Robust Estimation for the Mean and Covariance Matrix for High Dimensional Time Series

Danna Zhang

Department of Mathematics, University of California, San Diego

Supplementary Material

In this Supplementary Material, we provide the technical proofs for all results presented in the main body of the paper.

S1 Proof of Theorem 2.1

We first establish two lemmas (cf. Lemma S1.1 and Lemma S1.2) that are very useful in the proof of Theorem 2.1.

Lemma S1.1. Let $Z_i = f(\varepsilon_i, \varepsilon_{i-1}, \ldots)$, $i \in \mathbb{Z}$, be real-valued random variables, where ε_i are independent random elements. Assume $|Z_i| \le M$ for all i. Let ε_i' be an i.i.d. copy of ε_i. Let $T_i = \sum_{j=1}^{i} Z_j$. For $j_1, j_2 \in \mathbb{Z}$ with $j_1 \ge j_2$, define

$$Z_{i,\{j_1,j_2\}} = f(\varepsilon_i, \ldots, \varepsilon_{j_1+1}, \varepsilon_{j_1}', \ldots, \varepsilon_{j_2}', \varepsilon_{j_2-1}, \ldots). \qquad \text{(S1.1)}$$

For any $a > 0$, we have

$$\mathbb{E}\exp(aT_n) \le \prod_{i=1}^{n} \mathbb{E}\exp(aZ_i) + a\exp(aMn)\sum_{i=2}^{n} \mathbb{E}|Z_i - Z_{i,\{i-1,-\infty\}}|, \quad \text{(S1.2)}$$

and

$$\prod_{i=1}^{n} \mathbb{E}\exp(aZ_i) \le \mathbb{E}\exp(aT_n) + a\exp(aMn)\sum_{i=2}^{n} \mathbb{E}|Z_i - Z_{i,\{i-1,-\infty\}}|. \quad \text{(S1.3)}$$

Proof of Lemma S1.1. We can write

$$\begin{aligned}\mathbb{E}\exp(aT_n) - \prod_{i=1}^{n} \mathbb{E}\exp(aZ_i) &= \sum_{i=2}^{n} \big[\mathbb{E}\exp(aT_i) - \mathbb{E}\exp(aT_{i-1})\mathbb{E}\exp(aZ_i)\big] \\ &\quad \cdot \prod_{j=i+1}^{n} \mathbb{E}\exp(aZ_j).\end{aligned}$$

Observe that $Z_{i,\{i-1,-\infty\}}$ has the same distribution as Z_i. Also notice that $Z_{i,\{i-1,-\infty\}}$ and T_{i-1} are independent. We have

$$\begin{aligned} & \mathbb{E}\exp(aT_n) - \prod_{i=1}^{n} \mathbb{E}\exp(aZ_i) \\ = & \sum_{i=2}^{n} \mathbb{E}\big[\exp(aT_{i-1})(\exp(aZ_i) - \exp(aZ_{i,\{i-1,-\infty\}}))\big] \prod_{j=i+1}^{n} \mathbb{E}\exp(aZ_j) \\ \le & \sum_{i=2}^{n} \exp\{aM(n-1)\}\mathbb{E}|\exp(aZ_i) - \exp(aZ_{i,\{i-1,-\infty\}})|.\end{aligned}$$

Hence, (S1.2) follows in view of the fact that

$$\mathbb{E}|\exp(aZ_i) - \exp(aZ_{i,\{i-1,-\infty\}})| \le e^{aM} a\mathbb{E}|Z_i - Z_{i,\{i-1,-\infty\}}|.$$

And (S1.3) can be derived similarly. □

We embed the process X_i, $i \in \mathbb{Z}$, into a continuous time process by defining $X_t = X_{\lceil t \rceil}$ for $t \in \mathbb{R}$. For notational simplicity, we abbreviate

$\delta_{i,2}$ defined in (2.6) when $q = 2$ to δ_i. Then We also embed the index set of dependence measure δ_i defined in (2.6) into continuous time by letting $\delta_t = \delta_{\lceil t \rceil}$, $t \in \mathbb{R}$. For a Borel set K, define

$$S_K = \int_K X_t dt.$$

Lemma S1.2. For $x \geq 0$ and $B \geq 2$, let $K_B \in (x, x+B]$ be a finite union of intervals. Assume $\mathbb{E}X_i = 0$ and $|X_i| \leq M$ for all i. Let $c_1 = (e^4 - 5)/4$. Then for any $t > 0$ such that $tM \leq 1 \wedge \sqrt{\log(\rho^{-1})/(2B)}$, we have

$$\log \mathbb{E} \exp(tS_{K_B}) \leq c_1 B \|X_.\|_2^2 t^2 + \frac{\|X_.\|_2 BMt^2}{\rho(1-\rho)} \rho^{(2tM)^{-1}}.$$

