Estimation of covariance and precision matrices for high-dimensional time series with long-memory

by

Qihu Zhang

(Under the Direction of Cheolwoo Park)

Abstract

The dissertation deals with the estimation of covariance and precision matrices for high-dimensional time series with long-memory. In Chapter 2, we generalize part of the results of [35] (i) from the spectral norm to the general vector norm induced matrix $\ell_{v,w}$ norm $\|\cdot\|_{(v,w)}$ for any $v, w \in [1, \infty]$, (ii) from the Frobenius norm to the general entrywise matrix $L^{v,w}$ norm $\|\cdot\|_{L^{v,w}}$ for any $v, w \in [1, \infty]$, and (iii) from $p \geq n^c$ for some constant $c > 0$ to $p \geq (n/g_2)^c$ for some constant $c > 0$, where g_2 is an upper bound of $\max_{1 \leq k \leq p} \left\| (\rho^{ij}_{[k]})_{n \times n} \right\|_2$. We also generalize their minimax result by removing the restriction of $p \geq n^\beta$ for some $\beta > 1$. In particular, we obtain the minimax result for the convergence rate of the precision matrix estimator proposed by [11]. In Chapter 3, based on the results of [35], we investigate the joint estimation of multiple precision matrices. We generalize the results of [27] from i.i.d. data to long-memory data. Especially, we obtain the estimation of the entrywise L^1 norm and the Frobenius norm of risk, and their expectations. Our numerical experiment results support our theory analysis. In Chapter 4, based on the results of [35], we investigate the joint estimation of weighted multiple precision matrices. We generalize the results of [16] from i.i.d. data to long-memory data. Especially, we obtain the estimation of the entrywise L^1 norm and the Frobenius norm of risk, and their expectations. Our numerical experiment results support our theory analysis. In Chapter 5, based on the results of [35], we introduce a new assumption to investigate the joint estimation of multiple precision matrices, and generalize the results of [27] from i.i.d. data to long-memory data. Especially, we obtain the estimation of the vector norm induced matrix ℓ_1 norm and the Frobenius norm of risk, and their expectations.

INDEX WORDS: High-dimensional data, Long-memory, Covariance matrices, Precision matrices, Convergence rates

Estimation of covariance and precision matrices for high-dimensional time series with long-memory

by

Qihu Zhang

B.A., Lanzhou University, China, 1994

M.S., Lanzhou University, China, 1997

Ph.D, Lanzhou University, China, 2002

M.S., University of Georgia, 2018

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2020

Estimation of covariance and precision matrices for high-dimensional time series with long-memory

by

Qihu Zhang

Major Professor: Cheolwoo Park

Committee: Shuyang Bai
Yuan Ke
Ping Ma

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
August 2020

Acknowledgments

I would like to express my deepest appreciation to my major Prof. Cheolwoo Park for his instruction and support of my Ph.D study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this dissertation.

Besides my major Prof., I would like to thank the rest of my advisory committee members: Prof. Shuyang Bai, Prof. Yuan Ke, and Prof. Ping Ma, for their encouragement, insightful comments, and hard questions.

I would also like to thank Prof. Hai Shu from New York University, for his patience and detailed instruction about his interesting paper, especially how to conduct data analysis about long-memory data.

I would also like to thank my classmates Jinyang Chen and Shengjie Min, they provide me technical support.

This work was partly supported by National Science Foundation (NSF IIS-1607919).

Table of Contents

List of Figures

List of Tables

Chapter 1

Overview

1.1 Introduction

Multivariate time series data can be obtained in various fields, for example, prices in stock markets, traffic flows on highways, outputs of solar power plants, and functional magnetic resonance imaging (fMRI) data. One of main interests in multivariate time series analysis is identifying relationships among multiple time series. For this purpose, the estimation of covariance and precision (inverse of covariance matrix) matrices plays a key role in many areas of statistical analysis such as graphical models, principal component analysis (PCA), discriminant analysis, and canonical correlation analysis.

Let $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ be a sample of p-dimensional random vectors with mean $\boldsymbol{\mu}_p$, covariance matrix $\boldsymbol{\Sigma}$ and precision matrix $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$. The estimation of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Omega}$ for high-dimensional data has gained much attention in statistical and machine learning communities (see [1, 9, 17, 23, 25, 30, 38]). However, most of the studies have been focused on independent and identically distributed (i.i.d.) observations, and few studies have been done for serially correlated data, especially for long-memory time series data. For example, for resting-state fMRI data, the number of voxels in a brain (p) is typically greater than the number of subjects (n). Also, some voxel time series in a resting brain are known to exhibit a long-memory property (see Figures 3.1-3.3); also see [10] and [32] for example.

Recently, Shu and Nan (see [35]) investigated the estimation of high-dimensional covariance and precision matrices from long-memory time series. They extended the definition of long-memory to multivariate time series, and proposed a new and straightforward temporal dependence measure that solely depends on the Frobenius norm and the spectral norm of the autocorrelation matrices of $\{\mathbf{X}_t\}_{t=1}^n$. Hence, long-memory multivariate time series can be well dealt with based on their relaxed assumption.

Motivated by [35], this dissertation focuses on the estimation of covariance and precision matrices from high-dimensional long-memory time series.

1.2 Useful bounds

We borrow formulation of the problems and the notation from [35] for easy comparisons with their results. The following norms are used in this dissertation. For a real matrix $\mathbf{A} = (a_{ij})$, the spectral norm is defined as $\|\mathbf{A}\|_2 = [\varphi_{\max}(\mathbf{A}^\top\mathbf{A})]^{1/2}$ where $\varphi_{\max}$ is the largest eigenvalue, and φ_k and $\varphi_{\min}$ are the k-th and the smallest eigenvalues of $\mathbf{A}$, respectively. The Frobenius norm is defined as $\|\mathbf{A}\|_F = (\sum_i \sum_j a_{ij}^2)^{1/2}$, the entrywise L^1 norm $|\mathbf{A}|_1 = \sum_{i,j} |a_{ij}|$, and its off-diagonal version $|\mathbf{A}|_{1,\text{off}} = \sum_{i\neq j} |a_{ij}|$. The entrywise $L^{v,w}$ norm is defined as $\|\mathbf{A}\|_{L^{v,w}} = \left(\sum_{j=1}^n \left(\sum_{i=1}^m |a_{ij}|^v\right)^{w/v}\right)^{1/w}$, and the entrywise L^∞ norm $|\mathbf{A}|_\infty = \max_{i,j} |a_{ij}|$. The vector norm induced matrix ℓ_w norm is defined as $\|\mathbf{A}\|_w = \sup\limits_{\mathbf{x}\neq\mathbf{0}} \frac{|\mathbf{A}\mathbf{x}|_w}{|\mathbf{x}|_w}$, and vector norm induced matrix $\ell_{v,w}$ norm $\|\mathbf{A}\|_{(v,w)} = \sup\limits_{\mathbf{x}\neq\mathbf{0}} \frac{|\mathbf{A}\mathbf{x}|_w}{|\mathbf{x}|_v}$. Also, denote $\text{vec}(\mathbf{A}) = (\mathbf{A}_1^\top, \cdots, \mathbf{A}_n^\top)^\top$ where $\mathbf{A}_j$ is the j-th column of $\mathbf{A}$, and denote $\mathbf{A} \succ 0$ if $\mathbf{A}$ is positive definite.

Let $\mathbf{X}_{p\times n} = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$. We assume that each column $\mathbf{x}_i$ follows a distribution with the same covariance matrix $\mathbf{\Sigma} = (\sigma_{kl})_{p\times p}$. Let $\mathbf{x}_{[1]}, \cdots, \mathbf{x}_{[p]}$ be the p row vectors of $\mathbf{X}_{p\times n}$, and $\mathbf{R}_{[k]} = (\rho_{[k]}^{ij})_{n\times n}$ be the correlation matrix of $\mathbf{x}_{[k]}$, i.e., the autocorrelation matrix of the kth time series. For all k, the following inequalities are valid:

$$1 \leq \frac{1}{n} \left\|\mathbf{R}_{[k]}\right\|_F^2 \leq \left\|\mathbf{R}_{[k]}\right\|_2 \leq \left\|\mathbf{R}_{[k]}\right\|_1 \leq n. \tag{1.2.1}$$

Define $g_F(n)$ and $g_2(n)$ as

$$\max_{1\leq k\leq p} \frac{1}{n} \left\|\mathbf{R}_{[k]}\right\|_F^2 \leq g_F(n), \;\; \max_{1\leq k\leq p} \left\|\mathbf{R}_{[k]}\right\|_2 \leq g_2(n). \tag{1.2.2}$$

We note that $1 \leq g_F(n) \leq g_2(n) \leq n$ by (1.2.1).

A multivariate time series is "long-memory" (see [35]) if and only if

$$\max_{1\leq k\leq p} \left\|\mathbf{R}_{[k]}\right\|_1 = \infty \text{ as } n \to \infty. \tag{1.2.3}$$

In [35], Shu and Nan considered a Polynomial-dominated decay (PDD) model with parameter $\alpha > 0$. The model is long-memory if and only if $\alpha \leq 1$. They showed that the convergence rates of the covariance/precision estimators they considered were characterized by the bounds $g_F(n)$ and $g_2(n)$. In this dissertation, we also consider the following long-memory model, which is not covered by PDD in [35].

Definition 1.2.1. *(Log-Polynomial-dominated decay (LPDD) model)* $\mathbf{X}_{p\times n}$ *has LPDD temporal dependence if*

$$\max_{1\leq k\leq p} \left|\rho_{[k]}^{ij}\right| \leq C_0(\log|i-j|)^{-\alpha} \textit{ for all } |i-j|>1, \alpha>0.$$

For the LPDD model, the generalized harmonic number (see [15]) is given as

$$H_n^{(\alpha)} = \sum_{k=2}^{n} (\log k)^{-\alpha} < 1 + \frac{n}{(\log n)^{\alpha}}. \tag{1.2.4}$$

Note that the model always allows at least one individual time series to be long-memory.

1.3 Data generating mechanism

As in [35], we assume that the vectorized data $\mathbf{X}_{pn} = \text{vec}(\mathbf{X}_{p\times n})$ are obtained through the following linear spatio-temporal model

$$\mathbf{X}_{pn} = \mathbf{H}\mathbf{z} + \boldsymbol{\mu}_{pn}, \tag{1.3.1}$$

where $\mathbf{H} = (h_{ij})_{pn\times m}$ is a real non-random matrix, $\boldsymbol{\mu}_{pn} = \mathbf{1}_n \otimes \boldsymbol{\mu}_p$ where $\mathbf{1}_n = (1, 1, \cdots, 1)^\top$ with length n, $\otimes$ denotes the Kronecker product, and $\mathbf{z} = (z_1, \cdots, z_m)^\top$ consists of m independent random variables with $\mathbb{E}(z_i) = 0$ and $\mathbb{E}(z_i^2) = 1$ for $i = 1, \cdots, m$. We use $\mathbf{X}_{p\times n}$ and $\mathbf{X}_{pn}$ interchangeably.

1.4 Three types of moment conditions

We consider the following three types of moment conditions for the random variables $z_1, \cdots, z_m$ in (1.3.1) as in [35]. Let W be a random variable, and K, ϑ and η_k be positive constants.

(C1) Sub-Gaussian tails: For all $k \geq 1$, $(\mathbb{E}|W|^k)^{1/k} \leq Kk^{1/2}$.

(C2) Generalized sub-exponential tails: For some $\vartheta \in (0, 2)$ and all $k \geq \vartheta$, $(\mathbb{E}|W|^k)^{1/k} \leq K(k/\vartheta)^{1/\vartheta}$.

(C3) Polynomial-type tails: For some $k \geq 4$, $(\mathbb{E}|W|^k)^{1/k} \leq \eta_k$.

1.5 The organization of this dissertation

In Chapter 2, we generalize part of the results in [35], (i) from the spectral norm to the general vector norm induced matrix $\ell_{v,w}$ norm $\|\cdot\|_{(v,w)}$ for any $v, w \in [1, \infty]$, (ii) from the Frobenius norm to the general entrywise matrix $L^{v,w}$ norm $\|\cdot\|_{L^{v,w}}$ for any $v, w \in [1, \infty]$, and (iii) from $p \geq n^c$ for some constant $c > 0$ to $p \geq (n/g_2)^c$ for some constant $c > 0$, where g_2 is an upper bound of $\max_{1 \leq k \leq p} \left\|(\rho^{ij}_{[k]})_{n \times n}\right\|_2$. We also generalize their minimax result by removing the restriction of $p \geq n^\beta$ for some $\beta > 1$. In particular, we obtain the minimax result for the convergence rate of the precision matrix estimator proposed by [11].

In Chapter 3, based on the results of [35], we investigate the joint estimation of multiple precision matrices. We generalize the results of [27] from i.i.d. data to long-memory data. Especially, we obtain the estimation of the entrywise L^1 norm and the Frobenius norm of risk, and their expectations. Our numerical experiment results support our theory analysis.

In Chapter 4, based on the results of [35], we investigate the joint estimation of weighted multiple precision matrices. We generalize the results of [16] from i.i.d. data to long-memory data. Especially, we obtain the estimation of the entrywise L^1 norm and the Frobenius norm of risk, and their expectations. Our numerical experiment results support our theory analysis.

In Chapter 5, based on the results of [35], we introduce a new assumption to investigate the joint estimation of multiple precision matrices, and generalize the results of [27] from i.i.d. data to long-memory data. Especially, we obtain the estimation of the vector norm induced matrix ℓ_1 norm and the Frobenius norm of risk, and their expectations.

Chapter 2

Estimation of covariance and precision matrices for high-dimensional time series with long-memory

In this chapter, we focus on the estimation of covariance and precision matrices for high-dimensional long-memory time series. In [18], Ding *et al.* proposed the rotational invariant $L^{2,1}$-norm (they call it R_1-norm), $\|\mathbf{M}\|_{R_1} = \|\mathbf{M}\|_{L^{2,1}} = \sum_j \left(\sum_i m_{ij}^2\right)^{1/2}$, for the objective functions of PCA, which makes R_1-PCA robust to outliers and rotational invariant. This motivates us to generalize part of the results in [35] to different matrix norms. In this chapter, we generalize part of the results of [35], (i) from spectral norm to general vector norm induced matrix $\ell_{v,w}$ norm $\|\cdot\|_{(v,w)}$ for any $v, w \in [1, \infty]$, (ii) from Frobenius norm to general entrywise matrix $L^{v,w}$ norm $\|\cdot\|_{L^{v,w}}$ for any $v, w \in [1, \infty]$, and (iii) from $p \geq n^c$ for some constant $c > 0$ to $p \geq (n/g_2)^c$ for some constant $c > 0$, where g_2 is an upper bound of $\max_{1\leq k\leq p} \left\|(\rho_{[k]}^{ij})_{n\times n}\right\|_2$. We also generalize their minimax result by removing the restriction of $p \geq n^\beta$ for some $\beta > 1$. In particular, we obtain the minimax result for the convergence rate of the precision matrix estimator proposed by [11].

2.1 Estimation of covariance and correlation matrices for sub-Gaussian data

In this section, we extend the results in [35] for the generalized thresholding covariance matrix estimators to various matrix norms for sub-Gaussian data in (C1).

As in [35], let us consider the ℓ_q-ball sparse covariance matrices

$$\mathcal{U}(q, c_p, \upsilon_0) = \left\{ \mathbf{\Sigma} : \max_{1\leq i\leq p} \sum_{j=1}^{p} |\sigma_{ij}|^q \leq c_p, \max_{1\leq i\leq p} \sigma_{ii} \leq \upsilon_0 \right\},$$

where $\upsilon_0 > 0$ and $0 \leq q < 1$. Let $\widehat{\mathbf{\Sigma}} := (\widehat{\sigma}_{ij})_{p\times p}$ be the sample covariance matrix of observation $\mathbf{X}_{p\times n}$ given by

$$\widehat{\mathbf{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T - \overline{\mathbf{x}}\overline{\mathbf{x}}^T \tag{2.1.1}$$

where $\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i$, $\mathbf{x}_i$ is the ith column of $\mathbf{X}$.

For the covariance matrix estimation in high dimensional analysis, [35] considered the generalized thresholding estimators of $\mathbf{\Sigma}$, $S_\tau(\widehat{\mathbf{\Sigma}}) = (s_\tau(\widehat{\sigma}_{ij}))_{p\times p}$, where $s_\tau : \mathbb{R} \to \mathbb{R}$ is a generalized thresholding function with thresholding parameter $\tau \geq 0$, and it satisfies the following conditions for all $z \in \mathbb{R}$: (i) $|s_\tau(z)| \leq |z|$; (ii) $s_\tau(z) = 0$ for $|z| \leq \tau$; (iii) $|s_\tau(z) - z| \leq \tau$.

Define

$$u_1 = \max\left\{(\log p) g_2(n)/n, \left[(\log p) g_F(n)/n\right]^{1/2}\right\}, \tag{2.1.2}$$

and we assume $u_1 \to 0$ as $n \to \infty$. The following theorem generalizes Theorem 1 of [35] from the spectral and the Frobenius norms to various matrix norms based upon a mild assumption $p \geq (n/g_2)^\varepsilon$ for some constant $\varepsilon > 0$.

Theorem 2.1.1. *Assume (i) $\mathbf{X}_{p\times n}$ is generated from (1.3.1) and all z_i satisfy condition $(C1)$ with the same K; (ii) $\mathbf{\Sigma} \in \mathcal{U}(q, c_p, \upsilon_0)$ and $\{\mathbf{R}_{[k]}\}_{k=1}^{p}$ subject to (1.2.2); (iii) $u_1 = o(1)$ with u_1 defined in (2.1.2). Then for sufficiently large constant $M_1 = M_1(\upsilon_0, K, \varepsilon, q) > 0$ with $\tau = M_1 u_1$, uniformly on $\mathbf{\Sigma} \in \mathcal{U}(q, c_p, \upsilon_0)$, we have*

(i) If $p \geq (n/g_2)^\varepsilon$ for some constant $\varepsilon > 0$, then with sufficiently large M_1 additionally depending on ε and q, we have

$$\begin{aligned}
\mathbb{E}\left(\left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^2\right) &= O(u_1^2),\\
\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{(v,w)}^2\right) &= O(c_p^2 u_1^{2-2q} \max\{p^{2/w-2/v}, 1\}), \forall v, w \in [1, \infty],\\
\frac{1}{p^{2/w}}\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{L^{v,w}}^2\right) &= O(c_p^{2/v} u_1^{2-2q/v}), \forall v, w \in [1, \infty].
\end{aligned}$$

(ii) If $p < (n/g_2)^\varepsilon$ for any constant $\varepsilon > 0$, then with sufficiently large M_1 additionally depending on c_* and q, we have

$$\begin{aligned} \mathbb{E}\left(\left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^2\right) &= O(p^{-c_*}), \\ \mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{(v,w)}^2\right) &= O(p^{-c_*}), \forall v, w \in [1,\infty], \\ \frac{1}{p^{2/w}}\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{L^{v,w}}^2\right) &= O(p^{-c_*}), \forall v, w \in [1,\infty]. \end{aligned}$$

Moreover, if we also have $p \geq f(n/g_2)$, where $f(\cdot)$ is positive nondecreasing and $f(t) \to \infty$ as $t \to \infty$, then

$$\begin{aligned} \mathbb{E}\left(\left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^2\right) &= O\left(\left(f(\frac{1}{u_1})\right)^{-c_*}\right), \\ \mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{(v,w)}^2\right) &= O\left(\left(f(\frac{1}{u_1})\right)^{-c_*}\right), \forall v, w \in [1,\infty], \\ \frac{1}{p^{2/w}}\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{L^{v,w}}^2\right) &= O\left(\left(f(\frac{1}{u_1})\right)^{-c_*}\right), \forall v, w \in [1,\infty]. \end{aligned}$$

Proof of Theorem 2.1.1. By Riesz–Thorin interpolation theorem (see [4, page 1], [12, page 2406]), we have

$$\|\mathbf{A}\|_w \leq \max\{\|\mathbf{A}\|_1, \|\mathbf{A}\|_2, \|\mathbf{A}\|_\infty\} \text{ for all } w \in [1,\infty).$$

If $\mathbf{A}$ is symmetric, we have

$$\|\mathbf{A}\|_w \leq \|\mathbf{A}\|_1 \text{ for all } w \in [1,\infty] \tag{2.1.3}$$

since $\|\mathbf{A}\|_1 = \|\mathbf{A}\|_\infty, \|\mathbf{A}\|_2 \leq \sqrt{\|\mathbf{A}\|_1 \|\mathbf{A}\|_\infty} = \|\mathbf{A}\|_1$ (see [21, Corollary 2.3.2]). Also, $|\mathbf{x}|_w \leq |\mathbf{x}|_v$ for any $v, w \in [1,\infty]$ with $v \leq w$.

If $\mathbf{A}_{p\times p}$ is symmetric, it can be shown by Hölder inequality that

$$\begin{aligned} \|\mathbf{A}\|_{(v,w)} &= \sup_{\mathbf{x}\in\mathbb{R}^p\setminus\{0\}} \frac{|\mathbf{A}\mathbf{x}|_w}{|\mathbf{x}|_v} = \sup_{\mathbf{x}\in\mathbb{R}^p\setminus\{0\}} \frac{|\mathbf{A}\mathbf{x}|_w}{|\mathbf{x}|_w}\frac{|\mathbf{x}|_w}{|\mathbf{x}|_v} \leq \|\mathbf{A}\|_w \sup_{\mathbf{x}\in\mathbb{R}^p\setminus\{0\}} \frac{|\mathbf{x}|_w}{|\mathbf{x}|_v} \\ &\leq \|\mathbf{A}\|_w \max\{p^{1/w-1/v}, 1\} \leq \|\mathbf{A}\|_1 \max\{p^{1/w-1/v}, 1\}, \forall v, w \in [1,\infty]. \end{aligned} \tag{2.1.4}$$

From the proof of Theorem 1 of [35],

$$\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_1 = O_P(c_p u_1^{1-q}).$$

Thus, we have

$$\begin{aligned}\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_{(v,w)} &\leq \left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_1 \max\{p^{1/w-1/v}, 1\} \\ &= O_P(c_p u_1^{1-q} \max\{p^{1/w-1/v}, 1\}), \forall v, w \in [1, \infty].\end{aligned}$$

We recap the proof of Theorem 1 of [35]. By [21, (2.3.12) and (2.3.7)],

$$\begin{aligned}\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_1 &\leq \sqrt{p}\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_2 \leq \sqrt{p}\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_F, \\ \left|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right|_\infty &\leq \left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_F.\end{aligned}$$

Similar to the proof of (S.35) of [36], we have

$$\begin{aligned}\mathbb{E}\left(\left|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right|_\infty^2\right) &\leq p\left[\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_F^4\right)\right]^{1/2} O(p^{\frac{-c_1}{2}}) + C_1^2 u_1^2, \\ \mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_1^2\right) &\leq p\left[\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_F^4\right)\right]^{1/2} O(p^{\frac{-c_1}{2}}) + (C_1 c_p u_1^{1-q})^2.\end{aligned}$$

By the above inequalities, (2.1.4) and (S.40) of [36]

$$\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_F^{2d}\right) = O(p^C) \text{ for } d = 1, 2, 3,$$

we have

$$\begin{aligned}\mathbb{E}\left(\left|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right|_\infty^2\right) &\leq O(p^{\frac{c_5-c_1}{2}}) + C_1^2 u_1^2, \\ \mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_{(v,w)}^2\right) &\leq \max\{p^{2/w-2/v}, 1\}\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\boldsymbol{\Sigma}}) - \boldsymbol{\Sigma}\right\|_1^2\right) \\ &\leq O(p^{\frac{c_5-c_1}{2}}) + \max\{p^{2/w-2/v}, 1\}(C_1 c_p u_1^{1-q})^2, \forall v, w \in [1, \infty],\end{aligned}$$

where $C_1 > 0$ is a constant only dependent on M_1, $c_5 > 0$ is a constant, and $c_1 > 0$ is a sufficiently large constant.

(i) Since $p \geq (n/g_2)^c$ for some constant $c > 0$, we have

$$p^{-1/c} \leq \frac{1}{n/g_2} \leq \frac{\log p}{n/g_2} \leq u_1.$$

Thus, for sufficiently large $c_1 > 0$, we have

$$\begin{aligned}\mathbb{E}\left(\left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^2\right) &\leq C_1^2 u_1^2, \\ \mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{(v,w)}^2\right) &\leq (C_1 c_p u_1^{1-q})^2 \max\{p^{2/w-2/v}, 1\}, \forall v, w \in [1, \infty].\end{aligned}$$

It is straightforward to check that

$$\begin{aligned}\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{L^{v,w}} &= \left(\sum_{j=1}^p \left(\sum_{i=1}^p |s_{\tau_1}(\widehat{\sigma}_{ij}) - \sigma_{ij}|^v\right)^{w/v}\right)^{1/w} \\ &\leq \left(\sum_{j=1}^p \left(\sum_{i=1}^p \max_{ij} |s_{\tau_1}(\widehat{\sigma}_{ij}) - \sigma_{ij}|^{v-1} |s_{\tau_1}(\widehat{\sigma}_{ij}) - \sigma_{ij}|\right)^{w/v}\right)^{1/w} \\ &\leq \max_{ij} |s_{\tau_1}(\widehat{\sigma}_{ij}) - \sigma_{ij}|^{(v-1)/v} \left(\sum_{j=1}^p \left(\sum_{i=1}^p |s_{\tau_1}(\widehat{\sigma}_{ij}) - \sigma_{ij}|\right)^{w/v}\right)^{1/w} \\ &\leq \left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^{(v-1)/v} \left(\sum_{j=1}^p \left(\max_j \sum_{i=1}^p |s_{\tau_1}(\widehat{\sigma}_{ij}) - \sigma_{ij}|\right)^{w/v}\right)^{1/w} \\ &= \left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^{(v-1)/v} \left(p\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_1\right)^{w/v}\right)^{1/w} \\ &= \left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^{(v-1)/v} p^{1/w} \left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_1^{1/v}, \qquad (2.1.5)\end{aligned}$$

then by Hölder inequality, we have

$$\begin{aligned}\frac{1}{p^{2/w}}\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{L^{v,w}}^2\right) &\leq \mathbb{E}\left(\left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^{2(v-1)/v} \left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_1^{2/v}\right) \\ &\leq \mathbb{E}\left(\left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^2\right)^{(v-1)/v} \mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_1^2\right)^{1/v} \\ &\leq O(c_p^{2/v} u_1^{2-2q/v}).\end{aligned}$$

(ii) Since $p < (n/g_2)^\varepsilon$ for any constant $\varepsilon > 0$, we have

$$p^{-1/(4\varepsilon)} \geq p^{-1/(2\varepsilon)} \geq p^{-1/\varepsilon} \log p \geq \frac{\log p}{n/g_2}, \text{ as } p \to \infty,$$

and

$$p^{-1/(4\varepsilon)} \geq \sqrt{\frac{\log p}{n/g_2}} \geq \sqrt{\frac{\log p}{n/g_F}}, \text{ as } p \to \infty,$$

then $p^{-1/(4\varepsilon)} \geq u_1$.

Note that $c_p \leq v_0 p$. Thus, for small enough $\varepsilon > 0$, we have

$$p^{-\frac{1-q}{2\varepsilon}+2} = p^2 \left(p^{-1/(4\varepsilon)}\right)^{(1-q)2} \geq (c_p u_1^{1-q})^2, \text{ as } p \to \infty.$$

Therefore, we have

$$\begin{aligned}
\mathbb{E}\left(\left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^2\right) &= O(p^{-c_*}), \\
\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{(v,w)}^2\right) &= O(p^{-c_*}), \forall v, w \in [1, \infty], \\
\frac{1}{p^{2/w}}\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{L^{v,w}}^2\right) &= O(p^{-c_*}), \forall v, w \in [1, \infty].
\end{aligned}$$

Moreover, if we also have $p \geq f(n/g_2)$. Note that $\frac{1}{u_1} \leq \frac{n/g_2}{\log p}$. We have

$$f\left(\frac{1}{u_1}\right) \leq f\left(\frac{n/g_2}{\log p}\right) \leq f\left(n/g_2\right) \leq p,$$

then

$$\begin{aligned}
\mathbb{E}\left(\left|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty^2\right) &= O\left(\left(f(\frac{1}{u_1})\right)^{-c_*}\right), \\
\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{(v,w)}^2\right) &= O\left(\left(f(\frac{1}{u_1})\right)^{-c_*}\right), \forall v, w \in [1, \infty], \\
\frac{1}{p^{2/w}}\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{L^{v,w}}^2\right) &= O\left(\left(f(\frac{1}{u_1})\right)^{-c_*}\right), \forall v, w \in [1, \infty].
\end{aligned}$$

The proof is completed. □

Remark 2.1.1. *(i)* $\|\cdot\|_{(1,1)} = \|\cdot\|_1$ *and* $\|\cdot\|_{L^{2,2}} = \|\cdot\|_F$.

(ii) If $p \geq [\log(n/g_2)]^c$ *for some constant* $c > 0$*, we have*

$$\left(\log\frac{1}{u_1}\right)^c \leq \left(\log\frac{n/g_2}{\log p}\right)^c \leq (\log n/g_2)^c \leq p.$$

Thus,

$$\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{(v,w)}^2\right) = O\left(\left(\log\frac{1}{u_1}\right)^{-c_*}\right), \forall v, w \in [1, \infty].$$

(iii) If $p \geq [\log\log(n/g_2)]^c$ *for some constant* $c > 0$*, we have*

$$\left(\log\log\frac{1}{u_1}\right)^c \leq \left(\log\log\frac{n/g_2}{\log p}\right)^c \leq (\log\log n/g_2)^c \leq p.$$

Thus,

$$\mathbb{E}\left(\left\|S_{\tau_1}(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|^2_{(v,w)}\right) = O\left(\left(\log\log\frac{1}{u_1}\right)^{-c_*}\right), \forall v, w \in [1, \infty].$$

(iv) For LPDD temporal dependence, we can set the bounds as

$$g_F = O(n/(\log n)^{2\alpha}), g_2 = O(n/(\log n)^{\alpha}),$$

then we have the following property

$$u_1 \lesssim (\log p)/(\log n)^{\alpha}, \alpha > 0,$$

by (1.2.4). Here, $x_n \lesssim y_n$ *if* $x_n = O(y_n)$ *as* $n \to \infty$*.*

Theorem 2.1.2. *If we replace* $\widehat{\mathbf{\Sigma}}$*,* $\mathbf{\Sigma}$*,* $\widehat{\sigma}_{ij}$*,* σ_{ij}*,* $\mathcal{U}(q, c_p, v_0)$*,* τ_1*, and* M_1 *by* $\widehat{\mathbf{R}}$*,* $\mathbf{R}$*,* $\widehat{\rho}_{ij}$*,* ρ_{ij}*,* $\mathcal{R}(q, c_p)$*,* τ_2*, and* M_2 *without depending on* v_0*, respectively, Theorems 2.1.1 holds.*

The following theorem generalizes Theorem 3 of [35] from the spectral and the Frobenius norms to various matrix norms based upon a mild assumption of c_p and also by removing the restriction of $p \geq n^{c_1}$ with some constant $c_1 > 1$. Denote $x_n \asymp y_n$ if $x_n = O(y_n)$ and $y_n = O(x_n)$.

Theorem 2.1.3. *(Minimax rates) Assume (i)* $\mathbf{X}_{p\times n}$ *is generated from (1.3.1) and all* z_i *satisfy (C1) with constant* $K \geq K_G$*, where* $K_G = \sup_{k\geq 1}\sqrt{2/k}\left[\Gamma(\frac{1+k}{2})/\sqrt{\pi}\right]^{1/k}$*, and* $\Gamma(\cdot)$ *is the gamma function; (ii)* $\mathbf{\Sigma} \in \mathcal{U}(q, c_p, v_0)$ *and* $\{\mathbf{R}_{[k]}\}_{k=1}^p$ *subject to (1.2.2), (iii)* $u_1 \leq \kappa\sqrt{(\log p)/n}$ *with* $\kappa \geq 1$*. Denote* $\mathcal{P}_1(q, c_p, v_0, g_F, g_2, K, \kappa)$ *as the set of distributions of* $\mathbf{X}_{p\times n}$*, and* $\mathfrak{D}$ *denote the distribution of* $\mathbf{X}_{p\times n}$*. If* $\sqrt{(\log p)/n} = o(1)$*,*

$$c_p = c_{n,p} \leq \min\left\{\frac{1}{2}\epsilon^q_{n,p}p^{\frac{1}{3\beta}}, \frac{C_*}{27}\frac{\log p}{n}\epsilon^{q-3}_{n,p}\right\} \tag{2.1.6}$$

with some constants $\beta > 1$, $C_* = \frac{\beta-1}{(3\beta-1)\beta}$, $\epsilon_{n,p} = \sqrt{(\log p)/n}$, *then for any estimator,* $\widetilde{\mathbf{\Sigma}}$, *we have*

$$\begin{aligned}
\inf_{\widetilde{\mathbf{\Sigma}}} \sup_{\mathfrak{D}\in\mathcal{P}_1} \mathbb{E}_{\mathbf{X}_{p\times n}|\mathfrak{D}} \left(\left| \widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma} \right|_\infty^2 \right) &\asymp u_1^2, \\
\inf_{\widetilde{\mathbf{\Sigma}}} \sup_{\mathfrak{D}\in\mathcal{P}_1} \mathbb{E}_{\mathbf{X}_{p\times n}|\mathfrak{D}} \left(\|\widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_{(v,w)}^2 \right) &\asymp c_p^2 u_1^{2(1-q)} p^{2/w-2/v}, \forall v, w \in [1,\infty] \text{ with } w \le v, \\
\inf_{\widetilde{\mathbf{\Sigma}}} \sup_{\mathfrak{D}\in\mathcal{P}_1} \frac{1}{p^{2/w}} \mathbb{E}_{\mathbf{X}_{p\times n}|\mathfrak{D}} \left(\left\| \widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma} \right\|_{L^{v,w}}^2 \right) &\asymp c_p^{2/v} u_1^{2-2q/v}, \forall v, w \in [1,\infty].
\end{aligned}$$

Additionally if $c_p > 1$, *then for any estimator* $\widetilde{\mathbf{R}}$ *we have*

$$\begin{aligned}
\inf_{\widetilde{\mathbf{R}}} \sup_{\mathfrak{D}\in\mathcal{P}_2} \mathbb{E}_{\mathbf{X}_{p\times n}|\mathfrak{D}} \left(\left| \widetilde{\mathbf{R}} - \mathbf{R} \right|_\infty^2 \right) &\asymp u_1^2, \\
\inf_{\widetilde{\mathbf{R}}} \sup_{\mathfrak{D}\in\mathcal{P}_2} \mathbb{E}_{\mathbf{X}_{p\times n}|\mathfrak{D}} \left(\|\widetilde{\mathbf{R}} - \mathbf{R}\|_{(v,w)}^2 \right) &\asymp c_p^2 u_1^{2(1-q)} p^{2/w-2/v}, \forall v, w \in [1,\infty] \text{ with } w \le v, \\
\inf_{\widetilde{\mathbf{R}}} \sup_{\mathfrak{D}\in\mathcal{P}_2} \frac{1}{p^{2/w}} \mathbb{E}_{\mathbf{X}_{p\times n}|\mathfrak{D}} \left(\left\| \widetilde{\mathbf{R}} - \mathbf{R} \right\|_{L^{v,w}}^2 \right) &\asymp c_p^{2/v} u_1^{2-2q/v}, \forall v, w \in [1,\infty].
\end{aligned}$$

Set $\tau_1 = M_1 u_1$ *and* $\tau_2 = M_2 u_1$ *with sufficiently large positive constants* M_1 *and* M_2, *then the generalized thresholding estimators* $S_{\tau_1}(\widehat{\mathbf{\Sigma}})$ *and* $S_{\tau_2}(\widehat{\mathbf{R}})$ *attain the above minimax optimal rates, respectively.*

Proof of Theorem 2.1.3. Our proof is similar to the proof of Theorem 3 of [35] and Theorem 2 of [12].

For the upper bound, we only need to prove that

$$\inf_{\widetilde{\Sigma}} \max_{\theta\in\Theta} \mathbb{E}_{\mathbf{X}|\theta} \left\| \widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma} \right\|_1^2 \le c_2 c_p^2 u_1^{2(1-q)}, \tag{2.1.7}$$

and (2.1.7) is valid by Theorem 2.1.1.

For the lower bound of $\inf_{\widetilde{\Sigma}} \max_{\theta\in\Theta} \mathbb{E}_{\mathbf{X}|\theta} \|\widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_{(v,w)}^2$ with $v, w \in [1,\infty]$, we only need to calculate $\inf_{\widetilde{\Sigma}} \max_{\theta\in\Theta} \mathbb{E}_{\mathbf{X}|\theta} \left\| \widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma} \right\|_{(v,w)}$ because $\mathbb{E}_{\mathbf{X}|\theta} \left\| \widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma} \right\|_{(v,w)} \le \left(\mathbb{E}_{\mathbf{X}|\theta} \|\widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_{(v,w)}^2 \right)^{1/2}$ by the Lyapunov inequality .

We recap the proof of Theorem 2 of [12]. By applying Lemma 3 of [12] with $s = 1$, we have

$$\inf_{\widetilde{\Sigma}} \max_{\theta\in\Theta} 2\mathbb{E}_{\mathbf{X}|\theta} \left\| \widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma} \right\|_{(v,w)} \ge \alpha \cdot \frac{r}{2} \cdot \min_{1\le i\le r} \left\| \overline{\mathbb{P}}_{i,0} \wedge \overline{\mathbb{P}}_{i,1} \right\|,$$

where $r = \lceil p/2 \rceil$, the smallest integer $\geq p/2$, and

$$\alpha = \min_{\{(\theta,\theta'):H(\gamma(\theta),\gamma(\theta'))\geq 1\}} \frac{\|\mathbf{\Sigma}(\theta) - \mathbf{\Sigma}(\theta')\|_{(v,w)}}{H(\gamma(\theta),\gamma(\theta'))},$$

and

$$\min_{1\leq i\leq r} \left\|\overline{\mathbb{P}}_{i,0} \wedge \overline{\mathbb{P}}_{i,1}\right\| \geq c_2 > 0 \tag{2.1.8}$$

which we will prove below. The notations $\Sigma(\theta), \overline{\mathbb{P}}_{i,0}, \overline{\mathbb{P}}_{i,1}$, H, and γ are the same as in the proof of Theorem 2 in [12].

