rho_{i,j} = -\frac{\Lambda_{i,j}}{\sqrt{\Lambda_{i,i}\Lambda_{j,j}}}$, so that zeros in the off-diagonal entries of Λ correspond to conditional independence restrictions (Pourahmadi 2013). We therefore use the concentration matrix, together with the associated partial correlations, to define an undirected conditional dependency graph.

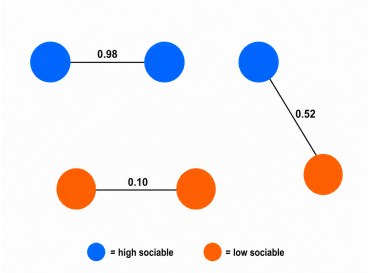
In this paper, we specify a generative model for the concentration matrix through a joint distribution $\Lambda \sim \mathbb{P}\left(\Lambda, A, \Pi, \vec{\varphi}\right)$, where A denotes a latent network, Π denotes a latent mixed-membership matrix, and $\vec{\varphi}$ denotes latent sociability. The joint distribution is represented by a Bayesian belief network that integrates the latent network, mixed-membership structure, and sociability structure into a unified model for the concentration matrix. This construction allows the latent structures to enter the panel through the innovation covariance, or equivalently through its inverse concentration matrix.

The first path of the model introduces mixed membership into the sparsity structure of the partial-correlation network. Let $\Pi \in \mathbb{R}^{N\times K}$ collect the unit-specific mixed-membership vectors $\pi_i \in \mathbb{R}^K$, where N is the number of units and K is the number of latent groups. Each π_i is drawn from a Dirichlet distribution with hyperparameter $\vec{\alpha}_i \in \mathbb{R}^K$, ensuring that its entries are nonnegative and sum to one. The matrix Π is the main object of interest in this paper. Conditional on Π, the adjacency $A \in \mathbb{R}^{N\times N}$ is generated from a Bernoulli model with edge-probability matrix – $\Theta\Pi B\Pi^{'}\Theta$, where $B \in \mathbb{R}^{K\times K}$ is the group-level connectivity matrix and $\Theta \in \mathbb{R}^{N\times N}$ captures node-specific heterogeneity. The diagonal entries of B govern within-group link probabilities, whereas the off-diagonal entries govern between-group link probabilities; in the assortative case, within-group probabilities are larger than between-group probabilities. This path corresponds to a standard mixed-membership network model (Fan et al. 2022, Jin et al. 2024). Figure 1a illustrates that the

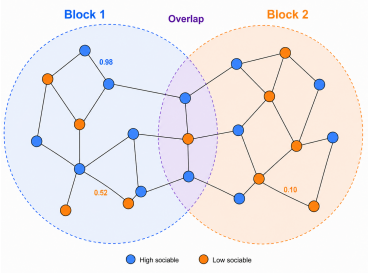
resulting adjacency matrix exhibits overlapping structure when units have mixed memberships. Under pure membership, the model reduces to a block structure without overlapping components.



(a) The first path governs the sparsity pattern. Because units can have mixed memberships, the resulting network exhibits overlap between latent groups.



(b) The second path governs connection intensity. Higher latent sociability leads to stronger interactions, while lower latent sociability leads to weaker interactions.



(c) Together, the two latent structures generate a concentration matrix whose connection weights are governed by vertex-specific latent sociability.

Figure 1: Let the innovation process satisfy $\epsilon_t \sim \mathcal{N}(0, \Lambda^{-1})$, where Λ is the concentration matrix. We model $(\Lambda, A, \Pi, \vec{\varphi})$ through a joint distribution represented by a Bayesian network. Here, A is a latent network, Π is the mixed-membership matrix, and $\vec{\varphi}$ captures unit-specific sociability. The model embeds these latent structures into Λ through two paths: one path governs the sparsity pattern through Π, and the other governs the intensity of nonzero conditional dependencies through $\vec{\varphi}$. The latent structures affect the observed panel $\{Y_t\}_{t=1}^T$ through the innovation process, and our main objective is to estimate Π from the panel data.

The second path introduces node-specific sociability into the intensity of conditional dependence. For each unit i, the sociability parameter $\varphi_i \in \mathbb{R}$ is drawn from a uniform distribution on $(\underline{u}_i, \overline{u}_i)$, where $\underline{u}_i \in \mathbb{R}$ and $\overline{u}_i \in \mathbb{R}$. This parameter measures the latent tendency of unit i to form strong conditional dependence with other units. As illustrated in Figure 1b, the connection intensity between two units is determined

by the interaction of their sociability parameters: pairs of highly sociable units generate larger partial-correlation weights, pairs of low-sociability units generate smaller weights, and mixed pairs generate intermediate weights.

The two paths jointly determine the structure of the concentration matrix Λ. The first path governs the sparsity pattern, while the second path governs the intensity of nonzero interactions. Conditional on the adjacency matrix A, the weight of each existing link between units i and j is drawn from a beta distribution whose parameters are $\varphi_i\varphi_j \in \mathbb{R}$ and $\sigma_{ij} \in \mathbb{R}$. The product $\varphi_i\varphi_j$ measures pairwise latent interaction intensity, and σ_{ij} acts as a pair-specific normalization parameter. Larger values of $\varphi_i\varphi_j$ generate stronger conditional dependence, reflected in larger entries of Λ, while the beta specification restricts the weights to $[0, 1]$. Figure 1c illustrates the concentration matrix generated by the 2PPCN model. The full generative mechanism is represented by a directed acyclic graph (DAG) in Figure 2, and the corresponding sequential distributional specification is given in Definition 1.

Furthermore, following the principles of Bayesian belief networks (Koller & Friedman 2009), the 2PPCN corresponding to Figure 2 can be represented by a sequence of

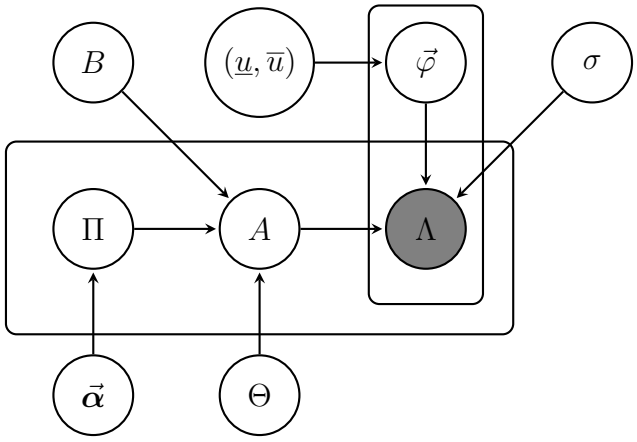


Figure 2: Graphical Representation of the 2PPCN Model. In the diagram, circles outside the plates represent hyper-parameters, while circles inside the plates denote latent or observed variables. Each plate corresponds to one of the two generative paths. Shaded (grey) nodes indicate observed quantities. The adjacency matrix A and the concentration matrix Λ are assumed to share the same sparsity pattern, reflecting the underlying structure of the partial correlation network.

simplified conditional distribution as follows:

$$\begin{aligned}
&\mathbb{P}\Big(\Lambda, A, \Pi, \vec{\varphi} \Big| B, \vec{\boldsymbol{\alpha}}, \Theta, \underline{u}, \overline{u}, \sigma\Big) \\
&= \left\{ \prod_{i,j}^{N} \mathbb{P}\Big(\Lambda_{ij} \Big| A_{ij}, \sigma_{ij}, \varphi_i, \varphi_j\Big) \right\} \\
&\left\{ \prod_{i}^{N} \mathbb{P}\Big(\varphi_i \Big| \underline{u}_i, \overline{u}_i\Big) \right\} \left\{ \prod_{i,j}^{N} \mathbb{P}\Big(A_{i,j} \Big| B, \pi_i, \pi_j, \Theta_i, \Theta_j\Big) \Big\{ \prod_{n}^{N} \mathbb{P}\Big(\pi_n \Big| \vec{\alpha}_n\Big) \Big\} \right\}.
\end{aligned} \tag{1}$$

Each conditional distribution in Equation (1) is specified by a specific conditional distribution in two paths above which is formally defined in Definition 1.

Remark 1 (Conditional independence assumptions)**.** *The factorization in Equation (1) follows from Bayes' rule together with the conditional independence assumptions encoded*

by the proposed probabilistic graphical model. Specifically, we assume that, conditional on its direct parent variables, each component is independent of the remaining variables in the model. These assumptions can be written as

$$\mathbb{P}(\Lambda \mid A, \Pi, \vec{\varphi}, B, \vec{\boldsymbol{\alpha}}, \Theta, \underline{u}, \overline{u}, \sigma) = \mathbb{P}(\Lambda \mid A, \sigma, \vec{\varphi}) = \prod_{i,j=1}^{N} \mathbb{P}(\Lambda_{ij} \mid \varphi_i, \varphi_j, \sigma_{ij}, A_{ij}), \quad (1)$$

$$\mathbb{P}(\vec{\varphi} \mid A, \Pi, B, \vec{\boldsymbol{\alpha}}, \Theta, \underline{u}, \overline{u}, \sigma) = \mathbb{P}(\vec{\varphi} \mid \underline{u}, \overline{u}) = \prod_{n=1}^{N} \mathbb{P}(\varphi_n \mid \underline{u}_n, \overline{u}_n), \quad (2)$$

$$\mathbb{P}(A \mid \Pi, B, \vec{\boldsymbol{\alpha}}, \Theta, \underline{u}, \overline{u}, \sigma) = \mathbb{P}(A \mid \Pi, B, \Theta) = \prod_{i,j=1}^{N} \mathbb{P}(A_{ij} \mid \pi_i, \pi_j, B, \Theta_i, \Theta_j), \quad (3)$$

$$\mathbb{P}(\Pi \mid B, \vec{\boldsymbol{\alpha}}, \Theta, u, \bar{u}, \sigma) = \mathbb{P}(\Pi \mid \vec{\boldsymbol{\alpha}}) = \prod_{n=1}^{N} \mathbb{P}(\pi_n \mid \vec{\alpha}_n). \quad (4)$$

These restrictions imply that Λ depends on the remaining model components only through $(A, \sigma, \vec{\varphi})$, the node-specific sociability parameters $\vec{\varphi}$ are conditionally independent given $(\underline{u}, \overline{u})$, the network A are conditionally independent given (Π, B, Θ), and the mixed membership vectors Π is conditionally independent given the Dirichlet parameter $\vec{\alpha}$. Under these assumptions, the joint distribution admits the factorization stated in Equation (1).

The standard 2PPCN is formally defined as follows:

Definition 1. *(Two-Path Partial Correlation Network). Given the hyper-parameter $\vec{\boldsymbol{\alpha}}$ and $(\underline{u}, \overline{u})$, normalized parameter for $Beta$ distribution σ, node heterogeneity matrix Θ and the community-specific probability matrix B, the concentration matrix is generated*

sequentially in the following way:

$$\pi_i \sim Dirichlet\Big(\vec{\alpha}_i\Big) \ for\ i = 1, \ldots, N.$$

$$A_{ij}\Big|\pi_i, \pi_j, B, \Theta_i, \Theta_j \sim Bernoulli\Big(\Theta_i \pi_i B \pi_j' \Theta_j\Big) \ for\ i, j = 1, \ldots, N.$$

$$\varphi_i \sim Uniform\Big(\underline{u}_i, \overline{u}_i\Big) \ for\ i = 1, \ldots, N.$$

$$\Lambda_{ij}\Big|A_{ij}, \sigma_{ij}, \varphi_i, \varphi_j = \begin{cases} Beta\Big(\varphi_i \varphi_j, \sigma_{ij}\Big) & if\ A_{ij} = 1 \\ 0 & if\ A_{ij} = 0 \end{cases} for\ i, j = 1, \ldots, N.$$

The following Theorem gives us the expectation of Λ defined in Definition 1 corresponding to Equation (1) and Figure 2.

Theorem 1. *Assuming a 2PPCN as defined in Definition 1 corresponding to Figure 2 under the assumption that* Π *is a set of fixed parameters,*

then, we have that:

$$\mathbb{E}(\Lambda) = \Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\Big) B \Big(\Pi' \Theta \boldsymbol{U} \frac{1}{2}\Big),$$

where $\Pi = \{\pi_i\}_{i=1}^N$, $\Theta = diag\Big(\{\Theta_i\}_{i=1}^N\Big)$, *and* $\boldsymbol{U} = diag\Big(\{\underline{u}_i + \overline{u}_i\}_{i=1}^N\Big)$.

Theorem 1 characterizes the population concentration matrix Λ generated by the proposed 2PPCN model. This characterization provides the theoretical foundation for the mixed-membership estimation algorithm developed below.

2.2 Algorithm

In this subsection, we propose a set of algorithms to estimate the mixed membership structure in a panel generated through the 2PPCN model. To this end, we extend the SCORE algorithm of Jin (2015) and its mixed-membership variant, Mixed-SCORE, proposed in Jin et al. (2024). For brevity, we assume that the number of communities is known and treated as a given input. Building on these foundations, we propose a new algorithm tailored to panel data, which we refer to as PartialCorr-mixed-SCORE:

We describe the PartialCorr-mixed-SCORE algorithm. Given panel data $\{Y_t\}_{t=1}^T$ and the number of communities K as inputs, we first estimate the sample covariance matrix via OLS: $\hat{\Sigma} = \frac{\boldsymbol{Y}'\boldsymbol{Y}}{T}$. We then calculate the sample concentration matrix $\hat{\Lambda} = \hat{\Sigma}^{-1}$. We set all diagonal entries of $\hat{\Lambda}$ to 0. Second, we normalize $\hat{\lambda}$ to obtain its symmetric normalized Laplacian: $\hat{L} = \hat{D}^{-1/2}\hat{\Lambda}\hat{D}^{-1/2}$ where $\hat{D}$ is the diagonal matrix of row (or column) sums of $\hat{\Lambda}$. This step follows the standard normalization in spectral graph theory (Von Luxburg 2007). The next step is the spectral decomposition of graph Laplacian. Let $\hat{\mathbf{X}} \in \mathbb{R}^{N\times K}$ contain the eigenvectors of $\hat{L}$ corresponding to its K largest eigenvalues. The rows of $\hat{\mathbf{X}}$ can be interpreted as a rotation of $\frac{1}{2}\hat{\boldsymbol{U}}\hat{\Theta}\hat{\Pi}$ with an unknown rotation matrix $\hat{P} \in \mathbb{R}^{K\times K}$. Next, for each node $i = 1, \ldots, N$, we calculate $\hat{r}_i = \left(\frac{\hat{\mathbf{X}}_{i,2}}{\hat{\mathbf{X}}_{i,1}}, \ldots, \frac{\hat{\mathbf{X}}_{i,K}}{\hat{\mathbf{X}}_{i,1}}\right) \in \mathbb{R}^{K-1}$, where $\hat{\mathbf{X}}_{i,k}$ is the k-th component of the i-th row of $\hat{\mathbf{X}}$. This step projects the mixed-membership structure onto a (K-1)-dimensional simplex, eliminating the influence of $\hat{\boldsymbol{U}}$ and $\hat{\Theta}$ while preserving relative positions of all $\hat{\pi}_i$'s in a K-dimensional simplex. Next, we identify the simplex corner

Algorithm 1 PartialCorr-mixed-SCORE

Input: Panel data $\{X_t\}_{t=1}^T$ and number of communities K.