Proof of Lemma S1.2. If $BMt \leq 4$, then we have $tS_{K_B} \leq 4$. Notice that the function $x \mapsto x^{-2}(e^x - x - 1)$ is increasing. We can obtain

$$\exp(tS_{K_B}) \leq 1 + tS_{K_B} + \frac{c_1}{4} t^2 S_{K_B}^2.$$

Since X_i is stationary and $\mathbb{E}X_i = 0$, it follows that $\mathbb{E}S_{K_B} = 0$ and

$$\mathbb{E}S_{K_B}^2 \leq B \sum_{k=-\infty}^{\infty} |\mathrm{Cov}(X_0, X_k)|.$$

By (2.11), we have

$$\mathbb{E} \exp(tS_{K_B}) \leq 1 + \frac{c_1}{2} B \|X_.\|_2^2 t^2. \tag{S1.4}$$

For $BMt > 4$, let $L = \lfloor BMt/2 \rfloor$. Divide the interval $(x, x+B]$ into $2L$ consecutive left-open and right-closed intervals of length $B/(2L)$, denoted by $I_1, \ldots, I_{2L}$. For $1 \le j \le L$, let $Z_j^o = S_{K_B \cap I_{2j-1}}$ and $Z_j^e = S_{K_B \cap I_{2j}}$. Define

$$S^o = \sum_{j=1}^{L} Z_j^o \quad \text{and} \quad S^e = \sum_{j=1}^{L} Z_j^e.$$

By the Cauchy-Schwarz inequality, we have

$$2 \log \mathbb{E} \exp(t S_{K_B}) \le \log \mathbb{E} \exp(2tS^o) + \log \mathbb{E} \exp(2tS^e).$$

For $1 \le j \le L$, denote $I_{2j-1} = (\ell_{2j-1}, u_{2j-1}]$. Observe that Z_j^o is measurable with respect to the σ-field of $\{\varepsilon_i : i \le \lceil u_{2j-1} \rceil\}$ and $B/(2L) \ge 1/(Mt) \ge 1$. By Lemma S1.1,

$$\mathbb{E} \exp(2tS^o) \le \prod_{j=1}^{L} \mathbb{E} \exp(2tZ_j^o) + 2t \exp(BMt) \sum_{j=2}^{L} \mathbb{E} |Z_j^o - Z^o_{j,\{\lceil u_{2j-3} \rceil, -\infty\}}|, \tag{S1.5}$$

where

$$Z^o_{j,\{\lceil u_{2j-3} \rceil, -\infty\}} = \int_{K_B \cap I_{2j-1}} X_{t,\{\lceil u_{2j-3} \rceil, -\infty\}} dt$$

with $X_{t,\{\lceil u_{2j-3} \rceil, -\infty\}} = X_{\lceil t \rceil,\{\lceil u_{2j-3} \rceil, -\infty\}}$. For $j \ge 2$, we have

$$\begin{aligned}
& |Z_j^o - Z^o_{j,\{\lceil u_{2j-3} \rceil, -\infty\}}| \\
\le & \int_{K_B \cap I_{2j-1}} |X_t - X_{t,\{\lceil u_{2j-3} \rceil, -\infty\}}| dt \\
\le & \int_{K_B \cap I_{2j-1}} |X_t - X_{t,\{\lceil u_{2j-3} \rceil, \lceil u_{2j-3} \rceil\}}| dt \\
& + \int_{K_B \cap I_{2j-1}} \sum_{m=0}^{\infty} |X_{t,\{\lceil u_{2j-3} \rceil, \lceil u_{2j-3} \rceil - m\}} - X_{t,\{\lceil u_{2j-3} \rceil, \lceil u_{2j-3} \rceil - m - 1\}}| dt.
\end{aligned}$$

Recall that for $t \in \mathbb{R}$, $\delta_t = \delta_{\lceil t \rceil}$. By the fact that $\mathbb{E}|X_t - X_{t,\{j,j\}}| \le \delta_{t-j}$ and $\mathbb{E}|X_{t,\{j,j'\}} - X_{t,\{j,j'-1\}}| \le \delta_{t-j'+1}$ for $t \in \mathbb{R}$ and $j, j' \in \mathbb{Z}$ with $j \ge j'$. Then we have

$$\begin{aligned}
&\mathbb{E}|Z_j^o - Z_{j,\{\lceil u_{2j-3}\rceil,-\infty\}}^o| \\
\leq\ & \int_{K_B\cap I_{2j-1}} \delta_{t-\lceil u_{2j-3}\rceil} dt + \int_{K_B\cap I_{2j-1}} \sum_{m=0}^{\infty} \delta_{t-\lceil u_{2j-3}\rceil+m+1} dt \\
=\ & \int_{K_B\cap I_{2j-1}} \sum_{m=0}^{\infty} \delta_{t-\lceil u_{2j-3}\rceil+m} dt \\
\leq\ & \|X_\cdot\|_2 \sum_{i=\lceil \ell_{2j-1}^+\rceil}^{\lceil u_{2j-1}\rceil} \rho^{i-\lceil u_{2j-3}\rceil} \leq (1-\rho)^{-1}\|X_\cdot\|_2 \rho^{\lceil \ell_{2j-1}^+\rceil - \lceil u_{2j-3}\rceil}. \quad \text{(S1.6)}
\end{aligned}$$