Similar to the proof of Lemma 5 of [12], we have

$$\|\mathbf{\Sigma}(\theta) - \mathbf{\Sigma}(\theta')\|_{(v,w)} \geq \frac{[H(\gamma(\theta),\gamma(\theta'))]^{1/w} \, k\epsilon_{n,p}}{p^{1/v}}.$$

Note that $H(\gamma(\theta),\gamma(\theta')) \leq p$. Thus

$$\alpha \geq \min_{\{(\theta,\theta'):H(\gamma(\theta),\gamma(\theta'))\geq 1\}} \frac{k\epsilon_{n,p}}{[H(\gamma(\theta),\gamma(\theta'))]^{1-1/w} \, p^{1/v}} \geq \frac{k\epsilon_{n,p}}{p^{1+1/v-1/w}}.$$

Note also that $k = \max\{\lceil \frac{1}{2} c_{n,p}\epsilon_{n,p}^{-q} \rceil - 1, 0\}$ and $v, w \geq 1$. Then,

$$\begin{aligned} \inf_{\widetilde{\Sigma}} \max_{\theta\in\Theta} 2\mathbb{E}_{\mathbf{X}|\theta} \left\|\widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma}\right\|_{(v,w)} &\geq \alpha \cdot \frac{r}{2} \cdot \min_{1\leq i\leq r} \left\|\overline{\mathbb{P}}_{i,0} \wedge \overline{\mathbb{P}}_{i,1}\right\| \\ &\geq \frac{k\epsilon_{n,p}}{p^{1+1/v-1/w}} \frac{p}{4} c_2 \\ &\geq \frac{1}{4} c_2 k\epsilon_{n,p} p^{1/w-1/v} = c_3 (c_p u_1^{1-q}) p^{1/w-1/v}. \end{aligned}$$

Therefore,

$$\inf_{\widehat{\Sigma}} \max_{\theta\in\Theta} \mathbb{E}_{\mathbf{X}|\theta} \left\|\widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma}\right\|_{(v,w)}^2 \geq c_3 c_{n,p}^2 u_1^{2(1-q)} p^{2/w-2/v} = c_3 c_p^2 u_1^{2(1-q)} p^{2/w-2/v}.$$

For the lower bound of $\inf_{\widehat{\Sigma}} \max_{\theta\in\Theta} \frac{1}{p^{2/w}} \mathbb{E}_{\mathbf{X}|\theta} \left\|\widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma}\right\|_{L^{v,w}}^2$, we have

$$\|\mathbf{\Sigma}(\theta) - \mathbf{\Sigma}(\theta')\|_{L^{v,w}} \geq [H(\gamma(\theta),\gamma(\theta'))]^{1/w} k^{1/v} \epsilon_{n,p},$$

by following the proof of Lemma 5 of [12]. Then

$$\inf_{\widetilde{\Sigma}} \max_{\theta\in\Theta} \frac{1}{p^{2/w}} \mathbb{E}_{\mathbf{X}|\theta} \left\|\widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma}\right\|_{L^{v,w}}^2 \geq \left(\inf_{\widetilde{\Sigma}} \max_{\theta\in\Theta} \frac{1}{p^{1/w}} \mathbb{E}_{\mathbf{X}|\theta} \left\|\widetilde{\mathbf{\Sigma}} - \mathbf{\Sigma}\right\|_{L^{v,w}} \right)^2$$

$$\begin{aligned}
&\geq \left(\frac{r\cdot c_2}{2p^{1/w}}\cdot\frac{p^{1/w}k^{1/v}\epsilon_{n,p}}{p}\right)^2\\
&\geq c_4\left(p^{1-1/w}\cdot\frac{c_{n,p}^{1/v}\epsilon_{n,p}^{1-q/v}}{p^{1-1/w}}\right)^2\\
&= c_4c_{n,p}^{2/v}u_1^{2-2q/v}.
\end{aligned}$$

Hence,

$$\inf_{\widetilde{\Sigma}}\max_{\theta\in\Theta}\frac{1}{p^{2/w}}\mathbb{E}_{\mathbf{X}|\theta}\left\|\widetilde{\mathbf{\Sigma}}-\mathbf{\Sigma}\right\|_{L^{v,w}}^2\geq c_4c_{n,p}^{2/v}u_1^{2-2q/v}=c_4c_p^{2/v}u_1^{2-2q/v}.$$

Similarly, for the lower bound of $\inf_{\widetilde{\Sigma}}\max_{\theta\in\Theta}\frac{1}{p}\mathbb{E}_{\mathbf{X}|\theta}\left|\widetilde{\mathbf{\Sigma}}-\mathbf{\Sigma}\right|_\infty^2$, we have

$$\left|\widetilde{\mathbf{\Sigma}}-\mathbf{\Sigma}\right|_\infty^2\geq\epsilon_{n,p}^2.$$

Thus,

$$\inf_{\widetilde{\Sigma}}\max_{\theta\in\Theta}\mathbb{E}_{\mathbf{X}|\theta}\left|\widetilde{\mathbf{\Sigma}}-\mathbf{\Sigma}\right|_\infty^2\geq\min_{\{(\theta,\theta'):H(\gamma(\theta),\gamma(\theta'))\geq1\}}\frac{\epsilon_{n,p}^2}{H(\gamma(\theta),\gamma(\theta'))}rc_2=c_2u_1^2.$$

Then, it only remains to prove (2.1.8), which can be proven by Lemma 8 of [12]. In order to check if Lemma 8 of [12] is valid in our setting, let us investigate Lemma 11 of [12]. According to the proof of Lemma 11 in [12], we only need to assume that

$$\begin{aligned}
2k\epsilon_{n,p} &< 1/3, &(2.1.9)\\
c_{n,p}\epsilon_{n,p}^{3-q} &\to 0 \text{ as } n\to\infty, &(2.1.10)\\
27nk\epsilon_{n,p}^3 &= 27nc_{n,p}\epsilon_{n,p}^{3-q}\leq C_*\log p, &(2.1.11)\\
(1+\frac{1}{\beta})\log p &= \frac{\beta+1}{\beta-1}(1-\frac{1}{\beta})\log p\leq\left(\frac{2\beta}{\beta-1}-1\right)\log\frac{p/8-1-k}{k^2}. &(2.1.12)
\end{aligned}$$

Note that $(\log p)/n\to0$ as $n\to\infty$. Here, (2.1.9) and (2.1.10) must be true if (2.1.11) is satisfied. Thus, we only need to assume that (2.1.11) and (2.1.12) are satisfied simultaneously, i.e.

$$\begin{cases} c_{n,p}\leq\frac{C_*}{27}\frac{\log p}{n}\epsilon_{n,p}^{q-3},\\ 4k\leq 2c_{n,p}\epsilon_{n,p}^{-q}\leq p^{\frac{1}{2\beta}},\end{cases}$$

which we can rewrite it as

$$\begin{cases} c_{n,p}\leq\frac{C_*}{27}\frac{\log p}{n}\epsilon_{n,p}^{q-3},\\ c_{n,p}\leq\frac{1}{2}\epsilon_{n,p}^{q}p^{\frac{1}{2\beta}}.\end{cases}$$

Then, we only need to assume

$$c_{n,p} \le \min\left\{\frac{1}{2}\epsilon_{n,p}^{q} p^{\frac{1}{2\beta}}, \frac{C_*}{27}\frac{\log p}{n}\epsilon_{n,p}^{q-3}\right\}.$$

Note that $u_1 \le \kappa\sqrt{(\log p)/n}$, and then

$$u_1 = \max\left\{(\log p) g_2(n)/n, [(\log p) g_F(n)/n]^{1/2}\right\} \asymp [(\log p)/n]^{1/2} = \epsilon_{n,p}.$$

If $p \ge n^{\beta}$ with some $\beta > 1$, we have $\frac{C_*}{27}\frac{\log p}{n}\epsilon_{n,p}^{q-3} \le \frac{1}{2}\epsilon_{n,p}^{q} p^{\frac{1}{2\beta}}$, and

$$c_{n,p} \le \frac{C_*}{27}\frac{\log p}{n}\epsilon_{n,p}^{q-3}.$$

If $p \le n^{\beta_0}$ with some positive $\beta_0 < \beta(> 1)$, we have $\frac{C_*}{27}\frac{\log p}{n}\epsilon_{n,p}^{q-3} \ge \frac{1}{2}\epsilon_{n,p}^{q} p^{\frac{1}{2\beta}}$ when n is large enough, and

$$c_{n,p} \le \frac{1}{2}\epsilon_{n,p}^{q} p^{\frac{1}{2\beta}}.$$

Note that $C_* = \frac{\beta-1}{(3\beta-1)\beta}$, then it is straightforward to have

$$\begin{aligned}
&\int_{t>\frac{2\beta}{\beta-1}} \exp\left[\log(2p) - (t-1)\log\frac{p/8-1-k}{k^2} + 27n(t+1)k\epsilon_{n,p}^3\right] dt \\
\le\ &\int_{t>\frac{2\beta}{\beta-1}} \exp\left[\log(2p) - (t-1)(1-\frac{1}{\beta})\log p + (t+1)C_*\log p\right] dt \\
=\ &2p^{1+2C_*}\int_{t>\frac{2\beta}{\beta-1}} \exp\left[-(t-1)(1-\frac{1}{\beta}-C_*)\log p\right] dt \\
=\ &2p^{1+2C_*}\frac{\exp\left[-(\frac{2\beta}{\beta-1}-1)\left(1-\frac{1}{\beta}-C_*\right)\log p\right]}{\left(1-\frac{1}{\beta}-C_*\right)\log p} \\
=\ &\frac{2}{\left(1-\frac{1}{\beta}-C_*\right)\log p} \to 0.
\end{aligned}$$

Therefore, Lemma 11 of [12] is valid under our assumptions.

Now, let us check the proof of Lemma 8 of [12]. Since (2.1.12) implies that

$$\frac{k^2}{p_{\lambda_{-1}} - k} \le \frac{k^2}{p/4 - k} \le \frac{1}{p^{1-1/\beta}},$$

and (2.1.9) implies that

$$j\epsilon_{n,p}^2 \le k\epsilon_{n,p}^2 \le \frac{1}{6}\epsilon_{n,p} \to 0 \text{ as } n \to \infty,$$

Lemma 8 of [12] is valid under our assumption.

The proof is completed. □

2.2 Estimation of precision matrix for sub-Gaussian data

In this section, we extend the results in [35] for the precision matrix estimator the constrained entrywise L^1 minimization for inverse matrix estimation (CLIME) proposed by [11], to various norms for sub-Gaussian data in (C1).

As in [35], we consider the following set of precision matrices

$$\mathcal{G}_1(q, c_{*p}, M_p, v_0)=\left\{\mathbf{\Omega} \succ 0 : \max_{1\leq i\leq p} \sum_{j=1}^{p} |\omega_{ij}|^q \leq c_{*p}, \|\mathbf{\Omega}\|_1 \leq M_p, \max_{1\leq i\leq p} \sigma_{ii} \leq v_0\right\},$$

where $0 \leq q < 1$.

To obtain the CLIME estimate, let $\widehat{\mathbf{\Theta}}_{\varepsilon,\lambda} := (\hat{\theta}_{ij}^{(\varepsilon,\lambda)})_{p\times p}$ be a solution to the following optimization:

$$\min |\mathbf{\Theta}|_1 \quad \text{subject to} \quad |\widetilde{\mathbf{\Sigma}}_\varepsilon \mathbf{\Theta} - \mathbf{I}|_\infty \leq \lambda, \quad \mathbf{\Theta} \in \mathbb{R}^{p\times p}, \tag{2.2.1}$$

where λ is a tuning parameter, $\widetilde{\mathbf{\Sigma}}_\varepsilon = \widehat{\mathbf{\Sigma}} + \varepsilon\mathbf{I}$, $\widehat{\mathbf{\Sigma}}$ is given in (2.1.1), $\varepsilon \geq 0$ is a perturbation parameter [11], and $\mathbf{I}$ is the $p\times p$ identity matrix. Because the solution to the above optimization problem does not guarantee symmetry, the final CLIME estimator $\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} := (\hat{\omega}_{ij}^{(\varepsilon,\lambda)})_{p\times p}$ is obtained by the following symmetrization step:

$$\hat{\omega}_{ij}^{(\varepsilon,\lambda)} = \hat{\omega}_{ji}^{(\varepsilon,\lambda)} = \hat{\theta}_{ij}^{(\varepsilon,\lambda)} \mathbb{1}(|\hat{\theta}_{ij}^{(\varepsilon,\lambda)}| \leq |\hat{\theta}_{ji}^{(\varepsilon,\lambda)}|) + \hat{\theta}_{ji}^{(\varepsilon,\lambda)} \mathbb{1}(|\hat{\theta}_{ij}^{(\varepsilon,\lambda)}| > |\hat{\theta}_{ji}^{(\varepsilon,\lambda)}|),$$

where $\mathbb{1}(A)$ is the indicator function of event A, $x_+ = x\mathbb{1}(x \geq 0)$ and $\text{sign}(x) = \mathbb{1}(x \geq 0) - \mathbb{1}(x \leq 0)$.

The following theorem generalizes Theorem 4 of [35] from the spectral and the Frobenius norms to various matrix norms based upon a mild assumption $p \geq (n/g_2)^c$ for some constant $c > 0$.

Theorem 2.2.1. *Assume (i) $\mathbf{X}_{p\times n}$ is generated from (1.3.1) and all z_i satisfy condition (C1) with the same K; (ii) $\mathbf{\Omega} \in \mathcal{G}_1(q, c_{*p}, M_p, v_0)$ and $\{\mathbf{R}_{[k]}\}_{k=1}^{p}$ subject to (1.2.2), (iii) $u_1 = o(1)$ with u_1 defined in (2.1.2); (iv) $p \geq (n/g_2)^c$ and $\min\{p^{-C}, u_1\} \leq \varepsilon \leq u_1$ for some positive constants c and C. Set $\lambda = MM_p u_1$ with sufficiently large constant $M = M(v_0, K, c, C, q) >$*

0, *uniformly on* $\mathbf{\Omega} \in \mathcal{G}_1(q, c_{*p}, M_p, v_0)$, *we have*

$$
\begin{aligned}
\mathbb{E}\left(|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}|_\infty^2\right) &= O\left((M_p^2 u_1)^2\right), \\
\mathbb{E}\left(\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_{(v,w)}^2\right) &= O\left(c_{*p}^2 (M_p^2 u_1)^{2-2q} \max\{p^{2/w-2/v}, 1\}\right), \forall v, w \in [1, \infty], \\
\frac{1}{p^{2/w}} \mathbb{E}\left(\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_{L^{v,w}}^2\right) &= O\left(c_{*p}^{2/v} (M_p^2 u_1)^{2-2q/v}\right), \forall v, w \in [1, \infty].
\end{aligned}
$$

Proof of Theorem 2.2.1. Similar to the proof of Theorem 4 of [35] we have

$$
\begin{aligned}
|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}|_\infty &= O_P(M_p^2 u_1), \\
\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_1 &= O_P\left(c_{*p} (M_p^2 u_1)^{1-q}\right),
\end{aligned}
$$

and then

$$
\begin{aligned}
\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_{(v,w)} &\leq \|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_1 \max\{p^{1/w-1/v}, 1\} \\
&= O_P\left(c_{*p} (M_p^2 u_1)^{1-q} \max\{p^{1/w-1/v}, 1\}\right), \forall v, w \in [1, \infty].
\end{aligned}
$$

Note that

$$
\begin{aligned}
\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}\|_1 &\leq \left|\widehat{\mathbf{\Theta}}_{\varepsilon,\lambda}\right|_1 \leq \left|\widetilde{\mathbf{\Sigma}}_{\varepsilon,\lambda}^{-1}\right|_1 \leq p \|\widetilde{\mathbf{\Sigma}}_{\varepsilon,\lambda}^{-1}\|_1 \leq p^2 \|\widetilde{\mathbf{\Sigma}}_{\varepsilon,\lambda}^{-1}\|_2 \\
&\leq p^2 \varphi_{\max}(\widetilde{\mathbf{\Sigma}}_{\varepsilon,\lambda}^{-1}) = p^2/\varphi_{\min}(\widetilde{\mathbf{\Sigma}}_{\varepsilon,\lambda}) = p^2/[\varphi_{\min}(\widehat{\mathbf{\Sigma}}) + \varepsilon] \\
&\leq p^2/\varepsilon.
\end{aligned}
$$

Similar to the proof of Theorem 2 of [11], we have

$$
\mathbb{E}\left(\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_1^2\right) = O\left(c_{*p}^2 (M_p^2 u_1)^{2-2q}\right) + O\left(p^4 (1/u_1)^2 p^{-\tau}\right).
$$

If $p \geq (n/g_2)^c$ for some $c > 0$, we have $p^4(1/u_1)^2 p^{-\tau} = O\left(c_{*p}^2 (M_p^2 u_1)^{2-2q}\right)$, then

$$
\mathbb{E}\left(\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_1^2\right) = O\left(c_{*p}^2 (M_p^2 u_1)^{2-2q}\right).
$$

By Riesz–Thorin interpolation theorem (see [4, page 1], [12, page 2406]),

$$
\mathbb{E}\left(\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_{(v,w)}^2\right) \leq \mathbb{E}\left(\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_w^2\right) \max\{p^{2/w-2/v}, 1\}
$$

$$
\begin{aligned}
&\leq \max\{p^{2/w-2/v},1\}\mathbb{E}\left(\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}\|_1^2\right)\\
&= O\left(c_{*p}^2(M_p^2u_1)^{2-2q}\max\{p^{2/w-2/v},1\}\right),\forall v,w\in[1,\infty].
\end{aligned}
$$

Note that $|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}|_\infty \leq \|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}\|_1$. Similarly, we have

$$
\mathbb{E}\left(|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}|_\infty^2\right) = O\left((M_p^2u_1)^2\right) + O\left(p^4(1/u_1)^2p^{-\tau}\right).
$$

If $p \geq (n/g_2)^c$ for some $c>0$, we have $p^4(1/u_1)^2p^{-\tau} = O\left(c_{*p}^2(M_p^2u_1)^{2-2q}\right)$, then

$$
\mathbb{E}\left(|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}|_\infty^2\right) = O\left((M_p^2u_1)^2\right).
$$

Similar to (2.1.5), by Hölder inequality, we have

$$
\begin{aligned}
\frac{1}{p^{2/w}}\mathbb{E}\left(\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}\|_{L^{v,w}}^2\right) &\leq \mathbb{E}\left(\left|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}\right|_\infty^{2(v-1)/v}\left\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}\right\|_1^{2/v}\right)\\
&\leq \mathbb{E}\left(\left|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}\right|_\infty^2\right)^{(v-1)/v}\mathbb{E}\left(\left\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda}-\mathbf{\Omega}\right\|_1^2\right)^{1/v}\\
&\leq O\left(c_{*p}^{2/v}(M_p^2u_1)^{2-2q/v}\right).
\end{aligned}
$$

The proof is completed. □

Based on Theorem 2.2.1, we have the following minimax result for the precision matrix estimation.

Theorem 2.2.2. *(Minimax rates) Assume (i)* $\mathbf{X}_{p\times n}$ *is generated from (1.3.1) and all* z_i *satisfy (C1) with same constant* $K \geq K_G$*, where* $K_G = \sup_{k\geq 1}\sqrt{2/k}\left[\Gamma(\frac{1+k}{2})/\sqrt{\pi}\right]^{1/k}$*, and* $\Gamma(\cdot)$ *is the gamma function; (ii)* $\mathbf{\Sigma}\in\mathcal{U}(q,c_p,v_0)$*,* $\mathbf{\Omega}\in\mathcal{G}_1(q,c_{*p},M_p,v_0)$ *and* $\{\mathbf{R}_{[k]}\}_{k=1}^p$ *subject to (1.2.2); (iii)* $u_1 \leq \kappa\sqrt{(\log p)/n}$*. Let* $\mathcal{P}_1(q,c_p,v_0,g_F,g_2,K,\kappa)$ *be the set of distributions of* $\mathbf{X}_{p\times n}$ *generated, and* $\mathfrak{D}$ *denote the distribution of* $\mathbf{X}_{p\times n}$*. If* $\sqrt{(\log p)/n} = o(1)$*,*

$$
c_p = c_{n,p} \leq \frac{1}{2}\epsilon_{n,p}^q p^{\frac{1}{3\beta}},
$$

with some constants $\beta>1$*,* $C_* = \frac{\beta-1}{(3\beta-1)\beta}$*,* $\epsilon_{n,p} = \sqrt{(\log p)/n}$*,* $p\leq n^\xi$ *with* $\xi < \frac{3\beta}{3\beta+1}$*, then for any estimator* $\widetilde{\mathbf{\Omega}}$ *and any* $v,w\in[1,\infty]$*, we have*

$$
u_1^2 \lesssim \inf_{\widetilde{\mathbf{\Omega}}}\sup_{\mathfrak{D}\in\mathcal{P}_1\cap\mathcal{G}_1}\mathbb{E}_{\mathbf{X}_{p\times n}|\mathfrak{D}}\left(\left|\widetilde{\mathbf{\Omega}}-\mathbf{\Omega}\right|_\infty^2\right) \lesssim (M_p^2u_1)^2,
$$

$$c_p^2 u_1^{2(1-q)} p^{2/w-2/v} \lesssim \inf_{\widetilde{\mathbf{\Omega}}} \sup_{\mathfrak{D}\in\mathcal{P}_1\cap\mathcal{G}_1} \mathbb{E}_{\mathbf{X}_{p\times n}|\mathfrak{D}} \left(\|\widetilde{\mathbf{\Omega}}-\mathbf{\Omega}\|_{(v,w)}^2\right) \lesssim c_{*p}^2 (M_p^2 u_1)^{2-2q} \max\{p^{2/w-2/v},1\},$$

$$c_p^{2/v} u_1^{2-2q/v} \lesssim \inf_{\widetilde{\mathbf{\Omega}}} \sup_{\mathfrak{D}\in\mathcal{P}_1\cap\mathcal{G}_1} \frac{1}{p^{2/w}} \mathbb{E}_{\mathbf{X}_{p\times n}|\mathfrak{D}} \left(\|\widetilde{\mathbf{\Omega}}-\mathbf{\Omega}\|_{L^{v,w}}^2\right) = O\left(c_{*p}^{2/v} (M_p^2 u_1)^{2-2q/v}\right).$$

Proof of Theorem 2.2.2. By Theorem 2.2.1, the upper bound is achieved.

For the lower bound of $\inf_{\widehat{\Sigma}}\max_{\theta\in\Theta}\mathbb{E}_{\mathbf{X}|\theta}\|\widetilde{\mathbf{\Omega}}-\mathbf{\Omega}\|_{(v,w)}^2$ with $v,w\in[1,\infty]$, we only need to calculate $\inf_{\widehat{\Sigma}}\max_{\theta\in\Theta}\mathbb{E}_{\mathbf{X}|\theta}\left\|\widetilde{\mathbf{\Omega}}-\mathbf{\Omega}\right\|_{(v,w)}$ because $\mathbb{E}_{\mathbf{X}|\theta}\left\|\widetilde{\mathbf{\Omega}}-\mathbf{\Omega}\right\|_{(v,w)} \le \left(\mathbb{E}_{\mathbf{X}|\theta}\|\widetilde{\mathbf{\Omega}}-\mathbf{\Omega}\|_{(v,w)}^2\right)^{1/2}$ by the Lyapunov inequality.

By the definition of $\mathbf{\Sigma}(\theta)$ in the proof of Theorem 2 in [12], we can partition $\mathbf{\Sigma}(\theta)$ as

$$\mathbf{\Sigma}(\theta) = \begin{pmatrix} \mathbf{I}_{p-r} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{I}_r \end{pmatrix},$$

where $\mathbf{A}_{21} = \mathbf{A}_{12}^T$. By (2.50) of [34], we have

$$\mathbf{\Sigma}(\theta)^{-1} = \begin{pmatrix} \mathbf{I}_{p-r} + \mathbf{A}_{12}\mathbf{B}^{-1}\mathbf{A}_{21} & -\mathbf{A}_{12}\mathbf{B}^{-1} \\ \mathbf{B}^{-1}\mathbf{A}_{21} & \mathbf{B}^{-1} \end{pmatrix},$$

where $\mathbf{B} = \mathbf{I}_r - \mathbf{A}_{21}\mathbf{A}_{12} = \mathbf{I}_r - \mathbf{A}_{12}^T\mathbf{A}_{12}$. Using the definition of $\mathbf{\Sigma}(\theta)$ again, we have

$$\mathbf{B}^{-1} = (\mathbf{I}_r - \mathbf{A}_{12}^T\mathbf{A}_{12})^{-1} = \mathbf{I}_r + \sum_{n=1}^{\infty}(\mathbf{A}_{12}^T\mathbf{A}_{12})^n.$$

It is straightforward to check that every element $a_{ij}^{(n)}$ of $(\mathbf{A}_{12}^T\mathbf{A}_{12})^n$ is nonnegative and

$$a_{ij}^{(n)} \le r^{n-1}\left(2k\epsilon_{n,p}^2\right)^n \le p^{n-1}\left(2k\epsilon_{n,p}^2\right)^n,$$

thus every element a_{ij} of $\sum_{n=1}^{\infty}(\mathbf{A}_{12}^T\mathbf{A}_{12})^n$ is nonnegative and

$$a_{ij} = \sum_{n=1}^{\infty} a_{ij}^{(n)} \le \sum_{n=1}^{\infty} p^{n-1}\left(2k\epsilon_{n,p}^2\right)^n = 2k\epsilon_{n,p}^2 \frac{1}{1-2pk\epsilon_{n,p}^2} \le 3k\epsilon_{n,p}^2,$$

provided

$$2pk\epsilon_{n,p}^2 = 2pc_{n,p}\epsilon_{n,p}^{2-q} = 2p^{\frac{3\beta+1}{3\beta}}\frac{\log p}{n} \to 0.$$

This is guarranted by the assmuption $p \le n^{\xi}$ with $\xi < \frac{3\beta}{3\beta+1}$. Since every element of $\mathbf{A}_{12}$ is nonnegative, we can drop $\sum_{n=1}^{\infty}(\mathbf{A}_{12}^T\mathbf{A}_{12})^n$ in the remaining lower bound computation, and take $\mathbf{I}_r$ as the $\mathbf{B}^{-1}$. Similar to the proof of Theorem 2.1.3 we can obtain this result.

For the lower bound of $\inf_{\widehat{\Sigma}} \max_{\theta \in \Theta} \mathbb{E}_{\mathbf{X}|\theta} \|\widetilde{\mathbf{\Omega}} - \mathbf{\Omega}\|_{L^{v,w}}^2$ with $v, w \in [1, \infty]$, we only need to compute $\inf_{\widehat{\Sigma}} \max_{\theta \in \Theta} \mathbb{E}_{\mathbf{X}|\theta} \left\| \widetilde{\mathbf{\Omega}} - \mathbf{\Omega} \right\|_{L^{v,w}}$ due to the Lyapunov inequality. Hence, we can complete the proof by following a similar path in the proof of Theorem 2.1.3. □

2.3 Extension to heavy tail data

In this section, we extend the results in [35] for the covariance and precision matrices estimators to various norms for heavy-tailed data in (C2) and (C3).

Define

$$u_2 = \max\left\{(\log p)^{1+2/\vartheta} g_2/n, (\log p)^{1+2/\vartheta} (g_F/n)^{1/2}\right\}, \tag{2.3.1}$$

and

$$u_3 = \max\left\{p^{(2+2C)/k} g_2/n, p^{(4+2C)/k} (g_F/n)^{1/2}\right\}. \tag{2.3.2}$$

Theorem 2.3.1. *(Generalized sub-exponential tails) Assume condition (C1), parameter K, and u_1 are replaced by condition (C2), parameters $\{K, \vartheta\}$, and u_2, respectively, then Theorems 2.1.1,2.1.2, and 2.2.1 hold.*

Theorem 2.3.2. *(Polynomial-type tails) Assume condition (C1), parameter K, and u_1 are replaced by condition (C3), parameters $\{k, \eta_k\}$, and u_3, respectively.*

(i) Under the conditions of Theorem 2.1.1, we have

$$\begin{aligned} \left|S_\tau(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right|_\infty &= O_P(u_3), \\ \left\|S_\tau(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{(v,w)} &= O_P(c_p u_3^{1-q} \max\{p^{1/w-1/v}, 1\}), \forall v, w \in [1, \infty], \\ \frac{1}{p^{2/w}} \left\|S_\tau(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}\right\|_{L^{v,w}}^2 &= O_P(c_p^{2/v} u_3^{2-2q/v}), \forall v, w \in [1, \infty]. \end{aligned}$$

(ii) Under the conditions of Theorem 2.2.1, we have

$$|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}|_\infty = O_P\left(M_p^2 u_3\right),$$

$$\|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_{(v,w)} = O_P\left(c_{*p} (M_p^2 u_3)^{1-q} \max\{p^{1/w-1/v}, 1\}\right), \forall v, w \in [1, \infty],$$

$$\frac{1}{p^{2/w}} \|\widehat{\mathbf{\Omega}}_{\varepsilon,\lambda} - \mathbf{\Omega}\|_{L^{v,w}}^2 = O_P\left(c_{*p}^{2/v} (M_p^2 u_3)^{2-2q/v}\right), \forall v, w \in [1, \infty].$$

The proofs of Theorems 2.3.1 and 2.3.2 are similar to those of the preceding theorems by using the corresponding results given in Lemmas A.1, A.2 and A.3 of [35] for conditions (C2) and (C3), respectively. Details are omitted.

Chapter 3

Joint estimation of multiple precision matrices with common structures for high-dimensional time series with long-memory

In [35], Shu and Nan investigated the precision matrix of long-memory rfMRI data for a single subject, provided by the WU-Minn Human Connectome Project (www.humanconnectome.org). Oftentimes, a group analysis rather than a single subject analysis is desirable in fMRI studies to draw a general conclusion. Hence, simultaneous estimation of multiple precision matrices from a group of subjects is of interest. In such case, it is reasonable to expect that the precision matrices of the subjects within a group would share some common characteristics, and also individual precision matrices would have their unique characteristics.

In this chapter, we study the joint estimation of multiple precision matrices rather than the separate estimation of individual precision matrices for long-memory data. In [27], Lee and Liu proposed the joint estimation of multiple precision matrices with common structure (JEMP), but they only obtained the entriwise L^∞ norm of the risk of the common structure for i.i.d. data. Motivated by [27] and [35], this chapter focuses on investigating the properties of the estimated precision matrices by JEMP for long-memory data. Our theory analysis generalizes part of the results in [27], (i) from i.i.d. data to long-memory data; (ii) from the entrywise L^∞ norm of $\widehat{\mathbf{C}} - \mathbf{C}_0$ to the entrywise L^1 norm and the Frobenius norm where $\mathbf{C}_0$ is the common structure of the target precision matrices and $\widehat{\mathbf{C}}$ is the estimator of $\mathbf{C}_0$; and (iii) obtain the expectations of $\widehat{\mathbf{C}}_\rho - \mathbf{C}_0$, where $\widehat{\mathbf{C}}_\rho$ is the modified estimator, which will be defined below. We also deal with the Polynomial-type tail on top of sub-Gaussian and sub-exponential tails.

We borrow formulation of the problems and the notation from [27] for easy comparisons with their results. We consider a data set with S different groups (e.g. subjects in our fMRI example), and each group has a different distribution. For the sth group ($s = 1, \cdots, S$), we denote $\{\mathbf{x}_1^{(s)}, \cdots, \mathbf{x}_{n_s}^{(s)}\}$ as a time series random sample of size n_s, where $\mathbf{x}_k^{(s)} = (x_{k1}^{(s)}, \cdots, x_{kp}^{(s)})^T$ is a p-dimensional random vector with the covariance matrix $\mathbf{\Sigma}_0^{(s)} = (\sigma_{ij,0}^{(s)})$, and precision matrix $\mathbf{\Omega}_0^{(s)} := (\mathbf{\Sigma}_0^{(s)})^{-1}$. For the sth group, $\widehat{\mathbf{\Sigma}}^{(s)} := \left(\widehat{\sigma}_{ij}^{(s)}\right)_{p\times p}$ denotes the sample covariance matrix of observations $\mathbf{X}^{(s)}$ ($p \times n_s$ dimension):

$$\widehat{\mathbf{\Sigma}}^{(s)} = \frac{1}{n_s}\sum_{i=1}^{n_s} \mathbf{x}_i^{(s)}\mathbf{x}_i^{(s)T} - \overline{\mathbf{x}}^{(s)}\overline{\mathbf{x}}^{(s)T},$$

where $\overline{\mathbf{x}}^{(s)} = \frac{1}{n_s}\sum_{i=1}^{n_s} \mathbf{x}_i^{(s)}$, $\mathbf{x}_i^{(s)}$ is the ith colomn of $\mathbf{X}^{(s)}$. Also denote $\mathbf{\Omega}_0^{(s)} = (\omega_{ij,0}^{(s)})$, $s = 1, \cdots, S$.

Define the common structure C_0 and the unique structure $\mathbf{U}_0^{(s)}$ as

$$\mathbf{C}_0 := \frac{1}{S}\sum_{s=1}^{S} \mathbf{\Omega}_0^{(s)}, \quad \text{and} \quad \mathbf{U}_0^{(s)} = \mathbf{\Omega}_0^{(s)} - \mathbf{C}_0; s = 1, \cdots, S.$$

Note that $\sum_{s=1}^{S} \mathbf{U}_0^{(s)} = 0$. In order to estimate $\{\mathbf{C}_0, \mathbf{U}_0^{(1)}, \cdots, \mathbf{U}_0^{(S)}\}$ for long-memory time series, we investigate the following constrained entrywise L^1 minimization problem [27]:

$$\min\left\{|\mathbf{C}|_1 + \nu\sum_{s=1}^{S}\left|\mathbf{U}^{(s)}\right|_1\right\}$$

$$\text{s.t} \left|\frac{1}{S}\sum_{s=1}^{S}\left\{\widehat{\mathbf{\Sigma}}^{(s)}(\mathbf{C}+\mathbf{U}^{(s)}) - \mathbf{I}\right\}\right|_\infty \le \lambda_1, \left|\widehat{\mathbf{\Sigma}}^{(s)}(\mathbf{C}+\mathbf{U}^{(s)}) - \mathbf{I}\right|_\infty \le \lambda_2, \sum_{s=1}^{S}\mathbf{U}^{(s)} = 0, \quad (3.0.1)$$

where ν is a prespecified weight, and (λ_1, λ_2) are tuning parameters. If $\lambda_1 > \lambda_2$, it is easy to see that the first inequality constraints does not work. Therefore, we always assume (λ_1, λ_2) satisfying $\lambda_1 \le \lambda_2$.

We can also consider the following vector optimal problem for $j = 1, \cdots, p$.

$$\min\left\{\left|\frac{1}{S}\sum_{s=1}^{S}\mathbf{c}_j^{(s)}\right|_1 + \nu\sum_{s=1}^{S}\left|\mathbf{c}_j^{(s)} - \frac{1}{S}\sum_{s=1}^{S}\mathbf{c}_j^{(s)}\right|_1\right\}$$

$$\text{s.t} \quad \left| \frac{1}{S} \sum_{s=1}^{S} \left\{ \widehat{\mathbf{\Sigma}}^{(s)} \mathbf{c}_j^{(s)} - \mathbf{e}_j \right\} \right|_\infty \leq \lambda_1, \left| \widehat{\mathbf{\Sigma}}^{(s)} \mathbf{c}_j^{(s)} - \mathbf{e}_j \right|_\infty \leq \lambda_2, \tag{3.0.2}$$

where $\mathbf{c}_j^{(s)}$ is a vector in $\mathbb{R}^p$, $\mathbf{e}_j$ is the jth colomn of the identity matrix $\mathbf{I}_p$.

By the proof of Lemma 1 of [11], it can be seen that (3.0.1) and (3.0.2) have the same solution set. This allows one to solve (3.0.1) by CLIME for data analysis.

Assume $\{\widehat{\mathbf{C}}_1, \widehat{\mathbf{U}}_1^{(1)}, \cdots, \widehat{\mathbf{U}}_1^{(S)}\}$ is a solution of (3.0.1), and denote $\widehat{\mathbf{\Omega}}_1^{(s)} := \widehat{\mathbf{C}}_1 + \widehat{\mathbf{U}}_1^{(s)}, s = 1, \cdots, S$. Because it is not guaranteed that $\{\widehat{\mathbf{\Omega}}_1^{(s)}, s = 1, \cdots, S\}$ is symmetric, [27] take the following symmerization step. We define our final JEMP estimator $\{\widehat{\mathbf{\Omega}}^{(1)}, \cdots, \widehat{\mathbf{\Omega}}^{(S)}\}$ of $\{\mathbf{\Omega}_0^{(1)}, \cdots, \mathbf{\Omega}_0^{(S)}\}$ as the symmetrized result of $\{\widehat{\mathbf{\Omega}}_1^{(1)}, \cdots, \widehat{\mathbf{\Omega}}_1^{(S)}\}$, and denote $\widehat{\mathbf{\Omega}}_1^{(s)} = (\widehat{\omega}_{ij,1}^{(s)})$, and $\widehat{\mathbf{\Omega}}^{(s)} = (\widehat{\omega}_{ij}^{(s)})$, for $s = 1, \cdots, S$, then

$$\widehat{\omega}_{ij}^{(s)} = \widehat{\omega}_{ij,1}^{(s)} \mathbb{1}\{\sum_{s=1}^{S} \left|\widehat{\omega}_{ij,1}^{(s)}\right| \leq \sum_{s=1}^{S} \left|\widehat{\omega}_{ji,1}^{(s)}\right|\} + \widehat{\omega}_{ji,1}^{(s)} \mathbb{1}\{\sum_{s=1}^{S} \left|\widehat{\omega}_{ij,1}^{(s)}\right| > \sum_{s=1}^{S} \left|\widehat{\omega}_{ji,1}^{(s)}\right|\}. \tag{3.0.3}$$

In what follows, we present the theoretical properties of the JEMP estimator. In this chapter, we consider the following class of matrices:

$$\mathcal{U} := \left\{ \mathbf{\Omega} : \mathbf{\Omega} \succ 0, \|\mathbf{\Omega}\|_1 \leq C_M, |\mathbf{\Omega}|_1 \leq C_M^* \right\}.$$

We assume $n = n_1 = \cdots = n_S$, and $\mathbf{\Omega}_0^{(s)} \in \mathcal{U}$ for all $s = 1, \cdots, S$. Denote $\mathbb{E}(\mathbf{x}^{(s)}) = (\mu_1^{(s)}, \cdots, \mu_p^{(s)})^T$.

3.1 Estimation of Precision matrices for sub-Gaussian data

In this section, we extend the results in [27] for the precision matrix estimator from i.i.d to long-memory for sub-Gaussian data in (C1).

Throughout this chapter, we define

$$\frac{1}{nS} \max_{1 \leq k \leq p} \left\| Cov(\mathbf{x}_{[k]}) \right\|_F^2 \leq g_F, \ \max_{1 \leq k \leq p} \left\| Cov(\mathbf{x}_{[k]}) \right\|_2 \leq g_2, \tag{3.1.1}$$

$$\frac{1}{n} \max_{1 \leq k \leq p} \left\| Cov(\mathbf{x}_{[k]}^{(s)}) \right\|_F^2 \leq g_F^{(s)}, \ \max_{1 \leq k \leq p} \left\| Cov(\mathbf{x}_{[k]}^{(s)}) \right\|_2 \leq g_2^{(s)}, \tag{3.1.2}$$

where $\mathbf{x}_{[k]}^{(s)}$ is the kth row vector of $\mathbf{X}^{(s)}$, $\mathbf{x}_{[k]}$ is the kth row vector of $\mathbf{X} = (\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(S)})$, and we also define

$$\widetilde{u}_1^{(s)} = \max\left\{c_2 K^2 (\log p) g_2^{(s)}/n, [c_2 K^4 (\log p) g_F^{(s)}/n]^{1/2}\right\}, \tag{3.1.3}$$

$$\tilde{u}_1 = \max\left\{c_2 K^2 (\log p) g_2/Sn, [c_2 K^4 (\log p) g_F/Sn]^{1/2}\right\}, \tag{3.1.4}$$

where $c_2 = (\tau + 2)/c_1$ with $\tau > 0$, $c_1 > 0$ is an absolute constant. We assume $\widetilde{u}_1^{(s)}, \tilde{u}_1 \to 0$ as $n \to \infty$. The following theorem generalizes Theorem 1 of [27] from i.i.d to long-memory observations.

Theorem 3.1.1. *Suppose that $\mathbf{X}_{p\times nS}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K. Let $\lambda_1 = \lambda_2 = 6C_M \max_s \tilde{u}_1^{(s)}$. Set $\nu = S^{-1}$. Then*

$$\max_{ij}\left(\frac{1}{S}\sum_{s=1}^{S}\left|\widehat{\omega}_{ij}^{(s)} - \omega_{ij,0}^{(s)}\right|\right) \leq 12C_M^2 \max_s \tilde{u}_1^{(s)},$$

with probability greater than $1 - 4Sp^{-\tau}$, where $\widetilde{u}_1^{(s)}$ is defined by (3.1.3).

Since $\lambda_1 = \lambda_2$ in this theorem, the first inequality constraint in (3.0.1) does not play any role. In order to prove Theorem 3.1.1, we need the following lemma, which generalizes Lemma 4 of [27] from i.i.d to long-memory observations.

Lemma 3.1.1. *Suppose that $\mathbf{X}_{p\times nS}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K. For any fixed $s = 1, \cdots, S$, we have*

$$\max_{ij}\left|\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)}\right| \leq 2\widetilde{u}_1^{(s)}$$

with probability greater than $1 - 4p^{-\tau}$, where $\widetilde{u}_1^{(s)}$ is defined by (3.1.3)

Proof of Lemma 3.1.1. By the proof of the first part of Lemma A2 of [35], we have

$$P\left[\left|\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}\right|_\infty \geq 2u\right] \;\leq\; 2p\exp\left\{-\frac{c_1 n u}{K^2 g_2^{(s)}}\right\}$$

$$+2p^2 \exp\left\{-c_1 \min\left(\frac{nu^2}{K^4 g_F^{(s)}}, \frac{nu}{K^2 g_2^{(s)}}\right)\right\},$$

where $c_1 > 0$ is an absolute constant. Let

$$u = \tilde{u}_1^{(s)} = \max\left\{c_2 K^2 (\log p) g_2^{(s)}/n, [c_2 K^4 (\log p) g_F^{(s)}/n]^{1/2}\right\},$$

with $c_2 = (\tau + 2)/c_1$, then

$$P\left[|\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}|_\infty \geq 2\tilde{u}_1^{(s)}\right] \leq 2p^{-(c_1 c_2 - 1)} + 2p^{-(c_1 c_2 - 2)} \leq 4p^{-\tau} = O(p^{-\tau}).$$

□

Proof of Theorem 3.1.1. We recap the proof of Theorem 1 of [27]. Lemma 3.1.1 implies that

$$|\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}|_\infty \leq 2\tilde{u}_1^{(s)} \text{ for all } s = 1, \cdots, S, \tag{3.1.5}$$

with probability greater than $1 - 4Sp^{-\tau}$. In the proof below, we assume (3.1.5) holds. It is straightforward to check that $\{\mathbf{C}_0, \mathbf{U}_0^{(1)}, \cdots, \mathbf{U}_0^{(S)}\}$ is a feasible solution of (3.0.1), because

$$\begin{aligned}
\left|\mathbf{I} - \widehat{\mathbf{\Sigma}}^{(s)}(\mathbf{C}_0 + \mathbf{U}_0^{(s)})\right|_\infty &= \left|(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\Omega_0^{(s)}\right|_\infty \\
&\leq \left\|\Omega_0^{(s)}\right\|_1 |\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}|_\infty \\
&\leq C_M 2\tilde{u}_1^{(s)} \leq \lambda_2/3.
\end{aligned}$$

Similar to the proof of Theorem 1 of [27], we have

$$\begin{aligned}
\sum_{s=1}^{S} |(\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j|_\infty &\leq \sum_{s=1}^{S} \left|\widehat{\mathbf{\Omega}}_1^{(s)}\mathbf{e}_j\right|_1 \lambda_2/3 + SC_M\lambda_2 \\
&\leq S\left\{|\widehat{\mathbf{c}}_j|_1 + S^{-1}\sum_{s=1}^{S}\left|\widehat{\mathbf{u}}_j^{(s)}\right|_1\right\}\lambda_2/3 + SC_M\lambda_2 \\
&\leq S\left\{|\mathbf{c}_{j,0}|_1 + S^{-1}\sum_{s=1}^{S}\left|\mathbf{u}_{j,0}^{(s)}\right|_1\right\}\lambda_2/3 + SC_M\lambda_2 \\
&\leq 2SC_M\lambda_2 = 12SC_M^2 \max_{1\leq s\leq S} \tilde{u}_1^{(s)},
\end{aligned}$$

where $\widehat{\mathbf{c}}_j$ is the jth column of $\widehat{\mathbf{C}}_1$, and $\widehat{\mathbf{u}}_j^{(s)}$ is the jth column of $\widehat{\mathbf{U}}_1^{(s)}$, $\mathbf{c}_{j,0}$ is the jth column of $\mathbf{C}_0$, $\mathbf{u}_{j,0}^{(s)}$ is the jth column of $\mathbf{U}_0^{(s)}$.

It is straightforward to check that

$$\max_{ij}\left(\frac{1}{S}\sum_{s=1}^{S}\left|\widehat{\omega}_{ij}^{(s)}-\omega_{ij,0}^{(s)}\right|\right)\le \max_{j}\frac{1}{S}\sum_{s=1}^{S}|(\widehat{\mathbf{\Omega}}_1^{(s)}-\mathbf{\Omega}_0^{(s)})\mathbf{e}_j|_\infty\le 12C_M^2\max_{1\le s\le S}\tilde{u}_1^{(s)}.$$

The proof is completed. □

The following theorem generalizes Theorem 2 of [27] from i.i.d to long-memory observations. In Theorem 3.1.2, we obtain a faster convergence rate than Theorem 3.1.1 by properly choosing λ_1 and using more structure information of the data.