1: Compute the sample covariance matrix $\hat{\Sigma}$ using OLS and the sample concentration matrix

$$\hat{\Lambda} = \hat{\Sigma}^{-1}.$$

2: Set the diagonal entries of $\hat{\Lambda}$ to zero.

3: Compute the normalized Laplacian

$$\hat{L} = \hat{D}^{-1/2}\hat{\Lambda}\hat{D}^{-1/2},$$

where $\hat{D}$ is the diagonal degree matrix associated with $\hat{\Lambda}$.

4: Let $\hat{X} = [\hat{\xi}_1, \ldots, \hat{\xi}_K] \in \mathbb{R}^{N\times K}$ be the matrix of the first K eigenvectors of $\hat{L}$ corresponding to its K largest eigenvalues.

5: Set $S = \log(N)$ and construct the entry-wise ratio matrix

$$\hat{R}_{ik} = \operatorname{sign}\left(\frac{\hat{X}_{i,k+1}}{\hat{X}_{i,1}}\right)\min\left\{\left|\frac{\hat{X}_{i,k+1}}{\hat{X}_{i,1}}\right|, S\right\},$$

for $1 \le i \le N$ and $1 \le k \le K-1$.

6: Find K corner points $\{\hat{v}_1, \ldots, \hat{v}_K\}$ from the rows of $\hat{R}$ using a corner-search algorithm.

7: For each $i = 1, \ldots, N$, solve

$$\hat{R}_i = \sum_{k=1}^{K} \hat{w}_{ik}\hat{v}_k, \qquad \sum_{k=1}^{K} \hat{w}_{ik} = 1.$$

8: For $1 \le i \le N$ and $1 \le k \le K$, compute

$$\hat{\pi}_{ik}^* = \max\left\{0, \frac{\hat{w}_{ik}}{\hat{b}_{1k}}\right\},$$

where

$$\hat{b}_{1k} = \left[\hat{\lambda}_1 + \hat{v}_k' \operatorname{diag}(\hat{\lambda}_2, \ldots, \hat{\lambda}_K)\hat{v}_k\right]^{-1/2}.$$

9: Normalize

$$\hat{\pi}_i = \frac{\hat{\pi}_i^*}{\|\hat{\pi}_i^*\|_1}, \qquad i = 1, \ldots, N.$$

Output: Estimated mixed membership matrix $\hat{\Pi} \in \mathbb{R}^{N\times K}$.

on which $\hat{r}_i$'s locate. We locate K anchor points $\{\hat{c}_k\}_{k=1}^K$ where $\hat{c}_k \in \mathbb{R}^K$ serve as approximate basis vertices spanning the simplex containing $\hat{r}_i$'s. We next compute weights $\hat{w}_i \in \mathbb{R}^K$ for $i = 1, \ldots, N$ such that $\hat{r}_i = \sum_{k=1}^K \hat{w}_{i,k}\hat{c}_k$ where the weights form a convex combination. Finally, we recover the mixed memberships by normalizing each $\hat{w}_i$ appropriately (using the normalization scheme in mixed-SCORE) to obtain the estimated mixed membership vector $\hat{\pi}_i$.

Our proposed PartiaCorr-mixed-SCORE differs from SCORE and mixed-SCORE in the following ways. First, the input data type is different. SCORE and mixed-SCORE operate on binary adjacency matrices, where entries are either 0 or 1. In contrast, PartialCorr-mixed-SCORE takes panel data as input, which contains continuous observations over time. Second, SCORE is designed for networks whose adjacency matrix A follows a Bernoulli distribution with parameters $\Theta Z B Z' \Theta$, i.e., $A \sim Bernoulli\Big(\Theta Z B Z' \Theta\Big)$ where Z contains pure community labels. Mixed-SCORE generalizes this to a mixed membership setting: $A \sim Bernoulli\Big(\Theta \Pi B \Pi' \Theta\Big)$ where Π contains the mixed membership structure. In contrast, PartialCorr-mixed-SCORE applies to panel data generated from panel generating process such that $Y_t = \epsilon_t$ where $\epsilon_t \sim \mathcal{N}\Big(0, \Lambda^{-1}\Big)$ and $\Lambda \sim 2PPCN\Big(\Lambda, A, \Pi, \varphi | B, \vec{\alpha}, \Theta, \underline{u}, \overline{u}, \sigma\Big)$. Third, the algorithm target is different. SCORE detects the pure membership structure Z from A. Mixed-SCORE estimates the mixed membership structure Π from A. PartialCorr-Mixed-SCORE, on the other hand, estimates the mixed membership structure Π from panel data Y_t, where the Π is enters into the panel through the innovation process.

2.3 Theory

This section establishes theoretical guarantees for the proposed PartialCorr-Mixed-SCORE algorithm when applied to large panels of time series. The main results provide bounds on the mean squared error between the estimated mixed-membership vector $\hat{\pi}_i$ and its population counterpart π_i. Because the model allows mixed membership, the usual misclassification rate used in pure-membership clustering is not directly applicable. Instead, we evaluate estimation accuracy through the MSE between the estimated and population membership vectors. The theoretical analysis proceeds in two steps. First, we show that PartialCorr-Mixed-SCORE recovers the mixed-membership structure at the population level, where the population graph is defined as in Theorem 1. Second, we derive error bounds for the MSE between the membership vectors estimated from the sample concentration matrix and their population counterparts. The sample concentration matrix is obtained from the panel data using OLS residuals. The identification of the model is discussed in Section B.1 of the supplementary material.

The following Lemma establishes that there exists a rotation matrix P such that $\mathbf{X} = \Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\Big)P$, where $\mathbf{X}$ denotes the matrix of eigenvectors corresponding to the top K eigenvalues of $\mathbb{E}\big(\Lambda\big)$. This result is crucial for estimating the latent mixed membership matrix Π in the 2PPCN model. Since Π is unobservable and the observed data consist only of the panel Y_t, we must recover information about Π indirectly. The key insight provided by Lemma 1 is that $\mathbf{X}$, which can be computed from observable

quantities (i.e., the sample partial correlation network), encodes the structure of Π up to an orthogonal transformation. This relationship enables us to estimate Π by analyzing the geometry of $\mathbf{X}$ and applying techniques such as those developed in the SCORE and mixed-SCORE frameworks (Jin 2015, Jin et al. 2024).

Lemma 1. *Let the 2PPCN defined in Definition 1 satisfying Assumption 1 and let* $\lambda_1, \cdots, \lambda_K$ *and* $\mathbf{x}_1, \cdots, \mathbf{x}_K$ *denote the top* K *eigenvalues and corresponding eigenvectors of* $\mathbb{E}\big(\Lambda\big)$ *as given in Theorem 1, ordered from largest to smallest. Then, there exists a* $K \times K$ *non-singular real matrix* P *such that* $\big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\big)P = \mathbf{X}$*, where* $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_K]'$*. Moreover, the matrix* P *is unique once* $\mathbf{X}$ *is fixed.*

Lemma 1 serves as the theoretical foundation for our algorithm. It establishes the critical connection between the observable eigenstructure of the expected concentration matrix $\mathbb{E}\big(\Lambda\big)$ and the latent mixed membership matrix Π. This result also paves the way for establishing the consistency of the proposed estimation algorithm. Building on Lemma 1, we now establish the consistency of the proposed algorithm.

2.3.1 Consistency in Population

The population analysis in this subsection heavily borrows insights from Jin et al. (2024). We begin by proving the Consistency of the PartialCorr-Mixed-SCORE algorithm under the 2PPCN model at the population level. It corresponds to the exact recovery of Π in $\mathbb{E}\big(\Lambda\big)$ under proposed algorithm. The following assumption is required for the proof of Lemma 2.

Assumption 1. *Given 2PPCN, define a $K\times K$ matrix $G = K\left\|\frac{1}{2}(\overline{u}-\underline{u})\theta\right\|^{-2}\left(\frac{1}{4}\Pi'\Theta^2\boldsymbol{U}^2\Pi\right)$, there exists a constant C such that $\Theta_{max} \leq C$, $\boldsymbol{U}_{max} \leq C$, $\|B\|_{max} \leq C$, $\|G\| \leq C$ and $\|G^{-1}\| \leq C$.*

Assumption 1 is to impose upper bounds on the hyper-parameters in the 2PPCN model, which are used for establishing Lemma 2 presented below.

Lemma 2. *Assuming a 2PPCN in Definition 1 satisfying Assumption 1 and 2 with $\mathbb{E}(\Lambda)$ as defined in Theorem 1.*
If B is non-singular and $B\left(\frac{1}{4}\Pi^{'}\Theta^2\boldsymbol{U}^2\Pi\right)$ is irreducible,
then, we have the following statements:

1. *All entries in first orthonormal eigenvector $\mathbf{X}_1$ of $\mathbf{X}$ is strictly positive.*
2. *There exists a K-vertex simplex $\mathcal{S} \in \mathbb{R}^{K-1}$ which is spanned by $v_1, \cdots, v_K$.*
3. *All R_i's locate on $\mathcal{S}$ and pure membership R_i on one of the vertices of $\mathcal{S}$ for $i = 1, \cdots, N$.*
4. *There exists a barycentric coordinate w_i of r_i in $\mathcal{S}$ such that $w_i = \left(\Pi_i \otimes b_1\right)/\|\Pi_i \otimes b_1\|_1$ where $b_{1,k} = \left(\lambda_1 + v_k^{'} diag(\lambda_2, \cdots, \lambda_K)v_k\right)^{-1/2}$, λ_k is k-th eigenvalue of $\mathbb{E}\left(\Lambda\right)$ and $\otimes$ is element-wise multiplication.*

Lemma 2 establishes that the rows of the mixed membership matrix Π lie within a simplicial cone with K supporting rays. Each pure node lies exactly on one of the supporting rays corresponding to its group, while mixed-membership nodes

are located in the interior of the cone. In Jin (2015), the SCORE normalization transforms this simplicial cone into a simplex, providing a direct geometric link between the simplex and the latent structure in Π. Later in Jin et al. (2024), the mixed-SCORE framework leverages this transformation to reconstruct Π from the eigenvector matrix $\mathbf{X}$ of the graph. Lemma 2 is a direct extension to Lemma 2.1 in Jin et al. (2024), which confirms that these geometric and spectral relationships between Π, the simplex, and $\mathbf{X}$ still remain valid under the 2PPCN model. This result is sufficient to show that Π can be recovered under the 2PPCN model by leveraging this relationship, which implies the consistency of the proposed algorithm on population graph $\mathbb{E}\big(\Lambda\big)$. Furthermore, $\mathbb{E}\big(\Lambda\big) = \big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\big)B\big(\Pi'\Theta\boldsymbol{U}\frac{1}{2}\big)$ can be seen as a special case of $\mathbb{E}\big(A\big) = \big(\tilde{\Theta}\Pi\big)B\big(\Pi'\tilde{\Theta}\big)$ in Jin et al. (2024) if we combine $\frac{1}{2}\boldsymbol{U}\Theta$ into $\tilde{\Theta}$. As a result, the proof of the simplex structure remains essentially unchanged.

2.3.2 Bounding the Deviation

In this subsection, we establish a bound on the difference between the estimated mixed membership and its population counterpart. Since the observable object in our setting is the panel data $\{Y_t\}$, rather than the graph itself, and our algorithm is applied directly to the panel, we must impose appropriate distributional and mixing assumptions on the data. For brevity, we assume that $Y_t = \epsilon_t$, then we remark that $\Sigma_Y = \Sigma$. In other words, Y_t and ϵ_t share the covariance structure. Next, we discuss the conditions under strongly mixing time series data. We heavily

borrow insight from Assumption 2 in Brownlees et al. (2022). Specifically, we require conditions that ensure the behavior for isotropic random vectors of the form $\Sigma_Y^{-1/2}Y_t$, where $\Sigma_Y = \mathbb{E}\Big[Y_t Y_t' | \mathcal{G}\Big]$, and $\mathcal{G}$ denotes the underlying random graph. Let $\mathcal{B}_{-\infty}^r$ and $\mathcal{B}_{r+m}^\infty$ be the σ-algebras generated by the sequences $\{\Sigma_Y^{-1/2}Y_t : -\infty \leq t \leq r\}$ and $\{\Sigma_Y^{-1/2}Y_t : r+m \leq t \leq \infty\}$, respectively. The α-mixing coefficient is then defined as $\alpha(m) = \sup_r \sup_{A \in \mathcal{B}_{-\infty}^r, B \in \mathcal{B}_{r+m}^\infty} |\mathbb{P}\Big(A \cap B | \mathcal{G}\Big) - \mathbb{P}\Big(A|\mathcal{G}\Big)\mathbb{P}\Big(B|\mathcal{G}\Big)|$. We impose the following assumptions on the panel data process $\{Y_t\}$:

Assumption 2. *Let $\{Y_t\}$ be a zero-mean stationary process with covariance matrix $\mathbb{E}\Big[Y_t Y_t'\Big]$, we impose several assumptions in the following:*

1. *$\{\Sigma_Y^{-1/2}Y_t\}$ is strongly mixing with coefficient satisfying $\alpha(m) \leq e^{-C_1 m^{\gamma_1}}$ where m is positive integer and γ_1 and C_1 are positive constants.*

2. *For any vector $\mathbf{x}$ satisfying $||\mathbf{x}|| = 1$ and for any $s > 0$, $\mathbb{P}\Big(\Big|\mathbf{x}'\Sigma_Y^{-1/2}Y_t\Big| > s\Big|\mathcal{G}\Big) \leq e^{1-(s/C_2)^{\gamma_2}}$, for all t and C_2 and γ_2 are positive constants.*

3. *$\gamma < 1$ and $1/\gamma = 1/\gamma_1 + 1/\gamma_2$.*

Assumption 2.1 requires that the α-mixing coefficients, which measure temporal dependence, decay sufficiently fast. Assumption 2.2 imposes generalized exponential tails on the distribution of $\Sigma_Y^{-1/2}Y_t$ (Fan et al. 2013). Finally, Assumption 2.3 specifies a bound for γ_1 and γ_2. Together, Assumptions 2.1–2.3 enable the use of the concentration inequality from Merlevède et al. (2011) which is instrumental in proving the following Theorem.