Since $\lceil \ell_{2j-1}^+\rceil - \lceil u_{2j-3}\rceil \geq B/(2L) - 1$, by (S1.5) and (S1.6), it follows that

$$\begin{aligned}
\mathbb{E}\exp(2tS^o) &\leq \prod_{j=1}^{L} \mathbb{E}\exp(2tZ_j^o) + \frac{\|X_\cdot\|_2 BMt^2}{\rho(1-\rho)} \exp(BMt)\rho^{B/(2L)} \\
&\leq \prod_{j=1}^{L} \mathbb{E}\exp(2tZ_j^o) + \frac{\|X_\cdot\|_2 BMt^2}{\rho(1-\rho)} \rho^{(2tM)^{-1}}, \quad \text{(S1.7)}
\end{aligned}$$

where the last step follows in view of the fact that for $0 < tM \leq \sqrt{\log(\rho^{-1})/(2B)}$,

$$\exp(BMt)\rho^{B/(2L)} \leq \exp(BMt)\rho^{(Mt)^{-1}} \leq \rho^{(2tM)^{-1}}.$$

By applying (S1.3) in Lemma S1.1, we can also obtain

$$\prod_{j=1}^{L} \mathbb{E}\exp(2tZ_j^o) \leq \mathbb{E}\exp(2tS^o) + \frac{\|X_\cdot\|_2 BMt^2}{\rho(1-\rho)} \rho^{(2tM)^{-1}}. \quad \text{(S1.8)}$$

Since $\mathbb{E}X_i = 0$, by Jensen's inequality, we have $\mathbb{E}\exp(2tS^o) \geq 1$ and $\mathbb{E}\exp(2tZ_j^o) \geq 1$ for each j. By the inequality $|\log x - \log y| \leq |x-y|$ for $x, y \geq 1$, (S1.7) and (S1.8), we have

$$\log \mathbb{E}\exp(2tS^o) \leq \sum_{j=1}^{L} \log \mathbb{E}\exp(2tZ_j^o) + \frac{\|X_\cdot\|_2 BMt^2}{\rho(1-\rho)} \rho^{(2tM)^{-1}}.$$

Notice that $2tZ_j^o \leq 2tMB/(2L)$ and $L = \lfloor MBt/2 \rfloor$. We have $2tZ_j^o \leq 4$. Similarly as (S1.4), we have

$$\sum_{j=1}^{L} \log \mathbb{E} \exp(2tZ_j^o) \leq \sum_{j=1}^{L} \log(1 + c_1 B \|X_{\cdot}\|_2^2 t^2 / L) \leq c_1 B \|X_{\cdot}\|_2^2 t^2.$$

Therefore, we can obtain (S1.2) by dealing with $\log \mathbb{E} \exp(2tS^e)$ similarly.

$\square$

Equipped with Lemma S1.1 and Lemma S1.2, Lemma S1.3 below can be proved without extra technical difficulties following the proof of Lemma 10 in Merlevède et al. (2009), which combines the idea of Bernstein big and small type argument with a twist, diadic recurrence and Cantor set construction. The proof is thus omitted here.

Lemma S1.3. Assume $\mathbb{E}X_i = 0$ and $|X_i| \leq M$ for all i. Let

$$c = \frac{\log(\rho^{-1})}{8} \wedge \sqrt{\frac{(\log 2)\log(\rho^{-1})}{4}}.$$

For every $A \geq 4 \vee (\log(\rho^{-1})/2)$, there exists a subset K_A of $(0, A]$, with Lebesgue measure larger than $A/2$ such that for any $t > 0$ such that $tM \leq c/\log A$, we have

$$\log \mathbb{E} \exp(tS_{K_A}) \leq c_1 A \|X_{\cdot}\|_2^2 t^2 + 8c_2 \|X_{\cdot}\|_2 A^{-1} M t^2. \tag{S1.9}$$

where $c_1 = (e^4 - 5)/4$ and $c_2 = [\rho(1-\rho)\log(\rho^{-1})]^{-1}$. Moreover, for any

$t > 0$ such that $tM \le 1 \wedge (\log(\rho^{-1})/2)$, we have

$$\log \mathbb{E} \exp(tS_{(0,A]}) \le c_1 A \|X.\|_2^2 t^2 + 4c_2 \|X.\|_2 AMt^2 + 2M^2 t^2 A \log A.$$