Theorem 3.1.2. *Suppose that (i) $\mathbf{X}_{p\times nS}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K, (ii) there exists $C_U>0$ such that $\left\|\mathbf{U}_0^{(s)}\right\|_1\le C_U$ for all $s=1,\cdots,S$ and $(\sum_{s=1}^S\left\|\mathbf{U}_0^{(s)}\right\|_1)\le C_U S^{1-\theta}$ where $\theta\in(0,1)$. Let*

$$\begin{aligned}\psi_1 &= C_M\left(\tilde{u}_1+\sum_{s=1}^{S}c_2K^2g_2^{(s)}\frac{\log p}{nS}\right)+C_US^{-\theta}\max_s 2\tilde{u}_1^{(s)},\\ \lambda_1 &= \psi_1,\lambda_2=6C_M\max_s\widetilde{u}_1^{(s)}.\end{aligned}$$

Set $\nu=S^{\theta-1}$. Then

$$\left|\widehat{\mathbf{C}}-\mathbf{C}_0\right|_\infty\le(2C_M+4C_U)\max\left\{\psi_1,S^{-\theta}C_M\max_s 2\widetilde{u}_1^{(s)}\right\}$$

with probability greater than $1-2(2+3S)p^{-\tau}$, where $\widetilde{u}_1^{(s)}$ and $\widetilde{u}_1$ are defined in (3.1.3) and (3.1.4).

In order to prove Theorem 3.1.2, we need the following lemma, which generalizes Lemma 5 of [27] from i.i.d to long-memory observations.

Lemma 3.1.2. *Suppose that $\mathbf{X}_{p\times nS}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K. With probability greater than $1-2(1+S)p^{-\tau}$ with $\tau>0$,*

$$\max_{ij}\left|\sum_{s=1}^{S}(\widehat{\sigma}_{ij}^{(s)}-\sigma_{ij,0}^{(s)})\right|\le S\tilde{u}_1+\sum_{s=1}^{S}c_2K^2g_2^{(s)}\frac{\log p}{n},$$

where $\tilde{u}_1$ is defined in (3.1.4).

Proof of Lemma 3.1.2. We assume $\mu_i^{(s)} = 0$ for all i and s. Denote $y_{kij}^{(s)} := x_{ik}^{(s)} x_{jk}^{(s)} - \mathbb{E}(x_{ik}^{(s)} x_{jk}^{(s)})$, and $\overline{x}_i^{(s)} = \sum_{k=1}^n x_{ik}^{(s)}/n$. Similar to the proof of Lemma 5 of [27], we have

$$\sum_{s=1}^S \left(\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)}\right) = \sum_{s=1}^S \left(\sum_{k=1}^n y_{kij}^{(s)}/n - \overline{x}_i^{(s)}\overline{x}_j^{(s)}\right). \tag{3.1.6}$$

We can define $\mathbf{P}_k^{(s)}$ such that $\mathbf{x}_{[k]}^{(s)} = \mathbf{P}_k^{(s)}\mathbf{H}\mathbf{z}$, which is the transpose of the kth row of $\mathbf{X}^{(s)}$, where $\mathbf{P}_k^{(s)}$ to be the $n \times pnS$ matrix with 1 in the $\big(j, k+(j-1)p\big)$ entries and 0 in all other entries, $j = 1, \cdots, n$.

Denote $\mathbf{x}_{[k]} = \mathbf{P}_k\mathbf{H}\mathbf{z} = \left(\left(\mathbf{x}_{[k]}^{(1)}\right)^T, \cdots, \left(\mathbf{x}_{[k]}^{(S)}\right)^T\right)^T$. It is easy to check that

$$Cov(\mathbf{x}_{[k]}) = Cov(\mathbf{P}_k\mathbf{H}\mathbf{z}) = \mathbf{P}_k\mathbf{H}\mathbf{H}^T\mathbf{P}_k^T.$$

By Lemma A1-(i) of [35] and similar to the proof of the first part of Lemma A2 of [35], we have

$$\begin{aligned} P\left[\max_{ij}\left|\frac{1}{Sn}\sum_{s=1}^S\sum_{k=1}^n y_{kij}^{(s)}\right| \geq u\right] \;\; &\leq \;\; 2p\exp\left\{-c_1\frac{nSu}{K^2g_2}\right\} \\ &\quad +2p^2\exp\left\{-c_1\min\left(\frac{nSu^2}{K^4g_F}, \frac{nSu}{K^2g_2}\right)\right\}, \end{aligned}$$

where $c_1 > 0$ is an absolute constant. Let

$$u = \tilde{u}_1 = \max\left\{c_2K^2(\log p)g_2/Sn, [c_2K^4(\log p)g_F/Sn]^{1/2}\right\},$$

with $c_2 = (\tau+2)/c_1$, then

$$P\left[\max_{ij}\left|\frac{1}{n}\sum_{s=1}^S\sum_{k=1}^n y_{kij}^{(s)}\right| \geq S\tilde{u}_1\right] \leq 2p^{-(c_1c_2-2)} + 2p^{-(c_1c_2-2)} \leq 4p^{-\tau} = O(p^{-\tau}). \tag{3.1.7}$$

By Lemma A1-(i) of [35] we have

$$\begin{aligned} P\left[\max_{ij}\left|\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq u^2\right] \;\; &= \;\; P\left[\max_i\left|\overline{x}_i^{(s)}\right|^2 \geq u^2\right] \\ &= \;\; P\left[\max_i\left|\overline{x}_i^{(s)}\right| \geq u\right] \\ &\leq \;\; 2\exp\left[-\frac{c_1u^2}{K^2\left\|\mathbf{b}\right\|_F^2\left\|\mathbf{P}_i^{(s)}\mathbf{H}\mathbf{H}^T\mathbf{P}_i^{(s)T}\right\|_2}\right] \end{aligned}$$

$$= \quad 2\exp\left[-\frac{c_1 n u^2}{K^2 g_2}\right].$$

Taking $u = \left(c_2 K^2 g_2^{(s)} \frac{\log p}{n}\right)^{1/2}$, we have

$$P\left[\max_{ij}\left|\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq c_2 K^2 g_2^{(s)}\frac{\log p}{n}\right] \leq 2p^{-\tau}.$$

Using this result, we have

$$\begin{aligned}
& P\left[\max_{ij}\left|\sum_{s=1}^{S}\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq \sum_{s=1}^{S} c_2 K^2 g_2^{(s)}\frac{\log p}{n}\right] \\
\leq \quad & P\left[\sum_{s=1}^{S}\max_{ij}\left|\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq \sum_{s=1}^{S} c_2 K^2 g_2^{(s)}\frac{\log p}{n}\right] \\
\leq \quad & \sum_{s=1}^{S} P\left[\max_{ij}\left|\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq c_2 K^2 g_2^{(s)}\frac{\log p}{n}\right] \leq 2Sp^{-\tau}. \qquad (3.1.8)
\end{aligned}$$

By (3.1.6), (3.1.7) and (3.1.8), we have

$$\begin{aligned}
& P\left[\max_{ij}\left|\sum_{s=1}^{S}\left(\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)}\right)\right| \geq S\tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 g_2^{(s)}\frac{\log p}{n}\right] \\
\leq \quad & P\left[\max_{ij}\left|\frac{1}{n}\sum_{s=1}^{S}\sum_{k=1}^{n} y_{kij}^{(s)}\right| \geq S\tilde{u}_1\right] \\
& + P\left[\max_{ij}\left|\sum_{s=1}^{S}\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq \sum_{s=1}^{S} c_2 K^2 g_2^{(s)}\frac{\log p}{n}\right] \\
\leq \quad & 2(2+S)p^{-\tau}.
\end{aligned}$$

The proof is completed. □

Proof of Theorem 3.1.2. We recap the proof of Theorem 2 of [27].

By Lemma 3.1.1 and Lemma 3.1.2, we have

$$\max_{ij}\left|\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)}\right| \leq 2\tilde{u}_1^{(s)}, \; \max_{ij}\left|\sum_{s=1}^{S}(\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)})\right| \leq S\tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 g_2^{(s)}\frac{\log p}{n}, \qquad (3.1.9)$$

for all $s = 1, \cdots, S$ with probability greater than $1 - 2(2+3S)p^{-\tau}$. In the proof below, we assume (3.1.9) holds.

It is straightforward to check that $\{\mathbf{C}_0, \mathbf{U}_0^{(1)}, \cdots, \mathbf{U}_0^{(S)}\}$ is a feasible solution of (3.0.1), because

$$\left|\mathbf{I} - \widehat{\mathbf{\Sigma}}^{(s)}(\mathbf{C}_0 + \mathbf{U}_0^{(s)})\right|_\infty \quad = \quad \left|(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\mathbf{\Omega}_0^{(s)}\right|_\infty$$

$$\leq \left\| \boldsymbol{\Omega}_0^{(s)} \right\|_1 |\widehat{\boldsymbol{\Sigma}}^{(s)} - \boldsymbol{\Sigma}_0^{(s)}|_\infty$$
$$\leq C_M 2\tilde{u}_1^{(s)} \leq \lambda_2/3,$$

and

$$\begin{aligned}
& \left| S^{-1} \sum_{s=1}^{S} \left\{ \mathbf{I} - \widehat{\boldsymbol{\Sigma}}^{(s)}(\mathbf{C}_0 + \mathbf{U}_0^{(s)}) \right\} \right|_\infty \\
\leq & \left| S^{-1} \sum_{s=1}^{S} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \mathbf{C}_0 \right|_\infty + \left| S^{-1} \sum_{s=1}^{S} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \mathbf{U}_0^{(s)} \right|_\infty \\
\leq & \|\mathbf{C}_0\|_1 \left| S^{-1} \sum_{s=1}^{S} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \right|_\infty + S^{-1} \sum_{s=1}^{S} \left\| \mathbf{U}_0^{(s)} \right\|_1 \left| \boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)} \right|_\infty \\
\leq & C_M \left(\tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 g_2^{(s)} \frac{\log p}{nS} \right) + C_U S^{-\theta} \max_s 2\tilde{u}_1^{(s)} \leq \lambda_1.
\end{aligned}$$

Now, we estimate $\left| S(\widehat{\mathbf{C}}_1 - \mathbf{C}_0)\mathbf{e}_j \right|_\infty = \left| \sum_{s=1}^{S} (\widehat{\boldsymbol{\Omega}}_1^{(s)} - \boldsymbol{\Omega}_0^{(s)})\mathbf{e}_j \right|_\infty$. We have

$$\begin{aligned}
\left| \sum_{s=1}^{S} (\widehat{\boldsymbol{\Omega}}_1^{(s)} - \boldsymbol{\Omega}_0^{(s)})\mathbf{e}_j \right|_\infty \leq & \left| \sum_{s=1}^{S} \boldsymbol{\Omega}_0^{(s)} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \widehat{\boldsymbol{\Omega}}_1^{(s)} \mathbf{e}_j \right|_\infty \\
& + \left| \sum_{s=1}^{S} \boldsymbol{\Omega}_0^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)} \widehat{\boldsymbol{\Omega}}_1^{(s)} - \mathbf{I}) \mathbf{e}_j \right|_\infty . \qquad (3.1.10)
\end{aligned}$$

Similar to the discussion of Theorem 2 of [27], we have

$$\begin{aligned}
& \left| \sum_{s=1}^{S} \boldsymbol{\Omega}_0^{(s)} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \widehat{\boldsymbol{\Omega}}_1^{(s)} \mathbf{e}_j \right|_\infty \\
\leq & \left| \sum_{s=1}^{S} \boldsymbol{\Omega}_0^{(s)} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \widehat{\mathbf{c}}_j \right|_\infty + \left| \sum_{s=1}^{S} \boldsymbol{\Omega}_0^{(s)} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \widehat{\mathbf{u}}_j^{(s)} \right|_\infty \\
\leq & \left| \sum_{s=1}^{S} \mathbf{C}_0 (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \widehat{\mathbf{c}}_j \right|_\infty + \left| \sum_{s=1}^{S} \mathbf{U}_0^{(s)} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \widehat{\mathbf{c}}_j \right|_\infty \\
& + \sum_{s=1}^{S} \left| \boldsymbol{\Omega}_0^{(s)} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \widehat{\mathbf{u}}_j^{(s)} \right|_\infty \\
\leq & C_M \left| \sum_{s=1}^{S} (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \right|_\infty |\widehat{\mathbf{c}}_j|_1 + \max_s \left| (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \right|_\infty \sum_{s=1}^{S} \left\| \mathbf{U}_0^{(s)} \right\|_1 |\widehat{\mathbf{c}}_j|_1 \\
& + C_M \max_s \left| (\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)}) \right|_\infty \sum_{s=1}^{S} \left| \widehat{\mathbf{u}}_j^{(s)} \right|_1
\end{aligned}$$

$$
\begin{aligned}
\leq\ & \left(C_M \left| \sum_{s=1}^{S} (\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}) \right|_\infty + \max_s \left| (\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}) \right|_\infty C_U S^{1-\theta} \right) |\widehat{\mathbf{c}}_j|_1 \\
& + S^{1-\theta} C_M \max_s \left| (\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}) \right|_\infty S^{\theta-1} \sum_{s=1}^{S} \left| \widehat{\mathbf{u}}_j^{(s)} \right|_1 \\
\leq\ & \max \left\{ C_M \left| \sum_{s=1}^{S} (\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}) \right|_\infty + \max_s \left| (\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}) \right|_\infty C_U S^{1-\theta}, \right. \\
& \left. S^{1-\theta} C_M \max_s \left| (\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}) \right|_\infty \right\} \times \left(|\widehat{\mathbf{c}}_j|_1 + S^{\theta-1} \sum_{s=1}^{S} \left| \widehat{\mathbf{u}}_j^{(s)} \right|_1 \right) \\
\leq\ & \max \left\{ C_M \left| \sum_{s=1}^{S} (\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}) \right|_\infty + \max_s \left| (\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}) \right|_\infty C_U S^{1-\theta}, \right. \\
& \left. S^{1-\theta} C_M \max_s \left| (\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}) \right|_\infty \right\} \times \left(|\mathbf{c}_{j,0}|_1 + S^{\theta-1} \sum_{s=1}^{S} \left| \mathbf{u}_{j,0}^{(s)} \right|_1 \right) \\
\leq\ & (C_M + C_U) \max \left\{ S\psi_1, S^{1-\theta} C_M \max_s 2\widetilde{u}_1^{(s)} \right\},
\end{aligned}
$$

where $\widehat{\mathbf{c}}_j$ is the jth column of $\widehat{\mathbf{C}}_1$, and $\widehat{\mathbf{u}}_j^{(s)}$ is the jth column of $\widehat{\mathbf{U}}_1^{(s)}$, $\mathbf{c}_{j,0}$ is the jth column of $\mathbf{C}_0$, $\mathbf{u}_{j,0}^{(s)}$ is the jth column of $\mathbf{U}_0^{(s)}$, and

$$
\begin{aligned}
& \left| \sum_{s=1}^{S} \mathbf{\Omega}_0^{(s)} (\widehat{\mathbf{\Sigma}}^{(s)} \widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I}) \mathbf{e}_j \right|_\infty \\
\leq\ & \left| \sum_{s=1}^{S} \mathbf{C}_0 (\widehat{\mathbf{\Sigma}}^{(s)} \widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I}) \mathbf{e}_j \right|_\infty + \left| \sum_{s=1}^{S} \mathbf{U}_0^{(s)} (\widehat{\mathbf{\Sigma}}^{(s)} \widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I}) \mathbf{e}_j \right|_\infty \\
\leq\ & \|\mathbf{C}_0\|_1 \left| \sum_{s=1}^{S} (\widehat{\mathbf{\Sigma}}^{(s)} \widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I}) \mathbf{e}_j \right|_\infty + \sum_{s=1}^{S} \left\| \mathbf{U}_0^{(s)} \right\|_1 \left| (\widehat{\mathbf{\Sigma}}^{(s)} \widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I}) \mathbf{e}_j \right|_\infty \\
\leq\ & C_M S \lambda_1 + C_U S^{1-\theta} \lambda_2 \\
\leq\ & C_M S \lambda_1 + C_U S^{1-\theta} 6 C_M \max_s \widetilde{u}_1^{(s)} \\
\leq\ & (C_M + 3C_U) \max \left\{ S\psi_1, S^{1-\theta} C_M \max_s 2\widetilde{u}_1^{(s)} \right\}.
\end{aligned}
$$

In sum, we have

$$
\left| \widehat{\mathbf{C}} - \mathbf{C}_0 \right|_\infty \leq \left| \widehat{\mathbf{C}}_1 - \mathbf{C}_0 \right|_\infty \leq (2C_M + 4C_U) \max \left\{ \psi_1, S^{-\theta} C_M \max_s 2\widetilde{u}_1^{(s)} \right\}.
$$

The proof is completed. □

Define a threshold estimator $\widetilde{\mathbf{\Omega}}^{(s)} = (\widetilde{\omega}_{ij}^{(s)})$ based on $\{\widehat{\mathbf{\Omega}}^{(1)}, \cdots, \widehat{\mathbf{\Omega}}^{(S)}\}$ as

$$\widetilde{\omega}_{ij}^{(s)} = \widehat{\omega}_{ij}^{(s)} \mathbb{1}\left\{ \left| \widehat{\omega}_{ij}^{(s)} \right| \geq \delta_n \right\},$$

where $\delta_n \geq 2C_M S \lambda_2$ and λ_2 is given in Theorem 3.1.1. Also define

$$\mathcal{S}_0 := \{(i,j,s) : \omega_{ij,0}^{(s)} \neq 0\}, \widehat{\mathcal{S}} := \{(i,j,s) : \widetilde{\omega}_{ij}^{(s)} \neq 0\} \text{ and } \theta_{\min} := \min_{(i,j,s) \in \mathcal{S}_0} \left| \omega_{ij,0}^{(s)} \right|.$$

Then, we have the following theorem which generalizes Theorem 3 of [27] from i.i.d to long-memory observations. The proof is straightforward. Details are omitted.

Theorem 3.1.3. *Suppose that (i) $\mathbf{X}_{p \times nS}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K, (ii) $\theta_{\min} > 2\delta_n$. Then*

$$P(\mathcal{S}_0 = \widehat{\mathcal{S}}) \geq 1 - 4Sp^{-\tau}.$$

3.2 Modified estimator $\widehat{\Omega}_\rho$

In this section, we focus on dealing with the expectation of the convergence rates, for example, $\mathbb{E}\left(\left| \widehat{\mathbf{C}} - \mathbf{C}_0 \right|_1^2 \right)$, motivated by [11]. Because the main interest is the inverse matrix, even the existence of $\mathbb{E}\left(\left| \widehat{\mathbf{C}} - \mathbf{C}_0 \right|_1^2 \right)$ is a concern. Hence, we modify the estimator $\widehat{\mathbf{\Omega}}$ to ensure the existence, and then consider the convergence rates.

Let $\{\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\}$ be the solution of the following optimization problem:

$$\min \left\{ \left| \mathbf{C} \right|_1 + \nu \sum_{s=1}^{S} \left| \mathbf{U}^{(s)} \right|_1 \right\}$$

$$\text{s.t } \left| \frac{1}{S} \sum_{s=1}^{S} \left\{ \widehat{\mathbf{\Sigma}}_\rho^{(s)} \mathbf{\Omega}^{(s)} - \mathbf{I} \right\} \right|_\infty \leq \lambda_1, \left| \widehat{\mathbf{\Sigma}}_\rho^{(s)} \mathbf{\Omega}^{(s)} - \mathbf{I} \right|_\infty \leq \lambda_2, \sum_{s=1}^{S} \mathbf{U}^{(s)} = 0, \qquad (3.2.1)$$

where $\widehat{\mathbf{\Sigma}}_\rho^{(s)} = \widehat{\mathbf{\Sigma}}^{(s)} + \rho \mathbf{I}$ with $\rho > 0$. Denote $\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} = \left(\omega_{1\rho ij}^{(s)} \right)$. Define the symmetrized estimator $\widehat{\mathbf{\Omega}}_\rho^{(s)}$ as in (3.0.3).

Theorem 3.2.1. *Suppose that $\mathbf{X}_{p \times nS}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K. Set $\nu = S^{\theta - 1}$.*

(i) *Let* $\lambda_1 = 2\psi_1, \lambda_2 = 12C_M \max_s \tilde{u}_1^{(s)}$, *where* $\tau > 0$. *If there exists* $C_U > 0$ *such that* $\left\|\mathbf{U}_0^{(s)}\right\|_1 \le C_U$ *for all* $s = 1, \cdots, S$ *and* $(\sum_{s=1}^S \left\|\mathbf{U}_0^{(s)}\right\|_1) \le C_U S^{1-\theta}$ *where* $\theta \in (0,1)$ *and*

$$0 \le \rho \le \min\left\{\tilde{u}_1 + \sum_{s=1}^S c_2 K^2 g_2^{(s)} \frac{\log p}{Sn}, \tilde{u}_1^{(1)}, \cdots, \tilde{u}_1^{(S)}\right\},$$

with $c > 0$, *if* $p \ge n_s^\xi$ *with* $\xi > 0$, *then*

$$\begin{aligned}\left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_\infty &= \left|\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0\right|_\infty \\ &\le 2(2C_M + 4C_U)\max\left\{\psi_1, S^{-\theta} C_M \max_s 2\widetilde{u}_1^{(s)}\right\}\end{aligned}$$

with probability greater than $1 - 2(1+3S)p^{-\tau}$.

(ii) *If there exists* $C_M^*, C_U^* > 0$ *such that* $\left|\mathbf{\Omega}^{(s)}\right|_1 \le C_M^*$, $\left|\mathbf{U}_0^{(s)}\right|_1 \le C_U^*$ *and*

$$\sum_{s=1}^S \left|\mathbf{U}_0^{(s)}\right|_1 \le S^{1-\theta} C_U^*.$$

Let

$$\begin{aligned}\psi_1^* &= 2C_M^*\left(\tilde{u}_1 + \sum_{s=1}^S c_2 K^2 g_2^{(s)} \frac{\log p}{nS}\right) + 2C_U^* S^{-\theta} \max_s 2\widetilde{u}_1^{(s)}, \\ \lambda_1 &= \psi_1^*, \lambda_2 = 12C_M^* \max_s \tilde{u}_1^{(s)},\end{aligned}$$

where $\tau > 0$. *Then*

$$\left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_1 \le 2\left|\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0\right|_1 \le 4p(C_M + C_U + C_M^* + 3C_U^*)\max\left\{\psi_1^*, S^{-\theta} C_M^* \max_s 2\widetilde{u}_1^{(s)}\right\}$$

with probability greater than $1 - 2(2+3S)p^{-\tau}$.

(iii) *If there exists* $C_U^{\#} > 0$ *such that*

$$\sum_{i=1}^p \sum_{j=1}^p \left|u_{ij,0}^{(s)}\right|^q \le C_U^{\#} \text{ and } \sum_{s=1}^S \sum_{i=1}^p \sum_{j=1}^p \left|u_{ij,0}^{(s)}\right|^q \le C_U^{\#} S^{1-\theta}$$

where $\theta \in (0,1)$, *then*

$$\frac{1}{S}\sum_{s=1}^S \left|\left(\widehat{\mathbf{U}}_\rho^{(s)} - \mathbf{U}_0^{(s)}\right)\right|_1 \le 4(1 + 2^{1-q} + 3^{1-q}) t^{1-q} C_U^{\#}/S^\theta + \frac{2}{S\nu}\left|(\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0)\right|_1$$

with probability greater than $1 - 2(2+3S)p^{-\tau}$, *where* $t = \max_{ijs}\left|\widehat{u}_{1\rho ij}^{(s)} - u_{ij,0}^{(s)}\right|$, $\widehat{\mathbf{U}}_{1\rho}^{(s)} = \left(\widehat{u}_{1\rho ij}^{(s)}\right)$, $\mathbf{U}_0^{(s)} = \left(u_{ij,0}^{(s)}\right)$.

(iv) If the conditions in (i) and (ii) are satisfied and $\lambda_1 = \psi_1^*, \lambda_2 = 12C_M^* \max_s \tilde{u}_1^{(s)}$, then

$$\left\|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right\|_F^2 \leq \left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_\infty \left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_1$$

with probability greater than $1 - 2(2+3S)p^{-\tau}$, and $\left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_\infty$ and $\left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_1$ with bound in (i) and (ii).

Proof of Theorem 3.2.1. (i) Similar to the proof of Theorem 3.1.2.

(ii) Similar to the proof of Theorem 3.1.2. Note that $\{\mathbf{C}_0, \mathbf{U}_0^{(1)}, \cdots, \mathbf{U}_0^{(S)}\}$ is a feasible solution of (3.2.1). Now, we estimate the upper bound of $\left|S(\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0)\mathbf{e}_j\right|_1 = \left|\sum_{s=1}^S(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j\right|_1$. Note that

$$\left|\sum_{s=1}^S(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j\right|_1 \leq \left|\sum_{s=1}^S \mathbf{\Omega}_0^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1 + \left|\sum_{s=1}^S \mathbf{\Omega}_0^{(s)}(\widehat{\mathbf{\Sigma}}_\rho^{(s)}\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} - \mathbf{I})\mathbf{e}_j\right|_1. \quad (3.2.2)$$

Similar to the discussion of Theorem 3.1.2, we have

$$\begin{aligned}
&\left|\sum_{s=1}^S \mathbf{\Omega}_0^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1 \\
\leq\ & \left|\sum_{s=1}^S \mathbf{C}_0(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\widehat{\mathbf{c}}_j\right|_1 + \left|\sum_{s=1}^S \mathbf{U}_0^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\widehat{\mathbf{c}}_j\right|_1 + \left|\sum_{s=1}^S \mathbf{\Omega}_0^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\widehat{\mathbf{u}}_j^{(s)}\right|_1 \\
\leq\ & |\mathbf{C}_0|_1 \left|\sum_{s=1}^S(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\right|_\infty |\widehat{\mathbf{c}}_j|_1 + \sum_{s=1}^S \left|\mathbf{U}_0^{(s)}\right|_1 \left|(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\right|_\infty |\widehat{\mathbf{c}}_j|_1 \\
& + \max_s \left|\mathbf{\Omega}_0^{(s)}\right|_1 \sum_{s=1}^S \left|(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\right|_\infty \left|\widehat{\mathbf{u}}_j^{(s)}\right|_1 \\
\leq\ & \max\left\{S\psi_1^*, S^{1-\theta}C_M^* \max_s 2\widetilde{u}_1^{(s)}\right\}\left(|\mathbf{c}_{j,0}|_1 + S^{\theta-1}\sum_{s=1}^S \left|\mathbf{u}_{j,0}^{(s)}\right|_1\right),
\end{aligned}$$

where $\widehat{\mathbf{c}}_j$ is the jth column of $\widehat{\mathbf{C}}_{1\rho}$, and $\widehat{\mathbf{u}}_j^{(s)}$ is the jth column of $\widehat{\mathbf{U}}_{1\rho}^{(s)}$ with $\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} = \widehat{\mathbf{C}}_{1\rho} + \widehat{\mathbf{U}}_{1\rho}^{(s)}$, $\mathbf{c}_{j,0}$ is the jth column of $\mathbf{C}_0$, $\mathbf{u}_{j,0}^{(s)}$ is the jth column of $\mathbf{U}_0^{(s)}$, and

$$\begin{aligned}
\left|\sum_{s=1}^S \mathbf{\Omega}_0^{(s)}(\widehat{\mathbf{\Sigma}}_\rho^{(s)}\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_1 &\leq C_M^* S\lambda_1 + S^{1-\theta}C_U^*\lambda_2 \\
&= (C_M^* + 3C_U^*)S\psi_1^*.
\end{aligned}$$

In sum, we have

$$\left\|\widehat{\mathbf{C}}_{1\rho}-\mathbf{C}_0\right\|_1 \le (C_M+C_U+C_M^*+3C_U^*)\max\left\{\psi_1^*, S^{-\theta}C_M^*\max_s 2\widetilde{u}_1^{(s)}\right\},$$

and then

$$\begin{aligned}\left|\widehat{\mathbf{C}}_{\rho}-\mathbf{C}_0\right|_1 &\le 2\left|\widehat{\mathbf{C}}_{1\rho}-\mathbf{C}_0\right|_1\\ &\le 2p\left\|\widehat{\mathbf{C}}_{1\rho}-\mathbf{C}_0\right\|_1\\ &\le 4p(C_M+C_U+C_M^*+3C_U^*)\max\left\{\psi_1^*, S^{-\theta}C_M^*\max_s 2\widetilde{u}_1^{(s)}\right\}.\end{aligned}$$

(iii) We assume the solution $\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}$ is combined from the column solution set $\{\widehat{\mathbf{c}}_{\rho j}^{(s)}\}$ of (3.0.2) with $\widehat{\mathbf{\Sigma}}^{(s)}$ being replaced by $\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}$. By the definition of $\{\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\}$, we have

$$\left|\widehat{\mathbf{C}}_{1\rho}\mathbf{e}_j\right|_1+\nu\sum_{s=1}^{S}\left|\widehat{\mathbf{U}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1 \le |\mathbf{C}_0\mathbf{e}_j|_1+\nu\sum_{s=1}^{S}\left|\mathbf{U}_0^{(s)}\mathbf{e}_j\right|_1 \text{ for } j=1,\cdots,p.$$

Denote

$$\begin{aligned}t &= \max_{ijs}\left|\widehat{u}_{1\rho ij}^{(s)}-u_{ij,0}^{(s)}\right|, \text{ where } \widehat{\mathbf{U}}_{1\rho}^{(s)}=\left(\widehat{u}_{1\rho ij}^{(s)}\right), \mathbf{U}_0^{(s)}=\left(u_{ij,0}^{(s)}\right),\\ \mathbf{L}^{(s)} &= \left(\widehat{u}_{1\rho ij}^{(s)}-u_{ij,0}^{(s)}\right),\\ \mathbf{L}_*^{(s)} &= \left(\widehat{u}_{1\rho ij}^{(s)}\mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right|\ge 2t)-u_{ij,0}^{(s)}\right),\\ \mathbf{L}_{**}^{(s)} &= \mathbf{L}^{(s)}-\mathbf{L}_*^{(s)}.\end{aligned}$$

Then

$$\begin{aligned}&\left|\widehat{\mathbf{C}}_{1\rho}\mathbf{e}_j\right|_1+\nu\sum_{s=1}^{S}\left(\left|\mathbf{U}_0^{(s)}\mathbf{e}_j\right|_1-\left|\mathbf{L}_*^{(s)}\mathbf{e}_j\right|_1+\left|\mathbf{L}_{**}^{(s)}\mathbf{e}_j\right|_1\right)\\ \le\ &\left|\widehat{\mathbf{C}}_{1\rho}\mathbf{e}_j\right|_1+\nu\sum_{s=1}^{S}\left(\left|\left(\mathbf{U}_0^{(s)}+\mathbf{L}_*^{(s)}\right)\mathbf{e}_j\right|_1+\left|\mathbf{L}_{**}^{(s)}\mathbf{e}_j\right|_1\right)\\ =\ &\left|\widehat{\mathbf{C}}_{1\rho}\mathbf{e}_j\right|_1+\nu\sum_{s=1}^{S}\left(\left|\widehat{\mathbf{U}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1\right) \le |\mathbf{C}_0\mathbf{e}_j|_1+\nu\sum_{s=1}^{S}\left(\left|\mathbf{U}_0^{(s)}\mathbf{e}_j\right|_1\right),\end{aligned}$$

which implies

$$\nu\sum_{s=1}^{S}\left|\mathbf{L}_{**}^{(s)}\mathbf{e}_j\right|_1 \le \nu\sum_{s=1}^{S}\left|\mathbf{L}_*^{(s)}\mathbf{e}_j\right|_1+\left|(\widehat{\mathbf{C}}_{1\rho}-\mathbf{C}_0)\mathbf{e}_j\right|_1.$$

Thus

$$\nu \sum_{s=1}^{S} \left|\mathbf{L}^{(s)}\mathbf{e}_j\right|_1 \le \nu \sum_{s=1}^{S} \left|\mathbf{L}_{**}^{(s)}\mathbf{e}_j\right|_1 + \nu \sum_{s=1}^{S} \left|\mathbf{L}_{*}^{(s)}\mathbf{e}_j\right|_1 \le 2\nu \sum_{s=1}^{S} \left|\mathbf{L}_{*}^{(s)}\mathbf{e}_j\right|_1 + \left|(\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0)\mathbf{e}_j\right|_1 .$$

Similar to the proof of Theorem 6 of [11], we have

$$\begin{aligned}
\sum_{s=1}^{S} \left|\mathbf{L}_{*}^{(s)}\mathbf{e}_j\right|_1 &= \sum_{s=1}^{S}\sum_{i=1}^{p} \left|\widehat{u}_{1\rho ij}^{(s)} \mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \ge 2t) - u_{ij,0}^{(s)}\right| \\
&\le \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)} \mathbb{1}(\left|u_{ij,0}^{(s)}\right| \le 2t)\right| \\
&\quad + \sum_{s=1}^{S}\sum_{i=1}^{p} \left|\widehat{u}_{1\rho ij}^{(s)} \mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \ge 2t) - u_{ij,0}^{(s)} \mathbb{1}(\left|u_{ij,0}^{(s)}\right| \ge 2t)\right| \\
&\le (2t)^{1-q} \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q + \sum_{s=1}^{S}\sum_{i=1}^{p} \left|\left(\widehat{u}_{1\rho ij}^{(s)} - u_{ij,0}^{(s)}\right) \mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \ge 2t)\right| \\
&\quad + \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)} \{\mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \ge 2t) - \mathbb{1}(\left|u_{ij,0}^{(s)}\right| \ge 2t)\}\right| \\
&\le (2t)^{1-q} \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q + t \sum_{s=1}^{S}\sum_{i=1}^{p} \mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \ge 2t) \\
&\quad + \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right| \left|\mathbb{1}\{\left|\left|u_{ij,0}^{(s)}\right| - 2t\right| \le \left|\widehat{u}_{1\rho ij}^{(s)} - u_{ij,0}^{(s)}\right|\}\right| \\
&\le (2t)^{1-q} \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q + t \sum_{s=1}^{S}\sum_{i=1}^{p} \mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \ge 2t) \\
&\quad + \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right| \mathbb{1}\{\left|u_{ij,0}^{(s)}\right| \le 3t\} \\
&\le (2t)^{1-q} \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q + t^{1-q} \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q + (3t)^{1-q} \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q \\
&\le (1 + 2^{1-q} + 3^{1-q}) t^{1-q} \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q .
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{1}{S} \sum_{s=1}^{S} \left|\mathbf{L}^{(s)}\mathbf{e}_j\right|_1 &\le \frac{2}{S} \sum_{s=1}^{S} \left|\mathbf{L}_{*}^{(s)}\mathbf{e}_j\right|_1 + \frac{1}{S\nu} \left|(\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0)\mathbf{e}_j\right|_1 \\
&\le 2(1 + 2^{1-q} + 3^{1-q}) t^{1-q} \sum_{s=1}^{S}\sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q / S + \frac{1}{S\nu} \left|(\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0)\mathbf{e}_j\right|_1 .
\end{aligned}$$

Therefore

$$\begin{aligned}\frac{1}{S}\sum_{s=1}^{S}\left|\mathbf{U}_{\rho}^{(s)}-\mathbf{U}_{0}^{(s)}\right|_1 &\le \frac{2}{S}\sum_{s=1}^{S}\left|\mathbf{U}_{1\rho}^{(s)}-\mathbf{U}_{0}^{(s)}\right|_1\\ &\le 4(1+2^{1-q}+3^{1-q})t^{1-q}C_U^{\#}/S^{\theta}+\frac{2}{S\nu}\left|(\widehat{\mathbf{C}}_{1\rho}-\mathbf{C}_0)\right|_1.\end{aligned}$$

(iv) It is straightforward by the definition of matrix norm.

The proof is completed. □

Furthermore, we can consider the expectation of convergence rates, which generalizes Theorems 2 and 5 of [11] (i) from i.i.d. to long-memory observations, (ii) from a single subject to a group of subjects.

Theorem 3.2.2. *Under the conditions of Theorem 3.2.1, and*

$$\begin{aligned}&\min\left\{\tilde{u}_1+\sum_{s=1}^{S}c_2K^2g_2^{(s)}\frac{\log p}{Sn},\tilde{u}_1^{(1)},\cdots,\tilde{u}_1^{(S)},p^{-c}\right\}\\ \le\ &\rho\le\min\left\{\tilde{u}_1+\sum_{s=1}^{S}c_2K^2g_2^{(s)}\frac{\log p}{Sn},\tilde{u}_1^{(1)},\cdots,\tilde{u}_1^{(S)}\right\},\end{aligned}$$

with $c>0$, if $p\ge n_s^{\xi}$ with $\xi>0$, we have

(i) $\mathbb{E}\left(\left|\widehat{\mathbf{C}}_{\rho}-\mathbf{C}_0\right|_{\infty}^2\right)=O\left((C_M+C_U)^2\max\left\{\psi_1,S^{-\theta}C_M\max_s 2\widetilde{u}_1^{(s)}\right\}^2\right)$,

(ii) $\mathbb{E}\left(\left|\widehat{\mathbf{C}}_{\rho}-\mathbf{C}_0\right|_{1}^2\right)=O\left(p^2(C_M^*+C_U^*)^2\max\left\{\psi_1^*,S^{-\theta}C_M^*\max_s 2\widetilde{u}_1^{(s)}\right\}^2\right)$,

(iii) $\mathbb{E}\left(\left\|\widehat{\mathbf{C}}_{\rho}-\mathbf{C}_0\right\|_{F}^2\right)\le\sqrt{\mathbb{E}\left(\left|\widehat{\mathbf{C}}_{\rho}-\mathbf{C}_0\right|_{\infty}^2\right)\mathbb{E}\left(\left|\widehat{\mathbf{C}}_{\rho}-\mathbf{C}_0\right|_{1}^2\right)}$.

Proof of Theorem 3.2.2. Note that $\{\left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)^{-1}\}$ is a feasible solution of (3.2.1).

Also, note that

$$\begin{aligned}&\max_j\frac{1}{S}\sum_{s=1}^{S}\left|\widehat{\mathbf{\Omega}}_{\rho}^{(s)}\mathbf{e}_j\right|_1\\ \le\ &\max_j\frac{1}{S}\sum_{s=1}^{S}\left|\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1\end{aligned}$$

$$
\begin{aligned}
&\leq \max_j \left(\left| \widehat{\mathbf{C}}_{1\rho}^{(s)} \mathbf{e}_j \right|_1 + \frac{1}{S} \sum_{s=1}^{S} \left| \widehat{\mathbf{U}}_{1\rho}^{(s)} \mathbf{e}_j \right|_1 \right) \\
&\leq \max_j \left(\left| \widehat{\mathbf{C}}_{1\rho}^{(s)} \mathbf{e}_j \right|_1 + \frac{1}{S^{1-\theta}} \sum_{s=1}^{S} \left| \widehat{\mathbf{U}}_{1\rho}^{(s)} \mathbf{e}_j \right|_1 \right) \\
&\leq \max_j \left(\left| \frac{1}{S} \sum_{s=1}^{S} \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} \mathbf{e}_j \right|_1 + \frac{1}{S^{1-\theta}} \sum_{s=1}^{S} \left| \left(\left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} - \frac{1}{S} \sum_{s=1}^{S} \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} \right) \mathbf{e}_j \right|_1 \right) \\
&\leq \max_j \left(\frac{1}{S} \sum_{s=1}^{S} \left| \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} \mathbf{e}_j \right|_1 + \frac{1}{S^{1-\theta}} \sum_{s=1}^{S} \left| \left(\left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} \right) \mathbf{e}_j \right|_1 \right. \\
&\qquad \left. + \frac{1}{S^{1-\theta}} \sum_{s=1}^{S} \left| \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} \mathbf{e}_j \right|_1 \right) \\
&\leq \max_j \frac{3}{S^{1-\theta}} \sum_{s=1}^{S} \left| \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} \mathbf{e}_j \right|_1 \\
&\leq \max_j \frac{3}{S^{1-\theta}} \sum_{s=1}^{S} \left\| \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} \right\|_1 \\
&\leq \frac{3p}{S^{1-\theta}} \sum_{s=1}^{S} \left\| \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} \right\|_2 \\
&\leq \frac{3p}{S^{1-\theta}} \sum_{s=1}^{S} \varphi_{\max} \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right)^{-1} \\
&= \frac{3p}{S^{1-\theta}} \sum_{s=1}^{S} 1/\varphi_{\min} \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)} \right) \\
&\leq \frac{3p}{S^{1-\theta}} \sum_{s=1}^{S} 1/\rho = S^{\theta} p/\rho.
\end{aligned}
$$

The proof is similar to the proof of Theorem 2 of [11]. □

For LPDD temporal dependence, we can set the bounds as

$$
g_F^{(s)} \asymp n_s/(\log n_s)^{2\alpha}, g_2^{(s)} \asymp n_s/(\log n_s)^{\alpha},
$$

then we have the following property

$$
u_1^{(s)} \asymp (\log p)/(\log n_s)^{\alpha}, \alpha > 0.
$$

For LPDD model, therefore, we only need to assume $p \geq (\log n_s)^{\xi}$ in the above theorem.

3.3 Extension to heavy tail data

In this section, we extend the previous results for the precision matrices estimators to heavy-tailed data in (C2) and (C3). According to the proof of Lemma A.2 of [36], similar to the proof of Lemma 3.1.2, define

$$\widetilde{u}_2^{(s)} = \max\left\{(c_2K(2/\vartheta)^{2/\vartheta}\log p)^{1+2/\vartheta}g_2^{(s)}/n, (c_2K(4/\vartheta)^{4/\vartheta}\log p)^{1+2/\vartheta}(g_F^{(s)}/n)^{1/2}\right\}, \quad (3.3.1)$$

$$\widetilde{u}_2 = \max\left\{(c_2K(2/\vartheta)^{2/\vartheta}\log p)^{1+2/\vartheta}g_2/nS, (c_2K(4/\vartheta)^{4/\vartheta}\log p)^{1+2/\vartheta}(g_F/nS)^{1/2}\right\},$$

and

$$\widetilde{u}_3^{(s)} = k\eta_k^2\max\left\{c_2^2p^{(2+2\tau)/k}g_2^{(s)}/n, c_2p^{(4+2\tau)/k}(g_F^{(s)}/n)^{1/2}\right\}, \quad (3.3.2)$$

$$\widetilde{u}_3 = k\eta_k^2\max\left\{c_2^2p^{(2+2\tau)/k}g_2/nS, c_2p^{(4+2\tau)/k}(g_F/nS)^{1/2}\right\},$$

where $c_2 = (\tau+2)/c_1$ with $\tau > 0$ and $c_1 > 0$ is an absolute constant, $g_2^{(s)}$ and $g_F^{(s)}$ are defined in (3.1.2), g_2 and g_F are defined in (3.1.1). We assume $\widetilde{u}_2^{(s)}, \tilde{u}_2, \widetilde{u}_3^{(s)}, \tilde{u}_3 \to 0$ as $n_s \to \infty$.