The following theorem establishes that the spectral norm difference between the estimated and the population eigenvector matrix is bounded with high probability. It is a direct extension to Theorem 1 in Brownlees et al. (2022). This result builds upon a bound for the discrepancy between the estimated eigenvectors of sample covariance and its population counterpart. It also serves as a key intermediate step in the proof of Theorem 3.

Theorem 2. *Assuming a 2PPCN in Definition 1 satisfying Assumption 1 and 2 and the panel Y_t generated from 2PPCN satisfying Assumption 3.*
Let $\mathbf{O}$ *be a* $k \times k$ *orthonormal rotation matrix,* $\underline{\Pi} = \arg\min_i \left\|\Pi_i\right\|^2$ *for* $i = 1, \cdots, N$ *and let* $\mathcal{X}$ *denotes eigenvector matrix of the population graph and* $\hat{X}$ *eigenvector matrix of observed graph, both corresponding to the largest* K *eigenvalues.*
Then, we have that:

$$\|\mathbf{O}\hat{X} - \mathcal{X}\| = O_p\left(\sqrt{\frac{log(N)}{\underline{\Pi} N B_N}} + \sqrt{\frac{N}{T}}\right).$$

The result of Theorem 2 is instrumental in establishing Theorem 3, where we aim to bound the deviation of the sample SCORE-R matrix from its population counterpart, where R is the stack of all r_i for $i = 1, \ldots, N$ as discussed above. These representations are derived from implementing SCORE step on sample and population eigenvector matrices, respectively. By leveraging the bounded difference established in Theorem 2, we ensure that the deviation between sample and pop-

ulation SCORE-R matrix remains controlled, providing a foundation for proving Theorem 4.

Theorem 3. *Assuming a 2PPCN in Definition 1 satisfying Assumption 1 and 2 and the panel data Y_t satisfying Assumption 3.*

Let R and $\hat{R}$ respectively denote SCORE-R of $\hat{X}$ and $\mathcal{X}$ as defined in Theorem 2.

Then, we have that:

$$\|\mathbf{L}\hat{R} - R\| = O_p\left(D\sqrt{K}\left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\right)\right),$$

where $D = \dfrac{\|\theta(\overline{u}-\underline{u})\| - \left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\right)}{C\left(\theta(\overline{u}-\underline{u})\right)_{min}}$.

Theorem 3 provides a high-probability bound on the deviation between the population SCORE-R matrix and its sample counterpart. The former is constructed from the population graph implied by the 2PPCN model, while the latter is constructed from the observed panel data.

Theorem 3 provides an intermediate bound for controlling $\|\hat{\Pi} - \Pi\|$. Specifically, it bounds the perturbation of the SCORE point cloud, but it does not by itself imply a membership error bound. To translate this perturbation bound into an error bound for mixed-membership vectors, we also require stability of the inverse map from SCORE coordinates to barycentric coordinates. When the population simplex is well conditioned and the vertex-search step recovers its vertices stably, small perturbations

in the SCORE coordinates lead to small errors in the estimated mixed-membership vectors. Given $r_i = \sum_{k=1}^{K} w_i(k)v_k$, since the membership reconstruction map

$$r_i \mapsto w_i \mapsto \frac{w_i \oslash b_1}{\|w_i \oslash b_1\|_1} = \pi_i, \ where \ \oslash denotes \ element-wise \ division,$$

is Lipschitz when b_1 is bounded away from zero and infinity, the same rate transfer from $\left(\mathbf{L}\hat{R} - R\right)$ to $\left(\hat{\Pi} - \Pi\right)$ with one more bound on $\|\mathbf{L}\hat{v} - v\|$. Lemma 2.2 implies that there exists vertices $v_1, \cdots, v_K \in \mathbb{R}^{K-1}$ such that each population SCORE-R point satisfies

$$r_i = \sum_{k=1}^{K} w_i(k)v_k, \ \ \sum_{k=1}^{K} w_i(k) = 1, \ \ w_i(k) \geq 0.$$

And denote the population simplex matrix and its transformation

$$\mathcal{V} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ v_1 & v_2 & \cdots & v_K \end{pmatrix} \in \mathbb{R}^{K \times K}, \quad \mathcal{V}w_i = \begin{pmatrix} 1 \\ r_i \end{pmatrix}.$$

The following assumptions are standard in Mixed-SCORE-type arguments.

Assumption 3. *(Well-conditioned simplex). There exists a constant $c_{\mathcal{V}} > 0$ such that $\sigma_{min}\left(\mathcal{V}\right) \geq c_{\mathcal{V}}$, where $\sigma_{min}\left(\mathcal{V}\right)$ denotes the smallest singular value of the matrix $\mathcal{V}$.*

Assumption 3 ensures that the map from SCORE coordinates r_i to barycentric coordinates w_i is well conditioned. Geometrically, the assumption rules out degeneracy of the population simplex. For instance, when $K = 3$, it requires the triangle spanned

by the simplex vertices to have non-negligible area, rather than being nearly collinear.

Assumption 4. *(Stable vertex search) . Let $\hat{v}_1, \cdots, \hat{v}_K$ be the estimated vertices produced by the vertex-search step, aligned by* $\mathbf{L}$*. Define $\Delta_V = \max_{1\le k\le K} \|\mathbf{L}\hat{v}_k - v_k\|_2$, there exists a constant $C_V > 0$ such that, with probability tending to one, $\Delta_V \le C_V \max_{1\le i\le N} \|\mathbf{L}\hat{r}_i - r_i\|_2$.*

Assumption 4 controls the error from the vertex-search algorithm. While Theorem 3 bounds the perturbation of the SCORE-R matrix, an additional condition is needed to ensure that the estimated simplex vertices are close to the population vertices. Assumption 4 provides this link by requiring the vertex search error, $\max_k \|\mathbf{L}\hat{v}_k - v_k\|$, to be controlled by the SCORE-coordinate error, $\max_i \|\mathbf{L}\hat{r}_i - r_i\|$. In this sense, the assumption imposes stability of the vertex-search step.

Assumption 5. *(Stable scaling vector). There exist constants $0 < c_b < C_b < \infty$ such that $c_b \le \min_k b_{1,k} \le \max_k b_{1,k} \le C_b$. Moreover, $\|\hat{b}_1 - b_1\|_\infty = O_p\left(D\sqrt{K}\left(\sqrt{\frac{log(N)}{\Pi N B_N}} + \sqrt{\frac{N}{T}}\right)\right)$ with D as defined in Theorem 3.*

Assumption 5 imposes lower and upper bounds on b_1. These bounds rule out degeneracy in the normalization step, ensuring that the map from barycentric coordinates w_i to mixed-membership weights π_i is Lipschitz continuous and therefore stable.

Then, we have the following Theorem:

Theorem 4. *Assuming a 2PPCN in Definition 1 satisfying Assumption 1, 2, 4 and 5*

and the panel data Y_t satisfying Assumption 3.

Let π_i and $\hat{\pi}_i$ respectively denote population mixed membership structure and estimated mixed membership structure, then, we have that:

$$\frac{1}{N}\sum_{i=1}^{N}\|\hat{\pi}_i - \pi_i\|^2 = O_p\left(D^2K\left(\frac{log(N)}{\underline{\Pi}NB_N} + \frac{N}{T}\right)\right),$$

where $D = \dfrac{\|\theta(\overline{u}-\underline{u})\| - \left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\right)}{C\left(\theta(\overline{u}-\underline{u})\right)_{min}}$.

Theorem 4 provides an bound on $\frac{1}{N}\sum_{i=1}^{N}\|\hat{\pi}_i - \pi_i\|^2$, for the proposed PartialCorr-Mixed-SCORE algorithm. The result requires N not to be too large relative to T, and the number of groups K to remain moderate.

2.4 Discussion

The condition $T > N$ is imposed to ensure that the sample covariance matrix can be inverted directly. When $T < N$, $\hat{\Sigma}$ can still be computed but is singular, and hence the unregularized estimator $\hat{\Sigma}^{-1}$ is not well defined. Thus, the restriction $T > N$ reflects the particular first-stage estimator used to construct the sample concentration matrix, rather than an inherent limitation of the PartialCorr-Mixed-SCORE framework. Extending the procedure to high-dimensional settings with $T < N$ is possible by replacing $\hat{\Sigma}^{-1}$ with a regularized precision matrix estimator, such as the graphical Lasso, CLIME, nodewise Lasso, or a shrinkage-based estimator.

Throughout the theoretical analysis, we impose the simplifying assumption $Y_t = \epsilon_t$. This assumption implies that the covariance matrix of the observed process is identical to the covariance matrix of the innovations. In a more general dynamic specification, the innovation would be defined as $\epsilon_t = Y_t - \mathbb{E}(Y_t \mid \mathcal{F}_{t-1})$, so that the conditional mean must be estimated before constructing the sample covariance or concentration matrix. Incorporating the resulting first-stage estimation error into the theoretical analysis is nontrivial. For this reason, we restrict attention to the baseline case $Y_t = \epsilon_t$ in the present paper.

Finally, the theoretical bound highlights the role of the number of latent groups K. Theorem 4 shows that the upper bound on $\frac{1}{N}\sum_{i=1}^{N} \|\hat{\pi}_i - \pi_i\|^2$ is increasing in K, implying that the estimator becomes less stable as the latent group structure becomes more complex. This dependence on K represents an important limitation of the proposed algorithm. Improving the robustness of PartialCorr-Mixed-SCORE in settings with many latent groups is therefore a useful direction for future research.

3 Simulation

3.1 Set-Up

In this section, we conduct a simulation study to examine the finite-sample performance of the proposed PartialCorr-mixed-SCORE algorithm. The simulation setup is as follows: The number of latent communities is fixed at K=3. The community in-

teraction probability matrix B is specified such that $b_{i,i} = b_{j,j}$ and $b_{i,j} = b_{j,i}$ for all i, j, reflecting symmetric intra- and inter-group interaction probabilities. We then consider three parameter settings for $(b_{i,i}, b_{i,j})$: $(0.25.0.01), (0.25, 0.05)$ and $(0.5, 0.01)$. The total number of series is set to $N = 120$, with the number of pure-type series (i.e., those entirely affiliated with a single community) denoted by N_0. Then, obviously, $N - N_0$ is the number of mixed membership series (i.e., those affiliated with more than one community). We vary N_0 across five settings: $(30, 50, 80, 100, 120)$. The influence parameters $(\underline{u}_i, \overline{u}_i)$ are fixed at $[0.3, 1.0]$ for all $i = 1, \ldots, N$. The node-specific heterogeneity parameters Θ_i for $i = 1, \ldots, N$ are linearly spaced from 0.9 to 1.0 in increments of 0.01. In this simulation, the mixed membership matrix Π is assumed to be taken as given. For pure nodes, membership vectors $\hat{\pi}_i$ are set to one of the canonical basis vectors: $[1, 0, 0]$, $[0, 1, 0]$ or $[0, 0, 1]$. For mixed nodes, we consider two configurations: $[0.5, 0.5, 0]$ and $[0, 0.5, 0.5]$. The concentration matrix Λ is sampled according to the 2PPCN model using the hyper-parameters described above.

To induce the mixed membership structure in the panel data, we employ a single-factor model. Specifically, we generate Y_t according to: $Y_t = \phi_1 Y_{t-1} + \epsilon_t$, where ϕ_1 is drawn independently from a Uniform distribution $U(0.3, 1.0)$. The innovation term ϵ_t is sampled from a multivariate normal distribution $\mathcal{N}(0, \Sigma)$ (Fan et al. 2016), where the concentration matrix Λ is the inverse of the covariance matrix Σ generated via the 2PPCN model, as previously described. In this case, we have $Y_t \sim \mathcal{N}(\phi_1 Y_{t-1}, \Sigma)$.

We repeat the experiment 500 times to assess the algorithm's finite-sample behavior. The estimation accuracy of the mixed membership vectors is evaluated using the average ℓ_2 error across all nodes: $1/N \sum_{i=1}^{N} \|\hat{\pi}_i - \pi_i\|^2$, with the reported value being the average over the 500 replications.

3.2 Main Result

In this subsection, we report the results of the simulated study.

The results from applying the PartialCorr-mixed-SCORE algorithm to panel data generated under the 2PPCN model are summarized in Table 1. Overall, the algorithm performs well in finite samples. Estimation accuracy improves when the within-group pair-specific probabilities $(b_{i,i}, b_{j,j})$ are large relative to the between-group probabilities $(b_{i,j}, b_{j,i})$, as shown across the rows of Table 1.

We also observe that precision improves as the number of time periods T increases. As the ratio $b_{i,i}b_{j,j}/b_{i,j}b_{j,i}$ increases, the underlying network exhibits stronger block structure, leading to more accurate estimation of mixed memberships. Moreover, the proportion of mixed-membership nodes affects estimation performance. As this proportion increases, the recovery becomes more difficult. For a fixed total number of series N, a higher number of pure-membership series leads to more distinct block formation and thus better estimation accuracy.

These findings consistently reinforce the conclusion that estimation quality improves

with clearer block structure and greater data availability. For conciseness, all graphical illustrations are reported in Section B.2 of the supplementary material.

Table 1: **Estimation errors of PartialCorr-Mixed-SCORE.** Each entry reports the average estimation error over 500 simulation replications, using the metric $\frac{1}{N}\sum_{i=1}^{N}\|\hat{\pi}_i - \pi_i\|^2$, where $\hat{\pi}_i$ denotes the estimated mixed membership vector and π_i is the true membership vector for i.