Corollary S1.4. Assume $\mathbb{E}X_i = 0$ and $|X_i| \le M$ for all i. Let c, c_1, c_2 be defined the same as in Lemma S1.3 and let $C = \max\{c_1, 8c_2\}$. For every $A \ge 4 \vee (\log(\rho^{-1})/2)$, there exists a subset K_A of $(0, A]$, with Lebesgue measure larger than $A/2$ such that for any $t > 0$ such that $tM < c/\log A$, we have

$$\log \mathbb{E} \exp(tS_{K_A}) \le \frac{CAt^2(\|X.\|_2 + \sqrt{\|X.\|_2 M}/A)^2}{1 - tM \log A/c}.$$

Moreover, for any $t > 0$ such that $tM < 1 \wedge (\log(\rho^{-1})/2)$, we have

$$\log \mathbb{E} \exp(tS_{(0,A]}) \le \frac{CAt^2 \log A(\|X.\|_2 + M)^2}{1 - tM/(1 \wedge (\log(\rho^{-1})/2))}.$$

Corollary S1.4 follows directly from Lemma S1.3. Now we are ready to prove Theorem 2.1.

Proof of Theorem 2.1. The proof follows the same spirit of the proof of Theorem 1 in Merlevède et al. (2009). Adopting the framework of functional dependence measure, we shall establish some useful moment inequalities (cf. Lemma S1.1 – S1.3) characterized by the parameter ρ and dependence adjusted moment $\|X.\|_2$. In our proof, we apply the newly established Lemma S1.1, Lemma S1.2, Lemma S1.3 and Corollary S1.4 in place of respectively Lemma 15, Lemma 8, Lemma 10 and Corollary 11 in the paper

Merlevède et al. (2009). We omit the complete derivation of the proof, but point out the main differences on the choices of quantities adapted to our setting. Using the same notations in the proof of Theorem 1 in their paper, we take

$$L = L_n = \inf\{j \in \mathbb{N} : A_j \leq 4 \vee (\log(\rho^{-1})/2)\}.$$

Hence, the inequality (4.21) in that paper becomes

$$L_n \leq \Big\lfloor \frac{\log(n) - \log(4 \vee (\log(\rho^{-1})/2))}{\log 2} \Big\rfloor + 1.$$

As one of the key steps, to apply Lemma 3 in Merlevède et al. (2011), for $0 \leq i \leq L-1$, we take

$$\kappa_i = M \log(n/2^j)/c, \quad \sigma_i = \sqrt{C}[(n/2^j)^{1/2}\|X.\|_2 + (n/2^j)^{-1/2}\sqrt{\|X.\|_2 M}],$$

where c is defined the same as in Lemma S1.3 and $C = \max\{(e^4-5)/4, [\rho(1-\rho)\log(\rho^{-1})]^{-1}\}$. And we also take

$$\kappa_L = M(1 \vee (\log(\rho^{-1})/8)), \quad \sigma_L = \sqrt{C}M(4 \vee (\log(\rho^{-1})/2)).$$

Then it follows that

$$\sum_{i=0}^{L} \kappa_i \leq M\Big[\frac{(\log n)^2}{c \log 2} + 1 \vee (\log(\rho^{-1})/8)\Big], \tag{S1.10}$$

and

$$\sum_{i=0}^{L} \sigma_i \quad \leq \quad \sqrt{C}[4\|X.\|_2\sqrt{n} + 2\sqrt{\|X.\|_2 M} + M(4 \vee (\log(\rho^{-1})/2))]$$

$$\leq \quad \sqrt{C_1/2}(\sqrt{n}\|X.\|_2 + M). \tag{S1.11}$$

Using the bounds (S1.10) and (S1.11) in place accordingly, we can obtain (2.8) without extra technical difficulties. The Bernstein-type inequality (2.9) follows by the Markov equality

$$\mathbb{P}(S_n \geq x) \leq \frac{\mathbb{E}\exp(tS_n)}{\exp(tx)} \leq \exp\left\{\frac{C_1 t^2(n\|X.\|_2^2 + M^2)}{1 - C_2 tM(\log n)^2} - tx\right\}$$

and letting $t = x/[2C_1(n\|X.\|_2^2 + M^2) + C_2M(\log n)^2 x]$. □

S2 Proof of Results in Section 3

Proof of Theorem 3.1. Let $S_{nj}(\theta) = \sum_{i=1}^n \varphi_\kappa(\mathbf{x}_{ij} - \theta)$ and

$$R_{nj}(\theta) = \sum_{i=1}^n [\varphi_\kappa(\mathbf{x}_{ij} - \theta) - \mathbb{E}\varphi_\kappa(\mathbf{x}_{ij} - \theta)]. \tag{S2.12}$$