Theorem 3.3.1. *(Generalized sub-exponential tails) Assume condition (C1), parameter K, $\sum_{s=1}^S c_2K^2g_2^{(s)}\log p$, $\widetilde{u}_1^{(s)}$ and $\widetilde{u}_1$ replaced by condition (C2), parameters $\{K,\vartheta\}$, $\sum_{s=1}^S(c_2K(2/\vartheta)^{2/\vartheta}\log p)^{1+2/\vartheta}g_2^{(s)}$, $\widetilde{u}_2^{(s)}$ and $\widetilde{u}_2$, respectively, then Theorems 3.1.1-3.2.2 hold.*

Theorem 3.3.2. *(Polynomial-type tails) Assume condition (C1), $\sum_{s=1}^S c_2K^2g_2^{(s)}\log p$, parameter K, $\widetilde{u}_1^{(s)}$ and $\widetilde{u}_1$ replaced by condition (C3), $\sum_{s=1}^S c_2^2k\eta_k^2p^{(2+2\tau)/k}g_2^{(s)}$, parameters $\{k,\eta_k\}$, $\widetilde{u}_3^{(s)}$ and $\widetilde{u}_3$, respectively, then Theorems 3.1.1-3.2.1 hold.*

The proofs of Theorems 3.3.1 and 3.3.2 are similar to the proofs of the preceding theorems by using the corresponding results given in Lemmas A.1, A.2 and A.3 of [35] for conditions (C2) and (C3), respectively. Details are omitted.

3.4 Computation

In this section, we solve the numerical solutions of the optimization problem in (3.0.1). We follow the numerical algorithm in [27]. In Section 3.4.1, we decompose the optimization problem (3.4.1) into p individual subproblems, and a linear programming approach is applied to solve them. In Section 3.4.2, we introduce the Gap-block data split for cross-validation proposed by Shu and Nan (see [35]). Section 3.5.2 explains the selection of tuning parameters.

3.4.1 Decomposition of (3.0.1) and linear programming approach

Note that the optimization problem (3.0.1) can be decomposed into p individual vector minimization problems, which is more doable for computation. Denote $\mathbf{e}_j$ as the jth column of $p \times p$ identity matrix $\mathbf{I}$. For $1 \leq j \leq p$, denote $\{\widehat{\mathbf{c}}_j, \widehat{\mathbf{u}}_j^{(1)}, \cdots, \widehat{\mathbf{u}}_j^{(S)}\}$ as the solution of the following vector optimization problem:

$$\min \left\{ |\mathbf{c}|_1 + \nu \sum_{s=1}^{S} \left|\mathbf{u}^{(s)}\right|_1 \right\}$$

$$\text{s.t } \left| \frac{1}{S} \sum_{s=1}^{S} \left\{ \widehat{\mathbf{\Sigma}}^{(s)} \left(\mathbf{c} + \mathbf{u}^{(s)} \right) - \mathbf{e}_j \right\} \right|_\infty \leq \lambda_1, \left| \widehat{\mathbf{\Sigma}}^{(s)} \left(\mathbf{c} + \mathbf{u}^{(s)} \right) - \mathbf{e}_j \right|_\infty \leq \lambda_2, \sum_{s=1}^{S} \mathbf{u}^{(s)} = 0, \tag{3.4.1}$$

where $\mathbf{c}, \mathbf{u}^{(1)}, \cdots, \mathbf{u}^{(S)}$ are vectors in $\mathbb{R}^p$.

We further reformulate the optimization problem in (3.4.1) as a linear programming problem, and use the simplex method to solve this problem (see [8]). For our simulation study and the rfMRI data analysis, we obtain the solution of (3.4.1) using the R-package *fastclime*, which provides a generic fast linear programming solver (see [31]).

In R-package *fastclime*, the standard linear programming in inequality form is as follows:

$$\max \mathbf{a}^T \mathbf{x} \text{ subject to: } \mathbf{M}\mathbf{x} \leq \mathbf{b},\ \mathbf{x} \geq 0,\ \mathbf{x} \in \mathbb{R}^n, \tag{3.4.2}$$

where $\mathbf{M} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{a} \in \mathbb{R}^n$ are given. In order to apply the R-package *fastclime*, we need to reformulate (3.4.1) into the above linear programming form. Let

$$\mathbf{c} = \mathbf{c}^+ - \mathbf{c}^-, \mathbf{u}^{(s)} = \mathbf{u}^{(s)+} - \mathbf{u}^{(s)-}$$

where $\mathbf{c}^+, \mathbf{c}^-, \mathbf{u}^{(s)+}, \mathbf{u}^{(s)-}$ are p-dimensional vectors. Denote

$$\mathbf{x}^+ = \begin{pmatrix} \mathbf{c}^+ \\ \mathbf{u}^{(1)+} \\ \vdots \\ \mathbf{u}^{(S)+} \end{pmatrix}, \mathbf{x}^- = \begin{pmatrix} \mathbf{c}^- \\ \mathbf{u}^{(1)-} \\ \vdots \\ \mathbf{u}^{(S)-} \end{pmatrix}, \mathbf{x} = \begin{pmatrix} \mathbf{x}^+ \\ \mathbf{x}^- \end{pmatrix},$$

where $\mathbf{x}^+$ and $\mathbf{x}^-$ are $(1+S)p$-dimensional vectors, and $\mathbf{x}$ is a $2(1+S)p$-dimensional vector. Then the equation (3.4.1) can be written as (3.4.2), with

$$\begin{aligned}
\mathbf{a} &= -\Big(\underbrace{1,\cdots,1}_{p},\underbrace{\nu,\cdots,\nu}_{p\times S},\underbrace{1,\cdots,1}_{p},\underbrace{\nu,\cdots,\nu}_{p\times S}\Big)^T, \\
\mathbf{M} &= \begin{pmatrix} \mathbf{A} \\ -\mathbf{A} \\ \mathbf{B} \\ -\mathbf{B} \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \\ -\mathbf{b}_3 \end{pmatrix},
\end{aligned}$$

with $\mathbf{A} = \left(\widetilde{\mathbf{A}}, -\widetilde{\mathbf{A}}\right)$ and $\mathbf{B} = \left(\widetilde{\mathbf{B}}, -\widetilde{\mathbf{B}}\right)$ where

$$\begin{aligned}
\widetilde{\mathbf{A}} &= \begin{pmatrix} \sum_{s=1}^S S^{-1}\widehat{\boldsymbol{\Sigma}}^{(s)} & S^{-1}\widehat{\boldsymbol{\Sigma}}^{(1)} & S^{-1}\widehat{\boldsymbol{\Sigma}}^{(2)} & \cdots & S^{-1}\widehat{\boldsymbol{\Sigma}}^{(S)} \\ \widehat{\boldsymbol{\Sigma}}^{(1)} & \widehat{\boldsymbol{\Sigma}}^{(1)} & \mathbf{O}_{p\times p} & \cdots & \mathbf{O}_{p\times p} \\ \widehat{\boldsymbol{\Sigma}}^{(2)} & \mathbf{O}_{p\times p} & \widehat{\boldsymbol{\Sigma}}^{(2)} & \cdots & \mathbf{O}_{p\times p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \widehat{\boldsymbol{\Sigma}}^{(S)} & \mathbf{O}_{p\times p} & \mathbf{O}_{p\times p} & \cdots & \widehat{\boldsymbol{\Sigma}}^{(S)} \end{pmatrix}, \\
\mathbf{b}_1 &= \begin{pmatrix} \lambda_1 + \mathbf{e}_i \\ \lambda_2 + \mathbf{e}_i \\ \vdots \\ \lambda_2 + \mathbf{e}_i \end{pmatrix}_{(S+1)p\times 1}, \quad \mathbf{b}_2 = \begin{pmatrix} \lambda_1 - \mathbf{e}_i \\ \lambda_2 - \mathbf{e}_i \\ \vdots \\ \lambda_2 - \mathbf{e}_i \end{pmatrix}_{(S+1)p\times 1}, \quad \mathbf{b}_3 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{p\times 1}, \\
\widetilde{\mathbf{B}} &= \Big(\mathbf{O}, \underbrace{\mathbf{I}_p, \cdots, \mathbf{I}_p}_{S}\Big).
\end{aligned}$$

3.4.2 Gap-block data split for cross-validation

Because of the long-memory property, we adopt the gap-block cross-validation method proposed in [35] to select tuning parameters. We illustrate the steps below.

- *Step 1.* We split the data $\mathbf{X}_{p\times n}^{(s)}$ into H_1 blocks $\mathbf{X}_{p\times n}^{(s)} = (\mathbf{X}_1^{(s)*}, \mathbf{X}_2^{(s)*}, \cdots, \mathbf{X}_{H_1}^{(s)*})$ ($H_1 \geq 4$); They are non-overlapping and all approximately equal-sized blocks. For every i, $\mathbf{X}_i^{(s)*}$ will be used as validation data, and it is dropped along with its neighboring blocks at both sides of $\mathbf{X}_i^{(s)*}$ from $\mathbf{X}_{p\times n}^{(s)}$. Denote the remaining data as $\mathbf{X}_i^{(s)**}$, which will be used as training data.

- *Step 2.* From $\mathbf{X}_{p\times n}^{(s)}$, randomly sample H_2 blocks $\mathbf{X}_{H_1+1}^{(s)*}, \cdots, \mathbf{X}_{H_1+H_2}^{(s)*}$; Every $\mathbf{X}_{H_1+j}^{(s)*}$ consists of $\lceil n/H_1 \rceil$ consecutive columns of $\mathbf{X}_{p\times n}^{(s)}$, $j = 1, \cdots, H_2$. For every $i = H_1 + 1, \cdots, H_1 + H_2$, $\mathbf{X}_i^{(s)*}$ will be used as validation data, and it is dropped along with its neighboring blocks at both sides of $\mathbf{X}_i^{(s)*}$ from $\mathbf{X}_{p\times n}^{(s)}$. Denote the remaining data as $\mathbf{X}_i^{(s)**}$, which will be used as training data.

In our numerical experiment, we select $H_1 = H_2 = 10$ by considering the long-memory characteristic and the 10-fold CV recommendation (see [20,24]).

3.5 Numerical study

3.5.1 Simulation settings

We generate the common structure $\mathbf{C}_0 = (c_{0ij})$ of precision matrices from one of the following two models, where Model 1 follows Model 4 of [35], and Model 2 follows Model 3 of [27]:

- *Model 1.* $c_{0ii} = 1$, $c_{0i(i+1)} = c_{0(i+1)i} = 0.6$, $c_{0i(i+2)} = c_{0(i+2)i} = 0.3$, and $c_{0ij} = 0$ for $|i - j| \geq 3$.

- *Model 2.* $\mathbf{C}_0 = \Gamma + \delta\mathbf{I}$, where every diagonal entry in Γ is 0, and every off-diagonal entry in Γ independently follows the Bernoulli distribution, the success probability is 0.02 and the success entry is valued as 0.5.

For every $\mathbf{U}_0^{(s)}$, we randomly select a pair of symmetric elements (off-diagonal), and assign a value randomly generated from the interval $[-1, -0.5] \cup [0.5, 1]$. We repeat this procedure until $\sum_{i<j} \mathbb{1}\left(\left|u_{0ij}^{(s)}\right| > 0\right) / \sum_{i<j} \mathbb{1}\left(|c_{0ij}| > 0\right) = \rho$, where $\mathbf{C}_0 = (c_{0ij})$ and $\mathbf{U}_0^{(s)} = (u_{0ij}^{(s)})$. The δ is selected sufficiently large such that $\mathbf{C}_0 + \mathbf{U}_0^{(s)}$ is positive definite. Finally, each matrix $\mathbf{\Omega}_0^{(s)} = \mathbf{C}_0 + \mathbf{U}_0^{(s)}$ will be divided by δ to unitization diagonals.

Following [36], we set $Corr(X_{ki}, X_{\ell j}) = \Lambda_{k\ell}^{ij}\rho_{k\ell}$, where

$$\Lambda_{k\ell}^{ij} = (|i-j|+1)^{-\alpha}, \quad 1 \le i, j \le n, \tag{3.5.1}$$

then $\rho_{[k]}^{ij} \sim |i-j|^{-\alpha}$. Due to the large dimension, we simulate the data by the method of [7] as in [36] so that (3.5.1) is approximately satisfied.

The simulations are conducted with group size $S = 4$, sample size $n = 200$, variable dimension p ranging from 50 to 100, $\alpha \in \{0.1, 0.25, 0.5, 1, 2\}$, $\rho \in \{0.1, 0.25, 1, 4\}$, and 50 replications under each setting.

3.5.2 Tuning parameter selection

We have two tuning parameters λ_1 and λ_2 satisfying $\lambda_1 \le \lambda_2$. We select the optimal pair of λ_1 and λ_2 by minimizing the following likelihood loss (LL) (see [27])

$$LL = \sum_{k=1}^{K}\sum_{s=1}^{S}\left[\mathrm{tr}(\widehat{\mathbf{\Sigma}}_{(k)}^{(s)*}\widehat{\mathbf{\Omega}}_{(k)}^{**(s)}) - \log\det\left(\widehat{\mathbf{\Omega}}_{(k)}^{**(s)}\right)\right],\ K = H_1 + H_2,$$

where the sample covariance matrix $\widehat{\mathbf{\Sigma}}_{(k)}^{(s)*}$ is computed from the validation set $\mathbf{X}_{(k)}^{(s)*}$. We compare the performance of two methods, separately estimate each group by CLIME (SCLIME) and JEMP. We compute SCLIME using the R package `flare` [28], and adopt the default perturbation $\varepsilon = n^{-1/2}$ of `flare`. Also, the optimal tuning parameter for SCLIME is chosen from 50 candidates as in [35]. The optimal tuning parameter (λ_1, λ_2) for JEMP is chosen from 6×6 candidates. The JEMP is computed by the R package `fastclime` [31].

3.5.3 Simulation results

Tables 3.1-3.2 report the results of the simulation. Tables 3.1 and 3.2 imply that, (i) if $\alpha \leq 0.25$, JEMP performs significantly better than SCLIME for both models in terms of the spectral and the Frobenius norms; (ii) if $\alpha = 0.5$ JEMP performs significantly better than SCLIME in Model 2; (iii) if $\alpha = 0.5$ and $\rho \in \{0.25, 1, 4\}$, JEMP performs better than SCLIME in Model 1; (iv) otherwise, SCLIME performs better than JEMP.Therefore, the stronger temporal dependence is, the better JEMP performs.

For Model 1 or Modle 2 with $p = 50$ or 100, we compute the computational time (hours) for the 20 combinantions of $\alpha \times \rho = \{0.1, 0.25, 0.5, 1, 2\} \times \{0.1, 0.25, 1, 4\}$. The computation time of JEMP is faster than SCLIME as can be seen in Tables 4.5 and 4.6 for $p = 50$ and $p = 100$. Note that `fastclime` performs $2(1+S)p$ dimensional parametric linear computation. Due to the very high dimension, JEMP will not work if λ_1 and λ_2 are not sufficiently large. Since λ_1 and λ_2 are too large, for example, for any combination of $\lambda_1 = \{0.3, 0.31, 0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4\}$ and $\lambda_2 = \{0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4, 0.41, 0.42\}$, `fastclime` will estimate many nonzero elements to be 0.

3.5.4 rfMRI data analysis

For the estimation of precision matrices and the brain functional connectivity, we analyze the rfMRI data of three healthy subjects ($S = 3$), provided by the WU-Minn Human Connectome Project (www.humanconnectome.org). The original data of each subject consist of $1,200$ temporal brain images, and each image contains about 2×10^5 brain voxels. Due to the high image dimension, we reduce it to 907 by the grid-based method [37]. Also, the first 10 images are discarded because of the early magnetization effect. We choose four subjects, Subjects 100307, 100408, 101006 and 101410 from the data, at the first step to test the stationarity, linearity, and Gaussianity for the reliability of data. Each of them has 907 functional brain nodes.

Table 3.1: Comparison of average (SD) matrix losses for $p = 50$

		SCLIME	JEMP	SCLIME	JEMP
α	ρ	Spectral norm		Frobenius norm	
		Model 1			
0.1	0.1	7.3422(0.3566)	4.0894(0.4229)	30.0442(0.9959)	12.5555(0.4134)
	0.25	7.4576(0.2451)	4.2788(0.3776)	31.0838(0.8742)	13.6734(0.4866)
	1	7.5935(0.3112)	4.0227(0.3617)	31.281(0.9004)	12.4848(0.4358)
	4	8.4898(0.7337)	3.3796(0.3129)	34.0082(2.694)	10.6545(0.4399)
0.25	0.1	2.5548(0.1166)	1.1445(0.05)	8.979(0.3933)	4.0661(0.0608)
	0.25	2.5351(0.1314)	1.0972(0.0638)	9.0665(0.4166)	3.6482(0.0989)
	1	2.6534(0.1397)	1.086(0.0654)	9.3876(0.466)	3.6719(0.0717)
	4	3.188(0.4022)	1.1329(0.0433)	10.6842(1.3126)	3.8565(0.0469)
0.5	0.1	1.2543(0.0693)	1.3681(0.0199)	4.0387(0.168)	4.1381(0.0342)
	0.25	1.2325(0.0785)	1.1035(0.0221)	3.9957(0.1707)	3.2564(0.0359)
	1	1.3004(0.0795)	1.213(0.0175)	4.1842(0.219)	3.7055(0.0352)
	4	1.6605(0.2025)	1.301(0.0133)	5.1548(0.5732)	4.3187(0.0371)
1	0.1	0.8006(0.0312)	1.4709(0.0161)	2.566(0.0517)	4.539(0.0237)
	0.25	0.7321(0.028)	1.2096(0.0156)	2.413(0.0369)	3.6048(0.0408)
	1	0.7623(0.0233)	1.3199(0.0121)	2.5553(0.0477)	4.1461(0.0335)
	4	1.1566(0.0769)	1.4015(0.0104)	3.5698(0.1661)	4.8118(0.0308)
2	0.1	0.8013(0.0311)	1.4918(0.0113)	2.3509(0.0553)	4.6469(0.0237)
	0.25	0.7229(0.0226)	1.2269(0.0154)	2.1873(0.0437)	3.7163(0.0301)
	1	0.7416(0.0216)	1.3395(0.0095)	2.2716(0.0355)	4.2496(0.0256)
	4	0.9755(0.0448)	1.4227(0.0092)	3.1362(0.1101)	4.9376(0.0272)
		Model 2			
0.1	0.1	7.4786(0.2101)	4.5717(0.3843)	31.8289(0.758)	15.0458(0.4386)
	0.25	7.5204(0.2372)	4.4613(0.3668)	31.9076(0.8567)	14.7588(0.4925)
	1	7.5323(0.2538)	4.4467(0.401)	31.6324(0.8335)	14.3266(0.4507)
	4	6.8425(0.7066)	2.9579(0.4899)	27.053(2.3511)	7.9602(1.0064)
0.25	0.1	2.5422(0.1049)	1.1459(0.0972)	9.3173(0.346)	3.2817(0.1565)
	0.25	2.547(0.1256)	1.1463(0.1021)	9.3313(0.3673)	3.2976(0.1486)
	1	2.5592(0.0984)	1.147(0.1031)	9.3329(0.3435)	3.3361(0.1197)
	4	2.742(0.2182)	1.4989(0.0339)	9.7537(0.7667)	4.0451(0.071)
0.5	0.1	1.2162(0.0715)	0.6056(0.0307)	3.9003(0.1677)	1.7048(0.0536)
	0.25	1.1843(0.0578)	0.6923(0.0453)	3.8548(0.1488)	1.8129(0.0631)
	1	1.2162(0.0642)	0.9067(0.029)	3.8981(0.1728)	2.4029(0.0542)
	4	1.6245(0.1341)	1.6221(0.0202)	5.4452(0.3647)	4.6693(0.1041)
1	0.1	0.5633(0.0356)	0.6489(0.0274)	1.8395(0.0732)	1.7925(0.0627)
	0.25	0.569(0.0312)	0.7688(0.027)	1.8464(0.0588)	1.9618(0.0509)
	1	0.6064(0.028)	1.0036(0.0169)	1.977(0.0494)	2.6913(0.047)
	4	1.0828(0.0455)	1.6728(0.0229)	3.7852(0.0725)	5.0789(0.0967)
2	0.1	0.4888(0.0214)	0.6722(0.0205)	1.4671(0.0301)	1.9788(0.044)
	0.25	0.5153(0.0267)	0.7901(0.0224)	1.5084(0.0297)	2.1447(0.0425)
	1	0.5617(0.0217)	1.021(0.015)	1.664(0.0333)	2.8188(0.0307)
	4	0.9993(0.0238)	1.6826(0.0208)	3.462(0.0354)	5.1763(0.1021)

Table 3.2: Comparison of average (SD) matrix losses for $p = 100$

		SCLIME	JEMP	SCLIME	JEMP
α	ρ	Spectral norm		Frobenius norm	
		Model 1			
0.1	0.1	8.2916(0.8387)	3.4815(0.3047)	49.9714(4.582)	13.5156(0.3612)
	0.25	8.7608(0.4786)	3.7798(0.3)	53.5359(2.4532)	14.7279(0.3482)
	1	8.066(0.4542)	3.4128(0.2821)	49.879(2.47)	13.1391(0.4187)
	4	7.9247(0.5105)	3.2535(0.2874)	48.9273(2.6849)	12.0524(0.3935)
0.25	0.1	2.7247(0.1433)	1.4208(0.0309)	13.3898(0.5393)	5.7686(0.0475)
	0.25	2.7988(0.1287)	1.1212(0.0527)	14.2116(0.5102)	4.7957(0.0734)
	1	2.7569(0.1043)	1.1599(0.023)	13.9332(0.5264)	5.0797(0.0466)
	4	2.6299(0.1222)	1.0925(0.0303)	13.1427(0.6099)	5.1277(0.0387)
0.5	0.1	1.315(0.0857)	1.6063(0.0252)	5.9231(0.2441)	6.3708(0.0736)
	0.25	1.3608(0.0577)	1.2841(0.0182)	6.1493(0.1972)	5.0088(0.0476)
	1	1.3513(0.0658)	1.3546(0.0143)	6.0731(0.2077)	5.7092(0.0564)
	4	1.3435(0.0509)	1.2825(0.0109)	6.0848(0.1834)	5.9603(0.0494)
1	0.1	0.8857(0.0307)	1.6677(0.0189)	3.9214(0.0473)	6.865(0.0576)
	0.25	0.7856(0.0246)	1.3653(0.0117)	3.6088(0.0429)	5.4635(0.0399)
	1	0.8319(0.0221)	1.4229(0.009)	3.8161(0.0366)	6.2067(0.0418)
	4	0.9099(0.0188)	1.3364(0.0104)	4.322(0.0239)	6.388(0.0533)
2	0.1	0.9118(0.0226)	1.677(0.0194)	3.7522(0.0757)	6.9839(0.0734)
	0.25	0.8097(0.0188)	1.3791(0.0098)	3.3759(0.0461)	5.6197(0.0359)
	1	0.853(0.0157)	1.4337(0.009)	3.5789(0.0435)	6.3182(0.0376)
	4	0.9353(0.0156)	1.3417(0.0063)	4.1351(0.0458)	6.4555(0.0371)
		Model 2			
0.1	0.1	9.2967(0.4369)	4.3392(0.3593)	56.4275(2.3172)	17.8846(0.4256)
	0.25	9.2136(0.4281)	4.2458(0.3316)	56.327(2.2826)	17.3155(0.5017)
	1	8.405(0.4129)	3.8897(0.3091)	51.818(2.3159)	14.9214(0.449)
	4	8.2434(0.6028)	3.6751(0.2552)	50.9657(3.2031)	14.0743(0.3692)
0.25	0.1	2.8867(0.1134)	1.1063(0.1111)	15.0803(0.5158)	4.1584(0.1287)
	0.25	2.8446(0.1146)	1.104(0.1017)	14.9087(0.5353)	4.2513(0.1137)
	1	2.8033(0.1292)	1.1249(0.0391)	14.3192(0.5911)	4.8262(0.0581)
	4	2.7552(0.1266)	1.089(0.0475)	13.8253(0.6054)	4.9262(0.0606)
0.5	0.1	1.3967(0.0565)	0.8057(0.0265)	6.2999(0.2315)	3.1416(0.0549)
	0.25	1.3711(0.0562)	0.8752(0.0226)	6.2353(0.2244)	3.5122(0.0649)
	1	1.3498(0.059)	1.302(0.0159)	6.0388(0.2058)	5.1866(0.0474)
	4	1.3634(0.0536)	1.2598(0.0138)	6.1103(0.2141)	5.4808(0.0464)
1	0.1	0.6559(0.0295)	0.8546(0.0219)	2.9758(0.0736)	3.4163(0.0532)
	0.25	0.6643(0.0237)	0.9297(0.0179)	3.0494(0.0585)	3.8524(0.0525)
	1	0.7278(0.0194)	1.3807(0.0119)	3.3172(0.0423)	5.7364(0.0395)
	4	0.8353(0.0211)	1.3161(0.0095)	3.9022(0.0344)	5.9419(0.0558)
2	0.1	0.6126(0.0202)	0.8725(0.016)	2.541(0.0323)	3.6463(0.043)
	0.25	0.6438(0.0139)	0.9432(0.0106)	2.6622(0.0346)	4.0495(0.0441)
	1	0.7223(0.0232)	1.397(0.0087)	3.01(0.055)	5.8797(0.0295)
	4	0.8619(0.0204)	1.3216(0.01)	3.6323(0.0427)	6.0349(0.0384)

As in [36], all node time series in the four subjects have been checked by the Priestley-Subba Rao test [33], the Hinich's bispectral test [26], and the generalized Jarque-Bera test [2], for stationarity, linearity, and Gaussianity, respectively. The nodes of each subject that do not pass any test are removed from all subjects to keep the same dimension and nodes for the four subjects. Therefore, the linear spatio-temporal model with sub-Gaussian tails, defined in Section 1.3, seems adequate for the data. Detected by the GPH test [22], there are $129, 83, 87$ and 132 long-memory time series in these four subjects, respectively. Finally, we obtain 861 brain nodes of each subject. All these tests are performed at a significant level (adjusted by [3]) of 0.05. Therefore, the models of [5, 6, 13] are not suitable for our data, because their models are about weak temporal dependence.

In order to reduce the dimension, we select 3 subjects, Subjects 100307, 100408 and 101410 to perform the real data analysis. We call them Group 1 to Group 3, respectively.

For every node k, its autocorrelation function $\rho_k(t) := \rho_{[k]}^{ij}$ $(t = |i - j|)$ is approximated by $\hat{\rho}_k(t)$, which is the sample autocorrelation function. Figures 3.1-3.3 are the sample autocorrelations of brain nodes of Group 1 to Group 3, respectively. In each figure, we plot two long-memory time series, the $\max_{1\leq k\leq p} |\widehat{\rho}_k(t)|$ time series, which is what we need to estimate to confirm whether it is long-memory, and $t^{-0.3}$, the upper bound function. Besides, we also plot two more time series to provide some rough understanding of the characteristic of long-memory time series in the brain for each group. Figures 3.1-3.3 show that the rfMRI data of these three groups all approximately satisfy the PDD model (see [35]) with $C_0 = 1$ and $\alpha = 0.30$ since $\max_{1\leq k\leq p} |\hat{\rho}_k(t)| \leq t^{-0.30}$ roughly holds. The figures also illustrate that, for each group, the estimated autocorrelation functions of the two selected brain nodes clearly have different patterns.

In order to further reduce the dimension, we use SCLIME to estimate the precision matrix of Group 1, and take the same nodes of three Groups of the first 50 hubs (high connectivity) of Group 1 by SCLIME as our data to perform SCLIME and JEMP and compare the effects. Next, we obtain the estimated precision matrices. We calculate the direct connectivity from

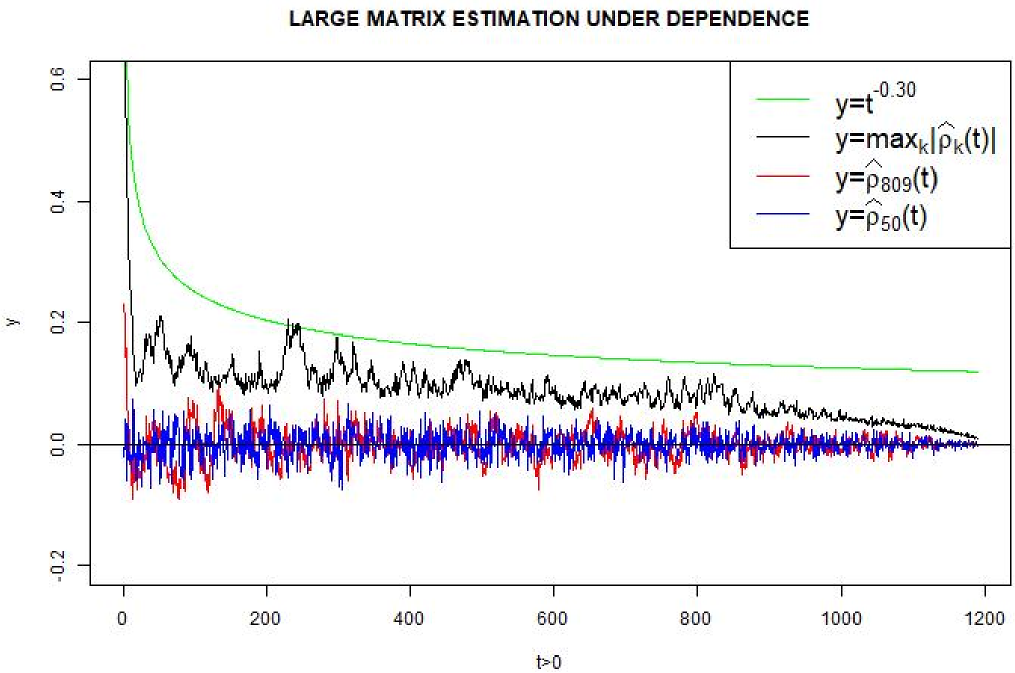

Figure 3.1: Sample autocorrelations of brain nodes of Group 1

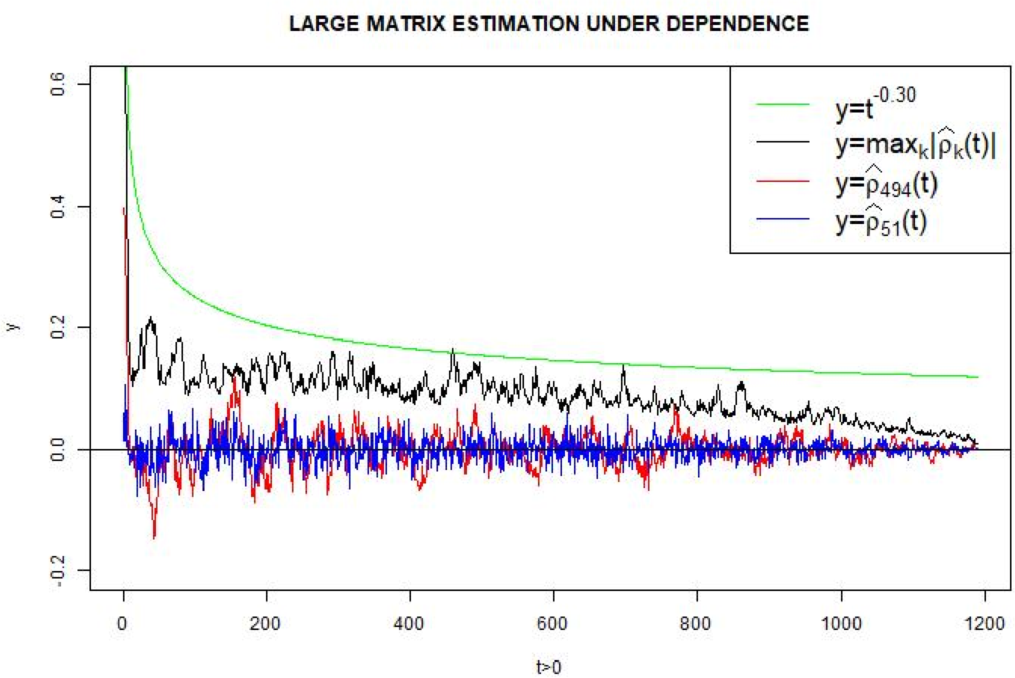

Figure 3.2: Sample autocorrelations of brain nodes of Group 2

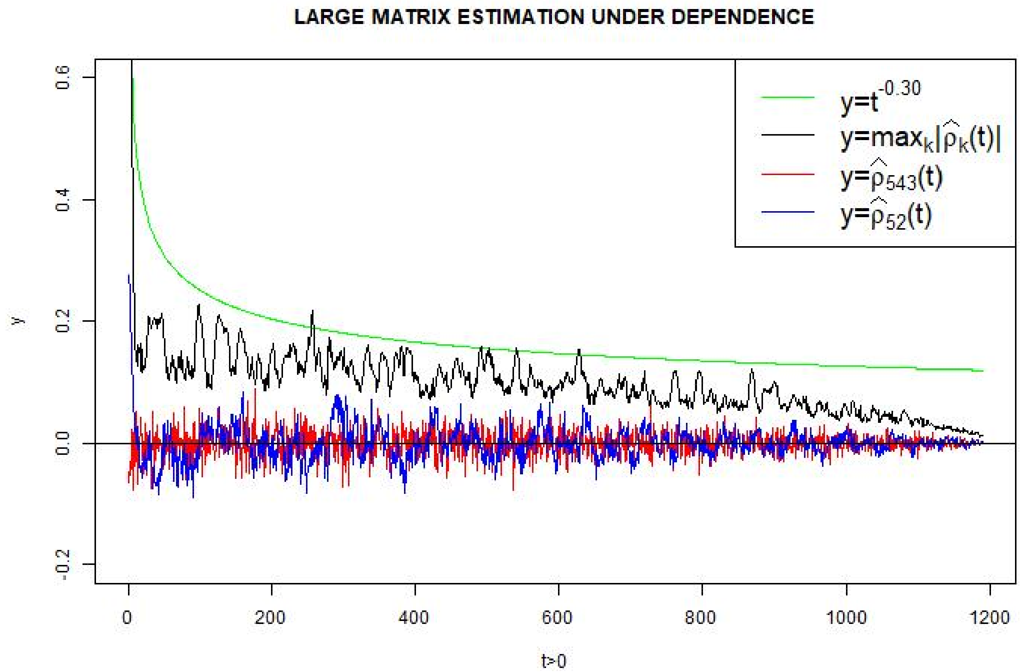

Figure 3.3: Sample autocorrelations of brain nodes of Group 3

the estimated precision matrix $\widehat{\mathbf{\Omega}}^{(s)} := (\hat{\omega}_{ij}^{(s)})_{p \times p}$ via JEMP. For these three groups, the minimum connectivity is 0, the maximum connectivity is 3, and the median is 0. Only 10 time series of Group 1 have nonzero connectivity, and they have 14 connections. On the other hand, for SCLIME, the minimum connectivity is 6, the maximum connectivity is 47, and the median is 43 for Group 1.

The time series indices of the top 10 hubs of three groups by JEMP are shown in Table 3.3. We can see that the six time series (1, 10, 17, 18, 24, and 38) are found in their top 10 hubs in all three groups. For example, time series 10 is ranked the first in all three groups. This implies that JEMP works well, and finds some common information across different subjects.

Table 3.3: The time series indices of the top 10 hubs for direct connectivity of Groups $1-3$ found by JEMP

Rank	Group 1	Group 2	Group 3
1	10	10	10
2	4	18	18
3	18	19	17
4	1	17	23
5	7	21	24
6	17	20	36
7	21	38	38
8	24	24	1
9	28	34	2
10	38	1	3

Table 3.4: Top 10 hubs for direct connectivity of Group 1 found by JEMP and comparison with SCLIME

JEMP Rank	Order in 50	JEMP Degree	SCLIME rank	SCLIME degree
1	10	3	3	46
2	4	2	8	44
3	18	2	10	44
4	1	1	7	44
5	7	1	5	45
6	17	1	19	43
7	21	1	2	47
8	24	1	13	44
9	28	1	16	44
10	38	1	38	42

Table 3.4 lists top 10 hubs of Group 1 found by JEMP and their degrees of connectivity. Also, they are compared with the results from SCLIME. It can be seen that top 10 hubs of Group 1 by JEMP has 6 overlaps with those of SCLIME: time series 1, 4, 7, 10, 18, 21. Moreover, time series 24 and 26 also have 44 degree, which is the same as that of time series 1, 4, or 18. Hence, that time series could be considered as in top 10 as well. It implies that the results of JEMP are similar to those of CLIME for this subject.

Chapter 4

Joint estimation of multiple precision matrices with weighted common structures for high-dimensional time series with long-memory

In Chapter 3, we obtain the precision matrices by JEMP that produces a group precision matrix through the common structure, which is captured by the simple average of individual precision matrices. This is particularly useful in fMRI studies because it can produce a representative brain network based on the estimated group precision matrix from multiple subjects. If some subjects show different brain network patterns from the majority, however, this may affect the construction of the common structure by JEMP because the simple average cannot handle these outlying subjects. In [16], the authors proposed a new joint method for estimating group and individual precision matrices by assigning different weights to different groups (subjects in fMRI examples). We note that, however, they developed the method under the i.i.d assumption. In this chapter, we extend the result of Chapter 3 to the weighted aggregation of individual precision matrices for long-memory data, motivated by [16], [27] and [35]. In particular, we remove the restriction of $n_1 = \cdots = n_S$ in theory of JEMP, and use the weights $w^{(s)}$ for the sth group to replace the uniform weight $\frac{1}{S}$ for all groups.

In this chapter, similar to Chapter 3, we consider a heterogeneous data set with S different groups. The notations $\{x_1^{(s)}, \cdots, x_{n_s}^{(s)}\}$, n_s, $\mathbf{x}_k^{(s)} = (x_{k1}^{(s)}, \cdots, x_{kp}^{(s)})^T$, p, $\boldsymbol{\Sigma}_0^{(s)} = (\sigma_{ij,0}^{(s)})$, $\mathbf{X}^{(s)}$, $\widehat{\boldsymbol{\Sigma}}^{(s)}$, $\boldsymbol{\Omega}_0^{(s)} = (\omega_{ij,0}^{(s)})$ and precision matrix $\boldsymbol{\Omega}_0^{(s)} := (\boldsymbol{\Sigma}_0^{(s)})^{-1}$ $(s = 1, \cdots, S)$ have the same meaning as in Chapter 3. Our aim is to estimate the precision matrices $\boldsymbol{\Omega}_0^{(1)}, \cdots, \boldsymbol{\Omega}_0^{(S)}$ for long-memory data by weighted-JEMP based on different weights $w^{(s)}$.

We consider the following class of matrices

$$\mathcal{U} := \left\{\mathbf{\Omega} : \mathbf{\Omega} \succ 0, \|\mathbf{\Omega}\|_1 \leq C_M, |\mathbf{\Omega}|_1 \leq C_M^*\right\},$$

and assume that $\mathbf{\Omega}_0^{(s)} \in \mathcal{U}$ for all $s = 1, \cdots, S$. Write $\mathbb{E}(\mathbf{x}^{(s)}) = (\mu_1^{(s)}, \cdots, \mu_p^{(s)})^T$.

Also, we define the common structure $\mathbf{C}_0$ and the unique structure $\mathbf{U}_0^{(s)}$ as

$$\mathbf{C}_0 := \sum_{s=1}^{S} w^{(s)} \mathbf{\Omega}_0^{(s)}, \mathbf{U}_0^{(s)} = \mathbf{\Omega}_0^{(s)} - \mathbf{C}_0, \sum_{s=1}^{S} w^{(s)} = 1, w^{(s)} \geq 0, s = 1, \cdots, S.$$

Note that $\sum_{s=1}^{S} w^{(s)} \mathbf{U}_0^{(s)} = 0$. In order to estimate $\{\mathbf{C}_0, \mathbf{U}_0^{(1)}, \cdots, \mathbf{U}_0^{(S)}\}$, we investigate the following constrained entrywise L^1 minimization problem [16]:

$$\min \left\{ |\mathbf{C}|_1 + \nu \sum_{s=1}^{S} w^{(s)} \left|\mathbf{U}^{(s)}\right|_1 \right\}$$

$$\text{s.t} \left| \sum_{s=1}^{S} w^{(s)} \left\{ \widehat{\mathbf{\Sigma}}^{(s)} (\mathbf{C} + \mathbf{U}^{(s)}) - \mathbf{I} \right\} \right|_\infty \leq \lambda_1, \left| \widehat{\Sigma}^{(s)} (\mathbf{C} + \mathbf{U}^{(s)}) - \mathbf{I} \right|_\infty \leq \lambda_2, \sum_{s=1}^{S} w^{(s)} \mathbf{U}^{(s)} = 0, \tag{4.0.1}$$

where ν is a prespecified weight, and (λ_1, λ_2) are tuning parameters with $\lambda_1 \leq \lambda_2$ as in Chapter 3.