$N = 120$						
N_0	$b_{11}, b_{22}/b_{12}, b_{21}$	$T = 100$	$T = 300$	$T = 500$	$T = 800$	$T = 1000$
30	$0.25/0.01$	0.190	0.168	0.158	0.150	0.145
	$0.50/0.01$	0.188	0.150	0.122	0.104	0.100
	$0.25/0.05$	0.191	0.172	0.169	0.164	0.162
50	$0.25/0.01$	0.214	0.187	0.170	0.154	0.142
	$0.50/0.01$	0.211	0.151	0.102	0.068	0.056
	$0.25/0.05$	0.215	0.195	0.189	0.182	0.177
80	$0.25/0.01$	0.252	0.195	0.154	0.103	0.090
	$0.50/0.01$	0.247	0.115	0.045	0.026	0.022
	$0.25/0.05$	0.253	0.221	0.203	0.181	0.170
100	$0.25/0.01$	0.275	0.200	0.124	0.069	0.056
	$0.50/0.01$	0.271	0.082	0.031	0.019	0.016
	$0.25/0.05$	0.278	0.235	0.202	0.165	0.147
120	$0.25/0.01$	0.300	0.186	0.089	0.048	0.040
	$0.50/0.01$	0.294	0.057	0.024	0.016	0.013
	$0.25/0.05$	0.302	0.242	0.188	0.128	0.108

4 Application

4.1 Learning the Mixed Memberships of U.S. States

We apply the proposed methodology to U.S. state-level payroll employment growth rates. The data consist of seasonally adjusted, annualized quarter-to-quarter growth rates in payroll employment from 1956 to 2007. Following Hamilton & Owyang (2012) and Brownlees et al. (2022), we exclude Alaska and Hawaii, yielding a panel with $N = 48$ states and $T = 207$ quarterly observations. The data are obtained from the U.S. Bureau of Labor Statistics. Our objective is to estimate the mixed-membership structure across states and compare the resulting latent affiliations with the group structures reported in earlier studies.

Hamilton & Owyang (2012) study business-cycle synchronicity across U.S. states using this dataset, while Brownlees et al. (2022) use it to detect pure-membership groups in panel data. We instead use the same panel to estimate latent mixed-membership structure. To contextualize our empirical findings, we briefly summarize the methodologies and main results of these two related studies.

Hamilton & Owyang (2012) apply a cross-validation procedure to estimate the number of groups and find evidence of three distinct communities in the same dataset. Following their result, we treat the number of groups as given and do not implement a separate procedure for estimating it. Methodologically, Hamilton & Owyang (2012) propose a Markov-switching model in which U.S. states are grouped according to

the timing of their business cycles. Using Bayesian inference, including Markov chain Monte Carlo (MCMC), they estimate the posterior distribution over group assignments. By contrast, Brownlees et al. (2022) assume that the partial-correlation matrix follows a stochastic block model structure and apply spectral clustering to recover discrete group memberships. Unlike the posterior-based approach of Hamilton & Owyang (2012), the method of Brownlees et al. (2022) produces point estimates of the community structure.

Table 2 reports the results from applying PartialCorr-Mixed-SCORE to the full sample. Following Hamilton & Owyang (2012) and Brownlees et al. (2022), we fix the number of communities at three. Hamilton & Owyang (2012) identify three groups of states based on business-cycle synchronicity: oil-producing, manufacturing, and financial states. Our results are broadly consistent with this classification but allow states to have partial affiliations with multiple groups. Based on the estimated membership weights, we label the three latent groups ex post as: (1) financial and other states, (2) oil-producing and agricultural states, and (3) manufacturing states. The labels are assigned using states with pure or nearly pure memberships as benchmarks, and mixed-membership states are interpreted according to their relative weights across these benchmark groups.

Several patterns emerge from the estimated memberships. First, most states exhibit mixed memberships, suggesting that state-level employment dynamics are associated with multiple latent economic groups rather than a single group. This

pattern is consistent with the diversified structure of state economies. Second, East and West Coast states tend to have higher membership weights in the group we label as financial and service oriented, consistent with the importance of finance, technology, and professional services in these regions. Third, several Midwestern and Southern states, including Texas and Louisiana, show strong affiliations with the oil-producing and agricultural group. Maine also displays a non-negligible loading on this group, which may reflect the role of resource-based industries in its employment dynamics. Finally, traditional manufacturing centers such as Indiana and Michigan have substantial membership in the manufacturing group, while also displaying meaningful affiliation with the financial and service group. This mixed pattern is consistent with gradual structural change in these regions from industrial production toward service-oriented economic activity. For conciseness, the full estimation results are reported in Section B.3 of the supplementary material.

Table 2: Mixed Membership Estimates for U.S. States

States	Financial and Others	Oil-Producing and Agricultural	Manufacturing
Massachusetts	0.736	0	0.264
New Jersey	0.777	0	0.223
New York	0.524	0	0.476
California	0.848	0.032	0.120
Texas	0.062	0.600	0.339
Florida	0.112	0.732	0.155
Missouri	0.129	0.871	0
Maine	0.046	0.954	0
Louisiana	0.247	0.480	0.273
Michigan	0.602	0	0.398
Indiana	0.340	0	0.660

5 Conclusion

We introduce a class of partial-correlation network models in which the concentration matrix is modeled as a random matrix generated from a joint distribution specified by a probabilistic graphical model. The framework can be embedded into general panel data-generating processes through the innovation structure. We develop PartialCorr-Mixed-SCORE, a spectral algorithm for estimating the latent mixed-membership structure from panels generated by this class of models. Consistency is established when N and T are large, N/T is sufficiently small, and the number of latent groups K remains moderate. An empirical application studies mixed affiliations of U.S. states with latent industry-related groups.

6 Disclosure statement

The authors declare that no conflicts of interest exists.

References

Airoldi, E. M., Blei, D. M., Fienberg, S. E. & Xing, E. P. (2008), 'Mixed membership stochastic blockmodels', Journal of machine learning research **9**(65), 1981–2014.

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003), 'Latent dirichlet allocation', Journal of machine Learning research **3**(Jan), 993–1022.

Bonhomme, S. & Manresa, E. (2015), 'Grouped patterns of heterogeneity in panel data', Econometrica **83**(3), 1147–1184.

Brownlees, C., Guðmundsson, G. S. & Lugosi, G. (2022), 'Community detection in partial correlation network models', Journal of Business & Economic Statistics **40**(1), 216–226.

Chen, C. Y.-H., Härdle, W. K. & Klochkov, Y. (2022), 'Sonic: Social network analysis with influencers and communities', Journal of Econometrics **228**(2), 177–220.

Chen, E. Y., Fan, J. & Zhu, X. (2023), 'Community network auto-regression for high-dimensional time series', Journal of Econometrics **235**(2), 1239–1256.

Dempster, A. P. (1972), 'Covariance selection', Biometrics pp. 157–175.

Fan, J., Fan, Y., Han, X. & Lv, J. (2022), 'Simple: Statistical inference on membership profiles in large networks', Journal of the Royal Statistical Society Series B: Statistical Methodology **84**(2), 630–653.

Fan, J., Liao, Y. & Liu, H. (2016), 'An overview of the estimation of large covariance and precision matrices', The Econometrics Journal **19**(1), C1–C32.

Fan, J., Liao, Y. & Mincheva, M. (2013), 'Large covariance estimation by thresholding principal orthogonal complements', Journal of the Royal Statistical Society Series B: Statistical Methodology **75**(4), 603–680.

Guðmundsson, G. S. (2026), 'Detecting giver and receiver spillover groups in large

vector autoregressions', Journal of Business & Economic Statistics **44**(1), 297–308.

Guðmundsson, G. S. & Brownlees, C. (2021), 'Detecting groups in large vector autoregressions', Journal of Econometrics **225**(1), 2–26.

Hamilton, J. D. & Owyang, M. T. (2012), 'The propagation of regional recessions', Review of Economics and Statistics **94**(4), 935–947.

Jin, J. (2015), 'Fast community detection by score', Annals of Statistics .

Jin, J., Ke, Z. T. & Luo, S. (2024), 'Mixed membership estimation for social networks', Journal of Econometrics **239**(2), 105369.

Karrer, B. & Newman, M. E. (2011), 'Stochastic blockmodels and community structure in networks', Physical Review E—Statistical, Nonlinear, and Soft Matter Physics **83**(1), 016107.

Koller, D. & Friedman, N. (2009), Probabilistic graphical models: principles and techniques, MIT press.

Merlevède, F., Peligrad, M. & Rio, E. (2011), 'A bernstein type inequality and moderate deviations for weakly dependent sequences', Probability Theory and Related Fields **151**, 435–474.

Munro, E. & Ng, S. (2022), 'Latent dirichlet analysis of categorical survey responses', Journal of Business & Economic Statistics **40**(1), 256–271.

Pourahmadi, M. (2013), 'High-dimensional covariance estimation: with high-dimensional data', John Wiley & Sons .

Rohe, K., Chatterjee, S. & Yu, B. (2011), 'Spectral clustering and the high-dimensional stochastic blockmodel', Annals of Statistics .

Su, L., Shi, Z. & Phillips, P. C. (2016), 'Identifying latent structures in panel data', Econometrica **84**(6), 2215–2264.

Von Luxburg, U. (2007), 'A tutorial on spectral clustering', Statistics and computing **17**, 395–416.

Zhang, Y., Levina, E. & Zhu, J. (2020), 'Detecting overlapping communities in networks using spectral methods', SIAM Journal on Mathematics of Data Science **2**(2), 265–283.

Zhang, Y., Wang, H. J. & Zhu, Z. (2019), 'Quantile-regression-based clustering for panel data', Journal of Econometrics **213**(1), 54–67.

Zhu, X., Chang, X., Li, R. & Wang, H. (2019), 'Portal nodes screening for large scale social networks', Journal of econometrics **209**(2), 145–157.

Zhu, X., Xu, G. & Fan, J. (2023), 'Simultaneous estimation and group identification for network vector autoregressive model with heterogeneous nodes', Journal of Econometrics p. 105564.

2PPCN Supplementry Material

Siao Xu*

June 11, 2026

Abstract

This document contains supplementary material for the two-path partial correlation network model and the PartialCorr-mixed-SCORE procedure.

Keywords: 2PPCN; partial correlation network; mixed membership; Mixed-SCORE; spectral methods.

*Graduate School of Economic and Social Sciences, University of Mannheim.
E-mail: siao.xu@students.uni-mannheim.de

SUPPLEMENTARY MATERIAL

2PPCN Supplementry Material: This supplementary material contains all theoretical proofs for the paper, "Mixed Membership Estimation in a Two-Path Partial Correlation Network Model." In addition, it discusses model identification, states the conditional-dependence assumption underlying the model, provides graphical illustrations for the simulation study, and reports supplementary results for the empirical application.

U.S. state-level employment growth rate dataset: The dataset is obtained from the U.S. Bureau of Labor Statistics.

Deidentified data have been made available at the following URL:

https://econweb.ucsd.edu/ jhamilto/software.htm.

1 Notation and the definition for Graph

Let N denote the number of time series and T represent the number of observations. We define N_0 as the number of time series with pure memberships and $N - N_0$ as the number of time series with mixed membership. Lowercase letters, such as y, denote scalars, while capital letters, like Y, represent matrices. For any matrix Y, we use $y_{i\cdot}$, $y_{\cdot j}$ and $y_{i,j}$ to denote its i-th row, j-th column and ij-th entry, respectively. For an arbitrary square matrix $Y \in \mathbb{R}^{n\times n}$, $\lambda_i(Y)$ refers to the i-th eigenvalue of Y with $|\lambda_1(Y)| \geq |\lambda_2(Y)| \geq \cdots \geq |\lambda_n(Y)|$. The spectral radius of a square matrix Y is defined as $\rho(Y) = |\lambda_1(Y)|$. For a symmetric matrix Y, $\lambda_{max}(Y)$ and $\lambda_{min}(Y)$ denote the largest and smallest eigenvalues respectively. The following matrix norms are used: ℓ_2-norm: $\|Y\|$ and Frobenius norm: $\|Y\|_F = Tr\Big(Y^{'}Y\Big)$.

A graph is defined as a triple: $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ where $\mathcal{V} = 1, \cdots, N$ stands for the vertices of the graph, $\mathcal{E} \in \mathcal{V} \times \mathcal{V}$ the set of edges, and, accordingly, $\mathcal{W} \in \mathbb{R}$ the weights of all edges if necessary. A is the adjacency matrix which is the matrix representation of the graph. In this work, I assume that the adjacency matrix A is a $N \times N$ matrix where the elements $[A]_{ij} = 1$ if there exists an edge connecting node j and node i and $[A]_{ij} = 0$ otherwise.