Denote $\tilde{\mathbf{x}}_{ij} = \mathbf{x}_{ij} - \mu_j$. Notice that the function $\theta \mapsto S_{nj}(\theta)$ is non-increasing and $\hat{\mu}_j^H$ is the solution to the equation $S_{nj}(\theta) = 0$. For $\Delta > 0$, we have

$$\mathbb{P}(\hat{\mu}_j^H - \mu_j \geq \Delta) \leq \mathbb{P}(S_{nj}(\mu_j + \Delta) \geq 0) = \mathbb{P}\Big(R_{nj}(\mu_j + \Delta) \geq -\sum_{i=1}^n \mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij} - \Delta)\Big). \tag{S2.13}$$

We first consider the term $-\sum_{i=1}^n \mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij} - \Delta)$ which can be written as

$$\sum_{i=1}^n [\mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij}) - \mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij} - \Delta) - \mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij})].$$

By elementary manipulation, we can obtain

$$\mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij}) - \mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij} - \Delta)$$

$$
\begin{aligned}
&= \int_{-\kappa}^{\kappa} \mathbb{P}(y \le \tilde{\mathbf{x}}_{ij} \le y + \Delta) dy \\
&= \Delta - \int_{\kappa}^{\kappa+\Delta} \mathbb{P}(\tilde{\mathbf{x}}_{ij} \ge y) dy - \int_{-\kappa}^{-\kappa+\Delta} \mathbb{P}(\tilde{\mathbf{x}}_{ij} \le y) dy.
\end{aligned}
$$

Under the condition $\kappa^{-1}\Delta \le 1/2$, which will be verified later, we have

$$
\begin{aligned}
& \int_{\kappa}^{\kappa+\Delta} \mathbb{P}(\tilde{\mathbf{x}}_{ij} \ge y) dy + \int_{-\kappa}^{-\kappa+\Delta} \mathbb{P}(\tilde{\mathbf{x}}_{ij} \le y) dy \\
\le & \int_{\kappa/2}^{\infty} \mathbb{E}\mathbf{1}\{|\tilde{\mathbf{x}}_{ij}| \ge y\} dy \le \int_{\kappa/2}^{\infty} \mathbb{E}\Big(\frac{|\tilde{\mathbf{x}}_{ij}|}{y}\Big)^2 dy = 2\kappa^{-1}\sigma_2^2. \quad \text{(S2.14)}
\end{aligned}
$$

Also, we have

$$
\begin{aligned}
|\mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij})| &\le |\mathbb{E}[(\tilde{\mathbf{x}}_{ij})\mathbf{1}\{|\tilde{\mathbf{x}}_{ij}| \le \kappa\}]| + \kappa|\mathbb{E}\mathbf{1}\{|\tilde{\mathbf{x}}_{ij}| > \kappa\}| \\
&\le \kappa^{-1}\mathbb{E}\Big(|\tilde{\mathbf{x}}_{ij}| \cdot \frac{|\tilde{\mathbf{x}}_{ij}|}{\kappa}\Big) + \kappa\mathbb{E}\Big(\frac{|\tilde{\mathbf{x}}_{ij}|}{\kappa}\Big)^2 \\
&\le 2\kappa^{-1}\sigma_2^2. \quad \text{(S2.15)}
\end{aligned}
$$

Then it follows that

$$
-\sum_{i=1}^{n} \mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij} - \Delta) \ge n(\Delta - 4\kappa^{-1}\sigma_2^2), \quad \text{(S2.16)}
$$

and by (S2.13),

$$
\mathbb{P}(\hat{\mu}_j^H - \mu_j \ge \Delta) \le \mathbb{P}(R_{nj}(\mu_j + \Delta) \ge n(\Delta - 4\kappa^{-1}\sigma_2^2)). \quad \text{(S2.17)}
$$

Letting $\theta = \mu_j + \Delta$, we shall apply Theorem 2.1 to $R_{nj}(\theta)$. By the Lipschitz continuity of the function φ_κ and the bound $|\varphi_\kappa(x)| \le \kappa$, we have, for all $\theta \in \mathbb{R}$,

$$
\|\varphi_\kappa(\mathbf{x}_{ij} - \theta) - \varphi_\kappa(\mathbf{x}_{ij,\{0\}} - \theta)\|_2 \le \|\mathbf{x}_{ij} - \mathbf{x}_{ij,\{0\}}\|_2 = \delta_{i,j,2}. \quad \text{(S2.18)}
$$

By Theorem 2.1, for $t > 0$,

$$\mathbb{P}(R_{nj}(\mu_j+\Delta) \geq t) \leq \exp\Big\{-\frac{t^2}{4C_1(n\|\mathbf{x}.\|_2^2+\kappa^2)+2C_2\kappa(\log n)^2 t}\Big\}. \quad \text{(S2.19)}$$