We can also consider the following vector optimal problem:

$$\min \left\{ \left| \sum_{s=1}^{S} w^{(s)} \mathbf{c}_j^{(s)} \right|_1 + \nu \sum_{s=1}^{S} w^{(s)} \left| \mathbf{c}_j^{(s)} - \sum_{s=1}^{S} w^{(s)} \mathbf{c}_j^{(s)} \right|_1 \right\}$$

$$\text{s.t} \left| \sum_{s=1}^{S} w^{(s)} \left\{ \widehat{\Sigma}^{(s)} \mathbf{c}_j^{(s)} - \mathbf{e}_j \right\} \right|_\infty \leq \lambda_1, \left| \widehat{\Sigma}^{(s)} \mathbf{c}_j^{(s)} - \mathbf{e}_j \right|_\infty \leq \lambda_2, \tag{4.0.2}$$

where $\mathbf{c}_j^{(s)}$ is a vector in $\mathbb{R}^p$, $\mathbf{e}_j$ is the jth colomn of the identity matrix $\mathbf{I}_p$. Again, by the proof of Lemma 1 of [11], it is easy to see that these two problems have the same solution set. We denote $N = \sum_{s=1}^{S} n_s$ in this chapter, and ν is set to be S^0 or S^θ for our theoretical results, so it will corresponding to $1/S$ or $\nu = 1/S^{1-\theta}$ (like in chapter 3) if we let $w^{(s)} \equiv 1/S$.

Assume $\{\widehat{\mathbf{C}}_1, \widehat{\mathbf{U}}_1^{(1)}, \cdots, \widehat{\mathbf{U}}_1^{(S)}\}$ is a soution of (4.0.1), and denote $\widehat{\mathbf{\Omega}}_1^{(s)} := \widehat{\mathbf{C}}_1 + \widehat{\mathbf{U}}_1^{(s)}, s = 1, \cdots, S$. Using the same reason in Chapter 3, we define our final weighted-JEMP estimator $\{\widehat{\mathbf{\Omega}}^{(1)}, \cdots, \widehat{\mathbf{\Omega}}^{(S)}\}$ of $\{\mathbf{\Omega}_0^{(1)}, \cdots, \mathbf{\Omega}_0^{(S)}\}$ as the symmetrized result of $\{\widehat{\mathbf{\Omega}}_1^{(1)}, \cdots, \widehat{\mathbf{\Omega}}_1^{(S)}\}$ by the following way. Denote $\widehat{\mathbf{\Omega}}_1^{(s)} = (\widehat{\omega}_{ij,1}^{(s)})$, and $\widehat{\mathbf{\Omega}}^{(s)} = (\widehat{\omega}_{ij}^{(s)})$, for $s = 1, \cdots, S$, then

$$\begin{aligned} \widehat{\omega}_{ij}^{(s)} &= \widehat{\omega}_{ij,1}^{(s)} \mathbb{1}\{\sum_{s=1}^{S} w^{(s)} \left|\widehat{\omega}_{ij,1}^{(s)}\right| \le \sum_{s=1}^{S} w^{(s)} \left|\widehat{\omega}_{ji,1}^{(s)}\right|\} \\ &\quad + \widehat{\omega}_{ji,1}^{(s)} \mathbb{1}\{\sum_{s=1}^{S} w^{(s)} \left|\widehat{\omega}_{ij,1}^{(s)}\right| > \sum_{s=1}^{S} w^{(s)} \left|\widehat{\omega}_{ji,1}^{(s)}\right|\}, \end{aligned} \tag{4.1.1}$$

for $s = 1, \cdots, S$. Then $\widehat{\mathbf{C}} = \sum_{s=1}^{S} w^{(s)} \widehat{\mathbf{\Omega}}^{(s)}$.

Define $\widehat{\mathbf{\Sigma}}^{(s)} := \left(\widehat{\sigma}_{ij}^{(s)}\right)_{p\times p}$ the sample covariance matrix given by

$$\widehat{\mathbf{\Sigma}}^{(s)} = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{x}_i^{(s)} \mathbf{x}_i^{(s)T} - \overline{\mathbf{x}}^{(s)} \overline{\mathbf{x}}^{(s)T}$$

with $\overline{\mathbf{x}}^{(s)} = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{x}_i^{(s)}$, $\mathbf{x}_i^{(s)}$ is the ith column of $\mathbf{X}^{(s)}$. The following theorem generalizes Theorem 1 of [27] from i.i.d to long-memory observations.

Theorem 4.1.1. *Suppose that $\mathbf{X}_{p\times N}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K. Let $\lambda_1 = \lambda_2 = 6C_M \max_s \tilde{u}_1^{(s)}$. Set $\nu = S^0 = 1$. Then*

$$\max_{ij} \left(\sum_{s=1}^{S} w^{(s)} \left|\widehat{\omega}_{ij}^{(s)} - \omega_{ij,0}^{(s)}\right| \right) \le 12 C_M^2 \max_s \tilde{u}_1^{(s)}$$

with probability greater than $1 - 4Sp^{-\tau}$, where $\tilde{u}_1^{(s)}$ is defined in (4.1.2).

In order to prove Theorem 4.1.1, we need the following lemma, which generalizes Lemma 4 in [27] from i.i.d to long-memory observations.

Lemma 4.1.1. *Suppose that $\mathbf{X}_{p\times N}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K. For any $s = 1, \cdots, S$, with probability greater than $1 - 4p^{-\tau}$,*

$$\max_{ij} \left|\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)}\right| \le 2\widetilde{u}_1^{(s)},$$

where K is given in Theorem 4.1.1, and $\tilde{u}_1^{(s)}$ is defined in (4.1.2).

Proof of Lemma 4.1.1. By the proof of Proposition 4 and the proof of the first part of Lemma A2 in [35] we have

$$\begin{aligned} P\left[\left|\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}\right|_\infty \geq 2u\right] &\leq 2p\exp\left\{-\frac{c_1 n_s u}{K^2 g_2^{(s)}}\right\} \\ &\quad +2p^2\exp\left\{-c_1\min\left(\frac{n_s u^2}{K^4 g_F^{(s)}}, \frac{n_s u}{K^2 g_2^{(s)}}\right)\right\}, \end{aligned}$$

where $c_1 > 0$ is an absolute constant. Let

$$u = \tilde{u}_1^{(s)} := \max\left\{c_2 K^2(\log p) g_2^{(s)}(n_s)/n_s, [c_2 K^4(\log p) g_F^{(s)}(n_s)/n_s]^{1/2}\right\}, \tag{4.1.2}$$

where $c_2 = (\tau + 2)/c_1$ with $\tau > 0$, and

$$\max_{1\leq k\leq p}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_2 \leq g_2^{(s)}(n_s), \max_{1\leq k\leq p}\frac{1}{n_s}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_F^2 \leq g_F^{(s)}(n_s), \tag{4.1.3}$$

where $\mathbf{x}_{[k]}^{(s)}$ is the k-th row vector of $\mathbf{X}^{(s)}$, $\mathbf{x}_{[k]}$ is the kth row vector of $\mathbf{X} = (\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(S)})$, then

$$P\left[|\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}|_\infty \geq 2\tilde{u}_1^{(s)}\right] \leq 2p^{-(c_1c_2-1)} + 2p^{-(c_1c_2-2)} \leq 4p^{-\tau} = O(p^{-\tau}).$$

The proof is completed. □

Proof of Theorem 4.1.1. We recap the proof of Theorem 1 in [27]. It is easy to see from Lemma 4.1.1 that

$$|\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}|_\infty \leq 2\tilde{u}_1^{(s)} \text{ for all } s = 1, \cdots, S, \tag{4.1.4}$$

with probability greater than $1 - 4Sp^{-\tau}$. In the rest of the proof, we assume (4.1.4) holds. Note that $\{\mathbf{C}_0, \mathbf{U}_0^{(1)}, \cdots, \mathbf{U}_0^{(S)}\}$ is a feasible solution of (4.0.1), because

$$\begin{aligned} \left|\mathbf{I} - \widehat{\mathbf{\Sigma}}^{(s)}(\mathbf{C}_0 + \mathbf{U}_0^{(s)})\right|_\infty &= \left|(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\mathbf{\Omega}_0^{(s)}\right|_\infty \\ &\leq \left\|\mathbf{\Omega}_0^{(s)}\right\|_1 |\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}|_\infty \\ &\leq C_M 2\tilde{u}_1^{(s)} \leq \lambda_2/3. \end{aligned}$$

Similar to the proof of Theorem 1 of [27], we have

$$\sum_{s=1}^S w^{(s)}|(\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j|_\infty \leq \sum_{s=1}^S w^{(s)}\left|\widehat{\mathbf{\Omega}}_1^{(s)}\mathbf{e}_j\right|_1 \lambda_2/3 + C_M\lambda_2$$

$$
\begin{aligned}
&\le \left\{|\widehat{\mathbf{c}}_j|_1 + \sum_{s=1}^{S} w^{(s)} \left|\widehat{\mathbf{u}}_j^{(s)}\right|_1\right\} \lambda_2/3 + C_M\lambda_2 \\
&\le \left\{|\mathbf{c}_{j,0}|_1 + \sum_{s=1}^{S} w^{(s)} \left|\mathbf{u}_{j,0}^{(s)}\right|_1\right\} \lambda_2/3 + C_M\lambda_2 \\
&\le 2C_M\lambda_2 = 12C_M^2 \max_s \tilde{u}_1^{(s)},
\end{aligned}
$$

where $\widehat{\mathbf{c}}_j$ is the jth column of $\widehat{\mathbf{C}}$, and $\widehat{\mathbf{u}}_j^{(s)}$ is the jth column of $\widehat{\mathbf{U}}^{(s)}$, $\mathbf{c}_{j,0}$ is the jth column of $\mathbf{C}_0$, $\mathbf{u}_{j,0}^{(s)}$ is the jth column of $\mathbf{U}_0^{(s)}$.

It is straightforward to check that

$$
\max_{ij}\left(\sum_{s=1}^{S} w^{(s)} \left|\widehat{\omega}_{ij}^{(s)} - \omega_{ij,0}^{(s)}\right|\right) \le \max_j \sum_{s=1}^{S} w^{(s)} |(\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j|_\infty \le 12C_M^2 \max_s \tilde{u}_1^{(s)}.
$$

Then, the proof is completed. □

The following theorem generalizes Theorem 2 of [27] from i.i.d to long-memory observations.

Theorem 4.1.2. *Suppose that (i) $\mathbf{X}_{p\times N}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K, (ii) there exists $C_U > 0$ such that $\left\|\mathbf{U}_0^{(s)}\right\|_1 \le C_U$ for all $s = 1, \cdots, S$ and $(\sum_{s=1}^S \left\|\mathbf{U}_0^{(s)}\right\|_1) \le C_U S^{1-\theta}$ where $\theta \in (0,1)$. Let*

$$
\begin{aligned}
\psi_1 &= C_M\left(\tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p\right) + C_U S^{1-\theta} \max_s 2w^{(s)}\tilde{u}_1^{(s)}, \\
\lambda_1 &= \psi_1, \lambda_2 = 6C_M \max_s \widetilde{u}_1^{(s)}.
\end{aligned}
$$

Set $\nu = S^\theta$. Then

$$
\left|\widehat{\mathbf{C}} - \mathbf{C}_0\right|_\infty \le 2(C_M + 2C_U)\max\{\psi_1, S^{-\theta}C_M \max_s 2\widetilde{u}_1^{(s)}\} \max_s Sw^{(s)}
$$

with probability greater than $1 - 2(2+3S)p^{-\tau}$, where $\tilde{u}_1^{(s)}$ is defined in (4.1.2), $\tilde{u}_1$ is given in (4.1.6), and $\xi_2^{(s)}$ is defined in (4.1.8).

In order to prove Theorem 4.1.2, we need the following lemma, which generalizes Lemma 5 in [27] from i.i.d to long-memory observations.

Lemma 4.1.2. *Suppose that $\mathbf{X}_{p\times N}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K. With probability greater than $1-2(2+S)p^{-\tau}$, the following holds*

$$\max_{ij}\left|\sum_{s=1}^{S} w^{(s)}(\widehat{\sigma}_{ij}^{(s)}-\sigma_{ij,0}^{(s)})\right| \le \tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p,$$

where $\tilde{u}_1$ is given in (4.1.6), and $\xi_2^{(s)}$ is defined in (4.1.8).

Proof of Lemma 4.1.2. We may assume that $\mu_i^{(s)}=0$ for all i and s. Denote

$$\begin{aligned} y_{kij}^{(s)} \quad &: \quad = w^{(s)}x_{ik}^{(s)}x_{jk}^{(s)} - \mathbb{E}(w^{(s)}x_{ik}^{(s)}x_{jk}^{(s)}) \\ &= \quad \left(\sqrt{w^{(s)}}x_{ik}^{(s)}\right)\left(\sqrt{w^{(s)}}x_{jk}^{(s)}\right) - \mathbb{E}\left(\left(\sqrt{w^{(s)}}x_{ik}^{(s)}\right)\left(\sqrt{w^{(s)}}x_{jk}^{(s)}\right)\right). \end{aligned}$$

Also denote $\overline{x}_i^{(s)} = \sum_{k=1}^{n_s} x_{ik}^{(s)}/n_s$. Similar to the proof of Lemma 5 of [27], we have

$$\sum_{s=1}^{S} w^{(s)}\left(\widehat{\sigma}_{ij}^{(s)}-\sigma_{ij,0}^{(s)}\right) = \sum_{s=1}^{S}\left(\sum_{k=1}^{n_s} y_{kij}^{(s)}/n_s - w^{(s)}\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right).$$

We can define $\mathbf{P}_k^{(s)}$ such that $\mathbf{x}_{[k]}^{*(s)} = \mathbf{P}_k^{(s)}\mathbf{H}\mathbf{z} = (\sqrt{w^{(s)}/n_s}\mathbf{X}_{[k]}^{(s)})^T$, which is the transpose of the kth row of $\mathbf{X}^{(s)}$ with the coefficient $\sqrt{w^{(s)}/n_s}$, where $\mathbf{P}_k^{(s)}$ is a $n_s \times pN$ matrix with $\sqrt{w^{(s)}/n_s}$ in the $\big(j, k+(j-1)p\big)$ entries and 0 in all other entries, $j=1,\cdots,n_s$, where $N=\sum_{s=1}^{S} n_s$. Denote $\mathbf{x}_{[k]}^{*} = \mathbf{P}_k\mathbf{H}\mathbf{z} = \left(\left(\mathbf{x}_{[k]}^{*(1)}\right)^T,\cdots,\left(\mathbf{x}_{[k]}^{*(S)}\right)^T\right)^T$, $s=1,\cdots,S$. Also, denote

$$\max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{*})\right\|_F^2 \le \xi_F(N), \ \max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{*})\right\|_2 \le \xi_2(N). \tag{4.1.5}$$

It is easy to check that

$$Cov(\mathbf{x}_{[k]}^{*}) = Cov(\mathbf{P}_k\mathbf{H}\mathbf{z}) = \mathbf{P}_k\mathbf{H}\mathbf{H}^T\mathbf{P}_k^T.$$

By Lemma A1 of [35] and similar to the proof of the first part of Lemma A2 of [35], we have

$$\begin{aligned} &P\left[\max_{ij}\left|\sum_{s=1}^{S}\sum_{k=1}^{n_s}\frac{1}{n_s}y_{kij}^{(s)}\right| \ge u\right] \\ \le \quad & 2p\exp\left\{-c_1\frac{u}{K^2\sqrt{\left\|\mathbf{P}_i\mathbf{H}\mathbf{H}^T\mathbf{P}_i^T\right\|_2\left\|\mathbf{P}_j\mathbf{H}\mathbf{H}^T\mathbf{P}_j^T\right\|_2}}\right\} \end{aligned}$$

$$
+2p^2 \exp\left\{-c_1 \min\left(\frac{u^2}{K^4 \left\|\mathbf{P}_i\mathbf{H}\mathbf{H}^T\mathbf{P}_i^T\right\|_F \left\|\mathbf{P}_j\mathbf{H}\mathbf{H}^T\mathbf{P}_j^T\right\|_F},\right.\right.
$$

$$
\left.\left.\frac{u}{K^2\sqrt{\left\|\mathbf{P}_i\mathbf{H}\mathbf{H}^T\mathbf{P}_i^T\right\|_2 \left\|\mathbf{P}_j\mathbf{H}\mathbf{H}^T\mathbf{P}_j^T\right\|_2}}\right)\right\}
$$

$$
\leq \quad 2p^2 \exp\left\{-c_1 \min\left(\frac{u^2}{K^4\xi_F}, \frac{u}{K^2\xi_2}\right)\right\} + 2p\exp\left\{-c_1\frac{u}{K^2\xi_2}\right\},
$$

where $c_1 > 0$ is an absolute constant.

Let

$$
u = \tilde{u}_1 = \max\left\{c_2K^2(\log p)\xi_2, [c_2K^4(\log p)\xi_F]^{1/2}\right\}, \tag{4.1.6}
$$

with $c_2 = (\tau+2)/c_1$ and $\tau > 0$, where ξ_2 and ξ_F are defined in (4.1.5), then

$$
P\left[\max_{ij}\left|\sum_{s=1}^{S}\sum_{k=1}^{n}\frac{1}{n_s}y_{kij}^{(s)}\right| \geq \tilde{u}_1\right] \leq 2p^{-(c_1c_2-2)} + 2p^{-(c_1c_2-2)} \leq 4p^{-\tau} = O(p^{-\tau}). \tag{4.1.7}
$$

By Lemma A1-(i) of [35] we have

$$
\begin{aligned}
P\left[\max_{ij} w^{(s)}\left|\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq u^2\right] &= P\left[\max_{i} w^{(s)}\left|\overline{x}_i^{(s)}\right|^2 \geq u^2\right] \\
&= P\left[\max_{i}\left|\sqrt{w^{(s)}}\overline{x}_i^{(s)}\right| \geq u\right] \\
&\leq 2p\exp\left[-\frac{c_1n_su^2}{K^2\left\|b\right\|_F^2\left\|\mathbf{P}_i^{(s)}\mathbf{H}\mathbf{H}^T\mathbf{P}_i^{(s)T}\right\|_2}\right] \\
&= 2p\exp\left[-\frac{c_1u^2}{K^2\xi_2^{(s)}}\right],
\end{aligned}
$$

where $c_1 > 0$ is an absolute constant and

$$
\max_{1\leq k\leq p}\left\|Cov(\mathbf{x}_{[k]}^{*(s)})\right\|_2 \leq \xi_2^{(s)}, \max_{1\leq k\leq p}\left\|Cov(\mathbf{x}_{[k]}^{*(s)})\right\|_F^2 \leq \xi_F^{(s)}. \tag{4.1.8}
$$

By taking $u = \left(c_2K^2\xi_2^{(s)}\log p\right)^{1/2}$ with $c_2 = (\tau+2)/c_1$ and $\tau > 0$, we have

$$
P\left[\max_{ij} w^{(s)}\left|\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq c_2K^2\xi_2^{(s)}\log p\right] \leq 2p^{-\tau}.
$$

Using this result, we have that

$$
P\left[\max_{ij}\left|\sum_{s=1}^{S} w^{(s)}\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq \sum_{s=1}^{S} c_2K^2\xi_2^{(s)}\log p\right]
$$

$$
\begin{aligned}
&\leq P\left[\sum_{s=1}^{S} w^{(s)} \max_{ij}\left|\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p\right] \\
&\leq \sum_{s=1}^{S} P\left[w^{(s)} \max_{ij}\left|\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq c_2 K^2 \xi_2^{(s)} \log p\right] \leq 2Sp^{-\tau}. \qquad (4.1.9)
\end{aligned}
$$

By (4.1.7) and (4.1.9), we have

$$
\begin{aligned}
& P\left[\max_{ij}\left|\sum_{s=1}^{S} w^{(s)}\left(\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)}\right)\right| \geq \tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p\right] \\
\leq\; & P\left[\max_{ij}\left|\sum_{s=1}^{S}\sum_{k=1}^{n_s} \frac{1}{n_s} y_{kij}^{(s)}\right| \geq \tilde{u}_1\right] \\
& + P\left[\max_{ij}\left|\sum_{s=1}^{S} w^{(s)}\overline{x}_i^{(s)}\overline{x}_j^{(s)}\right| \geq \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p\right] \\
\leq\; & 2(2+S)p^{-\tau}.
\end{aligned}
$$

The proof is completed. □

Proof of Theorem 4.1.2. We recap the proof of Theorem 2 of [27].

By Lemma 4.1.1 and Lemma 4.1.2, we have

$$
\max_{ij}\left|\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)}\right| \leq 2\tilde{u}_1^{(s)}, \; \max_{ij}\left|\sum_{s=1}^{S} w^{(s)}(\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)})\right| \leq \tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p, \qquad (4.1.10)
$$

for all $s = 1, \cdots, S$ with probability greater than $1 - 2(2+3S)p^{-\tau}$. In the rest of the proof, we assume (4.1.10) holds.

Note that $\{\mathbf{C}_0, \mathbf{U}_0^{(1)}, \cdots, \mathbf{U}_0^{(S)}\}$ is a feasible solution of (4.0.1), because

$$
\begin{aligned}
\left|\mathbf{I} - \widehat{\boldsymbol{\Sigma}}^{(s)}(\mathbf{C}_0 + \mathbf{U}_0^{(s)})\right|_\infty &= \left|(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\boldsymbol{\Omega}_0^{(s)}\right|_\infty \\
&\leq \left\|\boldsymbol{\Omega}_0^{(s)}\right\|_1 |\widehat{\boldsymbol{\Sigma}}^{(s)} - \boldsymbol{\Sigma}_0^{(s)}|_\infty \\
&\leq C_M 2\tilde{u}_1^{(s)} \leq \lambda_2/3,
\end{aligned}
$$

and

$$
\left|\sum_{s=1}^{S} w^{(s)}\left\{\mathbf{I} - \widehat{\boldsymbol{\Sigma}}^{(s)}(\mathbf{C}_0 + \mathbf{U}_0^{(s)})\right\}\right|_\infty
$$

$$
\begin{aligned}
&= \left|\sum_{s=1}^{S} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\mathbf{\Omega}_0^{(s)}\right|_\infty \\
&= \left|\sum_{s=1}^{S} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\mathbf{C}_0\right|_\infty + \left|\sum_{s=1}^{S} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\mathbf{U}_0^{(s)}\right|_\infty \\
&\leq \|\mathbf{C}_0\|_1 \left|\sum_{s=1}^{S} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\right|_\infty + \sum_{s=1}^{S} \left\|\mathbf{U}_0^{(s)}\right\|_1 \left|w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\right|_\infty \\
&\leq \|\mathbf{C}_0\|_1 \left|\sum_{s=1}^{S} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\right|_\infty + \max_s \left|w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\right|_\infty \sum_{s=1}^{S} \left\|\mathbf{U}_0^{(s)}\right\|_1 \\
&\leq C_M \left(\tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p\right) + C_U S^{1-\theta} \max_s 2w^{(s)}\tilde{u}_1^{(s)} \leq \lambda_1.
\end{aligned}
$$

Now, we estimate $\left|(\widehat{\mathbf{C}}_1 - \mathbf{C}_0)\mathbf{e}_j\right| = \left|\sum_{s=1}^{S} w^{(s)}(\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j\right|_\infty$. In particular, we use

$$
\begin{aligned}
\left|\sum_{s=1}^{S} w^{(s)}(\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j\right|_\infty &\leq \left|\sum_{s=1}^{S} w^{(s)}\mathbf{\Omega}_0^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\widehat{\mathbf{\Omega}}_1^{(s)}\mathbf{e}_j\right|_\infty \\
&\quad + \left|\sum_{s=1}^{S} \Omega_0^{(s)} w^{(s)}(\widehat{\mathbf{\Sigma}}^{(s)}\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty. \qquad (4.1.11)
\end{aligned}
$$

Similar to the discussion of Theorem 2 of [27], we have

$$
\begin{aligned}
&\left|\sum_{s=1}^{S} w^{(s)}\mathbf{\Omega}_0^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\widehat{\mathbf{\Omega}}_1^{(s)}\mathbf{e}_j\right|_\infty \\
\leq\ &\left|\sum_{s=1}^{S} w^{(s)}\mathbf{\Omega}_0^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\widehat{\mathbf{c}}_j\right|_\infty + \left|\sum_{s=1}^{S} w^{(s)}\mathbf{\Omega}_0^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\widehat{\mathbf{u}}_j^{(s)}\right|_\infty \\
\leq\ &\left|\sum_{s=1}^{S} \mathbf{C}_0 w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\widehat{\mathbf{c}}_j\right|_\infty + \left|\sum_{s=1}^{S} \mathbf{U}_0^{(s)} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\widehat{\mathbf{c}}_j\right|_\infty \\
&+ \sum_{s=1}^{S} \left|\mathbf{\Omega}_0^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})w^{(s)}\widehat{\mathbf{u}}_j^{(s)}\right|_\infty \\
\leq\ & C_M \left|\sum_{s=1}^{S} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\right|_\infty |\widehat{\mathbf{c}}_j|_1 + \max_s w^{(s)} \left|(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\right|_\infty \sum_{s=1}^{S} \left\|\mathbf{U}_0^{(s)}\right\|_1 |\widehat{\mathbf{c}}_j|_1 \\
&+ C_M \max_s \left|(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\right|_\infty \sum_{s=1}^{S} w^{(s)} \left|\widehat{\mathbf{u}}_j^{(s)}\right|_1 \\
\leq\ & \left(C_M \left|\sum_{s=1}^{S} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\right|_\infty + \max_s w^{(s)} \left|(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\right|_\infty C_U S^{1-\theta}\right) |\widehat{\mathbf{c}}_j|_1
\end{aligned}
$$

$$
\begin{aligned}
& +S^{-\theta}C_M \max_s \left|(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty S^\theta \sum_{s=1}^S w^{(s)} \left|\widehat{\mathbf{u}}_j^{(s)}\right|_1 \\
\leq\ & \max\left\{ C_M \left|\sum_{s=1}^S w^{(s)}(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty + \max_s w^{(s)} \left|(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty C_U S^{1-\theta}, \right. \\
& \left. S^{-\theta}C_M \max_s \left|(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty \right\} \times \left(|\widehat{\mathbf{c}}_j|_1 + S^\theta \sum_{s=1}^S w^{(s)} \left|\widehat{\mathbf{u}}_j^{(s)}\right|_1 \right) \\
\leq\ & \max\left\{ C_M \left|\sum_{s=1}^S w^{(s)}(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty + \max_s w^{(s)} \left|(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty C_U S^{1-\theta}, \right. \\
& \left. S^{-\theta}C_M \max_s \left|(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty \right\} \times \left(|\mathbf{c}_{j,0}|_1 + S^\theta \sum_{s=1}^S w^{(s)} \left|\mathbf{u}_{j,0}^{(s)}\right|_1 \right) \\
\leq\ & (C_M + C_U) \max\{\psi_1, S^{-\theta}C_M \max_s \widetilde{u}_1^{(s)}\} \max_s S w^{(s)},
\end{aligned}
$$

where $\widehat{\mathbf{c}}_j$ is the jth column of $\widehat{\mathbf{C}}_1$, and $\widehat{\mathbf{u}}_j^{(s)}$ is the jth column of $\widehat{\mathbf{U}}_1^{(s)}$, $\mathbf{c}_{j,0}$ is the jth column of $\mathbf{C}_0$, and $\mathbf{u}_{j,0}^{(s)}$ is the jth column of $\mathbf{U}_0^{(s)}$, and

$$
\begin{aligned}
& \left|\sum_{s=1}^S \boldsymbol{\Omega}_0^{(s)} w^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty \\
\leq\ & \left|\sum_{s=1}^S \mathbf{C}_0 w^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty + \left|\sum_{s=1}^S \mathbf{U}_0^{(s)} w^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty \\
\leq\ & \|\mathbf{C}_0\|_1 \left|\sum_{s=1}^S w^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty + \sum_{s=1}^S \left\|\mathbf{U}_0^{(s)}\right\|_1 \left|w^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty \\
\leq\ & C_M\lambda_1 + C_U S^{1-\theta} \max_s w^{(s)} \lambda_2 \\
\leq\ & (C_M + 3C_U) \max\{\psi_1, S^{-\theta}C_M \max_s 2\widetilde{u}_1^{(s)}\} \max_s S w^{(s)}.
\end{aligned}
$$

In sum, we have

$$
\left|\widehat{\mathbf{C}} - \mathbf{C}_0\right|_\infty \leq \left|\widehat{\mathbf{C}}_1 - \mathbf{C}_0\right|_\infty \leq 2(C_M + 2C_U) \max\{\psi_1, S^{-\theta}C_M \max_s 2\widetilde{u}_1^{(s)}\} \max_s S w^{(s)}.
$$

The proof is completed. □

Define a threshold estimator $\widetilde{\boldsymbol{\Omega}}^{(s)} = (\widetilde{\omega}_{ij}^{(s)})$ based on $\{\widehat{\boldsymbol{\Omega}}^{(1)}, \cdots, \widehat{\boldsymbol{\Omega}}^{(S)}\}$ as

$$
\widetilde{\omega}_{ij}^{(s)} = \widehat{\omega}_{ij}^{(s)} \mathbb{1}\left\{ w^{(s)} \left|\widehat{\omega}_{ij}^{(s)}\right| \geq \delta_n \right\},
$$

where $\delta_n \geq 2C_M\lambda_2$ and λ_2 is given in Theorem 4.1.1. Also, define

$$
\mathcal{S}_0 := \{(i,j,s) : \omega_{ij,0}^{(s)} \neq 0\}, \widehat{\mathcal{S}} := \{(i,j,s) : \widetilde{\omega}_{ij}^{(s)} \neq 0\} \text{ and } \theta_{\min} := \min_{(i,j,s)\in\mathcal{S}_0} w^{(s)} \left|\omega_{ij,0}^{(s)}\right|.
$$

The following theorem generalizes Theorem 3 in [27] from i.i.d to long-memory observations. The proof is straightforward. Details are omitted.

Theorem 4.1.3. *Suppose that (i) $\mathbf{X}_{p\times N}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K, (ii) $\theta_{\min} > 2\delta_n$. Then*

$$P(\mathcal{S}_0 = \widehat{\mathcal{S}}) \geq 1 - 4Sp^{-\tau}.$$

4.2 Modified estimator $\widehat{\Omega}_\rho$

This section focuses on dealing with the expectation of the convergence rates. In order to do that, we modify the estimator $\widehat{\mathbf{\Omega}}$ to ensure that the expectations exists. Let $\{\widehat{\mathbf{\Omega}}_{1\rho}\}$ be the solution of the following optimization problem:

$$\min\left\{|\mathbf{C}|_1 + \nu\sum_{s=1}^{S} w^{(s)}\left|\mathbf{U}^{(s)}\right|_1\right\}$$

$$\text{s.t}\ \left|\sum_{s=1}^{S} w^{(s)}\left\{\widehat{\mathbf{\Sigma}}_\rho^{(s)}\mathbf{\Omega}^{(s)} - \mathbf{I}\right\}\right|_\infty \leq \lambda_1, \left|\widehat{\mathbf{\Sigma}}_\rho^{(s)}\mathbf{\Omega}^{(s)} - \mathbf{I}\right|_\infty \leq \lambda_2, \sum_{s=1}^{S} w^{(s)}\mathbf{U}^{(s)} = 0, \qquad (4.2.1)$$

where $\widehat{\mathbf{\Sigma}}_\rho^{(s)} = \widehat{\mathbf{\Sigma}}^{(s)} + \rho\mathbf{I}$ with $\rho > 0$. Write $\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} = \left(\omega_{1\rho ij}^{(s)}\right)$. Define the symmetrized estimator $\widehat{\mathbf{\Omega}}_\rho^{(s)}$ as in (4.1.1).

Theorem 4.2.1. *Suppose that $\mathbf{X}_{p\times N}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K. Set $\nu = S^\theta$.*

(i) Let $\lambda_1 = 2\psi_1$, $\lambda_2 = 12C_M \max_s \widetilde{u}_1^{(s)}$, where $\tau > 0$. If there exists $C_U > 0$ such that $\left\|\mathbf{U}_0^{(s)}\right\|_1 \leq C_U$ for all $s = 1, \cdots, S$ and $(\sum_{s=1}^S \left\|\mathbf{U}_0^{(s)}\right\|_1) \leq C_U S^{1-\theta}$ where $\theta \in (0,1)$ and

$$0 \leq \rho \leq \min\left\{\tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p, \tilde{u}_1^{(1)}, \cdots, \tilde{u}_1^{(S)}\right\},$$

then

$$\left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_\infty = \left|\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0\right|_\infty \leq 4(C_M + 2C_U)\max\{\psi_1, S^{-\theta}C_M \max_s 2\widetilde{u}_1^{(s)}\} \max_s Sw^{(s)}$$

with probability greater than $1 - 2(2+3S)p^{-\tau}$.

(ii) If there exist $C_M^*, C_U^* > 0$ such that $\left|\boldsymbol{\Omega}^{(s)}\right|_1 \le C_M^*$, $\left|\mathbf{U}_0^{(s)}\right|_1 \le C_U^*$ and

$$\sum_{s=1}^{S} \left|\mathbf{U}_0^{(s)}\right|_1 \le S^{1-\theta} C_U^*.$$

Let

$$\begin{aligned}
\psi_1^* &= 2C_M^* \left(\tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p \right) + 2C_U^* S^{1-\theta} \max_s 2w^{(s)} \tilde{u}_1^{(s)}, \\
\lambda_1 &= \psi_1^*,\ \lambda_2 = 12 C_M^* \max_s \widetilde{u}_1^{(s)},
\end{aligned}$$

where $\tau > 0$. Then

$$\begin{aligned}
\left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_1 &\le 2\left|\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0\right|_1 \\
&\le 4p(C_M + C_U + C_M^* + 3C_U^*) \max\{\psi_1^*, C_M^* S^{-\theta} \max_s 2\tilde{u}_1^{(s)}\} \max_s S w^{(s)}
\end{aligned}$$

with probability greater than $1 - 2(2+3S)p^{-\tau}$.

(iii) If there exists $C_U^{\#} > 0$ such that

$$\max_j \sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q \le C_U^{\#} \text{ and } \max_j \sum_{s=1}^{S} \sum_{i=1}^{p} \left|u_{ij,0}^{(s)}\right|^q \le C_U^{\#} S^{1-\theta}$$

where $\theta \in (0,1)$. Then

$$\max_j \sum_{s=1}^{S} w^{(s)} \left|\left(\widehat{\mathbf{U}}_\rho^{(s)} - \mathbf{U}_0^{(s)}\right)\mathbf{e}_j\right|_1 \le 2(1+2^{1-q}+3^{1-q}) t^{1-q} C_U^{\#} \max_s S w^{(s)}/S^\theta + \frac{1}{\nu}\left|(\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0)\mathbf{e}_j\right|_1,$$

with probability greater than $1 - 4(2+3S)p^{-\tau}$, where $t = \max_{ijs}\left|\widehat{u}_{1\rho ij}^{(s)} - u_{ij,0}^{(s)}\right|$, $\widehat{\mathbf{U}}_{1\rho}^{(s)} = \left(\widehat{u}_{1\rho ij}^{(s)}\right)$, $\mathbf{U}_0^{(s)} = \left(u_{ij,0}^{(s)}\right)$.

(iv) If the conditions in (i) and (ii) are both satisfied, $\lambda_1 = \psi_1^*$, $\lambda_2 = 12C_M^* \max_s \widetilde{u}_1^{(s)}$, where $\tau > 0$. Then

$$\left\|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right\|_F^2 \le \left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_\infty \left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_1$$

with probability greater than $1 - 4(2+3S)p^{-\tau}$, and $\left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_\infty$ and $\left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_1$ with bound in (i) and (ii).

Proof of Theorem 4.2.1. (i) The proof is similar to the proof of Theorem 4.1.2.

(ii) The proof is similar to the proof of Theorem 4.1.2. Note that $\{\mathbf{C}_0, \mathbf{U}_0^{(1)}, \cdots, \mathbf{U}_0^{(S)}\}$ is a feasible solution of (4.2.1). Now, we estimate the upper bound of $\left|(\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0)\mathbf{e}_j\right|_1 = \left|\sum_{s=1}^S w^{(s)}(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j\right|_1$. Also note that

$$\begin{aligned}\left|\sum_{s=1}^S w^{(s)}(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j\right|_1 &\leq \left|\sum_{s=1}^S \mathbf{\Omega}_0^{(s)} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1 \\ &\quad + \left|\sum_{s=1}^S \mathbf{\Omega}_0^{(s)} w^{(s)}(\widehat{\mathbf{\Sigma}}^{(s)}\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} - \mathbf{I})\mathbf{e}_j\right|_1.\end{aligned}$$

Similar to the discussion of Theorem 3.2.1, we have

$$\begin{aligned}&\left|\sum_{s=1}^S \mathbf{\Omega}_0^{(s)} w^{(s)}(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}_\rho^{(s)})\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1 \\ \leq\ & 2(C_M + C_U)\max_s Sw^{(s)} \\ &\times \max\{C_M^*(2\tilde{u}_1 + \sum_{s=1}^S c_2 K^2 \xi_2^{(s)} \log p) + C_U^* S^{1-\theta}\max_s 2w^{(s)}\tilde{u}_1^{(s)}, C_M^* S^{-\theta}\max_s 2\tilde{u}_1^{(s)}\},\end{aligned}$$

and

$$\begin{aligned}&\left|\sum_{s=1}^S \mathbf{\Omega}_0^{(s)} w^{(s)}(\widehat{\mathbf{\Sigma}}_\rho^{(s)}\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} - \mathbf{I})\mathbf{e}_j\right|_1 \\ \leq\ & 2\left|C_0\right|_1 \left|\sum_{s=1}^S w^{(s)}(\widehat{\mathbf{\Sigma}}_\rho^{(s)}\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty + 2\sum_{s=1}^S \left|\mathbf{U}_0^{(s)}\right|_1 \left|w^{(s)}(\widehat{\mathbf{\Sigma}}_\rho^{(s)}\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty.\end{aligned}$$

In sum, we have

$$\left\|\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0\right\|_1 \leq 2(C_M + C_U + C_M^* + 3C_U^*)\max\{\psi_1^*, S^{-\theta}\max_s 2\tilde{u}_1^{(s)}\}\max_s Sw^{(s)},$$

and then

$$\begin{aligned}\left|\widehat{\mathbf{C}}_\rho - \mathbf{C}_0\right|_1 &\leq 2\left|\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0\right|_1 \\ &\leq 2p\left\|\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0\right\|_1 \\ &\leq 4p(C_M + C_U + C_M^* + 3C_U^*)\max\{\psi_1^*, S^{-\theta}\max_s 2\tilde{u}_1^{(s)}\}\max_s Sw^{(s)}.\end{aligned}$$

(iii) We assume the solution $\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}$ is obtained from the column solution set $\{\widehat{\mathbf{c}}_{\rho j}^{(s)}\}$ of (4.0.2) with $\widehat{\mathbf{\Sigma}}^{(s)}$ being replaced by $\widehat{\mathbf{\Sigma}}_\rho^{(s)}$. By the definition of $\{\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\}$, we have

$$\left|\widehat{\mathbf{C}}_{1\rho}\mathbf{e}_j\right|_1 + \nu\sum_{s=1}^S w^{(s)}\left|\widehat{\mathbf{U}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1 \leq \left|\mathbf{C}_0\mathbf{e}_j\right|_1 + \nu\sum_{s=1}^S w^{(s)}\left|\mathbf{U}_0^{(s)}\mathbf{e}_j\right|_1 \text{ for } j = 1, \cdots, p.$$

Denote

$$
\begin{aligned}
t &= \max_{ijs} \left| \widehat{u}_{1\rho ij}^{(s)} - u_{ij,0}^{(s)} \right|, \text{ where } \widehat{\mathbf{U}}_{1\rho}^{(s)} = \left(\widehat{u}_{1\rho ij}^{(s)} \right), \mathbf{U}_0^{(s)} = \left(u_{ij,0}^{(s)} \right), \\
\mathbf{L}^{(s)} &= \left(\widehat{u}_{1\rho ij}^{(s)} - u_{ij,0}^{(s)} \right), \\
\mathbf{L}_*^{(s)} &= \left(\widehat{u}_{1\rho ij}^{(s)} \mathbb{1}(\left| \widehat{u}_{1\rho ij}^{(s)} \right| \geq 2t) - u_{ij,0}^{(s)} \right), \\
\mathbf{L}_{**}^{(s)} &= \mathbf{L}^{(s)} - \mathbf{L}_*^{(s)}.
\end{aligned}
$$

Then

$$
\begin{aligned}
& \left| \widehat{\mathbf{C}}_{1\rho} \mathbf{e}_j \right|_1 + \nu \sum_{s=1}^S w^{(s)} \left(\left| \mathbf{U}_0^{(s)} \mathbf{e}_j \right|_1 - \left| \mathbf{L}_*^{(s)} \mathbf{e}_j \right|_1 + \left| \mathbf{L}_{**}^{(s)} \mathbf{e}_j \right|_1 \right) \\
\leq \; & \left| \widehat{\mathbf{C}}_{1\rho} \mathbf{e}_j \right|_1 + \nu \sum_{s=1}^S w^{(s)} \left(\left| \left(\mathbf{U}_0^{(s)} + \mathbf{L}_*^{(s)} \right) \mathbf{e}_j \right|_1 + \left| \mathbf{L}_{**}^{(s)} \mathbf{e}_j \right|_1 \right) \\
= \; & \left| \widehat{\mathbf{C}}_{1\rho} \mathbf{e}_j \right|_1 + \nu \sum_{s=1}^S w^{(s)} \left| \widehat{\mathbf{U}}_{1\rho}^{(s)} \mathbf{e}_j \right|_1 \leq \left| \mathbf{C}_0 \mathbf{e}_j \right|_1 + \nu \sum_{s=1}^S w^{(s)} \left| \mathbf{U}_0^{(s)} \mathbf{e}_j \right|_1,
\end{aligned}
$$

which implies

$$
\nu \sum_{s=1}^S w^{(s)} \left| \mathbf{L}_{**}^{(s)} \mathbf{e}_j \right|_1 \leq \nu \sum_{s=1}^S w^{(s)} \left| \mathbf{L}_*^{(s)} \mathbf{e}_j \right|_1 + \left| (\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0) \mathbf{e}_j \right|_1 .
$$

Thus

$$
\begin{aligned}
\nu \sum_{s=1}^S w^{(s)} \left| \mathbf{L}^{(s)} \mathbf{e}_j \right|_1 &\leq \nu \sum_{s=1}^S w^{(s)} \left| \mathbf{L}_{**}^{(s)} \mathbf{e}_j \right|_1 + \nu \sum_{s=1}^S w^{(s)} \left| \mathbf{L}_*^{(s)} \mathbf{e}_j \right|_1 \\
&\leq 2\nu \sum_{s=1}^S w^{(s)} \left| \mathbf{L}_*^{(s)} \mathbf{e}_j \right|_1 + \left| (\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0) \mathbf{e}_j \right|_1 .
\end{aligned}
$$

Similar to the proof of Theorem 6 of [11], we have

$$
\begin{aligned}
& \sum_{s=1}^S w^{(s)} \left| \mathbf{L}_*^{(s)} \mathbf{e}_j \right|_1 \\
= \; & \sum_{s=1}^S w^{(s)} \sum_{i=1}^p \left| \widehat{u}_{1\rho ij}^{(s)} \mathbb{1}(\left| \widehat{u}_{1\rho ij}^{(s)} \right| \geq 2t) - u_{ij,0}^{(s)} \right| \\
\leq \; & \sum_{s=1}^S w^{(s)} \sum_{i=1}^p \left| u_{ij,0}^{(s)} \mathbb{1}(\left| u_{ij,0}^{(s)} \right| \leq 2t) \right| \\
& + \sum_{s=1}^S w^{(s)} \sum_{i=1}^p \left| \widehat{u}_{1\rho ij}^{(s)} \mathbb{1}(\left| \widehat{u}_{1\rho ij}^{(s)} \right| \geq 2t) - u_{ij,0}^{(s)} \mathbb{1}(\left| u_{ij,0}^{(s)} \right| \geq 2t) \right|
\end{aligned}
$$

$$
\begin{aligned}
&\leq (2t)^{1-q}\sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\right|^{q} + \sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|\left(\widehat{u}_{1\rho ij}^{(s)} - u_{ij,0}^{(s)}\right)\mathbb{1}\left(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \geq 2t\right)\right| \\
&\quad + \sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\{\mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \geq 2t) - \mathbb{1}(\left|u_{ij,0}^{(s)}\right| \geq 2t)\}\right| \\
&\leq (2t)^{1-q}\sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\right|^{q} + t\sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \geq 2t) \\
&\quad + \sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\right|\left|\mathbb{1}\{\left|\left|u_{ij,0}^{(s)}\right| - 2t\right| \leq \left|\widehat{u}_{ij1\rho}^{(s)} - u_{ij,0}^{(s)}\right|\}\right| \\
&\leq (2t)^{1-q}\sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\right|^{q} + t\sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\mathbb{1}(\left|\widehat{u}_{1\rho ij}^{(s)}\right| \geq 2t) \\
&\quad + \sum_{s=1}^{S}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\right|\mathbb{1}\{\left|u_{ij,0}^{(s)}\right| \leq 3t\} \\
&\leq (2t)^{1-q}\sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\right|^{q} + t^{1-q}\sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\right|^{q} \\
&\quad + (3t)^{1-q}\sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\right|^{q} \\
&\leq (1 + 2^{1-q} + 3^{1-q})t^{1-q}\sum_{s=1}^{S} w^{(s)}\sum_{i=1}^{p}\left|u_{ij,0}^{(s)}\right|^{q} \\
&\leq (1 + 2^{1-q} + 3^{1-q})t^{1-q}C_U^{\#}\max_s Sw^{(s)}/S^{\theta}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\sum_{s=1}^{S} w^{(s)}\left|(\widehat{\mathbf{U}}_{1\rho}^{(s)} - \mathbf{U}_0^{(s)})\mathbf{e}_j\right|_1 &\leq 2\sum_{s=1}^{S} w^{(s)}\left|\mathbf{L}_*^{(s)}\mathbf{e}_j\right|_1 + \frac{1}{\nu}\left|(\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0)\mathbf{e}_j\right|_1 \\
&\leq 2(1 + 2^{1-q} + 3^{1-q})t^{1-q}C_U^{\#}\max_s Sw^{(s)}/S^{\theta} \\
&\quad + \frac{1}{\nu}\left|(\widehat{\mathbf{C}}_{1\rho} - \mathbf{C}_0)\mathbf{e}_j\right|_1.
\end{aligned}
$$

(iv) It can been proven from the definition of the matrix norm.

