S1 Notations

We introduce some notations here which will be used in the proofs. For any \((j-1)\)-dimensional vector \(a = (a_1, \ldots, a_{j-1})^T \in \mathbb{R}^{j-1}\), \(B_{k,j}(a)\) is defined by \(B_{k,j}(a) := (b_i = a_{j-1-k+i}, 1 \leq i \leq k)\). Let \(\Omega_{0,n}\) be the true precision matrix and \(\Omega_{0,n} = (I_p - A_{0,n})^T D_{0,n}^{-1} (I_p - A_{0,n})\) be its modified Cholesky decomposition with \(A_{0,n} = (a_{0,j})\) and \(D_{0,n} = \text{diag}(d_{0,j})\). It is easy to check that the explicit forms of \(a_{0,j} = (a_{0,j1}, \ldots, a_{0,j,j-1})^T\) and \(d_{0,j}\) are

\[
a_{0,j} = \text{Var}(X_{1,1:(j-1)})^{-1} \text{Cov}(X_{1,1:(j-1)}, X_{1j}),
\]

\[
d_{0,j} = \text{Var}(X_{1j}) - \text{Cov}(X_{1j}, X_{1,1:(j-1)}) \text{Var}(X_{1,1:(j-1)})^{-1} \text{Cov}(X_{1,1:(j-1)}, X_{1j}),
\]

(S1.1)

where \(X_{1,a:b} = (X_{1a}, \ldots, X_{1b})^T\) denotes the sub-vector of the first observation \(X_1 = (X_{11}, \ldots, X_{1p})^T \in \mathbb{R}^p\) for any positive integers \(1 \leq a \leq b \leq p\). Since we assume \(X_1, \ldots, X_n \overset{iid}{\sim} N_p(0, \Omega_{0,n}^{-1})\), \(X_{1,a:b}\) can be replaced by \(X_{i,a:b}\).
for any $i = 2, \ldots, n$. For more details on the above expression (S1.1), refer to Bickel and Levina (2008) (pages 202 and 221).

For a given $k$, we denote

$$a^{(k)}_{0,j} = \text{Var}(X_{1,(j-k):(j-1)})^{-1}\text{Cov}(X_{1,(j-k):(j-1)}, X_{1j}),$$

$$d_{0,jk} = \text{Var}(X_{1j}) - \text{Cov}(X_{1j}, X_{1,(j-k):(j-1)})\text{Var}(X_{1,(j-k):(j-1)})^{-1}\text{Cov}(X_{1,(j-k):(j-1)}, X_{1j}).$$

We denote the empirical estimators by

$$\hat{\text{Var}}(X_{1,(j-k):(j-1)}) = n^{-1}X_{:, (j-k):(j-1)}^TX_{:, (j-k):(j-1)}$$

and

$$\hat{\text{Cov}}(X_{1,(j-k):(j-1)}, X_{1j}) = n^{-1}X_{:, (j-k):(j-1)}^TX_{:, j}$$

for any $j = 2, \ldots, p$. For any $j = 1, \ldots, p$, we define

$$\hat{\text{Var}}(X_{1j}) = n^{-1}\|X_{:, j}\|_2^2.$$

## S2 Proofs

### S2.1 Proof of Proposition 1

**Proof.** First, we prove only the exponentially decreasing case, $\gamma(k) = Ce^{-\beta k}$ for some $\beta > 0$ and $C > 0$, because the proposition trivially holds for the exact banding case.

Suppose $\Omega_{0,n} = (\omega_{0,ij}) \in \mathcal{U}($ε₀, γ) and let $\Omega_{0,n} = (I_p - A_{0,n})^TD_{0,n}^{-1}(I_p -
where \( A_{0,n} = (a_{0,ij}) \) and \( D_{0,n} = \text{diag}(d_{0,j}) \). Note that

\[
\|D_{0,n}^{-1}\| = \max_j d_{0,j}^{-1} \\
= \max_j \left\| \text{Var}^{1/2}(X_{1,1:j}) \begin{pmatrix} -a_{0,j} \\ 1 \end{pmatrix} \right\|^{-2}_2 \\
\leq \max_j \lambda_{\min}(\text{Var}(X_{1,1:j}))^{-1} \cdot \left\| \begin{pmatrix} -a_{0,j} \\ 1 \end{pmatrix} \right\|^{-2}_2 \\
\leq \max_j \lambda_{\min}(\text{Var}(X_{1,1:j}))^{-1} \\
\leq \lambda_{\min}(\Omega_{0,n})^{-1} \leq \epsilon_0^{-1}
\]