Let

$$t = \sqrt{C_1 n\|\mathbf{x}.\|_2^2\log(1/x)}+(\sqrt{C_1}+C_2)\kappa(\log n)^2\log(1/x), \text{ where } 0 < x \leq 1/e.$$

and $n(\Delta - 4\kappa^{-1}\sigma_2^2) = t$. By (S2.17) and (S2.19), it follows that

$$\mathbb{P}(\hat{\mu}_j^H - \mu_j \geq \Delta) \leq e^{-1/4}x.$$

Choosing κ as in (3.17), condition (3.18) implies $\kappa^{-1}\Delta \leq 1/2$ and we also have $\Delta \leq \Delta_n(x)$, where $\Delta_n(x) = (\mathcal{C}_1\sigma^*\log n+\mathcal{C}_2\|\mathbf{x}.\|_2)n^{-1/2}\sqrt{\log(1/x)}$. On the other hand, we can derive $\mathbb{P}(\hat{\mu}_j^H - \mu_j \leq -\Delta_n(x)) \leq e^{-1/4}x$ similarly. Then (3.19) follows. And (3.20) is an immediate consequence by letting $x = p^{-\tau-1}$ and using the Bonferroni inequality. □

Proof of Theorem 3.2. We follow the proof of Theorem 3.1. Now we only assume $(1+\epsilon)$-th finite moment. The bounds in (S2.14) and (S2.15) become

$$\begin{aligned} &\int_{\kappa}^{\kappa+\Delta}\mathbb{P}(\tilde{\mathbf{x}}_{ij}\geq y)dy+\int_{-\kappa}^{-\kappa+\Delta}\mathbb{P}(\tilde{\mathbf{x}}_{ij}\leq y)dy \\ \leq\ &\int_{\kappa/2}^{\infty}\mathbb{E}\mathbf{1}\{|\tilde{\mathbf{x}}_{ij}|\geq y\}dy \leq \int_{\kappa/2}^{\infty}\mathbb{E}\Big[\Big(\frac{|\tilde{\mathbf{x}}_{ij}|}{y}\Big)^{1+\epsilon}\Big]dy = \epsilon^{-1}\sigma_{1+\epsilon}^{1+\epsilon}\big(\frac{2}{\kappa}\big)^{\epsilon}. \end{aligned}$$

with $\kappa^{-1}\Delta \leq 1/2$ and

$$|\mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij})| \quad\leq\quad |\mathbb{E}[(\tilde{\mathbf{x}}_{ij})\mathbf{1}\{|\tilde{\mathbf{x}}_{ij}|\leq\kappa\}]|+\kappa|\mathbb{E}\mathbf{1}\{|\tilde{\mathbf{x}}_{ij}|>\kappa\}|$$

$$
\begin{aligned}
&\leq \kappa^{-\epsilon}\mathbb{E}\Big[|\tilde{\mathbf{x}}_{ij}|\cdot\Big(\frac{|\tilde{\mathbf{x}}_{ij}|}{\kappa}\Big)^{\epsilon}\Big]+\kappa\mathbb{E}\Big[\Big(\frac{|\tilde{\mathbf{x}}_{ij}|}{\kappa}\Big)^{1+\epsilon}\Big] \\
&\leq 2\kappa^{-\epsilon}\sigma_{1+\epsilon}^{1+\epsilon}.
\end{aligned}
$$

Then it follows that

$$
-\sum_{i=1}^{n}\mathbb{E}\varphi_\kappa(\tilde{\mathbf{x}}_{ij}-\Delta)\geq n\big[\Delta-(2+2^\epsilon/\epsilon)\kappa^{-\epsilon}\sigma_{1+\epsilon}^{1+\epsilon}\big].
$$

The dependence measure in (S2.18) becomes

$$
\begin{aligned}
\|\varphi_\kappa(\mathbf{x}_{ij}-\theta)-\varphi_\kappa(\mathbf{x}_{ij,\{0\}}-\theta)\|_2^2 &\leq \mathbb{E}\min\{|\mathbf{x}_{ij}-\mathbf{x}_{ij,\{0\}}|^2,4\kappa^2\} \\
&= 4\kappa^2\mathbb{E}\min\Big\{\Big|\frac{\mathbf{x}_{ij}-\mathbf{x}_{ij,\{0\}}}{2\kappa}\Big|^2,1\Big\} \\
&\leq 4\kappa^2\mathbb{E}\Big|\frac{\mathbf{x}_{ij}-\mathbf{x}_{ij,\{0\}}}{2\kappa}\Big|^{1+\epsilon}=(2\kappa)^{1-\epsilon}\delta_{i,1+\epsilon,j}^{1+\epsilon}.
\end{aligned}
$$

As a result,

$$
\sup_{m\geq 0}\rho^{-m}\sum_{i=m}^{\infty}\|\varphi_\kappa(\mathbf{x}_{ij}-\theta)-\varphi_\kappa(\mathbf{x}_{ij,\{0\}}-\theta)\|_2\leq(2\kappa)^{(1-\epsilon)/2}\|\mathbf{x}_\cdot^*\|_{1+\epsilon}.
$$