The proof is completed. □

Furthermore, we can consider the convergence rates for expectation, which generalizes Theorems 2 and 5 in [11] from (i) i.i.d. to long-memory observations, and (ii) from a single subject to a group.

Theorem 4.2.2. *Under the conditions of Theorem 4.2.1, and*

$$\min\left\{\tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p, \tilde{u}_1^{(1)}, \cdots, \tilde{u}_1^{(S)}, p^{-c}\right\}$$
$$\leq \quad \rho \leq \min\left\{\tilde{u}_1 + \sum_{s=1}^{S} c_2 K^2 \xi_2^{(s)} \log p, \tilde{u}_1^{(1)}, \cdots, \tilde{u}_1^{(S)}\right\},$$

with $c > 0$*, if* $p \geq n_s^{\xi}$ *with* $\xi > 0$*, we have*

(i) $\mathbb{E}\left(\left|\widehat{\mathbf{C}}_{\rho} - \mathbf{C}_0\right|_{\infty}^2\right) = O\left((C_M + C_U)\max\{\psi_1, S^{-\theta} C_M \max_s 2\tilde{u}_1^{(s)}\}\max_s S w^{(s)}\right)^2,$

(ii) $\mathbb{E}\left(\left|\widehat{\mathbf{C}}_{\rho} - \mathbf{C}_0\right|_{1}^2\right) = O\left(p(C_M^* + C_U^*)\max\{\psi_1^*, S^{-\theta} C_M^* \max_s 2\tilde{u}_1^{(s)}\}\max_s S w^{(s)}\right)^2,$

(iii) $\mathbb{E}\left\|\widehat{\mathbf{C}}_{\rho} - \mathbf{C}_0\right\|_F^2 \leq \sqrt{\mathbb{E}\left(\left|\widehat{\mathbf{C}}_{\rho} - \mathbf{C}_0\right|_{\infty}^2\right)\mathbb{E}\left(\left|\widehat{\mathbf{C}}_{\rho} - \mathbf{C}_0\right|_{1}^2\right)}.$

Proof of Theorem 4.2.2. Note that $\{\left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)^{-1}\}$ is a feasible solution of (4.2.1). Also, note that

$$\begin{aligned}
& \max_j \sum_{s=1}^{S} w^{(s)} \left|\widehat{\mathbf{\Omega}}_{\rho}^{(s)} \mathbf{e}_j\right|_1 \\
\leq \; & \max_j \sum_{s=1}^{S} w^{(s)} \left|\widehat{\mathbf{\Omega}}_{1\rho}^{(s)} \mathbf{e}_j\right|_1 \\
\leq \; & \max_j \left(\left|\widehat{\mathbf{C}}_{1\rho}^{(s)} \mathbf{e}_j\right|_1 + \sum_{s=1}^{S} w^{(s)} \left|\widehat{\mathbf{U}}_{1\rho}^{(s)} \mathbf{e}_j\right|_1\right) \\
\leq \; & \max_j \left(\left|\widehat{\mathbf{C}}_{1\rho}^{(s)} \mathbf{e}_j\right|_1 + S^{\theta} \sum_{s=1}^{S} w^{(s)} \left|\widehat{\mathbf{U}}_{1\rho}^{(s)} \mathbf{e}_j\right|_1\right) \\
\leq \; & \max_j \left(\left|\sum_{s=1}^{S} w^{(s)} \left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)^{-1} \mathbf{e}_j\right|_1 + S^{\theta} \sum_{s=1}^{S} w^{(s)} \left|\left(\left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)^{-1} - \sum_{s=1}^{S} w^{(s)} \left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)^{-1}\right) \mathbf{e}_j\right|_1\right) \\
\leq \; & \max_j \left(\sum_{s=1}^{S} w^{(s)} \left|\left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)^{-1} \mathbf{e}_j\right|_1 + S^{\theta} \sum_{s=1}^{S} \left|w^{(s)} \left(\left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)^{-1}\right) \mathbf{e}_j\right|_1 \right. \\
& \left. + S^{\theta} \sum_{s=1}^{S} \left|w^{(s)} \left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)^{-1} \mathbf{e}_j\right|_1\right) \\
\leq \; & \max_j 3 S^{\theta} \sum_{s=1}^{S} w^{(s)} \left|\left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)^{-1} \mathbf{e}_j\right|_1
\end{aligned}$$

$$
\begin{aligned}
&\leq \quad 3S^{\theta}\sum_{s=1}^{S} w^{(s)} \left\| \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)}\right)^{-1} \right\|_{1} \\
&\leq \quad 3pS^{\theta}\sum_{s=1}^{S} w^{(s)} \left\| \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)}\right)^{-1} \right\|_{2} \\
&\leq \quad 3pS^{\theta}\sum_{s=1}^{S} w^{(s)} \varphi_{\max} \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)}\right)^{-1} \\
&= \quad 3pS^{\theta}\sum_{s=1}^{S} w^{(s)} 1/\varphi_{\min} \left(\widehat{\boldsymbol{\Sigma}}_{\rho}^{(s)}\right) \\
&\leq \quad 3pS^{\theta}\sum_{s=1}^{S} w^{(s)}/\rho = 3pS^{\theta}/\rho.
\end{aligned}
$$

The proof is similar to the proof of Theorem 2 of [11]. □

For LPDD temporal dependence, we can set the bounds as

$$g_F^{(s)} \asymp n_s/(\log n_s)^{2\alpha}, g_2^{(s)} \asymp n_s/(\log n_s)^{\alpha},$$

then we have the following property

$$u_1^{(s)} \asymp (\log p)/(\log n_s)^{\alpha}, \alpha > 0.$$

For LPDD model, therefore, we only need to assume $p \geq (\log n_s)^{\xi}$ in the above theorem.

4.3 Extension to heavy tail data

In this section, we extend the previous results for the precision matrices estimators to heavy-tailed data in (C2) and (C3). According to the proof of Lemma A.2 of 36, similar to the proof of Lemma 4.1.2, we define

$$\widetilde{u}_2^{(s)} = \max\left\{[c_2K(2/\vartheta)^{2/\vartheta}\log p]^{1+2/\vartheta} g_2^{(s)}/n_s, [c_2K^2(4/\vartheta)^{4/\vartheta}\log p]^{1+2/\vartheta}(g_F^{(s)}/n_s)^{1/2}\right\}, \quad (4.3.1)$$

$$\widetilde{u}_2 = \max\left\{[c_2K(2/\vartheta)^{2/\vartheta}\log p]^{1+2/\vartheta}\xi_2, [c_2K^2(4/\vartheta)^{4/\vartheta}\log p]^{1+2/\vartheta}(\xi_F)^{1/2}\right\},$$

and

$$\widetilde{u}_3^{(s)} = k\eta_k^2 \max\left\{c_2^2 p^{(2+2\tau)/k} g_2^{(s)}/n_s, c_2 p^{(4+2\tau)/k}(g_F^{(s)}/n_s)^{1/2}\right\}, \quad (4.3.2)$$

$$\widetilde{u}_3 = k\eta_k^2 \max\left\{c_2^2 p^{(2+2\tau)/k}\xi_2, c_2 p^{(4+2\tau)/k}(\xi_F)^{1/2}\right\},$$

where $c_2 = (\tau+2)/c_1$ with $\tau > 0$ and $c_1 > 0$ is an absolute constant, $g_2^{(s)}$ and $g_F^{(s)}$ are defined in (4.1.3), $\xi_2^{(s)}$ and $\xi_F^{(s)}$ are defined in (4.1.8), ξ_2 and ξ_F are defined in (4.1.5). We assume $\widetilde{u}_2^{(s)}, \tilde{u}_2, \widetilde{u}_3^{(s)}, \tilde{u}_3 \to 0$ as $n_s \to \infty$.

Theorem 4.3.1. *(Generalized sub-exponential tails) Assume condition (C1), parameter K, $\widetilde{u}_1^{(s)}$, $\widetilde{u}_1$ and $\sum_{s=1}^S c_2K^2\xi_2^{(s)}\log p$ replaced by condition (C2), parameters $\{K,\vartheta\}$, $\widetilde{u}_2^{(s)}$, $\widetilde{u}_2$ and $\sum_{s=1}^S (c_2K(2/\vartheta)^{2/\vartheta}\log p)^{1+2/\vartheta}\xi_2^{(s)}$, respectively, then Theorems 4.1.1-4.2.2 hold.*

Theorem 4.3.2. *(Polynomial-type tails) Assume condition (C1), $\sum_{s=1}^S c_2K^2\xi_2^{(s)}\log p$, parameter K, $\widetilde{u}_1^{(s)}$ and $\widetilde{u}_1$ replaced by condition (C3), $\sum_{s=1}^S c_2^2k\eta_k^2p^{(2+2\tau)/k}\xi_2^{(s)}$,parameters $\{k,\eta_k\}$, $\widetilde{u}_3^{(s)}$ and $\widetilde{u}_3$, respectively, then Theorems 4.1.1-4.2.1 hold.*

The proofs of Theorems 4.3.1 and 4.3.2 are similar to the proofs of the preceding theorems by using the corresponding results given in Lemmas A.1, A.2 and A.3 of [35] for conditions (C2) and (C3), respectively. Details are omitted.

4.4 Computation

We can reformulate (4.0.1) into the linear programming form (3.4.2). The numerical algorithm is similar to that of Chapter 3. The only difference is about the weight $w^{(s)}$, $s = 1, \cdots, S$. In order to do that, we need to use the following vector $\mathbf{a}$ and matrix $\widetilde{\mathbf{A}}$

$$\begin{aligned}
\widetilde{\mathbf{a}} &= \Big(\underbrace{1,\cdots,1}_{p},\underbrace{\nu w^{(1)},\cdots,\nu w^{(1)}}_{p},\cdots,\underbrace{\nu w^{(S)},\cdots,\nu w^{(S)}}_{p}\Big),\\
\mathbf{a} &= -\left(\widetilde{\mathbf{a}},\widetilde{\mathbf{a}}\right)^T,\\
\widetilde{\mathbf{A}} &= \begin{pmatrix}
\sum_{s=1}^S w^{(s)}\widehat{\boldsymbol{\Sigma}}^{(s)} & w^{(1)}\widehat{\boldsymbol{\Sigma}}^{(1)} & w^{(2)}\widehat{\boldsymbol{\Sigma}}^{(2)} & \cdots & w^{(S)}\widehat{\boldsymbol{\Sigma}}^{(S)}\\
\widehat{\boldsymbol{\Sigma}}^{(1)} & \widehat{\boldsymbol{\Sigma}}^{(1)} & \mathbf{O}_{p\times p} & \cdots & \mathbf{O}_{p\times p}\\
\widehat{\boldsymbol{\Sigma}}^{(2)} & \mathbf{O}_{p\times p} & \widehat{\boldsymbol{\Sigma}}^{(2)} & \cdots & \mathbf{O}_{p\times p}\\
\vdots & \vdots & \vdots & \vdots & \vdots\\
\widehat{\boldsymbol{\Sigma}}^{(S)} & \mathbf{O}_{p\times p} & \mathbf{O}_{p\times p} & \cdots & \widehat{\boldsymbol{\Sigma}}^{(S)}
\end{pmatrix},
\end{aligned}$$

and we need to use (4.1.1) to symmeterize matrices. The case in Chapter 3 is a special case of the uniform weights $w^{(s)} \equiv \frac{1}{S}$.

4.4.1 Weight selection

We can take weights $w^{(s)}$ $(s = 1, \cdots, S)$ as tuning parameters and select them in the following way. We set the tuning parameters $\{w^{(s)}\}$ to make the upper bound of $\left|\widehat{\mathbf{C}} - \mathbf{C}_0\right|_\infty$ achieve its minima. This will enforce `fastclime` to work with smaller λ_1 and λ_2, and produce more nonzero estimated elements, which in turn would achieve higher TPR, and faster computation.

Divide the interval $[0.1, 0.7]$ into 60 equal-sized subintervals, so the step equal to 0.01, and we can get an estimated optimal weight with error less than 0.01. Select every $w^{(s)}$ from $\{0.1 + (m-1) \times 0.01 : m = 1, \cdots, 61\}$ with $\sum_{s=1}^S w^{(s)} = 1$. Denote $C_{MU} = \max\{C_U, C_M\}$. Similar to the proof of Theorem 4.1.2, we can estimate the upper bound of $\left|\widehat{\mathbf{C}} - \mathbf{C}_0\right|_\infty$ as follows:

$$\begin{aligned}\left|\sum_{s=1}^S w^{(s)}(\widehat{\boldsymbol{\Omega}}_1^{(s)} - \boldsymbol{\Omega}_0^{(s)})\mathbf{e}_j\right|_\infty \leq & \left|\sum_{s=1}^S \boldsymbol{\Omega}_0^{(s)} w^{(s)}(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\widehat{\boldsymbol{\Omega}}_1^{(s)}\mathbf{e}_j\right|_\infty \\ & + \left|\sum_{s=1}^S \boldsymbol{\Omega}_0^{(s)} w^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty,\end{aligned}$$

where

$$\begin{aligned}& \left|\sum_{s=1}^S \boldsymbol{\Omega}_0^{(s)} w^{(s)}(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\widehat{\boldsymbol{\Omega}}_1^{(s)}\mathbf{e}_j\right|_\infty \\ \leq\ & \max\left\{C_M \left|\sum_{s=1}^S w^{(s)}(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty + \max_s w^{(s)} \left|(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty C_U S^{1-\theta},\right. \\ & \left. S^{-\theta} C_M \max_s \left|(\boldsymbol{\Sigma}_0^{(s)} - \widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty\right\} \times \left(|\mathbf{c}_{j,0}|_1 + S^\theta \sum_{s=1}^S w^{(s)} \left|\mathbf{u}_{j,0}^{(s)}\right|_1\right) \\ =\ & C_{MU} \max\{\psi_1, S^{-\theta} C_M \max_s 2\widetilde{u}_1^{(s)}\}(1 + S^\theta),\end{aligned}$$

and

$$\left|\sum_{s=1}^S \boldsymbol{\Omega}_0^{(s)} w^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty$$

$$\begin{aligned}
&\le \|\mathbf{C}_0\|_1 \left|\sum_{s=1}^{S} w^{(s)}(\widehat{\mathbf{\Sigma}}^{(s)}\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty + \sum_{s=1}^{S}\left\|\mathbf{U}_0^{(s)}\right\|_1 \left|w^{(s)}(\widehat{\mathbf{\Sigma}}^{(s)}\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty \\
&\le C_M \left|\sum_{s=1}^{S} w^{(s)}(\widehat{\mathbf{\Sigma}}^{(s)}\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty + C_U\lambda_2 \\
&\le C_M\psi_1 + C_U 6 C_M \max_s 2\widetilde{u}_1^{(s)}.
\end{aligned}$$

Thus

$$\begin{aligned}
\left|\sum_{s=1}^{S} w^{(s)}(\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j\right|_\infty &\le C_{MU}\max\{\psi_1, S^{-\theta}C_M \max_s 2\widetilde{u}_1^{(s)}\}(1+S^\theta) \\
&\quad + C_M\psi_1 + C_U 6 C_M \max_s 2\widetilde{u}_1^{(s)}.
\end{aligned}$$

Note that $\widetilde{u}_1^{(s)}$ is independent of $w^{(s)}$, for $s = 1, \cdots, S$. Hence, we only need to minimize ψ_1. Note that every term of ψ_1 has K^2, so we can assume $K^2 = 1$. We set $c_2 = 16$. By unifying C_U and C_M to be C_{MU}, we only need to minimize

$$\psi_1^{\#} = \tilde{u}_1 + \sum_{s=1}^{S} c_2 \xi_2^{(s)} \log p + S^{1-\theta}\max_s 2 w^{(s)}\tilde{u}_1^{(s)}.$$

We may assume that different groups are independent. Therefore, by the definitions of $g_2^{(s)}, g_F^{(s)}, \xi_2^{(s)}, \xi_F^{(s)}, \xi_2, \xi_F$ in (4.1.3), (4.1.8), and (4.1.5), we have

$$\begin{aligned}
g_2^{(s)} &= \max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_2, \\
g_F^{(s)} &= \frac{1}{n_s}\max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_F^2, \\
\xi_2^{(s)} &= \frac{w^{(s)}}{n_s}\max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_2 = \frac{w^{(s)}}{n_s} g_2^{(s)}, \\
\xi_F^{(s)} &= \frac{\left(w^{(s)}\right)^2}{n_s^2}\max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_F^2 = \frac{\left(w^{(s)}\right)^2}{n_s} g_F^{(s)}, \\
\xi_2 &= \max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{*})\right\|_2 = \max_{1\le k\le p}\max_{1\le s\le S}\frac{w^{(s)}}{n_s}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_2 = \max_{1\le s\le S}\frac{w^{(s)}}{n_s}\max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_2, \\
\xi_F &= \max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{*})\right\|_F = \max_{1\le k\le p}\sum_{s=1}^{S}\frac{\left(w^{(s)}\right)^2}{n_s^2}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_F^2,
\end{aligned}$$

and then

$$\widetilde{u}_1^{(s)} = \max\left\{c_2(\log p)\, g_2^{(s)}/n_s, \left[c_2(\log p)\, g_F^{(s)}/n_s\right]^{1/2}\right\}$$

$$
\begin{aligned}
&= \max\left\{c_2\,(\log p)\max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_2/n_s,\ \left[c_2\,(\log p)\max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_F^2/n_s^2\right]^{1/2}\right\},\\
\widetilde{u}_1 &= \max\left\{c_2\,(\log p)\,\xi_2, \left[c_2\,(\log p)\,\xi_F\right]^{1/2}\right\}\\
&= \max\left\{c_2\,(\log p)\max_{1\le s\le S}\frac{w^{(s)}}{n_s}\max_{1\le k\le p}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_2,\right.\\
&\quad\left.\left[c_2\,(\log p)\max_{1\le k\le p}\sum_{s=1}^{S}\frac{\left(w^{(s)}\right)^2}{n_s^2}\left\|Cov(\mathbf{x}_{[k]}^{(s)})\right\|_F^2\right]^{1/2}\right\}.
\end{aligned}
$$

Note that $\tilde{u}_1$ and $\widetilde{u}_1^{(s)}$ only depend on $g_2^{(s)}, g_F^{(s)}, \xi_2^{(s)}, \xi_F^{(s)}, \xi_2, \xi_F$, which only dependent on covariance matrices.

We will conduct the computation in three steps. *Step 1.* Compute $g_2^{(s)}, g_F^{(s)}, \xi_2^{(s)}, \xi_F^{(s)}, \xi_2, \xi_F$; *Step 2.* Select the weight $\{w^{(s)}\}$ by minimizing the upper bound of $\left|\widehat{\mathbf{C}} - \mathbf{C}_0\right|_\infty$ as above; *Step 3.* Compute the precision matrix by the vector optimization problem. We evaluate the numerical performance of the weighted-JEMP (Weighted) estimators for high dimensional precision matrices.

4.5 Numerical experiments

4.5.1 Results of simulation study

Since this chapter is about weighted-JEMP, we study the weight estimation first. We sample $\mathbf{X}_{p\times n}^{(s)}$ $(s = 1, \cdots, S)$ 20 times with distribution $\mathbf{\Sigma}^{(s)}$ and $\alpha \in \{0.1, 0.25, 0.5, 1, 2\}$. Then compute weight $\{w^{(s)}\}$ by the above computation steps, and repeat 50 times. Because each group is generated in the same way, the optimal weight is the uniform weight, $1/4 = 0.25$.

Tables 4.1-4.2 report the estimated weights for Groups 1-4. Overall, the weights are substantially different with the uniform weights in some settings, and the higher the dimension is, the closer the estimated weights are to the uniform weights. In Table 4.1 ($p = 50$), for Model 1 and $\rho = 0.1$ or 4 and for Model 2 and $\rho = 4$, the difference between the estimated and the uniform weights is large. For Model 1 with $\alpha = 1$ or 2 and $\rho = 1$ and for Model 2 with $\alpha = 2$ and $\rho = 1$, the estimated weights are somewhat close to the uniform weights, but difference still exists. In the other settings, the difference is marginal, and the estimated weights are close to 0.25. In Table 4.2 ($p = 100$), for Model 1 and $\rho = 0.1$, and for Model 2 and $\alpha = 1$ or 2 and $\rho = 1$ or 4, the difference between the estimated and the uniform weights is large. For Model 1 with $\alpha = 1$ or 2 and $\rho = 0.25$, the estimated weights are somewhat close to the uniform weights, but difference still exists. In the other settings, the difference is marginal, and the estimated weights are close to 0.25.

In what follows, we compare the performance between ordinary JEMP and weighted-JEMP. The following tables report the results on the comparison of average (SD) matrix losses, and computation time of SCLIME, ordinary JEMP and weighted-JEMP for $p = 50$ and $p = 100$. In the tables, "Weighted" stands for weighted-JEMP.

Tables 4.3-4.7 show the results of the simulation study. From Tables 4.3 and 4.4, we can see that the average (SD) matrix losses for $p = 50$ and $p = 100$ of ordinary JEMP and weighted-JEMP are almost the same. Tables 4.5 and 4.6 imply that the computation time of weighted-JEMP is lower than the other two methods. Moreover, the tuning parameters (λ_1, λ_2) chosen in weighted-JEMP is smaller by 0.01 than those of ordinary JEMP. Table 4.7 shows that, SCLIME succeeds 200 times in all settings, i.e., the computational success ratio of SCLIME is 100%. In comparison, ordinary JEMP only succeeds 172 times in all settings (the number of failure is 28), i.e., the computational success ratio of ordinary JEMP is 86%, and weighted-JEMP succeeds 182 times in all settings (the number of failure is 18), i.e., the computational success ratio of weighted-JEMP is 90.5%. This also illustrates that weighted-JEMP is more stable than ordinary JEMP in terms of computation.

4.5.2 rfMRI data analysis

We analyze the same rfMRI data set in Chapter 3 for the estimation of precision matrices and brain functional connectivity by our weighted-JEMP. Based on the result in Chapter 3, we estimate α to be 0.3. We estimate Σ by $\widehat{\Sigma}$, the sample covariance matrix. Following the method introduced in Section 4.5.1, we sample 20 times, and obtain the estimated

Table 4.1: Average (SD) weight for $p = 50$

α	ρ	weight 1	weight 2	weight 3	weight 4
		Model 1			
0.1	0.1	0.2948(0.0306)	0.1908(0.0251)	0.2416(0.0311)	0.2728(0.0307)
	0.25	0.2584(0.0257)	0.2476(0.0252)	0.2404(0.0241)	0.2536(0.0222)
	1	0.2456(0.0219)	0.2662(0.0228)	0.2268(0.0298)	0.2614(0.0263)
	4	0.539(0.2586)	0.1(0)	0.261(0.2586)	0.1(0)
0.25	0.1	0.2946(0.0266)	0.1894(0.0271)	0.2378(0.0284)	0.2782(0.028)
	0.25	0.2538(0.0249)	0.237(0.0238)	0.251(0.0199)	0.2582(0.0262)
	1	0.241(0.0247)	0.2632(0.026)	0.2264(0.0338)	0.2694(0.0236)
	4	0.5368(0.2619)	0.1(0)	0.2632(0.2619)	0.1(0)
0.5	0.1	0.3024(0.025)	0.1854(0.0239)	0.2308(0.0251)	0.2814(0.0243)
	0.25	0.2556(0.0238)	0.2376(0.0224)	0.2502(0.0212)	0.2566(0.0251)
	1	0.2396(0.0241)	0.264(0.0252)	0.2218(0.0278)	0.2746(0.0207)
	4	0.5814(0.2397)	0.1(0)	0.2186(0.2397)	0.1(0)
1	0.1	0.3214(0.0221)	0.184(0.0156)	0.2246(0.0168)	0.27(0.0169)
	0.25	0.2574(0.0165)	0.243(0.0216)	0.2428(0.0158)	0.2568(0.0168)
	1	0.2484(0.0146)	0.2676(0.0151)	0.205(0.0164)	0.279(0.0146)
	4	0.652(0.1644)	0.1(0)	0.148(0.1644)	0.1(0)
2	0.1	0.3214(0.012)	0.1858(0.0099)	0.222(0.009)	0.2708(0.0099)
	0.25	0.2572(0.0097)	0.2428(0.0136)	0.2406(0.0115)	0.2594(0.011)
	1	0.2492(0.0092)	0.2692(0.0107)	0.1942(0.0114)	0.2874(0.0099)
	4	0.7(0)	0.1(0)	0.1(0)	0.1(0)
		Model 2			
0.1	0.1	0.2518(0.0225)	0.2462(0.0199)	0.2546(0.0207)	0.2474(0.026)
	0.25	0.2532(0.0203)	0.2482(0.0233)	0.2438(0.0205)	0.2548(0.0183)
	1	0.2544(0.0248)	0.2436(0.0251)	0.237(0.0309)	0.265(0.0293)
	4	0.1(0)	0.1(0)	0.1(0)	0.7(0)
0.25	0.1	0.2468(0.0226)	0.248(0.0202)	0.2516(0.018)	0.2536(0.0195)
	0.25	0.2494(0.0203)	0.248(0.0192)	0.2538(0.0188)	0.2488(0.0199)
	1	0.2464(0.0198)	0.2558(0.0255)	0.2386(0.0247)	0.2592(0.0248)
	4	0.1(0)	0.1(0)	0.1(0)	0.7(0)
0.5	0.1	0.2452(0.0201)	0.2514(0.0171)	0.2532(0.0181)	0.2502(0.0174)
	0.25	0.252(0.0174)	0.2494(0.0178)	0.252(0.0178)	0.2466(0.0178)
	1	0.251(0.0202)	0.2504(0.0236)	0.2336(0.0268)	0.265(0.023)
	4	0.1(0)	0.1(0)	0.1(0)	0.7(0)
1	0.1	0.2566(0.0122)	0.2478(0.0134)	0.2466(0.0144)	0.249(0.0128)
	0.25	0.2564(0.0177)	0.2428(0.0151)	0.2512(0.0144)	0.2496(0.0118)
	1	0.2454(0.0168)	0.249(0.0168)	0.2292(0.0188)	0.2764(0.0171)
	4	0.1(0)	0.1(0)	0.1(0)	0.7(0)
2	0.1	0.2496(0.0088)	0.251(0.0071)	0.2466(0.008)	0.2528(0.0078)
	0.25	0.2618(0.0085)	0.2346(0.0073)	0.2504(0.006)	0.2532(0.0068)
	1	0.2408(0.0078)	0.2474(0.008)	0.2234(0.0102)	0.2884(0.0089)
	4	0.1(0)	0.1(0)	0.1(0)	0.7(0)

Table 4.2: Average (SD) weight for $p = 100$

α	ρ	weight 1	weight 2	weight 3	weight 4
			Model 1		
0.1	0.1	0.138(0.0225)	0.1348(0.0287)	0.3316(0.0319)	0.3956(0.0383)
	0.25	0.2712(0.0254)	0.2632(0.0285)	0.224(0.028)	0.2416(0.028)
	1	0.263(0.024)	0.2568(0.0249)	0.2406(0.0261)	0.2396(0.0224)
	4	0.2486(0.0252)	0.2594(0.0249)	0.2442(0.0247)	0.2478(0.0215)
0.25	0.1	0.139(0.023)	0.131(0.0267)	0.3366(0.0335)	0.3934(0.0391)
	0.25	0.2766(0.0234)	0.2602(0.0225)	0.2192(0.0311)	0.244(0.0282)
	1	0.2704(0.0254)	0.2578(0.027)	0.2376(0.0261)	0.2342(0.0261)
	4	0.2466(0.0232)	0.2488(0.0259)	0.2454(0.0224)	0.2592(0.0255)
0.5	0.1	0.1348(0.0197)	0.1252(0.0232)	0.3356(0.0298)	0.4044(0.034)
	0.25	0.2796(0.0234)	0.262(0.0214)	0.215(0.0292)	0.2434(0.0279)
	1	0.2696(0.0208)	0.2574(0.02)	0.2376(0.026)	0.2354(0.0236)
	4	0.2444(0.0192)	0.2492(0.0204)	0.2454(0.0238)	0.261(0.0234)
1	0.1	0.1278(0.0125)	0.1094(0.0108)	0.3452(0.0164)	0.4176(0.0264)
	0.25	0.2874(0.0182)	0.2664(0.0174)	0.2018(0.018)	0.2444(0.0184)
	1	0.265(0.0149)	0.2658(0.0133)	0.2358(0.0173)	0.2334(0.0167)
	4	0.2574(0.0186)	0.2486(0.0159)	0.2394(0.0156)	0.2546(0.0164)
2	0.1	0.1264(0.0063)	0.1062(0.0067)	0.3508(0.0126)	0.4166(0.0165)
	0.25	0.3016(0.0127)	0.259(0.0109)	0.1982(0.0112)	0.2412(0.0096)
	1	0.2698(0.0098)	0.266(0.0081)	0.2426(0.0103)	0.2216(0.0102)
	4	0.2548(0.0107)	0.251(0.0079)	0.2382(0.0106)	0.256(0.0086)
			Model 2		
0.1	0.1	0.2554(0.0193)	0.2452(0.0203)	0.2516(0.0217)	0.2478(0.0217)
	0.25	0.2526(0.022)	0.2564(0.0228)	0.244(0.0237)	0.247(0.0253)
	1	0.2356(0.027)	0.2584(0.0267)	0.2214(0.0333)	0.2846(0.0243)
	4	0.2528(0.0233)	0.2632(0.0233)	0.252(0.026)	0.232(0.0257)
0.25	0.1	0.2496(0.0213)	0.2478(0.0223)	0.2568(0.0174)	0.2458(0.0208)
	0.25	0.2554(0.017)	0.2502(0.0208)	0.2402(0.0174)	0.2542(0.0202)
	1	0.2398(0.0282)	0.2536(0.0301)	0.221(0.0231)	0.2856(0.0281)
	4	0.256(0.0221)	0.2714(0.0217)	0.2416(0.0247)	0.231(0.0236)
0.5	0.1	0.2488(0.0191)	0.249(0.0162)	0.2554(0.0182)	0.2468(0.0196)
	0.25	0.253(0.0174)	0.2538(0.0185)	0.2424(0.0191)	0.2508(0.0171)
	1	0.2364(0.023)	0.253(0.0247)	0.2204(0.0229)	0.2902(0.0254)
	4	0.2574(0.0203)	0.269(0.0191)	0.2444(0.0237)	0.2292(0.0212)
1	0.1	0.2482(0.0141)	0.2492(0.0138)	0.255(0.0169)	0.2476(0.0132)
	0.25	0.2554(0.0122)	0.255(0.0134)	0.2412(0.0142)	0.2484(0.0125)
	1	0.2334(0.0187)	0.2566(0.0195)	0.194(0.0181)	0.316(0.0181)
	4	0.2622(0.0149)	0.281(0.0179)	0.2472(0.0153)	0.2096(0.0131)
2	0.1	0.2462(0.0083)	0.2544(0.0079)	0.259(0.0079)	0.2404(0.0114)
	0.25	0.2572(0.0086)	0.2566(0.008)	0.2398(0.0091)	0.2464(0.0069)
	1	0.2294(0.011)	0.2638(0.0118)	0.1852(0.0133)	0.3216(0.0106)
	4	0.264(0.0101)	0.2876(0.0089)	0.2492(0.0105)	0.1992(0.0097)

Table 4.3: Comparison of average (SD) matrix losses for $p = 50$

		JEMP	WEIGHTED	JEMP	WEIGHTED
α	ρ	Spectral norm		Frobenius norm	
		Model 1			
0.1	0.1	4.0894(0.4229)	4.2301(0.4145)	12.5555(0.4134)	13.2803(0.4481)
	0.25	4.2788(0.3776)	4.5247(0.4248)	13.6734(0.4866)	14.5708(0.4848)
	1	4.0227(0.3617)	4.2107(0.3508)	12.4848(0.4358)	13.303(0.3987)
	4	3.3796(0.3129)	3.6694(0.2727)	10.6545(0.4399)	11.9447(0.4253)
0.25	0.1	1.1445(0.05)	1.1458(0.0576)	4.0661(0.0608)	4.0971(0.0674)
	0.25	1.0972(0.0638)	1.1315(0.0641)	3.6482(0.0989)	3.7337(0.1009)
	1	1.086(0.0654)	1.1171(0.0706)	3.6719(0.0717)	3.7353(0.0736)
	4	1.1329(0.0433)	1.127(0.0473)	3.8565(0.0469)	3.8827(0.0588)
0.5	0.1	1.3681(0.0199)	1.3552(0.0197)	4.1381(0.0342)	4.0953(0.0391)
	0.25	1.1035(0.0221)	1.0941(0.0207)	3.2564(0.0359)	3.2363(0.0356)
	1	1.213(0.0175)	1.2049(0.0198)	3.7055(0.0352)	3.6738(0.0345)
	4	1.301(0.0133)	1.272(0.0165)	4.3187(0.0371)	4.1873(0.041)
1	0.1	1.4709(0.0161)	1.4643(0.0164)	4.539(0.0237)	4.5177(0.0281)
	0.25	1.2096(0.0156)	1.2049(0.0154)	3.6048(0.0408)	3.6072(0.0336)
	1	1.3199(0.0121)	1.3173(0.0113)	4.1461(0.0335)	4.133(0.0331)
	4	1.4015(0.0104)	1.3791(0.0101)	4.8118(0.0308)	4.6904(0.0383)
2	0.1	1.4918(0.0113)	1.4873(0.0116)	4.6469(0.0237)	4.6341(0.0239)
	0.25	1.2269(0.0154)	1.2277(0.0142)	3.7163(0.0301)	3.7289(0.0262)
	1	1.3395(0.0095)	1.3396(0.0082)	4.2496(0.0256)	4.2495(0.0247)
	4	1.4227(0.0092)	1.4056(0.011)	4.9376(0.0272)	4.8432(0.0427)
		Model 2			
0.1	0.1	4.5717(0.3843)	4.7924(0.3995)	15.0458(0.4386)	15.8879(0.4337)
	0.25	4.4613(0.3668)	4.6725(0.3717)	14.7588(0.4925)	15.573(0.4585)
	1	4.4467(0.401)	4.6726(0.4138)	14.3266(0.4507)	15.151(0.4662)
	4	2.9579(0.4899)	3.316(0.5714)	7.9602(1.0064)	9.2158(1.4351)
0.25	0.1	1.1459(0.0972)	1.1906(0.0962)	3.2817(0.1565)	3.4479(0.1659)
	0.25	1.1463(0.1021)	1.1868(0.1012)	3.2976(0.1486)	3.4352(0.1527)
	1	1.147(0.1031)	1.1866(0.1063)	3.3361(0.1197)	3.4714(0.1293)
	4	1.4989(0.0339)	1.4321(0.0592)	4.0451(0.071)	3.976(0.0806)
0.5	0.1	0.6056(0.0307)	0.602(0.028)	1.7048(0.0536)	1.72(0.0491)
	0.25	0.6923(0.0453)	0.6891(0.0401)	1.8129(0.0631)	1.8317(0.0516)
	1	0.9067(0.029)	0.8956(0.0313)	2.4029(0.0542)	2.3947(0.0508)
	4	1.6221(0.0202)	1.5754(0.0499)	4.6693(0.1041)	4.4717(0.1952)
1	0.1	0.6489(0.0274)	0.659(0.0227)	1.7925(0.0627)	1.8635(0.0556)
	0.25	0.7688(0.027)	0.7755(0.0265)	1.9618(0.0509)	2.0277(0.0711)
	1	1.0036(0.0169)	0.9957(0.0164)	2.6913(0.047)	2.7041(0.0498)
	4	1.6728(0.0229)	1.6379(0.0366)	5.0789(0.0967)	4.9208(0.1708)
2	0.1	0.6722(0.0205)	0.6767(0.0196)	1.9788(0.044)	1.9908(0.0467)
	0.25	0.7901(0.0224)	0.7965(0.0194)	2.1447(0.0425)	2.1772(0.0457)
	1	1.021(0.015)	1.0164(0.0146)	2.8188(0.0307)	2.8493(0.033)
	4	1.6826(0.0208)	1.6565(0.0245)	5.1763(0.1021)	5.0767(0.1089)

Table 4.4: Comparison of average (SD) matrix losses for $p = 100$

		JEMP	WEIGHTED	JEMP	WEIGHTED
α	ρ	Spectral norm		Frobenius norm	
		Model 1			
0.1	0.1	3.4815(0.3047)	3.666(0.305)	13.5156(0.3612)	14.3779(0.3752)
	0.25	3.7798(0.3)	3.9296(0.2875)	14.7279(0.3482)	15.5837(0.3931)
	1	3.4128(0.2821)	3.5977(0.3341)	13.1391(0.4187)	13.9195(0.4154)
	4	3.2535(0.2874)	3.4258(0.2915)	12.0524(0.3935)	12.7715(0.411)
0.25	0.1	1.4208(0.0309)	1.3932(0.0496)	5.7686(0.0475)	5.7335(0.073)
	0.25	1.1212(0.0527)	1.1238(0.0591)	4.7957(0.0734)	4.8631(0.076)
	1	1.1599(0.023)	1.1424(0.0258)	5.0797(0.0466)	5.0879(0.0479)
	4	1.0925(0.0303)	1.0795(0.034)	5.1277(0.0387)	5.1122(0.038)
0.5	0.1	1.6063(0.0252)	1.6024(0.0364)	6.3708(0.0736)	6.3258(0.1198)
	0.25	1.2841(0.0182)	1.2646(0.0191)	5.0088(0.0476)	4.9606(0.0479)
	1	1.3546(0.0143)	1.3387(0.0137)	5.7092(0.0564)	5.6348(0.0477)
	4	1.2825(0.0109)	1.2725(0.0112)	5.9603(0.0494)	5.894(0.0433)
1	0.1	1.6677(0.0189)	1.6917(0.0482)	6.865(0.0576)	6.9591(0.206)
	0.25	1.3653(0.0117)	1.3511(0.0126)	5.4635(0.0399)	5.4406(0.0425)
	1	1.4229(0.009)	1.4143(0.0075)	6.2067(0.0418)	6.1739(0.0399)
	4	1.3364(0.0104)	1.3412(0.0084)	6.388(0.0533)	6.3996(0.0442)
2	0.1	1.677(0.0194)	1.6986(0.0396)	6.9839(0.0734)	7.0682(0.1697)
	0.25	1.3791(0.0098)	1.3629(0.0091)	5.6197(0.0359)	5.5715(0.0278)
	1	1.4337(0.009)	1.4279(0.0079)	6.3182(0.0376)	6.2861(0.0324)
	4	1.3417(0.0063)	1.3465(0.0069)	6.4555(0.0371)	6.4675(0.0423)
		Model 2			
0.1	0.1	4.3392(0.3593)	4.3399(0.342)	17.8846(0.4256)	17.9009(0.4218)
	0.25	4.2458(0.3316)	4.2814(0.2959)	17.3155(0.5017)	17.3752(0.4738)
	1	3.8897(0.3091)	3.8919(0.2886)	14.9214(0.449)	14.9163(0.4516)
	4	3.6751(0.2552)	3.6556(0.2674)	14.0743(0.3692)	14.0111(0.383)
0.25	0.1	1.1063(0.1111)	1.1059(0.1109)	4.1584(0.1287)	4.1445(0.1439)
	0.25	1.104(0.1017)	1.0979(0.1031)	4.2513(0.1137)	4.2391(0.1208)
	1	1.1249(0.0391)	1.123(0.0384)	4.8262(0.0581)	4.8333(0.0623)
	4	1.089(0.0475)	1.0864(0.0461)	4.9262(0.0606)	4.9263(0.056)
0.5	0.1	0.8057(0.0265)	0.8135(0.0242)	3.1416(0.0549)	3.1794(0.0567)
	0.25	0.8752(0.0226)	0.8793(0.0211)	3.5122(0.0649)	3.5519(0.0632)
	1	1.302(0.0159)	1.3023(0.0162)	5.1866(0.0474)	5.191(0.0491)
	4	1.2598(0.0138)	1.2618(0.0139)	5.4808(0.0464)	5.4903(0.0447)
1	0.1	0.8546(0.0219)	0.8729(0.0223)	3.4163(0.0532)	3.5677(0.0817)
	0.25	0.9297(0.0179)	0.9434(0.0184)	3.8524(0.0525)	3.9591(0.0691)
	1	1.3807(0.0119)	1.3814(0.0116)	5.7364(0.0395)	5.7592(0.0406)
	4	1.3161(0.0095)	1.3233(0.0096)	5.9419(0.0558)	5.9864(0.051)
2	0.1	0.8725(0.016)	0.8842(0.0155)	3.6463(0.043)	3.743(0.0663)
	0.25	0.9432(0.0106)	0.9532(0.0113)	4.0495(0.0441)	4.1428(0.0521)
	1	1.397(0.0087)	1.3965(0.009)	5.8797(0.0295)	5.9(0.0311)
	4	1.3216(0.01)	1.3272(0.0085)	6.0349(0.0384)	6.0771(0.037)

Table 4.5: Computational time of SCLIME, JEMP and Weighted for $p = 50$

	Model 1	Model 2
SCLIME	0.6723(0.2735)	0.5566(0.0221)
JEMP	0.6222(0.2434)	0.5469(0.2127)
Weighted	0.5026(0.1276)	0.5126(0.2089)

Table 4.6: Computational time of SCLIME, JEMP and Weighted for $p = 100$

	Model 1	Model 2
SCLIME	4.9713(1.8507)	3.9432(0.7965)
JEMP	3.2713(1.1372)	3.0396(1.0667)
Weighted	2.2929(0.2301)	2.0825(0.2916)

Table 4.7: Number of computation successes in 100 repetitions

Model	p	SCLIME	JEMP	Weighted
1	50	50	40	45
	100	50	45	45
2	50	50	46	46
	100	50	41	46

$(w^{(1)}, w^{(2)}, w^{(3)})$, which is $(0.43, 0.33, 0.24)$. For these 3 groups by weighted-JEMP, the minimum connectivity is 0, the maximum connectivity is 10, and the median is 0. Only 23 time series of Group 1 have nonzero connectivity, and they have 60 connections. It is apparently higher than that of ordinary JEMP.