and

\[
\|A_{0,n}\|_{\max} \leq \max_j \|a_{0,j}\|_2 = \max_j \|\text{Var}(X_{1,1:(j-1)})^{-1}\text{Cov}(X_{1,1:(j-1)}, X_{1,j})\|_2 \\
\leq \max_j \|\text{Var}(X_{1,1:(j-1)})^{-1}\| \|\text{Cov}(X_{1,1:(j-1)}, X_{1,j})\|_2 \\
\leq \max_j \|\text{Var}(X_{1,1:(j-1)})^{-1}\| \|\text{Var}(X_{1,1:j})\| \leq \epsilon_0^{-2}
\]

by the assumption \( \epsilon_0 \leq \lambda_{\min}(\Omega_{0,n}) \leq \lambda_{\max}(\Omega_{0,n}) \leq \epsilon_0^{-1} \).

Furthermore,

\[
\|A_{0,n} - B_k(A_{0,n})\|_{\infty} = \max_i \sum_{j \leq i-k} |a_{0,ij}| \leq \gamma(k), \quad (S2.2)
\]

\[
\|A_{0,n} - B_k(A_{0,n})\|_1 = \max_j \sum_{i > j+k} |a_{0,ij}| \leq \sum_{m=k}^{\infty} \gamma(m) \leq C' \gamma(k), \quad (S2.3)
\]

for some \( C' > 1 \) because \( \gamma(k) = C e^{-\beta k} \). Note that \( \omega_{0,pp} = d_{0,p}^{-1} \) and

\[
\omega_{0,ij} = -d_{0,j}^{-1}a_{0,ji} + \sum_{l=j+1}^{p} d_{0,l}^{-1}a_{0,li}a_{0,lj} \quad \text{for any } 1 \leq i < j \leq p. \quad (S2.4)
\]
Then for $1 \leq i < p$, define $k$ so that $i = p - k - 1$. Then, $k \geq 0$ and

$$|\omega_{0,ip}| = d_{0,p}^{-1}|a_{0,pi}|$$

$$\leq \epsilon_0^{-1}\gamma(k)$$

by (S2.2). On the other hand, for $1 \leq i < j \leq p - 1$, define $k$ so that $j - i = k + 1$. Then, $k \geq 0$ and

$$|\omega_{0,ij}| = |-d_{0,j}a_{0,ji} + \sum_{l=j+1}^{p} d_{0,l}^{-1}a_{0,li}a_{0,lj}|$$

$$\leq d_{0,j}^{-1}|a_{0,ji}| + \sum_{l=j+1}^{p} d_{0,l}^{-1}|a_{0,li}a_{0,lj}|$$

$$\leq \epsilon_0^{-3}\left(|a_{0,ji}| + \sum_{l=j+1}^{p} |a_{0,li}|\right)$$

$$= \epsilon_0^{-3}\sum_{l=j}^{p} |a_{0,li}| \leq \epsilon_0^{-3}\gamma(k)$$

by (S2.3). Thus, we have

$$\|\Omega_{0,n} - B_k(\Omega_{0,n})\|_\infty = \max_i \sum_{j:|i-j|>k} |\omega_{0,ij}|$$

$$\leq \max_i \sum_{j>i+k} |\omega_{0,ij}| + \max_i \sum_{j<i-k} |\omega_{0,ji}|$$

$$\leq 2\epsilon_0^{-3}\gamma(k)\sum_{m=k}^{\infty} \gamma(m) \leq C''\gamma(k)$$

for some constant $C'' > 0$. This proves the first inequality.