Appendix A. Proof

Proof of Theorem 1. We instead have the joint distribution of $\mathbb{P}\Big(\Lambda, A, \Pi, \vec{\varphi}\Big|B, \vec{\alpha}, \Theta, \underline{u}, \overline{u}, \sigma\Big)$. Then,

$$
\begin{aligned}
&\mathbb{E}\Big(\Lambda\Big) \\
&= \sum_{\Lambda} \Lambda\Bigg(\sum_{A} \int_{\Pi} \int_{\vec{\varphi}} \mathbb{P}\Big(\Lambda, A, \Pi, \vec{\varphi}\Big|B, \vec{\alpha}, \Theta, \underline{u}, \overline{u}, \sigma\Big)\, \mathrm{d}_{\Pi}\, \mathrm{d}_{\vec{\varphi}} \Bigg), \\
&\textit{since } \mathbb{P}\Big(\Lambda\Big) = \sum_{A} \int_{\Pi} \int_{\vec{\varphi}} \mathbb{P}\Big(\Lambda, A, \Pi, \vec{\varphi}\Big|B, \vec{\alpha}, \Theta, \underline{u}, \overline{u}, \sigma\Big)\, \mathrm{d}_{\Pi}\, \mathrm{d}_{\vec{\varphi}}, \\
&\textit{and since } A \textit{ is discrete and } \Pi \textit{ and } \vec{\varphi} \textit{ are constinuous}, \\
&= \sum_{\Lambda} \sum_{A} \int_{\Pi} \int_{\vec{\varphi}} \Lambda \Bigg\{ \prod_{i,j}^{N} \mathbb{P}\Big(\Lambda_{ij}\Big|A_{ij}, \sigma_{ij}, \varphi_i, \varphi_j\Big) \Bigg\} \Bigg\{ \prod_{i}^{N} \mathbb{P}\Big(\varphi_i\Big|\underline{u}_i, \overline{u}_i\Big) \Bigg\} \\
&\Bigg\{ \prod_{i,j}^{N} \mathbb{P}\Big(A_{i,j}\Big|B, \vec{\pi}_i, \vec{\pi}_j, \Theta\Big) \Big\{ \prod_{n}^{N} \mathbb{P}\Big(\vec{\pi}_n\Big|\vec{\alpha}\Big) \Big\} \Bigg\}\, \mathrm{d}_{\Pi}\, \mathrm{d}_{\vec{\varphi}}, \\
&\textit{based on } fundamental\ of\ Bayesian\ network\ corresponding\ to\ Figure \textbf{??}, \\
&= \sum_{A} \int_{\Pi} \int_{\vec{\varphi}} \vec{\varphi} A \vec{\varphi} \Bigg\{ \prod_{i}^{N} \mathbb{P}\Big(\varphi_i\Big|\underline{u}_i, \overline{u}_i\Big) \Bigg\} \\
&\Bigg\{ \prod_{i,j}^{N} \mathbb{P}\Big(A_{i,j}\Big|B, \vec{\pi}_i, \vec{\pi}_j, \Theta\Big) \Big\{ \prod_{n}^{N} \mathbb{P}\Big(\vec{\pi}_n\Big|\vec{\alpha}\Big) \Big\} \Bigg\}\, \mathrm{d}_{\Pi}\, \mathrm{d}_{\vec{\varphi}}, \\
&\textit{since } \Lambda_{ij}\Big|A_{ij}, \sigma_{ij}, \varphi_i, \varphi_j \sim Beta\Big(\varphi_i\varphi_j, \sigma_{ij}\Big), \\
&\textit{if } A_{ij} = 1;\ \textit{otherwise } 0, \\
&= \sum_{A} \int_{\Pi} \frac{1}{2}\boldsymbol{U} A \boldsymbol{U} \frac{1}{2} \Bigg\{ \prod_{i,j}^{N} \mathbb{P}\Big(A_{i,j}\Big|B, \vec{\pi}_i, \vec{\pi}_j, \Theta\Big) \Big\{ \prod_{n}^{N} \mathbb{P}\Big(\vec{\pi}_n\Big|\vec{\alpha}\Big) \Big\} \Bigg\}\, \mathrm{d}_{\Pi}, \\
&\textit{since } \varphi_i|\underline{u}_i, \overline{u}_i \sim Uniform\Big(\underline{u}_i, \overline{u}_i\Big), \\
&\textit{where } \boldsymbol{U} = diag\Big(\{\underline{u}_i + \overline{u}_i\}_{i=1}^{N}\Big), \\
&= \int_{\Pi} \frac{1}{2}\boldsymbol{U}\Theta\Pi B \Pi^{'}\Theta\boldsymbol{U}\frac{1}{2} \Bigg\{ \Big\{ \prod_{n}^{N} \mathbb{P}\Big(\vec{\pi}_n\Big|\vec{\alpha}\Big) \Big\} \Bigg\}\, \mathrm{d}_{\Pi}, \\
&\textit{since } A_{ij}|B, \vec{\pi}_i, \vec{\pi}_j, \Theta_i, \Theta_j \sim Bernoulli\Big(\Theta_i \vec{\pi}_i B \vec{\pi}_j^{'} \Theta_j\Big), \\
&= \frac{1}{2}\boldsymbol{U}\Theta\Pi B \Pi^{'}\Theta\boldsymbol{U}\frac{1}{2},\ \ \textit{under the assumption that } \Pi \textit{ is parameter.} \quad \square
\end{aligned}
$$

Proof of Lemma 1. It suffices to prove that $Span\Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\Big) = Span\Big(\mathbf{X}\Big)$. By the assumption that, for each group, there exists at least one pure node inside, then, $Rank(\Pi) = K$. Since $\boldsymbol{U}$ and Θ are both non-singular, we have $\Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\Big)B\Big(\Pi\Theta\boldsymbol{U}\frac{1}{2}\Big)$ with rank of K.

Furthermore, by definition,

$$\frac{1}{2}\boldsymbol{U}\Theta\Pi\Big(B\frac{1}{2}\Pi\Theta\boldsymbol{U}\mathbf{x}_k^{-1}\Big) = \lambda_k\mathbf{x}_k$$

for $k = 1, \cdots, K$. Obviously, column space of $Span\Big(\mathbf{X}\Big)$ is inside column space of $\Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\Big)$ and both sides have rank of K, then, two column spaces are the same. $\square$

Lemma A.1. *Given a 2PPCN in Definition 1 satisfying Assumption 1 and define a $K \times K$ matrix $G = K\Big\|\frac{1}{2}(\overline{u}+\underline{u})\theta\Big\|^{-2}\Big(\frac{1}{4}\Pi'\Theta^2\boldsymbol{U}^2\Pi\Big)$. Let $\lambda_1, \cdots, \lambda_K$ and $\mathbf{x}_1, \cdots, \mathbf{x}_K$ be the eigenvalues and eigenvectors of $\Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\Big)B\Big(\frac{1}{2}\Pi\Theta\boldsymbol{U}\Big)$ in descending order. We have a set of Lemmas as follows:*

1. *For $1 \le k \le K$, let α_k be the k-th largest eigenvalue of BG and its corresponding non-zero eigenvalues of $\mathbb{E}(\Lambda)$ are $(K)^{-1}\Big\|\frac{1}{2}(\overline{u}+\underline{u})\theta\Big\|^2\alpha_k$ for $k = 1, \cdots, K$.*

2. *For $1 \le k \le K$, let γ_k be the k-th column of P, then, γ_k is the k-th eigenvector of BG with respect to α_k.*

3. *$\lambda_1 > 0$ and it has a multiplicity 1.*

4. *If all entries in $\mathbf{x}_1$ are positive, then all entries of its associated γ_1 is also positive.*

Proof of Lemma A1.

1. Assume that B is symmetric and G is positive definite. Define $G^{1/2}$ the square root of G, since the eigenvalues of MN are equal to the eigenvalues of NM for $N \in \mathbb{R}^{n,m}$ and $M \in \mathbb{R}^{m,n}$. Then the eigenvalues of BG are the same as that of $G^{1/2}BG^{1/2}$. It means that

α_k for $i = 1, \cdots, K$ are real. On the other hand, it also implies that the eigenvalues of $\mathbb{E}\big(\Lambda\big) = \frac{1}{2}\boldsymbol{U}\Theta\Pi\Big(B\frac{1}{2}\Pi\Theta\boldsymbol{U}\Big)$ are the same as that of $(K)^{-1}\Big\|\frac{1}{2}(\overline{u}+\underline{u})\theta\Big\|^2 BG$. The claim follows up after using Lemma 2.2 in Jin et al. (2024) but now we have $\frac{1}{2}\theta_i(\overline{u}_i + \underline{u}_i) = \mathbf{x}_1(i)/(\Pi_i^{'}\gamma_1)$ instead.

2. We have that $\mathbb{E}(\Lambda)\mathbf{x}_k = \lambda_k\mathbf{x}_k$ and $\mathbf{x}_k = \frac{1}{2}\Pi'\Theta\boldsymbol{U}\gamma_k$. Define $\tilde{G} = \frac{1}{4}\Pi^{'}\Theta^2\boldsymbol{U}^2\Pi$, then we have

$$\begin{aligned}\Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi B\frac{1}{2}\Pi\Theta\boldsymbol{U}\Big)\mathbf{x}_k &= \lambda_k\mathbf{x}_k \\ \Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi B\frac{1}{2}\Pi\Theta\boldsymbol{U}\Big)\Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\gamma_k\Big) &= \lambda_k\Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\gamma_k\Big)\end{aligned} \tag{A7}$$

Multiplying both sides on (A7) by $\frac{1}{2}\Pi^{'}\Theta\boldsymbol{U}$, then, we have

$$\tilde{G}B\tilde{G}\gamma_k = \lambda_k\tilde{G}\gamma_k.$$

Since $\tilde{G}$ is non-singular, we have $B\tilde{G}\gamma_k = \lambda_k\gamma_k$. After plugging $\tilde{G} = \frac{1}{4}\Pi^{'}\Theta^2\boldsymbol{U}^2\Pi$ and $\lambda_k = (K)^{-1}\Big\|\frac{1}{2}\sqrt{M}(\overline{u}+\underline{u})\theta\Big\|^2\alpha_k$, the claim follows.

3. It suffices to prove that α_1 is positive and has multiplicity of 1 since $\lambda_1 = (K)^{-1}\Big\|\frac{1}{2}(\overline{u}+\underline{u})\theta\Big\|^2\alpha_1$. Under the assumption that BG is irreducible and by Perron-Frobenius theorem, the claim follows.

4. By Perron-Frobenius theorem, there exists a strictly positive γ_1 corresponding to α_1. We have that $\mathbf{x}_1 = \frac{1}{2}\Pi'\Theta\boldsymbol{U}\gamma_1$ and $\boldsymbol{U}\Theta\Pi$ is positive. If $\mathbf{x}_1$ is positive, then γ_1 is strictly positive. □

Proof of Lemma 2. By Lemma 1., we know that $\frac{1}{2}\boldsymbol{U}\Theta\Pi P = \mathbf{X}$ for a non-singular matrix $P = \Big[p_1, \cdots, p_K\Big] \in \mathbb{R}^{K\times K}$. And with Lemma A.1, we directly extend the proof of Lemma 2.1 in (Jin et al., 2024) to attain the claim. □

Lemma A.2. *We aim to bound* $\Big\|\hat{\Lambda} - \mathbb{E}\big(\Lambda\big)\Big\|$ *where* $\hat{\Lambda}$ *and* $\mathbb{E}\big(\Lambda\big)$ *represent the estimated partial*

correlation network and population concentration matrix, respectively. We prove the Lemma by establishing the following inequality:

$$\left\|\hat{\Lambda} - \mathbb{E}\left(\Lambda\right)\right\| \leq \left\|\hat{\Lambda} - \Lambda\right\| + \left\|\mathbb{E}\left(\Lambda\right) - \Lambda\right\|.$$

We prove that $\left\|\hat{\Lambda} - \mathbb{E}\left(\Lambda\right)\right\|$ *is bounded above by applying Lemma A.3 and Lemma A.4, which are defined below.*

Lemma A.3. *Given the 2PPCN that satisfies Assumption 3, where the pairwise probability matrix* B *satisfies Assumption 2, let* $\mathbb{E}\left(\Lambda\right)$ *and* Λ *denote the sample and population concentration matrix. Define* $\underline{\Pi} = \arg\min_i \left\|\Pi_i\right\|^2$ *for* $i = 1, \cdots, N$*. Then, the following conclusion is derived:*

$$\left\|\mathbb{E}\left(\Lambda\right) - \Lambda\right\| \leq O_p\left(\sqrt{\frac{log(N)}{\underline{\Pi} N B_N}}\right).$$

Proof of Lemma A.3. I prove this by applying the conclusion of Lemma A.1 of Guðmundsson and Brownlees (2021) which states:

Let $\overline{d}_{min} = \min_i d_i$ denote the minimum expected degree of the graph and define $v = 2\max_{ij}\overline{w}_{ij} + \max_{ij}(\overline{w}_{ij}^2/\underline{w}_{ij}^2)$, where $\overline{w}$ and $\underline{w}$ represent the lower and upper bounds for sampling the edge weights respectively. Then, for a constant $c > 0$, there exists another constant $C > 0$ independent of number of nodes, N, and the edge probabilities. Moreover, if $\overline{d}_{\min} \geq Clog(N)$ and for all $n^{-c} \leq \delta \leq 1/2$, we have:

$$\mathbb{P}\left(\left\|L - \mathcal{L}\right\| \leq 14\sqrt{\frac{vlog(2N/\delta)}{\overline{d}_{min}}}\right) \geq 1 - \delta.$$

Our goal is to prove that the current model setup satisfies the conditions required for Lemma A.1 to hold.

We have $\mathbb{E}\left(\Lambda\right) = \left(\frac{1}{2}\boldsymbol{U}\Theta\Pi\right)B\left(\Pi\Theta\boldsymbol{U}\frac{1}{2}\right)$ as defined in Section 2.2. Furthermore, Define $\mathcal{L} =$

$\mathcal{D}^{-1/2}\mathbb{E}\Big(\Lambda\Big)\mathcal{D}^{-1/2}$ and $l = D^{-1/2}CD^{-1/2}$ where $\mathcal{D} = \sum_{i=1}^{N}\Big[\mathbb{E}\Big(\Lambda\Big)\Big]_{i,j}$ and $D = \sum_{i=1}^{N}\Big[\Lambda\Big]_{i,j}$. Then,

$$\mathcal{D}_{i,i} = \sum_{j=1}^{N}\Big[\mathbb{E}\Big(\Lambda\Big)\Big]_{i,j} = \Big(\frac{1}{4}\Big)\Big[\boldsymbol{U}\Big]_{i,i}\Big[\Theta\Big]_{i,i}\sum_{j=1}^{N}\Big[\boldsymbol{U}\Big]_{j,j}\Big[\Theta\Big]_{j,j}\Pi_i B\Pi_j.$$

By Assumption 2, the minimal degree satisfies

$d_{min} = \min_i\Big[\mathcal{D}\Big]_{ii} \geq \underline{\Pi}^2\sum_{j=1}^{N}\Big[B\Big]_{i,j} = \underline{\Pi}^2\boldsymbol{\Omega}\Big(\sum_{j=1}^{N}\Big[B\Big]_{i,j}\Big) = \boldsymbol{\Omega}\Big(\underline{\Pi}^2 N B_N\Big) = \boldsymbol{\Omega}\Big(\underline{\Pi}^2 log(N)\Big)$,

where $\underline{\Pi} = \arg\min_i \Big\|\Pi_i\Big\|^2$ for $i = 1, \cdots, N$.