The top 10 hubs by weighted-JEMP are reported in Table 4.8. We can see that the six time series (4, 10, 14, 17, 18, and 21) are found in their top 10 hubs in all three groups.

Table 4.8: Top 10 hubs for direct connectivity of Groups $1-3$ found by weighted-JEMP

Rank	Group 1	Group 2	Group 3
1	10	19	10
2	18	18	18
3	17	10	21
4	21	20	4
5	28	17	14
6	4	21	17
7	5	14	38
8	14	22	1
9	38	4	15
10	20	24	16

Table 4.9: Top 10 hubs for direct connectivity of Group 1 found by weighted-JEMP

Weighted Rank	Order in 50	Weighted Degree	SCLIME rank	SCLIME degree
1	10	10	3	46
2	18	7	10	44
3	17	5	19	43
4	21	4	2	47
5	28	4	16	44
6	4	3	8	44
7	5	3	4	45
8	14	3	18	43
9	38	3	38	42
10	20	2	1	47

This implies that weighted-JEMP works well, and finds some common information across different subjects.

Table 4.9 lists top 10 hubs of Group 1 found by weighted-JEMP and their degrees of connectivity. Also, they are compared with the results from SCLIME. It can be seen that top 10 hubs of Group 1 by JEMP has 6 overlaps with those of SCLIME: time series 4, 5, 10, 18, 20, and 21. Moreover, time series 28 also has 44 degree, which is the same as that of time series 4 or 18. Hence, that time series could be considered as in top 10 as well. It implies that the results of weighted-JEMP are similar to those of CLIME for this subject.

Chapter 5

Joint estimation of multiple precision matrices with decay ℓ_1 norm for high-dimensional time series with long-memory

In Chapters 3 and 4, we consider the estimation of precision matrix of high-dimensional time series with long-memory, and use the entrywise $|\cdot|_\infty$, $|\cdot|_1$ and $\|\cdot\|_F$ norm of the risk of the estimation of the common part. However, those methods do not work for the vector norm induced matrix ℓ_1 norm $\|\cdot\|_1$. In this chapter, based on the decay ℓ_1 norm assumption, i.e., $\sum_{s=1}^{S} \left\|\mathbf{\Omega}_0^{(s)}\right\|_1 \le C_M S^{1-\theta}$ where $\theta \in (0,1)$, we propose the optimal problem (5.0.1) with the matrix norm $\|\cdot\|_1$, we propose the modified weighted JEMP estimator, called "weighted-Joint".

For simplicity, we assume that $n = n_1 = \cdots = n_S$. We consider the following class of matrices

$$\mathcal{U} := \left\{\mathbf{\Omega} : \mathbf{\Omega} \succ 0, \|\mathbf{\Omega}\|_1 \le C_M\right\},$$

and assume that $\mathbf{\Omega}_0^{(s)} \in \mathcal{U}$ for all $s = 1, \cdots, S$. Based on this assumption, the true precision matrices should be sparse and have many small entries. Write $\mathbb{E}(\mathbf{x}^{(s)}) = (\mu_1^{(s)}, \cdots, \mu_p^{(s)})^T$.

In our joint estimation method, we assume the precision matrices have the decay ℓ_1 norm, i.e., the decreasing $\left\|\mathbf{\Omega}_0^{(s)}\right\|_1$. To estimate $\{\mathbf{\Omega}_0^{(1)}, \cdots, \mathbf{\Omega}_0^{(S)}\}$, we propose the following constrained ℓ_1 minimization criterion:

$$\min \left\{ \max_j \sum_{s=1}^{S} w^{(s)} \left|\mathbf{\Omega}^{(s)} \mathbf{e}_j\right|_1 \right\}$$

$$\text{s.t} \left|\widehat{\mathbf{\Sigma}}^{(s)} \mathbf{\Omega}^{(s)} - \mathbf{I}\right|_\infty \le \lambda_2, \tag{5.0.1}$$

where $\mathbf{e}_j$ is the jth column of identity matix $\mathbf{I}_p$, then $\mathbf{\Omega}^{(s)}\mathbf{e}_j$ is the jth column of $\mathbf{\Omega}^{(s)}$, λ_2 is tuning parameter, $w^{(s)}$ are positive weights with $\sum_{s=1}^{S} w^{(s)} = 1$.

We can also consider the following vector optimization problem:

$$\min\left\{\sum_{s=1}^{S} w^{(s)}\left|\mathbf{c}_j^{(s)}\right|_1\right\}$$

$$\text{s.t } \left|\widehat{\mathbf{\Sigma}}^{(s)}\mathbf{c}_j^{(s)} - \mathbf{e}_j\right|_\infty \leq \lambda_2. \tag{5.0.2}$$

where $\mathbf{c}_j^{(s)}$ is a vector in $\mathbb{R}^p$, $\mathbf{e}_j$ is the jth colomn of the identity matrix $\mathbf{I}_p$, $w^{(s)}$ is prespecified weight.

Similar to the proof of Lemma 1 of [11], it can be seen that the solution $\{\widehat{\mathbf{c}}_1^{(s)}, \cdots, \widehat{\mathbf{c}}_p^{(s)}\}$ $(s = 1, \cdots, S)$ of the vector optimization problem (5.0.2) is a solution of the above matrix optimization problem in (5.0.1).

Assume $\{\widehat{\mathbf{\Omega}}_1^{(1)}, \cdots, \widehat{\mathbf{\Omega}}_1^{(S)}\}$ is a solution of (5.0.1). We define our final Joint estimator $\{\widehat{\mathbf{\Omega}}^{(1)}, \cdots, \widehat{\mathbf{\Omega}}^{(S)}\}$ of $\{\mathbf{\Omega}_0^{(1)}, \cdots, \mathbf{\Omega}_0^{(S)}\}$ as the symmetrized result of $\{\widehat{\mathbf{\Omega}}_1^{(1)}, \cdots, \widehat{\mathbf{\Omega}}_1^{(S)}\}$ in the following way. Denote $\widehat{\mathbf{\Omega}}_1^{(s)} = (\widehat{\omega}_{ij,1}^{(s)})$, and $\widehat{\mathbf{\Omega}}^{(s)} = (\widehat{\omega}_{ij}^{(s)})$, for $s = 1, \cdots, S$, then

$$\widehat{\omega}_{ij}^{(s)} = \widehat{\omega}_{ij,1}^{(s)}\mathbb{1}\{\sum_{s=1}^{S} w^{(s)}\left|\widehat{\omega}_{ij,1}^{(s)}\right| \leq \sum_{s=1}^{S} w^{(s)}\left|\widehat{\omega}_{ji,1}^{(s)}\right|\} + \widehat{\omega}_{ji,1}^{(s)}\mathbb{1}\{\sum_{s=1}^{S} w^{(s)}\left|\widehat{\omega}_{ij,1}^{(s)}\right| > \sum_{s=1}^{S} w^{(s)}\left|\widehat{\omega}_{ji,1}^{(s)}\right|\}. \tag{5.0.3}$$

5.1 Estimation of precision matrices for sub-Gaussian data

In this section, we consider the precision matrix estimator of long-memory data for sub-Gaussian data in (C1). Denote $\widehat{\mathbf{\Sigma}}^{(s)} := \left(\widehat{\sigma}_{ij}^{(s)}\right)_{p\times p}$ as the sample covariance matrix of observations $\mathbf{X}^{(s)}$ ($p \times n_s$ dimension) given by

$$\widehat{\mathbf{\Sigma}}^{(s)} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i^{(s)}\mathbf{x}_i^{(s)T} - \overline{\mathbf{x}}^{(s)}\overline{\mathbf{x}}^{(s)T},$$

where $\overline{\mathbf{x}}^{(s)} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i^{(s)}$, $\mathbf{x}_i^{(s)}$ is the ith colomn of the sth group $\mathbf{X}^{(s)}$. Define

$$\tilde{u}_1^{(s)} = \max\left\{c_2K^2(\log p)g_2^{(s)}/n, [c_2K^4(\log p)g_F^{(s)}/n]^{1/2}\right\}, \tag{5.1.1}$$

$$\tilde{u}_1 := \max\left\{c_2K^2(\log p)g_2/Sn, [c_2K^4(\log p)g_F/Sn]^{1/2}\right\}, \tag{5.1.2}$$

where $c_2 = (\tau+2)/c_1$ with $\tau > 0$ and $c_1 > 0$ is an absolute constant, $g_2^{(s)}, g_F^{(s)}, g_2$, and g_F are defined in (3.1.1) and (3.1.2).

We assume $\tilde{u}_1^{(s)}, \tilde{u}_1 \to 0$ as $n \to \infty$.

Theorem 5.1.1. *Suppose that* $\mathbf{X}_{p\times nS}$ *is generated from (1.3.1) with all* z_i *satisfying condition* $(C1)$ *with the same* K. *Let* $\lambda_2 = 6C_M \max_s \tilde{u}_1^{(s)}$. *Then*

$$\max_{ij}\left(\sum_{s=1}^{S} w^{(s)}\left|\widehat{\omega}_{ij}^{(s)} - \omega_{ij,0}^{(s)}\right|\right) \le 8C_M^2 \max_s \tilde{u}_1^{(s)}$$

with probability greater than $1-4Sp^{-\tau}$, *where* $\tilde{u}_1^{(s)}$ *is defined in (5.1.1).*

To prove Theorem 5.1.1, we need the following lemma.

Lemma 5.1.1. *Suppose that* $\mathbf{X}_{p\times nS}$ *is generated from (1.3.1) with all* z_i *satisfying condition* $(C1)$ *with the same* K. *For any* $s = 1,\cdots,S$, *with probability greater than* $1-4p^{-\tau}$,

$$\max_{ij}\left|\widehat{\sigma}_{ij}^{(s)} - \sigma_{ij,0}^{(s)}\right| \le 2\tilde{u}_1^{(s)},$$

where $\tilde{u}_1^{(s)}$ *and* K *are given in Theorem 5.1.1.*

Proof of Lemma 5.1.1**.** By the proof of the first part of Lemma A2 of [35], we have

$$\begin{aligned} P\left[\left|\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}^{(s)}\right|_\infty \ge 2u\right] \;\le\; & 2p\exp\left\{-\frac{c_1 nu}{K^2 g_2^{(s)}}\right\} \\ & +2p^2\exp\left\{-c_1\min\left(\frac{nu^2}{K^4 g_F^{(s)}}, \frac{nu}{K^2 g_2^{(s)}}\right)\right\}, \end{aligned}$$

where $c_1 > 0$ is an absolute constant. Let

$$u = \tilde{u}_1^{(s)} = \max\left\{c_2K^2(\log p)g_2^{(s)}/n, [c_2K^4(\log p)g_F^{(s)}/n]^{1/2}\right\},$$

with $c_2 = (\tau+2)/c_1$, then

$$P\left[|\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}|_\infty \ge 2\tilde{u}_1^{(s)}\right] \le 2p^{-(c_1c_2-1)} + 2p^{-(c_1c_2-2)} \le 4p^{-\tau} = O(p^{-\tau}).$$

The proof is completed. $\square$

Proof of Theorem 5.1.1. We recap the proof of Theorem 1 of [27]. It is easy to see from Lemma 5.1.1 that

$$|\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}|_\infty \leq 2\tilde{u}_1^{(s)} \text{ for all } s = 1, \cdots, S, \tag{5.1.3}$$

with probability greater than $1 - 4Sp^{-\tau}$. In the rest of the proof, we assume (5.1.3) holds. Note that $\{\mathbf{\Omega}_0^{(1)}, \cdots, \mathbf{\Omega}_0^{(S)}\}$ is a feasible solution of (5.0.1) since

$$\begin{aligned} \left|\mathbf{I} - \widehat{\mathbf{\Sigma}}^{(s)}\mathbf{\Omega}_0^{(s)}\right|_\infty &= \left|(\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)})\mathbf{\Omega}_0^{(s)}\right|_\infty \\ &\leq \left\|\mathbf{\Omega}_0^{(s)}\right\|_1 |\widehat{\mathbf{\Sigma}}^{(s)} - \mathbf{\Sigma}_0^{(s)}|_\infty \\ &\leq C_M 2\tilde{u}_1^{(s)} \leq \lambda_2/3. \end{aligned}$$

Similar to the proof of Theorem 1 of [27], we have

$$\begin{aligned} \sum_{s=1}^{S} w^{(s)}|(\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j|_\infty &\leq \sum_{s=1}^{S} w^{(s)} \left|\widehat{\mathbf{\Omega}}_1^{(s)}\mathbf{e}_j\right|_1 \lambda_2/3 + C_M\lambda_2 \\ &\leq \max_j \sum_{s=1}^{S} w^{(s)} \left|\mathbf{\Omega}^{(s)}\mathbf{e}_j\right|_1 \lambda_2/3 + C_M\lambda_2 \\ &\leq 4C_M\lambda_2/3 = 8C_M^2 \max_s \tilde{u}_1^{(s)}. \end{aligned}$$

Hence, we have the inequality

$$\max_{ij}\left(\sum_{s=1}^{S} w^{(s)}\left|\widehat{\omega}_{ij}^{(s)} - \omega_{ij,0}^{(s)}\right|\right) \leq \max_j \sum_{s=1}^{S} w^{(s)}|(\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j|_\infty \leq 8C_M^2 \max_s \tilde{u}_1^{(s)}.$$

The proof is completed. $\square$

Theorem 5.1.2. *Suppose that (i) $\mathbf{X}_{p\times nS}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K, (ii) $\sum_{s=1}^{S}\left\|\mathbf{\Omega}_0^{(s)}\right\|_1 \leq C_M S^{1-\theta}$ where $\theta \in (0,1)$. Let $\lambda_2 = 6C_M \max_s \widetilde{u}_1^{(s)}$. Then*

$$\left|\sum_{s=1}^{S} \max_s w^{(s)}\left(\widehat{\mathbf{\Omega}}^{(s)} - \mathbf{\Omega}_0^{(s)}\right)\right|_\infty \leq 8C_M^2 \max_s \tilde{u}_1^{(s)} \max_s Sw^{(s)}/S^\theta$$

with probability greater than $1 - 4Sp^{-\tau}$, where $\tilde{u}_1^{(s)}$ and $\tilde{u}_1$ are defined in (5.1.1) and (5.1.2).

Proof of Theorem 5.1.2. By Lemma 5.1.1, we have

$$\max_{ij}\left|\widehat{\sigma}_{ij}^{(s)}-\sigma_{ij,0}^{(s)}\right| \le 2\tilde{u}_1^{(s)}, \tag{5.1.4}$$

for all $s=1,\cdots,S$ with probability greater than $1-4Sp^{-\tau}$. In the rest of the proof, we assume (5.1.4) holds.

Note that $\{\boldsymbol{\Omega}_0^{(1)},\cdots,\boldsymbol{\Omega}_0^{(S)}\}$ is a feasible solution of (5.0.1), because

$$\begin{aligned}\left|\mathbf{I}-\widehat{\boldsymbol{\Sigma}}^{(s)}\boldsymbol{\Omega}_0^{(s)}\right|_\infty &= \left|(\boldsymbol{\Sigma}_0^{(s)}-\widehat{\boldsymbol{\Sigma}}^{(s)})\boldsymbol{\Omega}_0^{(s)}\right|_\infty\\ &\le \left\|\boldsymbol{\Omega}_0^{(s)}\right\|_1 |\widehat{\boldsymbol{\Sigma}}^{(s)}-\boldsymbol{\Sigma}_0^{(s)}|_\infty\\ &\le C_M 2\tilde{u}_1^{(s)} \le \lambda_2/3.\end{aligned}$$

Now, we estimate the upper bound of $\left|\sum_{s=1}^S w^{(s)}(\widehat{\boldsymbol{\Omega}}_1^{(s)}-\boldsymbol{\Omega}_0^{(s)})\mathbf{e}_j\right|_\infty$. Note that

$$\begin{aligned}\left|\sum_{s=1}^S w^{(s)}(\widehat{\boldsymbol{\Omega}}_1^{(s)}-\boldsymbol{\Omega}_0^{(s)})\mathbf{e}_j\right|_\infty &\le \left|\sum_{s=1}^S \boldsymbol{\Omega}_0^{(s)} w^{(s)}(\boldsymbol{\Sigma}_0^{(s)}-\widehat{\boldsymbol{\Sigma}}^{(s)})\widehat{\boldsymbol{\Omega}}_1^{(s)}\mathbf{e}_j\right|_\infty\\ &\quad+\left|\sum_{s=1}^S \boldsymbol{\Omega}_0^{(s)} w^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)}-\mathbf{I})\mathbf{e}_j\right|_\infty.\end{aligned} \tag{5.1.5}$$

We have

$$\begin{aligned}\left|\sum_{s=1}^S \boldsymbol{\Omega}_0^{(s)} w^{(s)}(\boldsymbol{\Sigma}_0^{(s)}-\widehat{\boldsymbol{\Sigma}}^{(s)})\widehat{\boldsymbol{\Omega}}_1^{(s)}\mathbf{e}_j\right|_\infty &\le \sum_{s=1}^S\left|\boldsymbol{\Omega}_0^{(s)} w^{(s)}(\boldsymbol{\Sigma}_0^{(s)}-\widehat{\boldsymbol{\Sigma}}^{(s)})\widehat{\boldsymbol{\Omega}}_1^{(s)}\mathbf{e}_j\right|_\infty\\ &\le \sum_{s=1}^S\left|\boldsymbol{\Omega}_0^{(s)} w^{(s)}(\boldsymbol{\Sigma}_0^{(s)}-\widehat{\boldsymbol{\Sigma}}^{(s)})\right|_\infty\left|\widehat{\boldsymbol{\Omega}}_1^{(s)}\mathbf{e}_j\right|_1\\ &\le \sum_{s=1}^S\left\|\boldsymbol{\Omega}_0^{(s)}\right\|_1\left|\boldsymbol{\Sigma}_0^{(s)}-\widehat{\boldsymbol{\Sigma}}^{(s)}\right|_\infty w^{(s)}\left|\widehat{\boldsymbol{\Omega}}_1^{(s)}\mathbf{e}_j\right|_1\\ &\le \max_s\left\|\boldsymbol{\Omega}_0^{(s)}\right\|_1\left|\boldsymbol{\Sigma}_0^{(s)}-\widehat{\boldsymbol{\Sigma}}^{(s)}\right|_\infty\sum_{s=1}^S w^{(s)}\left|\boldsymbol{\Omega}_0^{(s)}\mathbf{e}_j\right|_1\\ &\le S^{1-\theta}C_M^2 2\max_s\tilde{u}_1^{(s)}\max_s w^{(s)},\end{aligned}$$

and

$$\left|\sum_{s=1}^S \boldsymbol{\Omega}_0^{(s)} w^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)}-\mathbf{I})\mathbf{e}_j\right|_\infty \le \sum_{s=1}^S\left|\boldsymbol{\Omega}_0^{(s)} w^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)}\widehat{\boldsymbol{\Omega}}_1^{(s)}-\mathbf{I})\mathbf{e}_j\right|_\infty$$

$$\leq \sum_{s=1}^{S} w^{(s)} \left\|\mathbf{\Omega}_0^{(s)}\right\|_1 \left|(\widehat{\mathbf{\Sigma}}^{(s)}\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{I})\mathbf{e}_j\right|_\infty$$
$$\leq S^{1-\theta} C_M \lambda_2 \max_s w^{(s)} = 6S^{1-\theta} C_M^2 \max_s \tilde{u}_1^{(s)} \max_s w^{(s)}.$$

In sum, we have

$$\left|\sum_{s=1}^{S} w^{(s)} \left(\widehat{\mathbf{\Omega}}^{(s)} - \mathbf{\Omega}_0^{(s)}\right)\right|_\infty \leq 8C_M^2 \max_s \tilde{u}_1^{(s)} \max_s S w^{(s)}/S^\theta.$$

The proof is completed. □

Define a threshold estimator $\widetilde{\mathbf{\Omega}}^{(s)} = (\widetilde{\omega}_{ij}^{(s)})$ based on $\{\widehat{\mathbf{\Omega}}^{(1)}, \cdots, \widehat{\mathbf{\Omega}}^{(S)}\}$ as

$$\widetilde{\omega}_{ij}^{(s)} = \widehat{\omega}_{ij}^{(s)} \mathbb{1}\left\{w^{(s)} \left|\widehat{\omega}_{ij}^{(s)}\right| \geq \delta_n\right\},$$

where $\delta_n \geq 2C_M\lambda_2$ and λ_2 is given in Theorem 5.1.1. Also, define

$$\mathcal{S}_0 := \{(i,j,s) : \omega_{ij,0}^{(s)} \neq 0\}, \widehat{\mathcal{S}} := \{(i,j,s) : \widetilde{\omega}_{ij}^{(s)} \neq 0\} \text{ and } \theta_{\min} := \min_{(i,j,s)\in\mathcal{S}_0} w^{(s)} \left|\omega_{ij,0}^{(s)}\right|.$$

Similar to the Theorem 3 of [27], we have the following theorem. The proof is straightforward. Details are omitted.

Theorem 5.1.3. *Suppose that (i) $\mathbf{X}_{p\times nS}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K, (ii) $\theta_{\min} > 2\delta_n$. Then*

$$P(\mathcal{S}_0 = \widehat{\mathcal{S}}) \geq 1 - 4Sp^{-\tau}.$$

5.2 Modified estimator $\widehat{\Omega}_\rho$

This section focuses on dealing with the expectation of the convergence rates, for example, $\mathbb{E}\left(\max_j \sum_{s=1}^S w^{(s)} \left|\mathbf{\Omega}^{(s)}\mathbf{e}_j\right|_1\right)^2$. In order to do that, we modify the estimator $\widehat{\mathbf{\Omega}}$ to ensure that the expectations exist. Let $\{\widehat{\mathbf{\Omega}}_{1\rho}\}$ be the solution of the following optimization problem

$$\min\left\{\max_j \sum_{s=1}^{S} w^{(s)} \left|\mathbf{\Omega}^{(s)}\mathbf{e}_j\right|_1\right\}$$

$$\text{s.t } \left|\widehat{\mathbf{\Sigma}}_\rho^{(s)}\mathbf{\Omega}^{(s)} - \mathbf{I}\right|_\infty \leq \lambda_2, \tag{5.2.1}$$

where $\widehat{\boldsymbol{\Sigma}}_\rho^{(s)} = \widehat{\boldsymbol{\Sigma}}^{(s)} + \rho \mathbf{I}$ with $\rho > 0$. Write $\widehat{\Omega}_{1\rho}^{(s)} = \left(\omega_{1\rho ij}^{(s)}\right)$. Define the symmetrized estimator $\widehat{\boldsymbol{\Omega}}_\rho^{(s)}$ as in (5.0.3). Clearly, $\{(\widehat{\boldsymbol{\Sigma}}_\rho^{(s)})^{-1}\}$ is a feasible point.

Theorem 5.2.1. *Suppose that $\mathbf{X}_{p\times nS}$ is generated from (1.3.1) with all z_i satisfying condition $(C1)$ with the same K. Let $\lambda_2 = 12 C_M \max_s \tilde{u}_1^{(s)}$, where $\tau > 0$. If $\rho \le \min_s \tilde{u}_1^{(s)}$, then*

(i) If $\sum_{s=1}^S \left\|\boldsymbol{\Omega}_0^{(s)}\right\|_1 \le C_M S^{1-\theta}$ where $\theta \in (0,1)$, then

$$\left|\sum_{s=1}^S w^{(s)}\left(\widehat{\boldsymbol{\Omega}}_\rho^{(s)} - \boldsymbol{\Omega}_0^{(s)}\right)\right|_\infty \le \sum_{s=1}^S w^{(s)}\left|\widehat{\boldsymbol{\Omega}}_{1\rho}^{(s)} - \boldsymbol{\Omega}_0^{(s)}\right|_\infty \le 16 C_M^2 \max_s \tilde{u}_1^{(s)} \max_s S w^{(s)}/S^\theta$$

with probability greater than $1 - 4Sp^{-\tau}$.

(ii) If there exists $s(p) > 0$ such that

$$\max_j \sum_{i=1}^p \left|\omega_{ij,0}^{(s)}\right|^q \le s(p) \text{ and } \max_j \sum_{s=1}^S \sum_{i=1}^p \left|\omega_{ij,0}^{(s)}\right|^q \le s(p) S^{1-\theta}$$

where $\theta \in (0,1)$, then

$$\begin{aligned} &\max_j \sum_{s=1}^S w^{(s)}\left|\left(\widehat{\boldsymbol{\Omega}}_\rho^{(s)} - \boldsymbol{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_1 \\ \le\ & 2(1 + 2^{1-q} + 3^{1-q})\left(16 C_M^2 \frac{\max_s \tilde{u}_1^{(s)}}{\min_s w^{(s)}}\right)^{1-q} s(p) \max_s S w^{(s)}/S^\theta \end{aligned}$$

with probability greater than $1 - 4Sp^{-\tau}$.

(iii) If the conditions in (i) and (ii) are both satisfied, then

$$\begin{aligned} &\frac{1}{p}\left\|\sum_{s=1}^S w^{(s)}\left(\widehat{\boldsymbol{\Omega}}_\rho^{(s)} - \boldsymbol{\Omega}_0^{(s)}\right)\right\|_F^2 \\ \le\ & \left|\sum_{s=1}^S w^{(s)}\left(\widehat{\boldsymbol{\Omega}}_\rho^{(s)} - \boldsymbol{\Omega}_0^{(s)}\right)\right|_\infty \max_j \sum_{s=1}^S w^{(s)}\left|\left(\widehat{\boldsymbol{\Omega}}_\rho^{(s)} - \boldsymbol{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_1 \\ \le\ & 2(1 + 2^{1-q} + 3^{1-q})\left(16 C_M^2 \frac{\max_s \tilde{u}_1^{(s)}}{\min_s w^{(s)}}\right)^{2-q} s(p)(\max_s S w^{(s)})^2/S^{2\theta} \end{aligned}$$

with probability greater than $1 - 4Sp^{-\tau}$.

Proof of Theorem 5.2.1. (i) It is easy to check that

$$\begin{aligned}\left|\mathbf{I}-\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\mathbf{\Omega}_0^{(s)}\right|_{\infty} &= \left|\left(\mathbf{\Sigma}_0^{(s)}-\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right)\mathbf{\Omega}_0^{(s)}\right|_{\infty}\\ &\le \left\|\mathbf{\Omega}_0^{(s)}\right\|_1\left|\mathbf{\Sigma}_0^{(s)}-\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right|_{\infty}\\ &\le \left(\max_s\left\|\mathbf{\Omega}_0^{(s)}\right\|_1\right)\left|\mathbf{\Sigma}_0^{(s)}-\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\right|_{\infty}\\ &\le \lambda_2. \qquad (5.2.2)\end{aligned}$$

Thus $\{\mathbf{\Omega}_0^{(1)},\cdots,\mathbf{\Omega}_0^{(S)}\}$ is a feasible solution of (5.2.1). Therefore,

$$\max_j\sum_{s=1}^{S}w^{(s)}\left|\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1\le\max_j\sum_{s=1}^{S}w^{(s)}\left|\mathbf{\Omega}_0^{(s)}\mathbf{e}_j\right|_1. \qquad (5.2.3)$$

Note that

$$\begin{aligned}&\left|\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\left(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\right|_{\infty}\\ \le\ &\left|\left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{I}\right)\right|_{\infty}+\left|\left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\mathbf{\Omega}_0^{(s)}-\mathbf{I}\right)\right|_{\infty}\\ \le\ &2\lambda_2. \qquad (5.2.4)\end{aligned}$$

It follows that

$$\begin{aligned}&\left|\sum_{s=1}^{S}w^{(s)}\left(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_{\infty}\\ \le\ &\sum_{s=1}^{S}w^{(s)}\left|\left(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_{\infty}\\ \le\ &\sum_{s=1}^{S}w^{(s)}\left|\mathbf{\Omega}_0^{(s)}\mathbf{\Sigma}_0^{(s)}\left(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_{\infty}\\ \le\ &\sum_{s=1}^{S}w^{(s)}\left\|\mathbf{\Omega}_0^{(s)}\right\|_1\left|\mathbf{\Sigma}_0^{(s)}\left(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_{\infty}\\ \le\ &\sum_{s=1}^{S}w^{(s)}\left\|\mathbf{\Omega}_0^{(s)}\right\|_1\left(\left|\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}\left(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_{\infty}+\left|\left(\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}-\mathbf{\Sigma}_0^{(s)}\right)\left(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_{\infty}\right)\\ \le\ &\sum_{s=1}^{S}w^{(s)}\left\|\mathbf{\Omega}_0^{(s)}\right\|_1\left(2\lambda_2+\left|\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}-\mathbf{\Sigma}_0^{(s)}\right|_{\infty}\left|\left(\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_1\right)\\ \le\ &\sum_{s=1}^{S}w^{(s)}\left\|\mathbf{\Omega}_0^{(s)}\right\|_1 2\lambda_2+\sum_{s=1}^{S}\left\|\mathbf{\Omega}_0^{(s)}\right\|_1\left|\widehat{\mathbf{\Sigma}}_{\rho}^{(s)}-\mathbf{\Sigma}_0^{(s)}\right|_{\infty}\left(w^{(s)}\left|\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1+w^{(s)}\left|\mathbf{\Omega}_0^{(s)}\mathbf{e}_j\right|_1\right)\end{aligned}$$

$$
\begin{aligned}
&\le \sum_{s=1}^{S} w^{(s)} \left\|\boldsymbol{\Omega}_0^{(s)}\right\|_1 2\lambda_2 + C_M \max_s \left|\widehat{\boldsymbol{\Sigma}}_\rho^{(s)} - \boldsymbol{\Sigma}_0^{(s)}\right|_\infty \sum_{s=1}^{S} \left(w^{(s)} \left|\widehat{\boldsymbol{\Omega}}_{1\rho}^{(s)} \mathbf{e}_j\right|_1 + w^{(s)} \left|\boldsymbol{\Omega}_0^{(s)} \mathbf{e}_j\right|_1 \right) \\
&= \sum_{s=1}^{S} w^{(s)} \left\|\boldsymbol{\Omega}_0^{(s)}\right\|_1 2\lambda_2 \\
&\quad + C_M \max_s \left|\widehat{\boldsymbol{\Sigma}}_\rho^{(s)} - \boldsymbol{\Sigma}_0^{(s)}\right|_\infty \left(\max_j \sum_{s=1}^{S} w^{(s)} \left|\widehat{\boldsymbol{\Omega}}_{1\rho}^{(s)} \mathbf{e}_j\right|_1 + \max_j \sum_{s=1}^{S} w^{(s)} \left|\boldsymbol{\Omega}_0^{(s)} \mathbf{e}_j\right|_1 \right) \\
&\le \sum_{s=1}^{S} w^{(s)} \left\|\boldsymbol{\Omega}_0^{(s)}\right\|_1 2\lambda_2 + C_M \max_s \left|\widehat{\boldsymbol{\Sigma}}_\rho^{(s)} - \boldsymbol{\Sigma}_0^{(s)}\right|_\infty 2 \sum_{s=1}^{S} w^{(s)} \left\|\boldsymbol{\Omega}_0^{(s)}\right\|_1 \\
&\le 16 C_M^2 \min_s \tilde{u}_1^{(s)} \max_s S w^{(s)} / S^\theta.
\end{aligned}
$$

By the definition of $\widehat{\boldsymbol{\Omega}}_\rho^{(s)}$, we have

$$
\left|\sum_{s=1}^{S} w^{(s)} \left(\widehat{\boldsymbol{\Omega}}_\rho^{(s)} - \boldsymbol{\Omega}_0^{(s)}\right)\right|_\infty \le \left|\sum_{s=1}^{S} w^{(s)} \left(\widehat{\boldsymbol{\Omega}}_{1\rho}^{(s)} - \boldsymbol{\Omega}_0^{(s)}\right)\right|_\infty \le 16 C_M^2 \min_s \tilde{u}_1^{(s)} \max_s S w^{(s)} / S^\theta.
$$

(ii) Now, let us consider the L^1 norm.

We assume the solution $\widehat{\boldsymbol{\Omega}}_\rho^{(s)}$ is combined from the column solution set $\{\widehat{\mathbf{c}}_{j\rho}^{(s)}\}$ of (5.0.2) with $\widehat{\boldsymbol{\Sigma}}^{(s)}$ being replaced by $\widehat{\boldsymbol{\Sigma}}_\rho^{(s)}$. By the definition of $\{\widehat{\boldsymbol{\Omega}}_\rho^{(s)}\}$=$\{\left(\widehat{\omega}_{\rho ij}^{(s)}\right)\}$, we have

$$
\sum_{s=1}^{S} w^{(s)} \left|\widehat{\boldsymbol{\Omega}}_\rho^{(s)} \mathbf{e}_j\right|_1 \le \sum_{s=1}^{S} w^{(s)} \left|\widehat{\boldsymbol{\Omega}}_{1\rho}^{(s)} \mathbf{e}_j\right|_1 \le \sum_{s=1}^{S} w^{(s)} \left|\boldsymbol{\Omega}_0^{(s)} \mathbf{e}_j\right|_1 \quad \text{for } j = 1, \cdots, p.
$$

Denote

$$
\begin{aligned}
t &= \max_{ijs} \left|\widehat{\omega}_{\rho ij}^{(s)} - \omega_{ij,0}^{(s)}\right|, \text{ where } \boldsymbol{\Omega}_0^{(s)} = \left(\omega_{ij,0}^{(s)}\right), \\
\mathbf{L}^{(s)} &= \left(\widehat{\omega}_{\rho ij}^{(s)} - \omega_{ij,0}^{(s)}\right), \\
\mathbf{L}_*^{(s)} &= \left(\widehat{\omega}_{\rho ij}^{(s)} \mathbb{1}(\left|\widehat{\omega}_{\rho ij}^{(s)}\right| \ge 2t) - \omega_{ij,0}^{(s)}\right), \\
\mathbf{L}_{**}^{(s)} &= \mathbf{L}^{(s)} - \mathbf{L}_*^{(s)}.
\end{aligned}
$$

Then

$$
\begin{aligned}
&\sum_{s=1}^{S} w^{(s)} \left(\left|\boldsymbol{\Omega}_0^{(s)} \mathbf{e}_j\right|_1 - \left|\mathbf{L}_*^{(s)} \mathbf{e}_j\right|_1 + \left|\mathbf{L}_{**}^{(s)} \mathbf{e}_j\right|_1 \right) \\
\le &\sum_{s=1}^{S} w^{(s)} \left(\left|\left(\boldsymbol{\Omega}_0^{(s)} + \mathbf{L}_*^{(s)}\right) \mathbf{e}_j\right|_1 + \left|\mathbf{L}_{**}^{(s)} \mathbf{e}_j\right|_1 \right)
\end{aligned}
$$

$$= \sum_{s=1}^{S} w^{(s)} \left(\left| \widehat{\mathbf{\Omega}}_{\rho}^{(s)} \mathbf{e}_j \right|_1 \right) \leq \sum_{s=1}^{S} w^{(s)} \left(\left| \mathbf{\Omega}_{0}^{(s)} \mathbf{e}_j \right|_1 \right),$$

which implies

$$\sum_{s=1}^{S} w^{(s)} \left| \mathbf{L}_{**}^{(s)} \mathbf{e}_j \right|_1 \leq \sum_{s=1}^{S} w^{(s)} \left| \mathbf{L}_{*}^{(s)} \mathbf{e}_j \right|_1 .$$

Thus

$$\sum_{s=1}^{S} w^{(s)} \left| \mathbf{L}^{(s)} \mathbf{e}_j \right|_1 \leq \sum_{s=1}^{S} w^{(s)} \left| \mathbf{L}_{**}^{(s)} \mathbf{e}_j \right|_1 + \sum_{s=1}^{S} w^{(s)} \left| \mathbf{L}_{*}^{(s)} \mathbf{e}_j \right|_1 \leq 2 \sum_{s=1}^{S} w^{(s)} \left| \mathbf{L}_{*}^{(s)} \mathbf{e}_j \right|_1 .$$

Similar to the proof of Theorem 6 of [11], we have

$$\begin{aligned}
\sum_{s=1}^{S} w^{(s)} \left| \mathbf{L}_{*}^{(s)} \mathbf{e}_j \right|_1 &= \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \widehat{\omega}_{\rho ij}^{(s)} \mathbb{1}(\left| \widehat{\omega}_{\rho ij}^{(s)} \right| \geq 2t) - \omega_{ij,0}^{(s)} \right| \\
&\leq \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \mathbb{1}(\left| \omega_{ij,0}^{(s)} \right| \leq 2t) \right| \\
&\quad + \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \widehat{\omega}_{\rho ij}^{(s)} \mathbb{1}(\left| \widehat{\omega}_{\rho ij}^{(s)} \right| \geq 2t) - \omega_{ij,0}^{(s)} \mathbb{1}(\left| \omega_{ij,0}^{(s)} \right| \geq 2t) \right| \\
&\leq (2t)^{1-q} \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \right|^q + \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \left(\widehat{\omega}_{\rho ij}^{(s)} - \omega_{ij,0}^{(s)} \right) \mathbb{1}(\left| \widehat{\omega}_{\rho ij}^{(s)} \right| \geq 2t) \right| \\
&\quad + \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \{ \mathbb{1}(\left| \widehat{\omega}_{\rho ij}^{(s)} \right| \geq 2t) - \mathbb{1}(\left| \omega_{ij,0}^{(s)} \right| \geq 2t) \} \right| \\
&\leq (2t)^{1-q} \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \right|^q + t \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \mathbb{1}(\left| \widehat{\omega}_{\rho ij}^{(s)} \right| \geq 2t) \\
&\quad + \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \right| \left| \mathbb{1}\{ \left| \left| \omega_{ij,0}^{(s)} \right| - 2t \right| \leq \left| \widehat{\omega}_{\rho ij}^{(s)} - \omega_{ij,0}^{(s)} \right| \} \right| \\
&\leq (2t)^{1-q} \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \right|^q + t \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \mathbb{1}(\left| \widehat{\omega}_{\rho ij}^{(s)} \right| \geq 2t) \\
&\quad + \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \right| \mathbb{1}\{ \left| \omega_{ij,0}^{(s)} \right| \leq 3t \} \\
&\leq (2t)^{1-q} \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \right|^q + t^{1-q} \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \right|^q \\
&\quad + (3t)^{1-q} \sum_{s=1}^{S} \sum_{i=1}^{p} w^{(s)} \left| \omega_{ij,0}^{(s)} \right|^q
\end{aligned}$$

$$
\begin{aligned}
&\leq (1+2^{1-q}+3^{1-q})t^{1-q}\sum_{s=1}^{S}\sum_{i=1}^{p} w^{(s)}\left|\omega_{ij,0}^{(s)}\right|^q \\
&\leq (1+2^{1-q}+3^{1-q})t^{1-q}s(p)S^{1-\theta}\max_s w^{(s)}.
\end{aligned}
$$

By Theorem 5.1.1, it is easy to see that $t \leq 16C_M^2 \max_s \tilde{u}_1^{(s)}/\min_s w^{(s)}$. Thus,

$$
\begin{aligned}
& \max_j \sum_{s=1}^{S} w^{(s)}\left|\left(\widehat{\mathbf{\Omega}}_\rho^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_1 \\
= & \max_j \sum_{s=1}^{S} w^{(s)}\left|\mathbf{L}^{(s)}\mathbf{e}_j\right|_1 \\
\leq & 2\max_j \sum_{s=1}^{S} w^{(s)}\left|\mathbf{L}_*^{(s)}\mathbf{e}_j\right|_1 \\
\leq & 2(1+2^{1-q}+3^{1-q})t^{1-q}s(p)\max_s Sw^{(s)}/S^\theta \\
\leq & 2(1+2^{1-q}+3^{1-q})\left(16C_M^2\frac{\max_s \tilde{u}_1^{(s)}}{\min_s w^{(s)}}\right)^{1-q} s(p)\max_s Sw^{(s)}/S^\theta.
\end{aligned}
$$

(iii) It is easy to check that

$$
\begin{aligned}
\frac{1}{p}\left\|\sum_{s=1}^{S} w^{(s)}\left(\widehat{\mathbf{\Omega}}_\rho^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\right\|_F^2 &\leq \left|\sum_{s=1}^{S} w^{(s)}\left(\widehat{\mathbf{\Omega}}_\rho^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\right|_\infty \max_j \sum_{s=1}^{S} w^{(s)}\left|\left(\widehat{\mathbf{\Omega}}_\rho^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_1 \\
&\leq 2(1+2^{1-q}+3^{1-q})\left(16C_M^2\frac{\max_s \tilde{u}_1^{(s)}}{\min_s w^{(s)}}\right)^{2-q} s(p)(\max_s Sw^{(s)})^2/S^{2\theta}.
\end{aligned}
$$

The proof is completed. □

Theorem 5.2.2. *Under the conditions of Theorem 5.2.1, and* $\min\{\min_s \tilde{u}_1^{(s)}, p^{-c}\} \leq \rho \leq \min_s \tilde{u}_1^{(s)}$ *with* $c > 0$*, if* $p \geq n_s^\xi$ *with* $\xi > 0$*, we have*

(i) $\mathbb{E}\left(\sum_{s=1}^S w^{(s)}\left|\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}-\mathbf{\Omega}_0^{(s)}\right|_\infty\right)^2 = O\left((C_M^2\max_s \tilde{u}_1^{(s)}\max_s Sw^{(s)})^2/S^{2\theta}\right)$,

(ii) $\mathbb{E}\left(\max_j\sum_{s=1}^S w^{(s)}\left|\left(\widehat{\mathbf{\Omega}}_\rho^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\mathbf{e}_j\right|_1\right)^2 = O\left((C_M^2\frac{\max_s \tilde{u}_1^{(s)}}{\min_s w^{(s)}})^{2(1-q)}s^2(p)(\max_s Sw^{(s)})^2/S^{2\theta}\right)$,

(iii) $\mathbb{E}\frac{1}{p}\left\|\sum_{s=1}^S w^{(s)}\left(\widehat{\mathbf{\Omega}}_\rho^{(s)}-\mathbf{\Omega}_0^{(s)}\right)\right\|_F^2 = O\left((C_M^2\frac{\max_s \tilde{u}_1^{(s)}}{\min_s w^{(s)}})^{(2-q)}s(p)(\max_s Sw^{(s)})^2/S^{2\theta}\right)$.