Suppose $\Omega_{0,n} \in U^*(\epsilon_0, \gamma)$. We need to prove that $\max_i \sum_{j<i-k} |a_{0,ij}| = \max_i \sum_{j=1}^{i-k-1} |a_{0,ij}| \leq C\gamma(k)$ for some constant $C > 0$. Note that from
(S2.4), we have
\[ d_{0,p}^{-1} \sum_{j=1}^{p-k-1} |a_{0,pj}| = \sum_{j=1}^{p-k-1} |\omega_{0,jp}| \leq \gamma(k), \] (S2.5)
for any \(0 \leq k \leq p-2\). We will show that
\[ d_{0,p-t}^{-1} \sum_{j=1}^{p-t-k-1} |a_{0,p-t,j}| \leq \gamma(k) + \epsilon_0^{-2} \sum_{m=1}^{t} (1 + \epsilon_0^{-2})^{m-1} \gamma(k + m) \] (S2.6)
for any \(1 \leq t \leq p-k-2\) for some \(0 \leq k \leq p-3\). Then, (S2.5) and (S2.6) imply \(\Omega_{0,n} \in \mathcal{U}(\epsilon_0, C'\gamma)\) for some \(C' > 0\) because \(\max_j d_{0,j} \leq \max_j \text{Var}(X_{1j}) \leq \epsilon_0^{-1}\) and we assume that \(\gamma(k) = Ce^{-\beta k}\) and \(\beta > \log(\epsilon_0^{-2} + 1)\).

By (S2.4) and the assumption \(\Omega_{0,n} \in \mathcal{U}^*(\epsilon_0, \gamma)\),
\[ \sum_{j=1}^{p-k-2} \left| -d_{0,p-1,j}^{-1}a_{0,p-1,j} + d_{0,p}^{-1}a_{0,pj}a_{0,p,p-1} \right| = \sum_{j=1}^{p-k-2} |\omega_{0,j,p-1}| \leq \gamma(k) \] (S2.7)
for any \(0 \leq k \leq p-3\). Thus, (S2.5) and (S2.7) imply that
\[ d_{0,p-1}^{-1} \sum_{j=1}^{p-k-2} |a_{0,p-1,j}| \leq \sum_{j=1}^{p-k-2} \left| -d_{0,p-1,j}^{-1}a_{0,p-1,j} + d_{0,p}^{-1}a_{0,pj}a_{0,p,p-1} + \sum_{j=1}^{p-k-2} |d_{0,p}^{-1}a_{0,pj}a_{0,p,p-1}| \right| \leq \gamma(k) + \epsilon_0^{-2} \gamma(k + 1) \]
because \(\Omega_{0,n} \in \mathcal{U}^*(\epsilon_0, \gamma)\) means \(|a_{0,p,p-1}| \leq \|A_{0,n}\|_{\max} \leq \epsilon_0^{-2}\). Thus, (S2.6) holds for \(t = 1\). Now assume that (S2.6) holds for \(t-1\) and consider for the case of \(t\). Note that
\[ \gamma(k) \geq \sum_{j=1}^{p-t-k-1} |\omega_{0,j,p-t}| \]
\[ = \sum_{j=1}^{p-t-k-1} \left| -d_{0,p-t,j}^{-1}a_{0,p-t,j} + \sum_{l=p-t+1}^{p} d_{0,l}^{-1}a_{0,lj}a_{0,l,p-t} \right|, \]
which implies that
\[
\begin{align*}
\sum_{j=1}^{p-t-k-1} d_{0,p-t}^{-1} |a_{0,p-t,j}| & \leq \gamma(k) + \sum_{j=1}^{p-t-k-1} \sum_{l=p-t+1}^{p} d_{0,l}^{-1} |a_{0,l,j} a_{0,t,p-t}| \\
& \leq \gamma(k) + \epsilon_0^{-2} \sum_{l=p-t+1}^{p} d_{0,l}^{-1} \sum_{j=1}^{p-t-k-1} |a_{0,l,j}| \\
& = \gamma(k) + \epsilon_0^{-2} \sum_{t_1=0}^{t-1} d_{0,p-t_1}^{-1} \sum_{j=1}^{p-t_1-(k+t-t_1)-1} |a_{0,p-t_1,j}| \\
& \leq \gamma(k) + \epsilon_0^{-2} \gamma(k + t) \\
& + \epsilon_0^{-2} \sum_{t_1=1}^{t-1} \left( \gamma(k + t - t_1) + \epsilon_0^{-2} \sum_{m=1}^{t_1} (1 + \epsilon_0^{-2})^{m-1} \gamma(k + t - t_1 + m) \right).
\end{align*}
\]