By Theorem 2.1, for $t>0$,

$$
\mathbb{P}(R_{nj}(\mu_j+\Delta)\geq t)\leq\exp\Big\{-\frac{t^2}{4C_1[n(2\kappa)^{1-\epsilon}\|\mathbf{x}_\cdot^*\|_{1+\epsilon}^2+\kappa^2]+2C_2\kappa(\log n)^2t}\Big\}.
$$

Let

$$
t=\sqrt{C_1n(2\kappa)^{1-\epsilon}\|\mathbf{x}_\cdot^*\|_{1+\epsilon}^2\log(1/x)}+(\sqrt{C_1}+C_2)\kappa(\log n)^2\log(1/x),
$$

where $0<x\leq 1/e$ and let $n\big[\Delta-(2+2^\epsilon/\epsilon)\kappa^{-\epsilon}\sigma_{1+\epsilon}^{1+\epsilon}\big]=t$. Choosing κ as in (3.21), $\kappa^{-1}\Delta<1/2$ is ensured by condition $(\tau+1)\mathcal{C}_3n^{-1}(\log n)^2\log p\leq 1/4$

and we have $\Delta \leq \Delta_n^*(x)$ with

$$\Delta_n^*(x) = 2K_\epsilon \left(\frac{\mathcal{C}_3 (\log n)^2 \log(1/x)}{n} \right)^{\frac{\epsilon}{1+\epsilon}}.$$

Hence, it follows that $\mathbb{P}(\hat{\mu}_j^H - \mu_j \geq \Delta_n^*(x)) \leq e^{-1/4} x$. The remaining is the same to the corresponding part in the proof of Theorem 3.1. $\square$

S3 Proof of Results in Section 4

Proof of Corollary 4.1. Let $x = p^{-\tau-2}$. Then

$$\Delta_n(p^{-\tau-2}) = \sqrt{\tau+2}(\mathcal{C}_1 \sigma^* \log n + \mathcal{C}_2 \|\mathbf{x}_\cdot\|_2) \sqrt{\frac{\log p}{n}}.$$

Notice that (3.18) satisfies with the assumption (4.24). By the Bonferroni inequality and Theorem 3.1,

$$\mathbb{P}\big(|\hat{\mu}^H - \mu|_\infty \geq \Delta_n(p^{-\tau-2})\big) \leq 2e^{-1/4} p^{-\tau-1}. \tag{S3.20}$$

Since

$$\max_{1 \leq j,k \leq p} |\hat{\mu}_j^H \hat{\mu}_k^H - \mu_j \mu_k| \leq 2\mu^o |\hat{\mu}^H - \mu|_\infty + |\hat{\mu}_j^H - \mu_j|_\infty^2,$$

by (S3.20), it follows that

$$\mathbb{P}\Big(\max_{1 \leq j,k \leq p} |\hat{\mu}_j^H \hat{\mu}_k^H - \mu_j \mu_k| \geq 2\mu^o \Delta_n(p^{-\tau-2}) + \Delta_n^2(p^{-\tau-2}) \Big) \leq 2e^{-1/4} p^{-\tau-1}. \tag{S3.21}$$

By the Lipschitz continuity of the function φ_κ, the triangle inequality and the Hölder inequality, we can compute the dependence measure of the process $\varphi_\kappa(\mathbf{x}_{ij}\mathbf{x}_{ik})$, $i \in \mathbb{Z}$, as

$$\begin{aligned}
& \|\varphi_\kappa(\mathbf{x}_{ij}\mathbf{x}_{ik}) - \varphi_\kappa(\mathbf{x}_{ij,\{0\}}\mathbf{x}_{ik,\{0\}})\|_2 \\
\leq \; & \|\mathbf{x}_{ij}\mathbf{x}_{ik} - \mathbf{x}_{ij,\{0\}}\mathbf{x}_{ik,\{0\}}\|_2 \\
\leq \; & \|\mathbf{x}_{ij}\|_4 \|\mathbf{x}_{ik} - \mathbf{x}_{ik,\{0\}}\|_4 + \|\mathbf{x}_{ik,\{0\}}\|_4 \|\mathbf{x}_{ij} - \mathbf{x}_{ij,\{0\}}\|_4 \\
= \; & \omega_4(\delta_{i,4,j} + \delta_{i,4,k}), \qquad \text{(S3.22)}
\end{aligned}$$

which implies

$$\sup_{m\geq 0} \rho^{-m} \sum_{i=m}^{\infty} \|\varphi_\kappa(\mathbf{x}_{ij}\mathbf{x}_{ik}) - \varphi_\kappa(\mathbf{x}_{ij,\{0\}}\mathbf{x}_{ik,\{0\}})\|_2 \leq 2\omega_4 \|\mathbf{x}_\cdot\|_4.$$

Let

$$\tilde{\Delta}_n(p^{-\tau-2}) = \sqrt{\tau+2}(\mathcal{C}_1 \omega^* \log n + \mathcal{C}_2 w_4 \|\mathbf{x}_\cdot\|_4) \sqrt{\frac{\log p}{n}}.$$