Proof of Theorem 5.2.2. Note that $\{\left(\widehat{\mathbf{\Sigma}}_\rho^{(s)}\right)^{-1}\}$ is a feasible solution of (5.2.1). Note also that

$$
\max_j \sum_{s=1}^{S} w^{(s)}\left|\widehat{\mathbf{\Omega}}_\rho^{(s)}\mathbf{e}_j\right|_1 \leq \max_j \sum_{s=1}^{S} w^{(s)}\left|\widehat{\mathbf{\Omega}}_{1\rho}^{(s)}\mathbf{e}_j\right|_1
$$

$$
\begin{aligned}
&\le \max_j \sum_{s=1}^{S} w^{(s)} \left| \left(\widehat{\boldsymbol{\Sigma}}_\rho^{(s)}\right)^{-1} \mathbf{e}_j \right|_1 \\
&\le \max_j \sum_{s=1}^{S} w^{(s)} \left\| \left(\widehat{\boldsymbol{\Sigma}}_\rho^{(s)}\right)^{-1} \right\|_1 \\
&\le p \sum_{s=1}^{S} w^{(s)} \left\| \left(\widehat{\boldsymbol{\Sigma}}_\rho^{(s)}\right)^{-1} \right\|_2 \\
&\le p \sum_{s=1}^{S} w^{(s)} \lambda_{\max} \left(\widehat{\boldsymbol{\Sigma}}_\rho^{(s)}\right)^{-1} \\
&= p \sum_{s=1}^{S} w^{(s)} / \lambda_{\min} \left(\widehat{\boldsymbol{\Sigma}}_\rho^{(s)}\right), \\
&\le p \sum_{s=1}^{S} w^{(s)} / \rho = p/\rho.
\end{aligned}
$$

Thus $\mathbb{E}\left(\max_j \sum_{s=1}^{S} w^{(s)} \left| \left(\widehat{\boldsymbol{\Omega}}_\rho^{(s)} - \boldsymbol{\Omega}_0^{(s)}\right) \mathbf{e}_j \right|_1\right)^2$ is well defined. The proofs similarly follow the proof of Theorems 2 and 5 of [11]. □

For LPDD model, similarly, we only need to assume $p \ge (\log n_s)^\xi$ in the above theorem.

5.3 Extension to heavy tail data

In this section, we extend the previous results for the precision matrices estimators to heavy-tailed data in (C2) and (C3). Define

$$
\widetilde{u}_2^{(s)} = \max\left\{[c_2 K (2/\vartheta)^{2/\vartheta} \log p]^{1+2/\vartheta} g_2^{(s)}/n_s, [c_2 K^2 (4/\vartheta)^{4/\vartheta} \log p]^{1+2/\vartheta} (g_F^{(s)}/n_s)^{1/2}\right\}, \quad (5.3.1)
$$

and

$$
\widetilde{u}_3^{(s)} = k\eta_k^2 \max\left\{c_2^2 p^{(2+2\tau)/k} g_2^{(s)}/n_s, c_2 p^{(4+2\tau)/k} (g_F^{(s)}/n_s)^{1/2}\right\}, \quad (5.3.2)
$$

where $c_2 = (\tau+2)/c_1$ with $\tau > 0$ and $c_1 > 0$ is an absolute constant, $g_2^{(s)}, g_F^{(s)}, g_2$, and g_F are defined in (3.1.1) and (3.1.2). We assume $\widetilde{u}_2^{(s)}, \widetilde{u}_3^{(s)} \to 0$ as $n_s \to \infty$.

Theorem 5.3.1. *(Generalized sub-exponential tails) Assume condition (C1), parameter K, and $\widetilde{u}_1^{(s)}$ replaced by condition (C2), parameters $\{K, \vartheta\}$, and $\widetilde{u}_2^{(s)}$, respectively, then Theorems 5.1.1-5.2.2 hold.*

Theorem 5.3.2. *(Polynomial-type tails) Assume condition (C1), parameter K, and $\widetilde{u}_1^{(s)}$ replaced by condition (C3), parameters $\{k, \eta_k\}$, and $\widetilde{u}_3^{(s)}$, respectively, then Theorems 5.1.1-5.2.1 hold.*

The proofs of Theorems 5.3.1 and 5.3.2 are similar to the proofs of the preceding theorems by using the corresponding results given in Lemmas A.1, A.2 and A.3 of [35] for conditions (C2) and (C3), respectively. Details are omitted.

5.4 Computation

We can reformulate (5.0.1) into the linear programming form (3.4.2). The numerical algorithm is similar to that of Chapter 3. In order to do that, we need to use the following vector $\mathbf{a}$, matrix $\mathbf{A}$, and vector $\mathbf{b}$

$$\begin{aligned}
\widetilde{\mathbf{a}} &= \left(\underbrace{w^{(1)}, \cdots, w^{(1)}}_{p}, \cdots, \underbrace{w^{(S)}, \cdots, w^{(S)}}_{p}\right), \\
\mathbf{a} &= -\left(\widetilde{\mathbf{a}}, \widetilde{\mathbf{a}}\right)^T, \\
\mathbf{M} &= \begin{pmatrix} \mathbf{A} \\ -\mathbf{A} \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix},
\end{aligned}$$

with $\mathbf{A} = \left(\widetilde{\mathbf{A}}, -\widetilde{\mathbf{A}}\right)$ where

$$\begin{aligned}
\widetilde{\mathbf{A}} &= \begin{pmatrix} \widehat{\mathbf{\Sigma}}^{(1)} & \mathbf{O}_{p\times p} & \cdots & \mathbf{O}_{p\times p} \\ \mathbf{O}_{p\times p} & \widehat{\mathbf{\Sigma}}^{(2)} & \cdots & \mathbf{O}_{p\times p} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{O}_{p\times p} & \mathbf{O}_{p\times p} & \cdots & \widehat{\mathbf{\Sigma}}^{(S)} \end{pmatrix}, \\
\mathbf{b}_1 &= \begin{pmatrix} \lambda_2 + \mathbf{e}_i \\ \vdots \\ \lambda_2 + \mathbf{e}_i \end{pmatrix}_{Sp\times 1}, \quad \mathbf{b}_2 = \begin{pmatrix} \lambda_2 - \mathbf{e}_i \\ \vdots \\ \lambda_2 - \mathbf{e}_i \end{pmatrix}_{Sp\times 1}.
\end{aligned}$$

and we need to use (5.0.3) to symmeterize matrices.

5.4.1 Weight selection

We can take weights $w^{(s)}$ $(s = 1, \cdots, S)$ as tuning parameters and select them in the following way. We set the tuning parameters $\{w^{(s)}\}$ to make the upper bound of $\max_{ij}\left(\sum_{s=1}^{S}\left|\widehat{\omega}_{ij}^{(s)} - \omega_{ij,0}^{(s)}\right|\right)$ achieve its minima. Similar to the proof of Theorem 1 of [27], we have

$$\begin{aligned}
\sum_{s=1}^{S}|(\widehat{\mathbf{\Omega}}_1^{(s)} - \mathbf{\Omega}_0^{(s)})\mathbf{e}_j|_\infty &\leq C_M\sum_{s=1}^{S}w^{(s)}\left|\widehat{\mathbf{\Omega}}_1^{(s)}\mathbf{e}_j\right|_1\left|\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}\right|_\infty/w^{(s)} + SC_M\lambda_2 \\
&\leq C_M\max_j\sum_{s=1}^{S}w^{(s)}\left|\mathbf{\Omega}^{(s)}\mathbf{e}_j\right|_1\max_s\left(\left|\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}\right|_\infty/w^{(s)}\right) + SC_M\lambda_2 \\
&\leq C_M^2\max_s\left|\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}\right|_\infty/w^{(s)} + SC_M\lambda_2 \\
&\leq C_M^2\max_s\left|\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}\right|_\infty/w^{(s)} + 6SC_M^2\max_s\tilde{u}_1^{(s)}.
\end{aligned}$$

Hence, we only need to minimize $\max_s\left|\mathbf{\Sigma}_0^{(s)} - \widehat{\mathbf{\Sigma}}^{(s)}\right|_\infty/w^{(s)}$.

We will conduct the computation in three steps. *Step 1.* Compute $g_2^{(s)}, g_F^{(s)}$; *Step 2.* Select the weight $\{w^{(s)}\}$ by minimizing the upper bound of $\max_{ij}\left(\sum_{s=1}^{S}\left|\widehat{\omega}_{ij}^{(s)} - \omega_{ij,0}^{(s)}\right|\right)$ as above; *Step 3.* Compute the precision matrix by the vector optimization problem. We evaluate the numerical performance of the Joint estimators for high dimensional precision matrices.

5.5 Numerical experiments

5.5.1 Results of simulation study

Since this chapter is about weighted-Joint, we study the weight estimation first. We sample $\mathbf{X}_{p\times n}^{(s)}$ $(s = 1, \cdots, S)$ 20 times with distribution $\mathbf{\Sigma}^{(s)}$ and $\alpha \in \{0.1, 0.25, 0.5, 1, 2\}$. Then compute weight $\{w^{(s)}\}$ by the above computation steps, and repeat 50 times. Because each group is generated in the same way, the optimal weight is the uniform weight, $1/4 = 0.25$.

Tables 5.1-5.2 report the estimated weights for Groups 1-4. Overall, the weights are substantially different with the uniform weights in some settings, and the higher the dimension is, the closer the estimated weights are to the uniform weights. In Table 5.1 ($p = 50$), for

Model 1 and $\rho = 0.1$ or 4 and for Model 2 and $\rho = 4$, the difference between the estimated and the uniform weights is large. For Model 1 with $\alpha = 1$ or 2 and $\rho = 1$ and for Model 2 with $\alpha = 2$ and $\rho = 1$, the estimated weights are somewhat close to the uniform weights, but difference still exists. In the other settings, the difference is marginal, and the estimated weights are close to 0.25. In Table 5.2 ($p = 100$), for Model 1 and $\rho = 0.1$, and for Model 2 and $\alpha = 1$ or 2 and $\rho = 1$ or 4, the difference between the estimated and the uniform weights is large. For Model 1 with $\alpha = 1$ or 2 and $\rho = 0.25$, the estimated weights are somewhat close to the uniform weights, but difference still exists. In the other settings, the difference is marginal, and the estimated weights are close to 0.25.

In what follows, we compare the performance between SCLIME and weighted-Joint. We prespecified λ_2 from 0.05 to 0.25, the step is 0.01. In the tables, "Joint" stands for the algorithm introduced in this chapter. Tables 5.3-5.4 show the results of the simulation study. Tables 5.3 and 5.4 imply that, (i) if $p = 50$, $\alpha = 0.1$, Joint performs apparently better than SCLIME for both models in terms of the Frobenius norms; (ii) if $p = 50$, $\alpha = 0.25$ or 0.5, and $\rho = 0.1, 0.25$, or 1, Joint performs apparently better than SCLIME for both models in terms of the Frobenius norms; (iii) if $p = 100$ and $\alpha = 0.1$ or 0.25, Joint performs apparently better than SCLIME for both models in terms of the Frobenius norms, sometimes significantly better. Except the above settings, occasionally, SCLIME performs significantly better than Joint, generally, no apparent difference.

5.5.2 rfMRI data analysis

We analyze the same rfMRI data set in Chapter 3 for the estimation of precision matrices and brain functional connectivity by our Joint. For these 3 groups by Joint, the minimum connectivity is 0, the maximum connectivity is 35, and the median is 13.5. It is apparently higher than that of ordinary JEMP and weighted-JEMP.

Table 5.1: Average (SD) weight for $p = 50$

α	ρ	weight 1	weight 2	weight 3	weight 4
		Model 1			
0.1	0.1	0.2084(0.0244)	0.3122(0.0313)	0.2546(0.0308)	0.2248(0.0276)
	0.25	0.2416(0.0213)	0.2518(0.0249)	0.2592(0.0239)	0.2474(0.0209)
	1	0.253(0.0223)	0.235(0.0213)	0.2722(0.0313)	0.2398(0.0258)
	4	0.101(0.0036)	0.626(0.0284)	0.1008(0.0034)	0.1722(0.0256)
0.25	0.1	0.2084(0.0197)	0.314(0.0337)	0.2568(0.0285)	0.2208(0.0245)
	0.25	0.2458(0.0219)	0.2638(0.0244)	0.2478(0.0195)	0.2426(0.0223)
	1	0.2578(0.0245)	0.2362(0.0236)	0.275(0.0378)	0.231(0.0208)
	4	0.1006(0.0024)	0.6306(0.0247)	0.1002(0.0014)	0.1686(0.0232)
0.5	0.1	0.2022(0.0178)	0.319(0.029)	0.2622(0.0257)	0.2166(0.0203)
	0.25	0.2454(0.0221)	0.261(0.0228)	0.2494(0.0194)	0.2442(0.0241)
	1	0.2578(0.0249)	0.237(0.0225)	0.2788(0.0321)	0.2264(0.0188)
	4	0.1(0)	0.6356(0.0193)	0.1(0)	0.1644(0.0193)
1	0.1	0.19(0.0147)	0.3192(0.0194)	0.2672(0.0178)	0.2236(0.0132)
	0.25	0.2436(0.016)	0.2566(0.0207)	0.2566(0.0148)	0.2432(0.0148)
	1	0.2478(0.0131)	0.2314(0.0148)	0.2986(0.0208)	0.2222(0.0133)
	4	0.1(0)	0.646(0.0095)	0.1(0)	0.154(0.0095)
2	0.1	0.19(0.009)	0.3158(0.0116)	0.2704(0.0109)	0.2238(0.0092)
	0.25	0.243(0.0081)	0.2574(0.0126)	0.2588(0.0115)	0.2408(0.0105)
	1	0.2456(0.0088)	0.2284(0.0089)	0.3108(0.0121)	0.2152(0.0071)
	4	0.1(0)	0.6488(0.0075)	0.1(0)	0.1512(0.0075)
		Model 2			
0.1	0.1	0.2484(0.0214)	0.2528(0.0205)	0.2458(0.0208)	0.253(0.0275)
	0.25	0.2464(0.0191)	0.2522(0.0227)	0.2566(0.0186)	0.2448(0.0164)
	1	0.245(0.0227)	0.255(0.0242)	0.2634(0.0307)	0.2366(0.0251)
	4	0.7(0)	0.1(0)	0.1(0)	0.1(0)
0.25	0.1	0.2528(0.0225)	0.2526(0.0185)	0.2478(0.0159)	0.2468(0.018)
	0.25	0.2502(0.0204)	0.2516(0.0186)	0.2468(0.0172)	0.2514(0.0195)
	1	0.2526(0.0201)	0.2436(0.0239)	0.2624(0.0258)	0.2414(0.0234)
	4	0.7(0)	0.1(0)	0.1(0)	0.1(0)
0.5	0.1	0.255(0.0195)	0.2494(0.0161)	0.2464(0.0161)	0.2492(0.0163)
	0.25	0.2484(0.0162)	0.25(0.0171)	0.2486(0.0164)	0.253(0.0178)
	1	0.2492(0.0201)	0.2486(0.0219)	0.2654(0.026)	0.2368(0.0206)
	4	0.7(0)	0.1(0)	0.1(0)	0.1(0)
1	0.1	0.2448(0.0109)	0.251(0.0123)	0.2534(0.0132)	0.2508(0.0128)
	0.25	0.244(0.0162)	0.2576(0.0162)	0.249(0.0136)	0.2494(0.0117)
	1	0.253(0.0159)	0.2504(0.0164)	0.2694(0.0187)	0.2272(0.0155)
	4	0.7(0)	0.1(0)	0.1(0)	0.1(0)
2	0.1	0.2498(0.0082)	0.249(0.0071)	0.2538(0.0075)	0.2474(0.0075)
	0.25	0.2384(0.0084)	0.2652(0.0076)	0.2494(0.0059)	0.247(0.0061)
	1	0.2572(0.0081)	0.2506(0.0079)	0.2766(0.0114)	0.2156(0.0064)
	4	0.7(0)	0.1(0)	0.1(0)	0.1(0)

Table 5.2: Average (SD) weight for $p = 100$

α	ρ	weight 1	weight 2	weight 3	weight 4
			Model 1		
0.1	0.1	0.3516(0.0306)	0.3596(0.0387)	0.1566(0.0211)	0.1322(0.0202)
	0.25	0.2296(0.0232)	0.2374(0.026)	0.2758(0.0306)	0.2572(0.027)
	1	0.238(0.0212)	0.2432(0.0221)	0.2596(0.0271)	0.2592(0.023)
	4	0.2518(0.026)	0.241(0.022)	0.2562(0.0251)	0.251(0.0195)
0.25	0.1	0.346(0.0334)	0.3694(0.0439)	0.153(0.0209)	0.1316(0.0194)
	0.25	0.2256(0.0212)	0.2386(0.0213)	0.2826(0.036)	0.2532(0.0292)
	1	0.2308(0.0216)	0.2418(0.0255)	0.2622(0.0261)	0.2652(0.0262)
	4	0.253(0.0218)	0.2502(0.023)	0.2552(0.0219)	0.2416(0.0245)
0.5	0.1	0.3478(0.0309)	0.3744(0.0386)	0.1512(0.0198)	0.1266(0.0165)
	0.25	0.2224(0.0205)	0.2368(0.0203)	0.286(0.0323)	0.2548(0.0272)
	1	0.2326(0.0184)	0.2424(0.019)	0.2618(0.0256)	0.2632(0.0236)
	4	0.2556(0.0209)	0.2508(0.0189)	0.2534(0.0238)	0.2402(0.0232)
1	0.1	0.3506(0.0133)	0.3942(0.0173)	0.139(0.0093)	0.1162(0.0109)
	0.25	0.2146(0.0147)	0.2322(0.0147)	0.3002(0.0203)	0.253(0.0181)
	1	0.2358(0.0139)	0.235(0.0115)	0.2628(0.0157)	0.2664(0.0163)
	4	0.2434(0.017)	0.2514(0.0159)	0.2598(0.0155)	0.2454(0.015)
2	0.1	0.3538(0.009)	0.3956(0.0105)	0.1352(0.0058)	0.1154(0.0068)
	0.25	0.2042(0.0095)	0.238(0.0097)	0.3032(0.0106)	0.2546(0.0095)
	1	0.2316(0.0077)	0.234(0.0073)	0.2558(0.0109)	0.2786(0.0107)
	4	0.2454(0.0101)	0.2488(0.0075)	0.2618(0.0114)	0.244(0.0078)
			Model 2		
0.1	0.1	0.245(0.0191)	0.2548(0.0201)	0.248(0.0202)	0.2522(0.0209)
	0.25	0.2472(0.0215)	0.2442(0.0207)	0.2552(0.0233)	0.2534(0.0255)
	1	0.2626(0.0302)	0.2392(0.0272)	0.28(0.0389)	0.2182(0.0198)
	4	0.2458(0.0233)	0.2378(0.0223)	0.248(0.0252)	0.2684(0.026)
0.25	0.1	0.2498(0.02)	0.2526(0.0203)	0.2436(0.0161)	0.254(0.0193)
	0.25	0.2448(0.0167)	0.2498(0.0208)	0.2588(0.0172)	0.2466(0.0194)
	1	0.259(0.0303)	0.245(0.0297)	0.278(0.0253)	0.218(0.023)
	4	0.2438(0.0216)	0.23(0.0195)	0.257(0.0247)	0.2692(0.0249)
0.5	0.1	0.25(0.0183)	0.251(0.0159)	0.2456(0.0172)	0.2534(0.0187)
	0.25	0.2466(0.0165)	0.2466(0.017)	0.2578(0.0189)	0.249(0.0157)
	1	0.261(0.0259)	0.245(0.0253)	0.279(0.0263)	0.215(0.0207)
	4	0.2416(0.019)	0.2326(0.0168)	0.255(0.0232)	0.2708(0.0236)
1	0.1	0.2516(0.0145)	0.2506(0.0136)	0.246(0.0165)	0.2518(0.0129)
	0.25	0.2446(0.0122)	0.2454(0.0139)	0.2584(0.0149)	0.2516(0.0125)
	1	0.2604(0.0183)	0.2374(0.0194)	0.3076(0.021)	0.1946(0.0125)
	4	0.237(0.0137)	0.2202(0.0156)	0.2498(0.0145)	0.293(0.0164)
2	0.1	0.2534(0.0082)	0.2452(0.0074)	0.242(0.0067)	0.2594(0.0111)
	0.25	0.243(0.0086)	0.2438(0.0075)	0.26(0.0093)	0.2532(0.0071)
	1	0.2618(0.0127)	0.2304(0.0109)	0.3166(0.0135)	0.1912(0.0077)
	4	0.2336(0.0101)	0.2146(0.0079)	0.2466(0.0096)	0.3052(0.0127)

Table 5.3: Comparison of average (SD) matrix losses for $p = 50$

		SCLIME	Joint	SCLIME	Joint
α	ρ	Spectral norm		Frobenius norm	
		Model 1			
0.1	0.1	7.3422(0.3566)	8.0262(0.7119)	30.0442(0.9959)	27.1075(1.7891)
	0.25	7.4576(0.2451)	8.4302(0.7103)	31.0838(0.8742)	28.9636(1.5121)
	1	7.5935(0.3112)	8.0862(0.5829)	31.281(0.9004)	27.4566(1.6587)
	4	8.4898(0.7337)	6.0131(0.6158)	34.0082(2.694)	20.109(2.2924)
0.25	0.1	2.5548(0.1166)	2.4192(0.1815)	8.979(0.3933)	7.7491(0.4788)
	0.25	2.5351(0.1314)	2.3972(0.1177)	9.0665(0.4166)	7.8077(0.2998)
	1	2.6534(0.1397)	2.4572(0.1761)	9.3876(0.466)	7.8991(0.4994)
	4	3.188(0.4022)	5.2706(0.5627)	10.6842(1.3126)	17.1432(1.6608)
0.5	0.1	1.2543(0.0693)	1.2062(0.0901)	4.0387(0.168)	3.8691(0.1702)
	0.25	1.2325(0.0785)	1.2002(0.0934)	3.9957(0.1707)	3.7744(0.1996)
	1	1.3004(0.0795)	1.2341(0.094)	4.1842(0.219)	3.9139(0.1827)
	4	1.6605(0.2025)	3.2807(0.342)	5.1548(0.5732)	9.9503(0.9592)
1	0.1	0.8006(0.0312)	0.8462(0.0357)	2.566(0.0517)	2.6626(0.0609)
	0.25	0.7321(0.028)	0.7508(0.0303)	2.413(0.0369)	2.4435(0.0398)
	1	0.7623(0.0233)	0.7929(0.0339)	2.5553(0.0477)	2.6226(0.0431)
	4	1.1566(0.0769)	1.4052(0.2078)	3.5698(0.1661)	4.1448(0.4616)
2	0.1	0.8013(0.0311)	0.8089(0.0448)	2.3509(0.0553)	2.3721(0.0966)
	0.25	0.7229(0.0226)	0.7406(0.0228)	2.1873(0.0437)	2.2261(0.0376)
	1	0.7416(0.0216)	0.7607(0.0239)	2.2716(0.0355)	2.3178(0.0497)
	4	0.9755(0.0448)	0.9235(0.0622)	3.1362(0.1101)	2.8834(0.09)
		Model 2			
0.1	0.1	7.4786(0.2101)	8.8921(0.6767)	31.8289(0.758)	31.3341(1.6403)
	0.25	7.5204(0.2372)	8.7997(0.6127)	31.9076(0.8567)	31.1166(1.5828)
	1	7.5323(0.2538)	8.9481(0.5451)	31.6324(0.8335)	30.8231(1.2818)
	4	6.8425(0.7066)	5.0507(0.558)	27.053(2.3511)	14.9335(0.8595)
0.25	0.1	2.5422(0.1049)	2.556(0.1656)	9.3173(0.346)	8.5754(0.3512)
	0.25	2.547(0.1256)	2.5617(0.1712)	9.3313(0.3673)	8.5574(0.3953)
	1	2.5592(0.0984)	2.5006(0.1345)	9.3329(0.3435)	8.3611(0.3485)
	4	2.742(0.2182)	5.7981(1.9469)	9.7537(0.7667)	18.8084(6.1082)
0.5	0.1	1.2162(0.0715)	1.2155(0.1073)	3.9003(0.1677)	3.6771(0.2334)
	0.25	1.1843(0.0578)	1.1566(0.0951)	3.8548(0.1488)	3.5876(0.2491)
	1	1.2162(0.0642)	1.2117(0.0894)	3.8981(0.1728)	3.6912(0.1548)
	4	1.6245(0.1341)	3.2028(0.4083)	5.4452(0.3647)	9.8753(1.1542)
1	0.1	0.5633(0.0356)	0.5491(0.0405)	1.8395(0.0732)	1.7226(0.0608)
	0.25	0.569(0.0312)	0.5603(0.0344)	1.8464(0.0588)	1.7451(0.0502)
	1	0.6064(0.028)	0.6101(0.0391)	1.977(0.0494)	1.9278(0.0699)
	4	1.0828(0.0455)	1.476(0.0679)	3.7852(0.0725)	4.2703(0.1108)
2	0.1	0.4888(0.0214)	0.4786(0.0231)	1.4671(0.0301)	1.3723(0.0333)
	0.25	0.5153(0.0267)	0.5095(0.0286)	1.5084(0.0297)	1.424(0.0317)
	1	0.5617(0.0217)	0.5587(0.0287)	1.664(0.0333)	1.6094(0.0418)
	4	0.9993(0.0238)	0.8554(0.0532)	3.462(0.0354)	2.4739(0.0831)

Table 5.4: Comparison of average (SD) matrix losses for $p = 100$

		SCLIME	Joint	SCLIME	Joint
α	ρ	Spectral norm		Frobenius norm	
		Model 1			
0.1	0.1	8.2916(0.8387)	8.0836(0.6491)	49.9714(4.582)	36.7674(2.9551)
	0.25	8.7608(0.4786)	8.7237(0.6116)	53.5359(2.4532)	41.1232(2.0235)
	1	8.066(0.4542)	8.3149(0.5366)	49.879(2.47)	38.1606(1.8042)
	4	7.9247(0.5105)	7.7772(0.5492)	48.9273(2.6849)	35.622(1.6146)
0.25	0.1	2.7247(0.1433)	2.4331(0.1614)	13.3898(0.5393)	10.5234(0.5714)
	0.25	2.7988(0.1287)	2.5451(0.1794)	14.2116(0.5102)	11.1625(0.483)
	1	2.7569(0.1043)	2.425(0.1601)	13.9332(0.5264)	10.464(0.4586)
	4	2.6299(0.1222)	2.2851(0.1379)	13.1427(0.6099)	9.7812(0.2747)
0.5	0.1	1.315(0.0857)	3.391(1.1075)	5.9231(0.2441)	14.9783(4.814)
	0.25	1.3608(0.0577)	1.9254(1.485)	6.1493(0.1972)	8.4979(6.7995)
	1	1.3513(0.0658)	2.8675(1.7255)	6.0731(0.2077)	12.8117(7.8052)
	4	1.3435(0.0509)	3.7323(1.7268)	6.0848(0.1834)	16.5892(7.5335)
1	0.1	0.8857(0.0307)	2.2098(0.4505)	3.9214(0.0473)	9.0145(1.8508)
	0.25	0.7856(0.0246)	0.9207(0.3529)	3.6088(0.0429)	3.9731(1.4808)
	1	0.8319(0.0221)	2.4453(0.5293)	3.8161(0.0366)	10.1235(2.0885)
	4	0.9099(0.0188)	2.3269(0.6403)	4.322(0.0239)	9.6357(2.4321)
2	0.1	0.9118(0.0226)	1.5141(0.126)	3.7522(0.0757)	5.8817(0.4262)
	0.25	0.8097(0.0188)	0.8429(0.0232)	3.3759(0.0461)	3.4847(0.0652)
	1	0.853(0.0157)	0.8918(0.0177)	3.5789(0.0435)	3.6985(0.0325)
	4	0.9353(0.0156)	1.0013(0.0229)	4.1351(0.0458)	4.3952(0.0872)
		Model 2			
0.1	0.1	9.2967(0.4369)	9.4896(0.6373)	56.4275(2.3172)	45.4309(2.2526)
	0.25	9.2136(0.4281)	9.3945(0.6509)	56.327(2.2826)	44.9379(2.2564)
	1	8.405(0.4129)	8.7049(0.4422)	51.818(2.3159)	40.0793(1.9151)
	4	8.2434(0.6028)	8.2042(0.4948)	50.9657(3.2031)	38.2577(1.8575)
0.25	0.1	2.8867(0.1134)	2.6877(0.156)	15.0803(0.5158)	12.0609(0.2124)
	0.25	2.8446(0.1146)	2.6645(0.1701)	14.9087(0.5353)	11.909(0.3842)
	1	2.8033(0.1292)	2.5589(0.1701)	14.3192(0.5911)	11.1749(0.5875)
	4	2.7552(0.1266)	2.4335(0.1542)	13.8253(0.6054)	10.4867(0.4645)
0.5	0.1	1.3967(0.0565)	1.2963(0.0643)	6.2999(0.2315)	5.2788(0.1328)
	0.25	1.3711(0.0562)	1.2568(0.0613)	6.2353(0.2244)	5.2397(0.1216)
	1	1.3498(0.059)	3.0057(1.8307)	6.0388(0.2058)	13.2595(8.3516)
	4	1.3634(0.0536)	3.1202(1.8731)	6.1103(0.2141)	13.9106(8.3962)
1	0.1	0.6559(0.0295)	0.6501(0.0317)	2.9758(0.0736)	2.8707(0.0518)
	0.25	0.6643(0.0237)	0.6677(0.0206)	3.0494(0.0585)	2.9807(0.0436)
	1	0.7278(0.0194)	1.8967(0.9179)	3.3172(0.0423)	7.9415(3.7564)
	4	0.8353(0.0211)	1.6304(0.8485)	3.9022(0.0344)	6.9268(3.3964)
2	0.1	0.6126(0.0202)	0.613(0.0188)	2.541(0.0323)	2.4848(0.0301)
	0.25	0.6438(0.0139)	0.6538(0.014)	2.6622(0.0346)	2.6359(0.0341)
	1	0.7223(0.0232)	0.7362(0.0275)	3.01(0.055)	3.0209(0.0804)
	4	0.8619(0.0204)	0.9136(0.0181)	3.6323(0.0427)	3.8003(0.0287)

Table 5.5: Top 10 hubs for direct connectivity of Groups $1-3$ found by Joint

Rank	Group 1	Group 2	Group 3
1	21	10	38
2	10	19	16
3	5	20	32
4	20	1	3
5	18	15	28
6	4	18	18
7	1	21	36
8	17	38	12
9	6	32	21
10	9	4	22

Table 5.6: Top 10 hubs for direct connectivity of Group 1 found by Joint

Joint Rank	Order in 50	Joint Degree	SCLIME rank	SCLIME degree
1	21	35	2	47
2	10	32	3	46
3	5	26	4	45
4	20	25	1	47
5	18	24	10	44
6	4	23	8	44
7	1	22	7	44
8	17	21	19	43
9	6	20	47	7
10	9	20	9	44

The top 10 hubs by Joint are reported in Table 5.5. We can see that the six time series (1, 4, 10, 18, 20, and 21) are found in their top 10 hubs of Group 1 and Group 2. This implies that Joint works well, and finds some common information between Group 1 and Group 2.

Table 5.6 lists top 10 hubs of Group 1 found by Joint and their degrees of connectivity. Also, they are compared with the results from SCLIME. It can be seen that top 10 hubs of Group 1 by Joint has 8 overlaps with those of SCLIME: time series 1, 4, 5, 9,10, 18, 20, and 21. It implies that the results of Joint are very similar to those of CLIME for this subject.

Chapter 6

Future work

Since the current R-package *fastclime* is not good at solving very high dimensional linear problems and the True-positive ratio is very low in the simulation study, we plan to conduct more simulation studies and real data analysis with high dimension. We are particularly interested in fMRI group precision matrices analysis, because the inference and visualization of brain network using a graphical model is widely used in fMRI analysis. The proposed methods of joint;y estimating individual and group precision matrices can provide a representative and robust estimation of brain networks.

Bibliography

[1] Athreya, K.B. and Lahiri, S.N. (2006). *Measure Theory and Probability Theory.* Springer, New York.

[2] Bai, J. and Ng, S. (2005). Tests for skewness, kurtosis, and normality for time series data. *J. Bus. Econom. Statist.* 23 49-60.

[3] Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* 29, 1165-1188.

[4] Bennett, G. (1977) An extension of the Riesz-Thorin theorem. In: Baker J., Cleaver C., Diestel J. (eds) *Banach Spaces of Analytic Functions*. Lecture Notes in Mathematics, vol 604. Springer, Berlin, Heidelberg.

[5] Bhattacharjee, M. and Bose, A. (2014). Consistency of large dimensional sample covariance matrix under weak dependence. *Stat. Methodol.* 20 11-26.

[6] Bhattacharjee, M. and Bose, A. (2014). Estimation of autocovariance matrices for infinite dimensional vector linear process. *J. Time Series Anal.* 35 262-281.

[7] Bochud, T. and Challet, D. (2007). Optimal approximations of power laws with exponentials: Application to volatility models with long memory. *Quant. Finance* 7 585-589.

[8] Boyd, S. and Vandenberghe L. (2004). *Convex Optimization*. Cambridge University Press, Cambridge.

[9] Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer, New York.

[10] Bullmore, E., Fadili, J., Breakspear, M., Salvador, R., Suckling, J., and Brammer, M. (2003). Wavelets and statistical analysis of functional magnetic resonance images of the human brain. *Stat. Methods Med. Res.* 12 375–399.

[11] Cai, T., Liu, W. and Luo, X. (2011). A constrained ℓ_1 minimization approach to sparse precision matrix estimation. *J. Amer. Statist. Assoc.* 106 594-607.

[12] Cai, T.T. and Zhou, H.H. (2012). Optimal rates of convergence for sparse covariance matrix estimation. *Ann. Statist.* 40 2389-2420.

[13] Chen, X., Xu, M. and Wu, W.B. (2013). Covariance and precision matrix estimation for high-dimensional time series. *Ann. Statist.* 41, 2994-3021.

[14] Chen, X., Xu, M. and Wu, W.B. (2016). Regularized estimation of linear functionals of precision matrices for high-dimensional time series. *IEEE Trans. Signal Process.* 64 6459-6470.

[15] Chlebus, E. (2009). An approximate formula for a partial sum of the divergent p-series. *Appl. Math. Lett.* 22 732–737.

[16] Chung, J., Park, C., McDowell, J.E., Joint Estimation and Regularized Aggregation of Brain Network in fMRI Data, *JSM 2019 Online Program* preprint.

[17] Ciuciu, P., Abry, P. and He, B.J. (2014). Interplay between functional connectivity and scale-free dynamics in intrinsic fMRI networks. *NeuroImage* 95 248-263.

[18] Ding, C., Zhou, D., He, X., Zha, H. (2006). R1-PCA: Rotational Invariant L1-norm Principal Component Analysis for Robust Subspace Factorization. *Proceedings of the 23rd International Conference on Machine Learning*. ICML '06. Pittsburgh, Pennsylvania, USA: ACM. pp. 281-288.

[19] Evgeniou, T. and Pontil, M. (2004). Regularized multitask learning. In *Proceed-ings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 109-117, Seattle, Washington.

[20] Fang, Y., Wang, B. and Feng, Y. (2016). Tuning-parameter selection in regularized estimations of large covariance matrices. *J. Stat. Comput. Simul.* 86 494-509.

[21] Golub, G.H. and Loan, C.F.V., *Matrix Computations* 4th edition, The Johns Hopkins University Press.

[22] Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *J. Time Series Anal.* 4 221-238.

[23] Harrison, L., Penny, W. D. and Friston, K. (2003). Multivariate autoregressive modeling of fMRI time series. *Neuroimage* 19 1477-1491.

[24] Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, New York.

[25] He, B.J. (2011). Scale-free properties of the functional magnetic resonance imaging signal during rest and task. *J. Neurosci.* 31 13786-13795.

[26] Hinich, M. J. (1982). Testing for Gaussianity and linearity of a stationary time series. *J. Time Series Anal.* 3 169-176.

[27] Lee, W., Liu, Y.F. (2015). Joint Estimation of Multiple Precision Matrices with Common Structures, *J. Mach. Learn. Res.* 16 (2015) 1035-1062.

[28] Li, X., Zhao T., Yuan, X. and Liu, H. (2015). The flare package for high dimensional linear regression and precision matrix estimation in R. *J. Mach. Learn. Res.* 16 553-557.

[29] Liu, H., Aue, A. and Paul, D. (2015). On the Marčenko-Pastur law for linear time series. *Ann. Statist.* 43 675-712.

[30] Palma, W. (2007). *Long-Memory Time Series: Theory and Methods.* Wiley-Interscience, Hoboken, NJ.

[31] Pang, H., Liu, H. and Vanderbei R. (2014). *fastclime: A fast solver for parameterized lp problems and constrained l_1-minimization approach to sparse precision matrix estimation.* R package version 1.2.4.

[32] Park, C., Lazar, N., Ahn, J., and Sornborger, A. (2010). A multiscale analysis of the temporal characteristics of resting-state fMRI data *J. Neurosci. Methods.* 193 334-342.

[33] Priestely, M. B. and Subba Rao, T. (1969). A test for non-stationarity of time-series. *J. Roy. Statist. Soc. Ser. B* 31 140-149.

[34] Rencher, A.C. and Schaalje, G.B. (2008), *Linear Models in Statistics*, 2nd edition, Wiley.

[35] Shu, H. and Nan, B. (2019). Estimation of large covariance and precision matrices from temporally dependent observations, *Ann. Statist.* 47 1321-1350.

[36] Shu, H. and Nan, B. (2019). Supplement to "Estimation of large covariance and precision matrices from temporally dependent observations." DOI:10.1214/18-AOS1716SUPP.

[37] Sripada, C., Angstadt, M., Kessler, D., Phan, K.L., Liberzon, I., Evans, G.W., Welsh, R.C., Kim, P. and Swain, J.E. (2014). Volitional regulation of emotions produces distributed alterations in connectivity between visual, attention control, and default networks. *NeuroImage* 89 110-121.

[38] Tagliazucchi, E., Von Wegner, F., Morzelewski, A., Brodbeck, V., Jahnke, K. and Laufs, H. (2013). Breakdown of long-range temporal dependence in default mode and attention networks during deep sleep. *Proc. Natl. Acad. Sci. USA* 110 15419-15424.