In (S2.8), one can check that the coefficient of $\gamma(k + t - t')$ is
\[
\epsilon_0^{-2} + \epsilon_0^{-4} \sum_{m=1}^{t-t'-1} (1 + \epsilon_0^{-2})^{m-1} = \epsilon_0^{-2} (1 + \epsilon_0^{-2})^{t-t'-1}
\]
for $0 \leq t' \leq t - 1$, and the coefficient of $\gamma(k)$ is 1. Thus,
\[
\sum_{j=1}^{p-t-k-1} d_{0,p-t}^{-1} |a_{0,p-t,j}| \leq \gamma(k) + \epsilon_0^{-2} \sum_{m=1}^{t} (1 + \epsilon_0^{-2})^{m-1} \gamma(k + m).
\]

This completes the proof by induction.

Now suppose that $\gamma(k) = Ck^{-\alpha}$ for some constant $C > 0$ and $\Omega_{0,n} \in U(\epsilon_0, \gamma)$. We will show that $\Omega_{0,n} \in U'(\epsilon_0, \gamma')$, where $\gamma'(k) = C'k^{1-\alpha}$ for some constant $C' > 0$. Let $Q = D_{0,n}^{-1/2}(I_p - A_{0,n})$, then by the proof of
Lemma 2 in Liu and Ren (2017),

\[ \| \Omega_{0,n} - B_k(\Omega_{0,n}) \|_\infty \leq \| (Q - B_k(Q))^T Q \|_\infty + \| Q^T (Q - B_k(Q)) \|_\infty + \| (Q - B_k(Q))^T (Q - B_k(Q)) \|_\infty \]  
(S2.9)

\[ + \| B_k[(Q - B_k(Q))^T B_k(Q)] \|_\infty + \| B_k[B_k(Q)^T (Q - B_k(Q))] \|_\infty \]  
(S2.10)

\[ + \| B_k[(Q - B_k(Q))^T (Q - B_k(Q))] \|_\infty. \]  
(S2.11)

The two terms in (S2.10) are bounded above by $C'k^{1-2\alpha}$ for some constant $C' > 0$ by the proof of Lemma 2 in Liu and Ren (2017). With a slightly modified version of Lemma 24 and Lemma 25 in Liu and Ren (2017) by considering $\| \cdot \|_1$ and $\| \cdot \|_\infty$ instead of $\| \cdot \|$, one can show that the term (S2.11) is also bounded above by $C'k^{1-2\alpha}$. Three terms in (S2.9) are bounded above by $C'k^{1-\alpha}$ by the modified version of Lemma 24 in Liu and Ren (2017) and Lemma 8. It completes the proof. \[ \square \]

S2.2 Proof of Minimax Lower Bounds: Theorem 1 and Theorem 3

Proof of Theorem 1. We follow closely the line of a proof in Cai et al. (2010). Consider the polynomially decreasing case, \( \gamma(k) = Ck^{-\alpha} \), first. Two parameter classes are considered depending on the relation between \( p \) and \( n \).
For \( \exp(n^{1/(2\alpha+1)}) \geq p \) case, we show that
\[
\inf_{\hat{\Omega}_n} \sup_{\Omega_{0,n} \in U_{11}} \mathbb{E}_0 n ||\hat{\Omega}_n - \Omega_{0,n}|| \gtrsim \min \left( n^{-\alpha/(2\alpha+1)}, \left( \frac{p}{n} \right)^{1/2} \right),
\]
(S2.12)
and for \( \exp(n^{1/(2\alpha+1)}) \leq p \) case, we show that
\[
\inf_{\hat{\Omega}_n} \sup_{\Omega_{0,n} \in U_{12}} \mathbb{E}_0 n ||\hat{\Omega}_n - \Omega_{0,n}|| \gtrsim \left( \log \frac{p}{n} \right)^{1/2}
\]
(S2.13)
for some \( U_{11} \cup U_{12} \subset U(\epsilon_0, \gamma) \). Then, it gives a lower bound for the parameter space \( U(\epsilon_0, \gamma) \),
\[
\inf_{\hat{\Omega}_n} \sup_{\Omega_{0,n} \in U(\epsilon_0, \gamma)} \mathbb{E}_0 n ||\hat{\Omega}_n - \Omega_{0,n}|| \geq \inf_{\hat{\Omega}_n} \sup_{\Omega_{0,n} \in U_{11} \cup U_{12}} \mathbb{E}_0 n ||\hat{\Omega}_n - \Omega_{0,n}|| \geq \inf_{\hat{\Omega}_n} \sup_{\Omega_{0,n} \in U_{11}} \mathbb{E}_0 n ||\hat{\Omega}_n - \Omega_{0,n}|| + \inf_{\hat{\Omega}_n} \sup_{\Omega_{0,n} \in U_{12}} \mathbb{E}_0 n ||\hat{\Omega}_n - \Omega_{0,n}|| I(\exp(n^{1/(2\alpha+1)}) \geq p) + \inf_{\hat{\Omega}_n} \sup_{\Omega_{0,n} \in U_{12}} \mathbb{E}_0 n ||\hat{\Omega}_n - \Omega_{0,n}|| I(\exp(n^{1/(2\alpha+1)}) < p) \gtrsim \min \left\{ \left( \frac{\log p}{n} \right)^{1/2} + n^{-\alpha/(2\alpha+1)}, \left( \frac{p}{n} \right)^{1/2} \right\},
\]
which is the desired result.