If we pick $\delta = N^{-c}$ for any positive c and N such that $N^{-c} \leq 1/2$, then, we have that

$$\mathbb{P}\left(\Big\|L - \mathcal{L}\Big\| \leq C\sqrt{\frac{log(N)}{\underline{\Pi}^2 N B_N}}\right) \leq N^{-c}.$$

Then, the claim follows. □

Lemma A.4. *Given the 2PPCN model as defined in Definition 1, satisfying Assumption 3, let $\hat{Q}$ denote the estimated partial correlation network and Λ the concentration matrix. Then, with high probability, we have*

$$\Big\|\hat{\Lambda} - \Lambda\Big\| \leq O_p\left(\sqrt{\frac{N}{T}}\right).$$

Proof of Lemma A.4. First step is to prove that $\Big\|\hat{\Lambda} - \Lambda\Big\|$ is bounded above.

$$
\begin{aligned}
&\left\|\hat{\Lambda}-\Lambda\right\| \\
&=\left\|\hat{\Lambda}\left(\Sigma-\hat{\Sigma}\right)\Lambda\right\| \\
&\leq\left\|\hat{\Lambda}\right\|\left\|\Sigma-\hat{\Sigma}\right\|\left\|\Lambda\right\| \textit{ by submultiplicative law.}
\end{aligned}
$$

First, $\hat{\Lambda}$ converges to Λ as T goes to infinity. Second, by applying Lemma A2 of Guðmundsson and Brownlees (2021), we have

$$
\left\|\Sigma-\hat{\Sigma}\right\|=O_p\left(\sqrt{\frac{N}{T}}\|\Sigma\|\right).
$$

By Lemma OA.13 in of Guðmundsson and Brownlees (2021), we have $\left\|\Lambda\right\|$ bounded by a absolute constant. Overall, $\left\|\Sigma-\hat{\Sigma}\right\|=O_p\left(\sqrt{\frac{N}{T}}\right)$. Then, we have $\left\|\hat{\Lambda}-\Lambda\right\|$ bounded above. □

Proof of Theorem 2. By directly applying the result in Lemma 3. and arguments of Appendix B of Rohe et al. (2011) and Davis-Kahan theorem, it is with high probability that

$$
\|\mathbf{O}\hat{X}-\mathcal{X}\|\leq O_p\left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}}+\sqrt{\frac{N}{T}}\right). \quad \square
$$

Next, the proof of Lemma A.4 is a direct extension of Lemma C.3 in Jin et al. (2024).

Lemma A.5. *Given a 2PPCN model as defined in Definition 1, satisfying Assumption 3. If the first eigenvector* $\mathbf{X}_1$ *with respect to 2PPCN satisfies the condition such that* $\sum_{i=1}^{N}\mathbf{X}_{1,i}>0$ *,then, we have the following statement:*

$$
\sqrt{C}^{-1}\Theta_i\boldsymbol{U}_i/\|\Theta\|\|\boldsymbol{U}\|\leq\mathbf{X}_{1,i}\leq\sqrt{C}\Theta_i\boldsymbol{U}_i/\|\Theta\|\|\boldsymbol{U}\| \textit{ for } i=1,\cdots,N.
$$

Proof of Lemma A.5. From Lemma A.2.4 in Jin et al. (2024), we are able to choose a set of $(\mathbf{X}_1, p_1)$ such that all the elements in $\mathbf{X_1}$ and p_1 are positive and $\sum_{i=1}^{N} \mathbf{X}_{1,i} > 0$. And We have a rotation matrix P such that $\mathbf{X} = \Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi\Big)P$ in which $\mathbf{X}_{1,i} = \frac{1}{2}\boldsymbol{U}_i\Theta_i \sum_{k=1}^{K} \Pi_{i,k}p_{i,k}$. Since Π_i for $i = 1, \cdots, N$ are all positive and $\sum_{k=1}^{K} \Pi_{i,k} = 1$, we have

$$\Big(\frac{1}{2}\boldsymbol{U}_i\Theta_i\Big) \min_{1\le k\le K} p_{1,k} \le \mathbf{X}_{1,i} \le \Big(\frac{1}{2}\boldsymbol{U}_i\Theta_i\Big) \max_{1\le k\le K} p_{1,k},$$

for $i = 1, \cdots, K$. We bound $\mathbf{X}_{1,i}$ by bounding $p_{1,k}$.

Given $G = K\Big\|\frac{1}{2}(\overline{u} - \underline{u})\theta\Big\|^{-2}\Big(\frac{1}{4}\Pi'\boldsymbol{U}^2\Theta^2\Pi\Big)$, we have

$$K^{-1}\Big\|\frac{1}{2}(\overline{u} - \underline{u})\theta\Big\|^2 G = \frac{1}{4}\Pi'\Theta^2\boldsymbol{U}^2\Pi \tag{A8}$$

Multiply (A8) by rotation matrix $P^{'}$ on left hand side and P on right hand side, we have

$$\Big(P^{'}\frac{1}{2}\Pi'\Theta\boldsymbol{U}\Big)\Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi P\Big) = I_K$$

since $\mathbf{X}$ is orthonormal. Then, it implies that

$$P\Big(P^{'}\frac{1}{2}\Pi'\Theta\boldsymbol{U}\Big)\Big(\frac{1}{2}\boldsymbol{U}\Theta\Pi P\Big)P^{'} = PP^{'}$$

Hence,

$$PP^{'} = K\Big\|\frac{1}{2}(\overline{u} - \underline{u})\theta\Big\|^2 G^{-1}$$

Since $PP^{'} = \sum_{k=1}^{K} p_k p_k^{'} \ge p_1 p_1^{'}$, then ,it is obvious that $\|p_1\|^2 \le \|P\|^2 \le K\Big\|\frac{1}{2}(\overline{u} - \underline{u})\theta\Big\|^2\|G^{-1}\|$. Under the assumption that $\|G^{-1}\| \le C$, we have $\|p_1\| \le \sqrt{C}\sqrt{K}\Big\|\frac{1}{2}(\overline{u} -$

$\underline{u})\theta\Big\|$. $\square$

Lemma A.6 is a direct extension of Lemma C.4 in Jin et al. (2024).

Lemma A.6. *Given a 2PPCN as defined in Definition 1, satisfying Assumption 3, and ideal simplex vertices v_i for $i = 1, \cdots, K$, then, we have the following statement:*

$$\max_{1\le k\le K} \|v_k\| \le \sqrt{K}C.$$

Proof of Lemma A.6. Define a rotation matrix $P = \Big(p_1, \cdots, p_K\Big)$, $\mathcal{S}^{ideal} \subset \mathbb{R}^{K-1}$ is a simplex spanned by a set of basis defined as $v_1, \cdots, v_K$ where $v_{k,l} = p_{l+1,k}/p_{1,k}$. Define a matrix $Q = \begin{bmatrix} 1 & \cdots & 1 \\ v_1 & \cdots & v_K \end{bmatrix}$. Then, define a function $T(w) : \mathbb{S}_0 \to \mathbb{S}^{ideal}$ where $\begin{bmatrix} 1 \\ T_2(w) \end{bmatrix} = Q \cdot w$. Then, obviously, $r_i = T(w_i)$ and $Q' = \Big(diag(p_1)\Big)^{-1} P$.

Then, $\max_{1\le k\le K} \|v_k\| \le \|Q\| \le \|\Big(diag(p_1)\Big)^{-1}\|\|P\| \le \sqrt{K}C$ by assumption 2. $\square$

Proof of Theorem 3. This is an extension of proof of Theorem 3.1 in Jin et al. (2024) using results of Theorem 2 in present paper.

By definition, the SCORE step is defined as follows: $r_i = \frac{1}{\mathbf{x}_{1,i}}\mathbf{X}_{i,0}$ where $\mathbf{X}_{i,0}$ is the i-th row of eigenvector matrix $\mathbf{X}_0$ excluding its first eigenvector. Let $\mathbf{x}_{1,i}$ denotes the k-th entry in the first eigenvector with respect to the first eigenvalue in 2PPCN. Furthermore, let $L = \mathbf{O}_{1,i}\mathbf{O}^\star$ where $\mathbf{O}$ is defined in Theorem 2 and $\mathbf{O}_1$ is the first row of $\mathbf{O}$. $\mathbf{O}_{1,i}$ is ith entry of $\mathbf{O}_1$. In addition, when we exclude the first eigenvalue and its corresponding eigenvector, we are still able to find a $(K-1)\times(K-1)$ orthonormal matrix $\mathbf{O}^\star$ as defined in Theorem 2 since it's like using $K-1$ eigenvalues instead.

Then, it follows that

$$
\begin{aligned}
& L\hat{r}_i - r_i \\
&= \frac{1}{\mathbf{O}_{1,i}\hat{\mathbf{x}}_{1,i}}\Big(\mathbf{O}^{\star'}\hat{\mathbf{X}}_{i,0} - \mathbf{X}_{i,0}\Big) + \Big(\frac{1}{\mathbf{O}_{1,i}\hat{\mathbf{x}}_{1,i}} - \frac{1}{\mathbf{x}_{1,i}}\Big)\mathbf{X}_{i,0} \\
&= \frac{1}{\mathbf{O}_{1,i}\hat{\mathbf{x}}_{1,i}}\Big(\mathbf{O}^{\star'}\hat{\mathbf{X}}_{i,0} - \mathbf{X}_{i,0}\Big) - \frac{\Big(\mathbf{O}_{1,i}\hat{\mathbf{x}}_{1,i} - \mathbf{x}_{1,i}\Big)}{\mathbf{O}_{1,i}\hat{\mathbf{x}}_{1,i}} r_i
\end{aligned}
\tag{A9}
$$

By Theorem 2, $\Big(\mathbf{O}^{\star'}\hat{\mathbf{X}}_{i,0} - \mathbf{X}_{i,0}\Big)$ is bounded above. $\Big(\mathbf{O}_{1,i}\hat{\mathbf{x}}_{1,i} - \mathbf{x}_{1,i}\Big)$ is also bounded above by Theorem 2, since it's a part of $\|\mathbf{O}\hat{X} - \mathcal{X}\|$ and thus bounded above by $\|\mathbf{O}\hat{X} - \mathcal{X}\|$. By Lemma A.5, $\|r_i\|$ is bounded above by $\sqrt{K}C$ since r_i is on the simplex and bounded above by an element on the simplex with maximal norm. Finally, by applying Lemma C.3 in Jin et al. (2024), we have that $\mathbf{x}_{1,i} > C\Big(\theta(\overline{u} - \underline{u})\Big)_{min} \Big/ \|\theta(\overline{u} - \underline{u})\|$. Then, we have that $\frac{C\Big(\theta(\overline{u}-\underline{u})\Big)_{min}}{\|\theta(\overline{u}-\underline{u})\| - \Big(\sqrt{\frac{log(N)}{\underline{\Pi} N B_N}} + \sqrt{\frac{N}{T}}\Big)} \leq \mathbf{x}_{1,i} - \sqrt{\frac{log(N)}{\underline{\Pi} N B_N}} - \sqrt{\frac{N}{T}} \leq \mathbf{O}_{1,i}\hat{\mathbf{x}}_{1,i}$.

After combining all together, we have:

$$
\|L\hat{r}_i - r_i\| \leq D\left(\|\mathbf{O}^{\star'}\hat{\mathbf{X}}_{i,0} - \mathbf{X}_{i,0}\| + \sqrt{K}\Big|\mathbf{O}_{1,i}\hat{\mathbf{x}}_{1,i} - \mathbf{x}_{1,i}\Big|\right),
$$

where $D = \frac{\|\theta(\overline{u}-\underline{u})\| - \Big(\sqrt{\frac{log(N)}{\underline{\Pi} N B_N}} + \sqrt{\frac{N}{T}}\Big)}{C\Big(\theta(\overline{u}-\underline{u})\Big)_{min}}$.

$$
\|\mathbf{L}\hat{R} - R\| = O_p\left(\sqrt{K}D\left(\sqrt{\frac{log(N)}{\underline{\Pi} N B_N}} + \sqrt{\frac{N}{T}}\right)\right). \quad \square
$$

Lemma A.7. *Given a 2PPCN as defined in Definition 1, if Theorem 3 holds, then*

$$
\max_{1\leq i\leq N} \|\mathbf{L}\hat{r}_i - r_i\|_2 = O_p\left(\sqrt{K}D\left(\sqrt{\frac{log(N)}{\underline{\Pi} N B_N}} + \sqrt{\frac{N}{T}}\right)\right).
$$

Proof of Lemma A.7. For each row i, $\left\|\mathbf{L}\hat{r}_i - r_i\right\|_2 = \left\|e_i\left(\mathbf{L}\hat{R} - R\right)\right\|_2 \leq \left\|e_i\right\|_2\left\|\mathbf{L}\hat{R} - R\right\| = \left\|\mathbf{L}\hat{R} - R\right\|$. Then taking the maximum over i and applying Theorem 3 gives the result. □

Lemma A.8. *Given a 2PPCN as defined in Definition 1, satisfying Assumption 1,2, 3, 4 and 5 and Theorem 3, then*

$$\Delta_V = O_p\left(\sqrt{K}D\left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\right)\right).$$

Proof of Lemma A.8. Bt Assumption 4, $\Delta_V \leq C_V \max_i \|\mathbf{L}\hat{r}_i - r_i\|_2$. Then, Lemma A.6 gives the desired order. □

Lemma A.9. *Suppose Assumption 3 holds, and $\Delta_V = o_p(1)$. Let*

$$\hat{\mathcal{V}} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \mathbf{L}\hat{v}_1 & \mathbf{L}\hat{v}_2 & \cdots & \mathbf{L}\hat{v}_K \end{pmatrix}.$$

Let $\hat{w}_i$ solves

$$\hat{\mathcal{V}}\hat{w}_i = \begin{pmatrix} 1 \\ \mathbf{L}\hat{r}_i \end{pmatrix}.$$

Then, with probability tending to one,

$$\|\hat{w}_i - w_i\|_2 \leq C_{\mathcal{V}}\left(\left\|\mathbf{L}\hat{r}_i - r_i\right\|_2 + \Delta_V\right),$$

uniformly over i, where $C_{\mathcal{V}}$ depends only on $c_{\mathcal{V}}$.