Applying Theorem 3.1 on the process $(\mathbf{x}_{ij}\mathbf{x}_{ik})_i$ and by the Bonferroni inequality again, we have

$$\mathbb{P}\Big(\max_{1\leq j,k\leq p} |\hat{\mu}^H_{jk} - \mu_{jk}| \geq \tilde{\Delta}_n(p^{-\tau-2}) \Big) \leq 2e^{-1/4} p^{-\tau}. \qquad \text{(S3.23)}$$

Then (4.25) follows from (S3.21) and (S3.23) in view of $p \geq 3$ and

$$|\hat{\Sigma}^H_{\mathbf{x}} - \Sigma_{\mathbf{x}}|_\infty \leq \max_{1\leq j,k\leq p} |\hat{\mu}^H_j \hat{\mu}^H_k - \mu_j\mu_k| + \max_{1\leq j,k\leq p} |\hat{\mu}^H_{jk} - \mu_{jk}|.$$

□

Proof of Corollary 4.2. The key step to compute the dependence measure of the process $\varphi_\kappa(\mathbf{x}_{ij}\mathbf{x}_{ik} - \theta)$, $i \in \mathbb{Z}$, is shown below. For any $\theta \in \mathbb{R}$,

$$\|\varphi_\kappa(\mathbf{x}_{ij}\mathbf{x}_{ik} - \theta) - \varphi_\kappa(\mathbf{x}_{ij,\{0\}}\mathbf{x}_{ik,\{0\}} - \theta)\|_2^2$$

$$
\begin{aligned}
&\leq \mathbb{E}\min\{|\mathbf{x}_{ij}\mathbf{x}_{ik} - \mathbf{x}_{ij,\{0\}}\mathbf{x}_{ik,\{0\}}|^2, 4\kappa^2\} \\
&\leq 4\kappa^2\mathbb{E}\Big|\frac{\mathbf{x}_{ij}\mathbf{x}_{ik} - \mathbf{x}_{ij,\{0\}}\mathbf{x}_{ik,\{0\}}}{2\kappa}\Big|^{1+\epsilon} \\
&= (2\kappa)^{1-\epsilon}\|\mathbf{x}_{ij}\mathbf{x}_{ik} - \mathbf{x}_{ij,\{0\}}\mathbf{x}_{ik,\{0\}}\|_{1+\epsilon}^{1+\epsilon}.
\end{aligned}
$$

By the triangle inequality and the Hölder inequality,

$$
\begin{aligned}
&\|\mathbf{x}_{ij}\mathbf{x}_{ik} - \mathbf{x}_{ij,\{0\}}\mathbf{x}_{ik,\{0\}}\|_{1+\epsilon} \\
\leq\ & \|\mathbf{x}_{ij}\|_{2+2\epsilon}\|\mathbf{x}_{ik} - \mathbf{x}_{ik,\{0\}}\|_{2+2\epsilon} + \|\mathbf{x}_{ik,\{0\}}\|_4\|\mathbf{x}_{ij} - \mathbf{x}_{ij,\{0\}}\|_{2+2\epsilon} \\
=\ & \omega_{2+2\epsilon}(\delta_{i,2+2\epsilon,j} + \delta_{i,2+2\epsilon,k}).
\end{aligned}
$$

We then have

$$
\sup_{m\geq 0}\rho^{-m}\sum_{i=m}^{\infty}\|\varphi_\kappa(\mathbf{x}_{ij}\mathbf{x}_{ik}) - \varphi_\kappa(\mathbf{x}_{ij,\{0\}}\mathbf{x}_{ik,\{0\}})\|_2 \leq (2\kappa)^{(1-\epsilon)/2}\omega_{2+2\epsilon}^{(1+\epsilon)/2}\|\mathbf{x}_{\cdot}^*\|_{1+\epsilon}.
$$

We can follow the proof of Corollary 4.1 to complete the remaining proof and thus it is omitted here. □

Proof of Theorem 4.3. The proof of Theorem 4.3 essentially follows from the error bound of Huber type covariance estimator (4.25) or (4.28) and the arguments in Cai et al. (2011) without extra technical difficulties. □

References

Cai, T., W. Liu, and X. Luo (2011). A constrained ℓ_1 minimization approach to sparse precision matrix estimation. *Journal of the American Statistical Association 106*(494), 594–607.

Merlevède, F., M. Peligrad, and E. Rio (2009). Bernstein inequality and moderate deviations under strong mixing conditions. In *High dimensional probability V: the Luminy volume*, pp. 273–292. Institute of Mathematical Statistics.

Merlevède, F., M. Peligrad, and E. Rio (2011). A bernstein type inequality and moderate deviations for weakly dependent sequences. *Probability Theory and Related Fields 151*(3-4), 435–474.