Consider \( \exp(n^{1/(2\alpha+1)}) \geq p \) case first. Without loss of generality, we assume \( k = \min(n^{1/(2\alpha+1)}, p) \) is an even number, and define a class of precision matrices
\[
U_{11} = \left\{ \Omega(\theta) \in \mathbb{R}^{p \times p} : \Omega(\theta) = (I_p - A(\theta))^T (I_p - A(\theta)), A(\theta) = -\tau a \sum_{m=1}^{k/2} \theta_m B(m,k), \theta = (\theta_m, 1 \leq m \leq k/2) \in \{0, 1\}^{k/2} \right\}
\]
where \( B(m,k) = (b_{ij} = I(i = m + 1, \ldots, k \text{ and } j = m), 1 \leq i, j \leq p) \) is a \( p \times p \) matrix and \( a = (nk)^{-1/2} \). If we choose sufficiently small constant
$\tau > 0$, it is easy to check that for any $\Omega(\theta) \in \mathcal{U}_{11}$, $\varepsilon_0 \leq \lambda_{\min}(\Omega(\theta)) \leq \lambda_{\max}(\Omega(\theta)) \leq \varepsilon_0^{-1}$ and $\|A(\theta) - B_{k_1}(A(\theta))\|_\infty \leq Ck_1^{-\alpha}$ for any $k_1 > 0$, so that $\mathcal{U}_{11} \subset \mathcal{U}(\varepsilon_0, \gamma)$ for all sufficiently large $n$.

We use the Assouad’s lemma (Assouad, 1983)

$$\inf_{\Omega_n} \sup_{\Omega(\theta) \in \mathcal{U}_{11}} 2\mathbb{E}_\theta \|\hat{\Omega}_n - \Omega(\theta)\| \geq \min_{H(\theta, \theta') \geq 1} \frac{H(\theta, \theta')}{2} \frac{k/2}{\min_{H(\theta, \theta') = 1} \|\mathbb{P}_\theta \wedge \mathbb{P}_{\theta'}\|}$$

where $H(\theta, \theta') = \sum_{m=1}^{k/2} |\theta_m - \theta'_m|$, $\|\mathbb{P}_\theta \wedge \mathbb{P}_{\theta'}\| = \int p_{\theta} \wedge p_{\theta'} d\mu$, and $p_{\theta}$ and $p_{\theta'}$ are the joint distribution function and density function, with respect to a dominating measure $\mu$, of observation $X_1, \ldots, X_n \overset{iid}{\sim} N(0, \Omega(\theta)^{-1})$, respectively. If we show that

$$\min_{H(\theta, \theta') \geq 1} \frac{H(\theta, \theta')}{2} \frac{k/2}{\min_{H(\theta, \theta') = 1} \|\mathbb{P}_\theta \wedge \mathbb{P}_{\theta'}\|} \geq a$$

(S2.14)

and

$$\min_{H(\theta, \theta') = 1} \|\mathbb{P}_\theta \wedge \mathbb{P}_{\theta'}\| \geq c$$