Proof of Lemma A.9. We have population and sample equations

$$\mathcal{V}w_i = \begin{pmatrix} 1 \\ r_i \end{pmatrix}, \quad \hat{\mathcal{V}}\hat{w}_i = \begin{pmatrix} 1 \\ \mathbf{L}\hat{r}_i \end{pmatrix}.$$

Subtracting gives

$$\hat{\mathcal{V}}\Big(\hat{w}_i - w_i\Big) + \Big(\hat{\mathcal{V}} - \mathcal{V}\Big)w_i = \begin{pmatrix} 0 \\ \mathbf{L}\hat{r}_i - r_i \end{pmatrix}.$$

Hence,

$$\Big(\hat{w}_i - w_i\Big) = \hat{\mathcal{V}}^{-1}\begin{pmatrix} 0 \\ \mathbf{L}\hat{r}_i - r_i \end{pmatrix} - \hat{\mathcal{V}}^{-1}\Big(\hat{\mathcal{V}} - \mathcal{V}\Big)w_i.$$

Taking norms,

$$\Big\|\hat{w}_i - w_i\Big\|_2 \leq \Big\|\hat{\mathcal{V}}^{-1}\Big\|\Bigg(\Big\|\mathbf{L}\hat{r}_i - r_i\Big\|_2 + \Big\|\Big(\hat{\mathcal{V}} - \mathcal{V}\Big)w_i\Big\|_2\Bigg).$$

Because w_i is a convex-weight vector,

$$\Big\|\Big(\hat{\mathcal{V}} - \mathcal{V}\Big)w_i\Big\|_2 \leq \max_k \|\mathbf{L}\hat{v}_k - v_k\|_2 = \Delta_V.$$

Assumption 3 implies that $\Big\|\hat{\mathcal{V}}^{-1}\Big\| \leq c_{\mathcal{V}}^{-1}$. Since $\Delta_V = o_p\Big(1\Big)$, Weyl's inequality gives $\sigma_{min}\Big(\hat{\mathcal{V}}\Big) \geq c_{\mathcal{V}}/2$ with probability tending to one. Therefore, $\Big\|\hat{\mathcal{V}}^{-1}\Big\| \leq 2/c_{\mathcal{V}}$ with probability tending to one. The result follows. □

Lemma A.10. *Let $T_b\Big(w\Big) = \frac{w_i \oslash b_1}{\|w_i \oslash b_1\|_1}$, where $\oslash$ denotes element-wise division. Suppose $\hat{w}$ is close to w, and b, $\hat{b}$ satisfy*

$$c_b \leq b_k \leq C_b, \quad \hat{b}_k \geq c_b/2$$

for all k. Then, there exists a constant C_T, depending only on c_b, C_b and K, such that

$$\Big\|T_{\hat{b}}\Big(\hat{w}\Big) - T_b\Big(w\Big)\Big\|_1 \leq C_T\Big(\|\hat{w} - w\|_1 + \|\hat{b} - b\|_\infty\Big).$$

Proof of Lemma A.10. Define $z = w \oslash b$ and $\hat{z} = \hat{w} \oslash \hat{b}$. Since w is probability-like vector and $b_k \leq C_b$, we have

$$\|z\|_1 = \sum_{k=1}^{K} \frac{w_k}{b_k} \geq \frac{1}{C_b}.$$

Also, for $\hat{w}$ and $\hat{b}$ sufficiently close to w and b, $\|\hat{z}\|_1 \geq \frac{2}{C_b}$. We first bound $\hat{z} - z$. For each k,

$$\left|\frac{\hat{w}_k}{\hat{b}_k} - \frac{w_k}{b_k}\right| \leq \frac{\left|\hat{w}_k - w_k\right|}{\hat{b}_k} + w_k\left|\frac{1}{\hat{b}_k} - \frac{1}{b_k}\right|.$$

Using $\hat{b}_k \geq c_b/2$ and $b_k, \hat{b}_k \geq c_b/2$,

$$\|\hat{z} - z\|_1 \leq \frac{2}{c_b}\|\hat{w} - w\| + \frac{2}{c_b^2}\|\hat{b} - b\|_\infty \sum_{k=1}^{K} w_k.$$

Since $\sum_{k=1}^{K} w_k = 1$, we have

$$\|\hat{z} - z\|_1 \leq C_T\Big(\|\hat{w} - w\|_1 + \|\hat{b} - b\|_\infty\Big).$$

Now use the elementary inequality,

$$\left\|\frac{\hat{z}}{\|\hat{z}\|_1} - \frac{z}{\|z\|_1}\right\| \leq \frac{2\|\hat{z} - z\|_1}{\|z\|_1}$$

is valid when $\|\hat{z}\|_1$ and $\|z\|_1$ are bounded away from zero. Since $\|z\|_1 \geq \frac{1}{C_b}$, combining all inequalities proves the claim. □

Proof of Theorem 4. By Lemma A.7,

$$\max_i \|\mathbf{L}\hat{r}_i - r_i\|_2 = O_p\left(\sqrt{K}D\left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\right)\right).$$

By Lemma A.8,

$$\Delta_V = O_p\left(\sqrt{K}D\left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\right)\right).$$

Then, Lemma A.9 gives, uniformly over i,

$$\left\|\hat{w}_i - w_i\right\|_2 \leq C_{\mathcal{V}}\Big(\left\|\mathbf{L}\hat{r}_i - r_i\right\|_2 + \Delta_V\Big).$$

Therefore,

$$\max_i \|\hat{w}_i - w_i\|_2 = O_p\left(\sqrt{K}D\left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\right)\right).$$

Since K is fixed or grows slowly enough that constants can be tracked,

$$\max_i \|\hat{w}_i - w_i\|_1 \leq \sqrt{K}\max_i \|\hat{w}_i - w_i\|_2 = O_p\left(\sqrt{K}D\left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\right)\right).$$

If K is fixed, this is simply $O_p\left(KD\left(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\right)\right)$. If K grows, additional $\sqrt{K}$ should be retained. By Assumption 5,

$$\|\hat{b}_1 - b_1\|_\infty = O_p\left(D\sqrt{K}\Big(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\Big)\right).$$

The mixed membership estimator uses the mapping

$$\hat{\pi}_i = T_{\hat{b}_1}\Big(\hat{w}_i\Big), \quad \pi_i = T_{b_1}\Big(w_i\Big).$$

Applying Lemma A.10,

$$\|\hat{\pi}_i - \pi_i\|_1 \leq C_T\Big(\|\hat{w}_i - w_i\|_1 + \|\hat{b}_1 - b_1\|_\infty\Big).$$

Taking the maximum over i yields

$$\max_i \|\hat{\pi}_i - \pi_i\|_1 = O_p\left(D\sqrt{K}\Big(\sqrt{\frac{log(N)}{\underline{\Pi}NB_N}} + \sqrt{\frac{N}{T}}\Big)\right),$$

when K is fixed. This proves the row-wise membership bound. For the Frobenius bound, use

$$\left\|\hat{\pi}_i - \pi_i\right\|_2 \leq \left\|\hat{\pi}_i - \pi_i\right\|_1.$$

Thus,

$$\frac{1}{N}\|\hat{\Pi} - \Pi\|_F^2 = \frac{1}{N}\sum_{i=1}^{N}\|\hat{\pi}_i - \pi_i\|_2^2 \leq \max_i \|\hat{\pi}_i - \pi_i\|_1^2 = O_p\left(D^2K\left(\sqrt{\frac{log(N)}{\underline{\Pi} N B_N}} + \sqrt{\frac{N}{T}}\right)^2\right).$$

Using $\left(x+y\right)^2 \leq 2x^2 + 2y^2$ gives

$$\frac{1}{N}\|\hat{\Pi} - \Pi\|_F^2 = O_p\left(D^2K\left(\frac{log(N)}{\underline{\Pi} N B_N} + \frac{N}{T}\right)\right). \quad \square$$

Proof of Theorem B.1. The object is to prove that given two sets of parameters with respect to 2PPCN such that $\left(\boldsymbol{U}_1, \Theta_1, \Pi_1, B_1\right)$ and $\left(\boldsymbol{U}_2, \Theta_2, \Pi_2, B_2\right)$,
if $\left(\frac{1}{2}\boldsymbol{U}_1\Theta_1\Pi_1\right)B_1\left(\frac{1}{2}\Pi_1\Theta_1\boldsymbol{U}_1\right) = \left(\frac{1}{2}\boldsymbol{U}_2\Theta_2\Pi_2\right)B_2\left(\frac{1}{2}\Pi_2\Theta_2\boldsymbol{U}_2\right)$,
then, we must have $\Theta_1 = \Theta_2$, $\boldsymbol{U}_1 = \boldsymbol{U}_2$, $B_1 = B_2$ and $\Pi_1 = \Pi_2$. This is sufficient for the identifiability of the model.

This is an extension of Theorem 2.1 in Zhang et al. (2020) using Lemma A.1 of Tang et al. (2013).

By using the Lemma A.1 of of Tang et al. (2013), we have that:

Let $X_1, X_2 \in \mathbb{R}^{N\times d}$ such that $d < N$ be full rank matrices and define $\boldsymbol{X}_1 = X_1X_1'$ and $\boldsymbol{X}_2 = X_2X_2'$, then there exists an orthonormal matrix $\boldsymbol{O}$ such that

$$\|X_1\boldsymbol{O} - X_2\|_F = \frac{\sqrt{d}\|\boldsymbol{X}_1 - \boldsymbol{X}_2\|\left(\sqrt{\|\boldsymbol{X}_1\|} + \sqrt{\|\boldsymbol{X}_2\|}\right)}{\lambda_{min}\left(\boldsymbol{X}_2\right)}, \tag{A1}$$

where $\lambda_{min}\left(\cdot\right)$ is the smallest positive eigenvalue.

First of all, by (A1), we have that

$$\big(\frac{1}{2}\boldsymbol{U}_1\Theta_1\Pi_1\sqrt{B_1}\big)\boldsymbol{O}_{1,2} = \frac{1}{2}\Pi_2\Theta_2\boldsymbol{U}_2\sqrt{B_2}, \tag{A2}$$

where $\boldsymbol{O}_{1,2}$ is an orthonormal matrix.

By Assumption B.1.2, there's at least one pure node in each community. The first step is to prove that the indices for pure nodes in Π_1 and Π_2 are the same. Let's define the set $\mathcal{L}_1 = \Big\{i : row_i\big(\Pi_1\big) = e_k \; where \; k \in K\Big\}$. It suffices to prove that for any $i \in \mathcal{L}_1$, then $row_i\big(\Pi_2\big)$ is also pure.

We prove it by contradiction. For $i \in \mathcal{L}_1$ and $row_i\big(\Pi_2\big)$ is not pure, then $row_i\big(\Pi_2\big)$ can be represented by a linear combination of other mixed memberships as follows:

$$row_i\big(\Pi_2\big) = \sum_{j=1}^{Z} \omega_j \cdot row_j\big(\Pi_2\big) \tag{A3}$$

where $\omega_1, \cdots, \omega_N \geq 0$ and $Z = \Big\{j : j \notin \mathcal{L}_1\Big\}$.

By (A.2), we can represent (A.3) as:

$$row_i\big(\Pi_1\big) = \sum_{j=1}^{Z} \omega_j \frac{\frac{1}{2}\big(\boldsymbol{U}_1\big)_{j,j}\big(\Theta_1\big)_{j,j}}{\frac{1}{2}\big(\boldsymbol{U}_2\big)_{j,j}\big(\Theta_2\big)_{j,j}} \cdot row_j\big(\Pi_1\big), \tag{A4}$$

then $row_i\big(\Pi_1\big)$ now can be represented as a non-negative linear combination of rows in Π. It means that $row_i\big(\Pi_1\big)$ is not pure node, which is a contradiction to our assumption. Hence, i must be in $\mathcal{L}_1$.

we have proven the identifiability for all pure nodes. As for the others, let's construct a sub-matrix $\tilde{\Pi}$ of Π, in which each element is one pure node in each community. We also construct corresponding sub-matrices $\tilde{\boldsymbol{U}}$ of $\boldsymbol{U}$ and $\tilde{\Theta}$ of Θ respectively with the same order

of new $\tilde{\Theta}$.

After replacing $\boldsymbol{U}$, Θ and Π in (A2) with $\tilde{\boldsymbol{U}}$, $\tilde{\Theta}$ and $\tilde{\Pi}$, since now $\tilde{\Pi}_1$ and $\tilde{\Pi}_2$ are identity matrix, we have

$$\left(\tilde{\boldsymbol{U}}_1\tilde{\Theta}_1\sqrt{B_1}\right)\boldsymbol{O}_{1,2} = \tilde{\boldsymbol{U}}_2\tilde{\Theta}_2\sqrt{B_2}. \tag{A5}$$

By Assumption B.1.2, both $\sqrt{B_1}\boldsymbol{O}_{1,2}$ and $\sqrt{B_2}$ have rows of norm 1. Hence, from (A5), we have that $\tilde{\Theta}_1\tilde{\boldsymbol{U}}_1 = \tilde{\Theta}_2\tilde{\boldsymbol{U}}_2$. Then, furthermore, we have:

$$\sqrt{B_1}\boldsymbol{O}_{1,2} = \sqrt{B_2}. \tag{A6}$$

Then, $B_1 = \sqrt{B_1}\boldsymbol{O}_{1,2}\left(\sqrt{B_1}\boldsymbol{O}_{1,2}\right)' = B_2$.

Next, (A5) and (A6) imply that $\Theta_1\boldsymbol{U}_1 = \Theta_2\boldsymbol{U}_2$. By Assumption B.1.4, we have that $\boldsymbol{U}_1 = \boldsymbol{U}_2$ and $\Theta_1 = \Theta_2$.

Finally, combining all together into (A2), we have $\Pi_1 = \Pi_2 \quad \square$.

Appendix B.

Appendix B.1 Identification

In general, identifiability is nontrivial to build for most overlapping community model since an edge can be explained by either their common memberships in the same communities or high probabilities of having links between different communities. In this subsection, we discuss the conditions required for full identifiability of the 2PPCN model. As is standard in models with group structure, identifiability is considered up to a permutation of group labels. That is, the model is deemed identifiable if any two equivalent parameterizations differ only by a relabeling of the communities. To examine the identifiability conditions for the 2PPCN model, we begin by analyzing the expected concentration matrix Λ. Specifically, under the Theorem 1, we know that the expectation of Λ is $\left(\frac{1}{2}\boldsymbol{U}\Theta\Pi\right)B\left(\Pi'\Theta\boldsymbol{U}\frac{1}{2}\right)$, where $\Pi = \{\vec{\pi}_i\}_{i=1}^N$, $\Theta = diag\Big(\{\Theta_i\}_{i=1}^N\Big)$, and $\boldsymbol{U} = diag\Big(\{\underline{u}_i + \overline{u}_i\}_{i=1}^N\Big)$. We first outline the identifiability conditions pertaining to these parameters before analyzing the identifiability of 2PPCN model.