(S2.15)

for some constant $c > 0$, it will complete the proof. To show (S2.14), define a $p$-dimensional vector $v = (I(k/2 \leq i \leq k), 1 \leq i \leq p)$. By the construction of $\Omega(\theta)$ and $v$, one can check that

$$((\Omega(\theta) - \Omega(\theta'))v)_l = \begin{cases} 
(\tau a)^2 \frac{k}{2}(\theta_{l\theta_{k/2}} - \theta'_{l\theta'_{k/2}}) + \tau a(\frac{k}{2} + 1)(\theta_{l\theta} - \theta'_{l\theta'}) & \text{if } 1 \leq l \leq \frac{k}{2} - 1 \\
(\tau a)^2 \frac{k}{2}(\theta_{k/2\theta_{k/2}} - \theta'_{k/2\theta'_{k/2}}) + \tau a \frac{k}{2}(\theta_{k/2\theta} - \theta'_{k/2\theta'}) & \text{if } l = \frac{k}{2} \\
\tau a(\theta_{k+1-l\theta} - \theta'_{k+1-l\theta'}) & \text{if } \frac{k}{2} + 1 \leq l \leq k \\
0 & \text{if } l \geq k + 1.
\end{cases}$$
Then, we have \( \| (\Omega(\theta) - \Omega(\theta'))v \|_2^2 \geq (\tau a)^2 (k/2)^2 H(\theta, \theta') \) and

\[
\| \Omega(\theta) - \Omega(\theta') \| \geq \frac{\| (\Omega(\theta) - \Omega(\theta'))v \|_2}{\| v \|_2} \geq \frac{\| (\Omega(\theta) - \Omega(\theta'))v \|_2}{\sqrt{k/2}} \geq \left( \frac{(k/2 \times \tau a)^2 H(\theta, \theta')}{k/2} \right)^{1/2} = \left( \frac{k/2}{H(\theta, \theta')} \right)^{1/2} \tau a H(\theta, \theta') \geq \tau a H(\theta, \theta').
\]

Thus, we have shown the first part.

To show (S2.15), note that

\[
\| P_{\theta} \land P_{\theta'} \| = \int_{p_{\theta} > p_{\theta'}} p_{\theta} d\mu + \int_{p_{\theta} \leq p_{\theta'}} p_{\theta} d\mu = \left( \frac{1}{2} - \frac{1}{2} \int_{p_{\theta} \leq p_{\theta'}} p_{\theta} d\mu + \frac{1}{2} \int_{p_{\theta} > p_{\theta'}} p_{\theta} d\mu \right) + \left( \frac{1}{2} - \frac{1}{2} \int_{p_{\theta} > p_{\theta'}} p_{\theta} d\mu + \frac{1}{2} \int_{p_{\theta} \leq p_{\theta'}} p_{\theta} d\mu \right) = 1 - \frac{1}{2} \int_{p_{\theta} > p_{\theta'}} (p_{\theta} - p_{\theta'}) d\mu - \frac{1}{2} \int_{p_{\theta} \leq p_{\theta'}} (p_{\theta'} - p_{\theta}) d\mu = 1 - \frac{1}{2} \int \left| p_{\theta} - p_{\theta'} \right| d\mu.
\]

Let \( \| P_{\theta} - P_{\theta'} \|_1 = \int \left| p_{\theta} - p_{\theta'} \right| d\mu \). Thus, it suffices to show that \( \| P_{\theta} - P_{\theta'} \|_1^2 \leq \)
$1/2$ when $H(\theta, \theta') = 1$. Also note that

$$\|P_{\theta} - P_{\theta'}\|_1 \leq 2K(P_{\theta'} \mid P_{\theta})$$

$$= n \left[ \text{tr}(\Omega(\theta')^{-1}\Omega(\theta)) - \log \det(\Omega(\theta')^{-1}\Omega(\theta)) - p \right]$$

$$= n \left[ \text{tr}(\Omega(\theta')^{-1}D_1) - \log \det(\Omega(\theta')^{-1}D_1 + I_p) \right]$$

$$= n \left[ \text{tr}(\Omega(\theta')^{-1/2}D_1\Omega(\theta')^{-1/2}) - \log \det(\Omega(\theta')^{-1/2}D_1\Omega(\theta')^{-1/2} + I_p) \right]$$

where $K(P_{\theta'} \mid P_{\theta}) = \int \log(\frac{p_{\theta'}}{p_{\theta}})p_{\theta'}d\mu$ is the Kullback-Leibler divergence and $D_1 = \Omega(\theta) - \Omega(\theta')$. Let $\Omega(\theta')^{-1} = UVU^T$ be the diagonalization of $\Omega(\theta')^{-1}$.