The following assumptions are sufficient for full identifiability of the 2PPCN model:

Assumption B.1. *Given a 2PPCN as defined in Definition 1:*

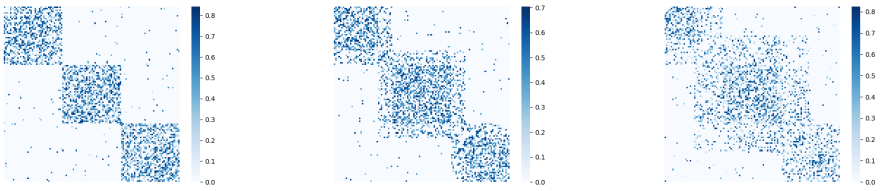
1. *B is full rank and strictly positive definite, with $B_{ii} = B_{jj} = 1$ for $i \neq j$.*
2. *There's at least one pure node in each community. That is, for each $k = 1, \cdots, K$, there at least exists one i such that $\pi_{i,k} = 1$.*
3. *Either the degree parameters $\theta_1, \cdots, \theta_N$ are all positive and $\frac{\sum_{i=1}^N \theta_i}{N} = 1$ or the influential effect parameters $\Big(\underline{u}_1 + \overline{u}_1\Big), \cdots, \Big(\underline{u}_N + \overline{u}_N\Big)$ are all positive and $\frac{\sum_{i=1}^N \Big(\underline{u}_i + \overline{u}_i\Big)}{N} = 1$.*
4. *We have $\frac{\sum_{i=1}^N \theta_i(\underline{u}_i + \overline{u}_i)}{N} = 1$.*

Under Assumption 1, we establish the following theorem regarding the identifiability of the 2PPCN model.

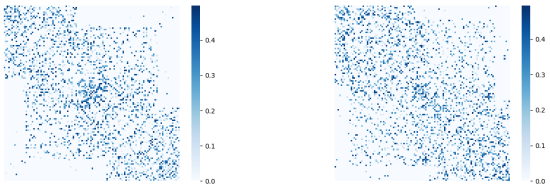
Theorem B.1. *If a 2PPCN defined in Definition 1 satisfies Assumption 1, then it is identifiable. That is, given* $\mathbb{E}(\Lambda)$ *under 2PPCN defined in Equation 1 with a set of parameter* $\left(\boldsymbol{U}, \Theta, \Pi, B\right)$ *satisfying Assumption 1, these parameters are unique up to a permutation of community labels.*

Theorem SM.1 establishes that the 2PPCN model is well-defined by ensuring its identifiability. The formal proof is provided in the Appendix. Intuitively, an edge between two nodes in the partial correlation network may arise either from their shared memberships in multiple communities or from a high probability of connection between distinct communities. This ambiguity—where multiple latent structures can explain the same observed link—is a challenge unique to mixed membership models, and does not arise in models with pure membership structures. Related work on this topic includes the identifiability of overlapping stochastic block models, as shown in Latouche et al. (2011), and the identifiability of mixed membership models under spectral conditions, as established in Zhang et al. (2020).

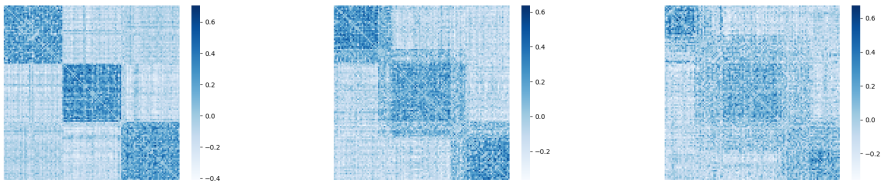
Appendix B.2 More on Simulation



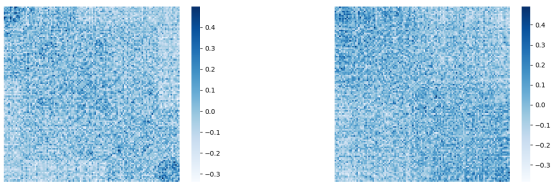
(a) Simulated Λ, $\frac{N_0}{N} = \frac{120}{120}$ (b) Simulated Λ, $\frac{N_0}{N} = \frac{100}{120}$ (c) Simulated Λ, $\frac{N_0}{N} = \frac{80}{120}$



(d) Simulated Λ, $\frac{N_0}{N} = \frac{50}{120}$ (e) Simulated Λ, $\frac{N_0}{N} = \frac{30}{120}$



(f) Estimated $\hat{Q}$, $\frac{N_0}{N} = \frac{120}{120}$ (g) Estimated $\hat{Q}$, $\frac{N_0}{N} = \frac{100}{120}$ (h) Estimated $\hat{Q}$, $\frac{N_0}{N} = \frac{80}{120}$



(i) Estimated $\hat{Q}$, $\frac{N_0}{N} = \frac{50}{120}$ (j) Estimated $\hat{Q}$, $\frac{N_0}{N} = \frac{30}{120}$

Figure 1: The comparisons between heatmaps of adjacency matrices of simulated and estimated correlation network. N_0 is the number of pure membership units which are 100% affiliated with a group. N is the total number of units in simulation. Obviously, as the number of mixed membership units increase, the block structure collapses gradually.

Appendix B.3 More on Application

We apply our methodology to a panel dataset of U.S. state-level payroll employment growth rates. The objective is to estimate the mixed membership structure of all states. The data consist of seasonally adjusted, annualized quarter-to-quarter growth rates in payroll employment. The sample spans from 1956:Q2 to 2007:Q4 and excludes Alaska and Hawaii. The dataset is sourced from the U.S. Bureau of Labor Statistics. In this work, $N = 48$ and $T = 207$. The dataset was originally constructed and used by Hamilton and Owyang (2012), and subsequently employed by Brownlees et al. (2022). The data are seasonally adjusted and annualized. See Hamilton and Owyang (2012) for more details.

Hamilton and Owyang (2012) used the dataset to study business cycle synchronicity, while Brownlees et al. (2022) employed it to investigate pure group structure in panel data. Similarly, our focus is on estimating latent mixed membership structures in the panel. Although their research questions differ from ours, we briefly summarize their methodologies and findings to provide context.

Hamilton and Owyang (2012) applied a cross-validation procedure to estimate the number of groups and found evidence supporting the presence of three distinct communities in the same dataset. Based on their result, we take the number of groups as given in this study and do not implement a separate procedure to estimate it. Methodologically, Hamilton and Owyang (2012) proposed a Markov-switching model in which U.S. states are grouped according to the timing of their business cycles. Using Bayesian inference techniques, such as Markov Chain Monte Carlo (MCMC), they estimated the posterior distribution over group assignments. In contrast, Brownlees et al. (2022) assumed that the partial correlation matrix follows a stochastic block model (SBM) structure and applied spectral clustering to identify discrete group memberships. Unlike the posterior-based approach, the method in Brownlees et al. (2022) yields point estimates of community structure.

In Brownlees et al. (2022), each state is exclusively assigned to one of three groups, reflecting a pure membership assumption. In contrast, Hamilton and Owyang (2012) assign to each state a posterior probability distribution over the three groups; notably, some states receive either high or low probabilities across all groups, reflecting uncertainty in classification. Our methodology, like Brownlees et al. (2022), produces point estimates. However, the structure of the results more closely resembles that of Hamilton and Owyang (2012), in the sense that we obtain mixed membership vectors, where each element lies within $[0, 1]$ and the entries sum to one. However, in Hamilton and Owyang (2012), the sum of posteriors across groups is not necessarily equal to one. The key distinction is that, unlike Brownlees et al. (2022), who assume each state belongs entirely to a single group, our model allows for partial group affiliation—for example, a state may belong 50% to one group and 50% to another. Compared to Hamilton and Owyang (2012), our approach does not yield posterior probabilities but instead directly estimates deterministic mixed membership vectors, which provide a clear and interpretable summary of each state's affiliation across groups.

We present the results of applying the PartialCorr-mixed-SCORE algorithm to the full sample in Figure Table 1. The number of communities is fixed at 3, following Hamilton and Owyang (2012) and Brownlees et al. (2022). Hamilton and Owyang (2012) identified three distinct groups of states based on business cycle synchronicity: oil-producing, manufacturing, and financial states. Brownlees et al. (2022), using a different methodological approach, uncovered a similar but not identical pattern: states were clustered into East and West Coast, oil-producing and agricultural, and Midwest groups. While their results overlap to some extent, they are not identical and can be interpreted as complementary perspectives on the latent structure of state-level economic dynamics. Our findings align with both of these earlier results but also exhibit meaningful differences. Table 1 shows that the states

cluster into three main groups: (1) Financial, service, and other diversified sectors, (2) Oil-producing and agricultural, and (3) Manufacturing. Several key patterns emerge: First, most states exhibit mixed memberships, indicating that they do not rely exclusively on a single industry. This observation is consistent with real-world economic complexity. For example, New York plays major roles in both the financial and manufacturing sectors, while Texas depends heavily on both oil and manufacturing. Second, East and West Coast states tend to have higher membership weights in the financial and service group, reflecting their dominance in finance, tech, and professional services. Third, Midwestern and Southern states such as Texas and Louisiana show strong affiliations with the oil-producing and agricultural group. Interestingly, Maine, in the East coast, which relies heavily on the fishing industry, also clusters with this group. Finally, traditional manufacturing centers like Indiana and Michigan have significant membership in the manufacturing group, but they also show notable affiliation with the financial and service group – highlighting a long-term transition in these regions from industrial production to service-oriented economic activities.

In summary, while our groupings share elements with those found in Hamilton and Owyang (2012) and Brownlees et al. (2022), the distinctions lie in our mixed membership formulation, which provides a richer, more flexible depiction of states' economic identities. These three sets of findings, taken together, offer complementary insights into the evolving industrial geography of the United States.

States	Group 1	Group 2	Group 3
Alabama	0.0	0.456	0.544
Arizona	0.0	0.0	1.0
Arkansas	0.215	0.281	0.504
California	0.848	0.032	0.12

Colorado	0.286	0.247	0.467
Connecticut	0	0.797	0.203
Delaware	0.357	0.494	0.15
Florida	0.112	0.732	0.155
Georgia	0.921	0	0.079
Idaho	0	1.0	0
Illinois	0	0	1.0
Indiana	0.34	0	0.66
Iowa	0.806	0.169	0.025
Kansas	0.789	0	0.211
Kentucky	0	0	1.0
Louisiana	0.247	0.48	0.273
Maine	0.046	0.954	0
Maryland	0	0.762	0.238
Massachusetts	0.736	0	0.264
Michigan	0.602	0	0.398
Minnesota	0.327	0.368	0.305
Mississippi	0.938	0.062	0
Missouri	0.129	0.871	0
Montana	0.758	0.174	0.068
Nebraska	0.477	0	0.523
Nevada	0	0.815	0.185
New Hampshire	0.121	0.879	0
New Jersey	0.777	0	0.223
New Mexico	0.025	0	0.975

New York	0.524	0	0.476
North Carolina	0	0.008	0.992
North Dakota	0.212	0	0.788
Ohio	0	1.0	0
Oklahoma	0.332	0.326	0.342
Oregon	0.011	0.218	0.772
Pennsylvania	1.0	0	0
Rhode Island	0	0.262	0.738
South Carolina	0.4	0.515	0.084
South Dakota	0.088	0.165	0.747
Tennessee	0.775	0.105	0.12
Texas	0.062	0.6	0.339
Utah	0.99	0.01	0
Vermont	0	0	1.0
Virginia	1.0	0	0
Washington	0.37	0	0.63
West Virginia	0.573	0	0.427
Wisconsin	0.667	0.306	0.027
Wyoming	0.404	0	0.596

Appendix B.4 Determining the Number of Communities

The algorithm presented in this paper requires specifying the number of communities. Therefore, we introduce a heuristic method to assist with this selection. However, a detailed theoretical analysis is beyond the scope of this work. This method is an extension of the non-backtracking approach proposed by Le and Levina (Le and Levina, 2015; Li et al., 2022). The overlapping part is treated approximately as a non-overlapping block detection problem, aiming to identify the number of significant diagonal blocks in the adjacency matrix.

Input: a panel, Y_t for $t = 1, \cdots, T$.

1. Estimate the partial correlation network $\hat{Q}$ using OLS.
2. Try different K from 2 onwards, apply K-means to $\hat{Q}$ and reorder series based on the results.
3. With a pre-specified threshold γ_K, get A_K by setting all elements of $\hat{Q} > \gamma_K$ to 1; otherwise, to 0.
4. Define $B_{nb,K} = \begin{bmatrix} 0 & D_K - I \\ -I & A_K \end{bmatrix}$ where A_K is the result from last step, D_K is the degrees of A_K and I is identity matrix.
5. Find the number of eigenvalues, $\hat{K}$, such that $|\lambda_i| > \|B_{nb}\|^{1/2}$ where λ_i is i-th eigenvalue of $\|B_{nb}\|^{1/2}$.
6. Check if or not the pre-specified K is equal to $\hat{K}$ and find the maximal $\hat{K}$ such that $\hat{K} \neq K$.

Output: return $\hat{K} - 1$, the estimated number of communities.

B.0 Data Availability Statement

References

Brownlees, C., Guðmundsson, G. S., and Lugosi, G. (2022). Community detection in partial correlation network models. Journal of Business & Economic Statistics, 40(1):216–226.

Guðmundsson, G. S. and Brownlees, C. (2021). Detecting groups in large vector autoregressions. Journal of Econometrics, 225(1):2–26.

Hamilton, J. D. and Owyang, M. T. (2012). The propagation of regional recessions. Review of Economics and Statistics, 94(4):935–947.

Jin, J., Ke, Z. T., and Luo, S. (2024). Mixed membership estimation for social networks. Journal of Econometrics, 239(2):105369.

Latouche, P., Birmelé, E., and Ambroise, C. (2011). Overlapping stochastic block models with application to the french political blogosphere. Annals of Applied Statistics.

Le, C. M. and Levina, E. (2015). Estimating the number of communities in networks by spectral methods. Electronic Journal of Statistics.

Li, T., Lei, L., Bhattacharyya, S., Van den Berge, K., Sarkar, P., Bickel, P. J., and Levina, E. (2022). Hierarchical community detection by recursive partitioning. Journal of the American Statistical Association, 117(538):951–968.

Rohe, K., Chatterjee, S., and Yu, B. (2011). Spectral clustering and the high-dimensional stochastic blockmodel. Annals of Statistics.

Tang, M., Sussman, D. L., and Priebe, C. E. (2013). Universally consistent vertex classification for latent positions graphs. Annals of Statistics.

Zhang, Y., Levina, E., and Zhu, J. (2020). Detecting overlapping communities in networks using spectral methods. SIAM Journal on Mathematics of Data Science, 2(2):265–283.