$U$ is a orthogonal matrix whose columns are the eigenvectors of $\Omega(\theta')^{-1}$, and $V$ is a diagonal matrix whose $i$th diagonal element is the eigenvalue of $\Omega(\theta')^{-1}$ corresponding to the $i$th column of $U$. It is easy to check that

$$\|\Omega(\theta')^{-1/2}D_1\Omega(\theta')^{-1/2}\|_F^2 = \|UV^{1/2}U^TD_1UV^{1/2}U^T\|_F^2$$

$$= \|V^{1/2}U^TD_1UV^{1/2}\|_F^2$$

$$\leq \|V\|^2\|U^TD_1U\|_F^2$$

$$= \|\Omega(\theta')^{-1}\|^2\|D_1\|_F^2$$

$$\leq Ck(\tau a)^2$$
for some constant $C > 0$ because $\| \Omega(\theta')^{-1} \| \leq C_0^{-1}$ and

$$
(\Omega(\theta))_{(i,j)} = \begin{cases}
1 + (\tau a)^2 \theta_i(k - i) & \text{if } 1 \leq i = j \leq \frac{k}{2} \\
(\tau a)\theta_i + (\tau a)^2 \theta_i(k - j) & \text{if } 1 \leq i \neq j \leq \frac{k}{2} \\
\tau a \theta_i & \text{if } 1 \leq i \leq \frac{k}{2}, \frac{k}{2} + 1 \leq j \leq k \\
\tau a \theta_j & \text{if } \frac{k}{2} + 1 \leq i \leq k, 1 \leq j \leq \frac{k}{2} \\
0 & \text{otherwise.}
\end{cases}
$$

Also note that, if $\lambda_1(\theta, \theta') \leq \cdots \leq \lambda_p(\theta, \theta')$ are the eigenvalues of $\Omega(\theta')^{-1/2} D_1 \Omega(\theta')^{-1/2}$, we have

$$
\sum_{j=1}^p \lambda_j(\theta, \theta')^2 = \| \Omega(\theta')^{-1/2} D_1 \Omega(\theta')^{-1/2} \|_F^2 \leq C k(\tau a)^2 = C \tau^2 / n,
$$

which implies $|\lambda_j(\theta, \theta')| \leq \sqrt{C} \tau / \sqrt{n}$ for all $1 \leq j \leq p$. Thus, $\Omega(\theta')^{-1/2} D_1 \Omega(\theta')^{-1/2} + I_p$ is a positive definite matrix for all large $n$. Since $\| \Omega(\theta')^{-1/2} D_1 \Omega(\theta')^{-1/2} \|_F^2$ is small, by Lemma C.2 in Lee and Lee (2018),

$$
\| P_\theta - P_{\theta'} \|_1 \leq n R_n
$$

where $R_n \leq C \| \Omega(\theta')^{-1/2} D_1 \Omega(\theta')^{-1/2} \|_F^2$ for some constant $C > 0$. Thus, we have $\| P_\theta - P_{\theta'} \|_1 \leq 1/2$ for some small $\tau > 0$ because $nka^2 = 1$.

Now consider $\exp(n^{1/2(2\alpha+1)}) \leq p$ case. To show (S2.13), define a class of diagonal precision matrices

$$
U_{12} = \left\{ \Omega_m \in \mathbb{R}^{p \times p} : \Omega_m = I_p + \tau \left( \frac{\log p}{n} \right)^{1/2} \left( I(i = j = m) \right), \ 0 \leq m \leq p \right\}
$$

for some small $\tau > 0$. Since $p \leq \exp(cn)$ for some constant $c > 0$, $U_{12} \subset U(\epsilon_0, \gamma)$ holds trivially. Let $r_{\min} = \inf_{1 \leq m \leq p} \| \Omega_0 - \Omega_m \|$. We use the Le
Cam’s lemma (LeCam, 1973)

\[
\inf_{\Omega_n} \sup_{\Omega_m \in \Omega_{12}} \mathbb{E}_m \|\hat{\Omega}_n - \Omega_m\| \geq \frac{1}{2} r_{\text{min}} \|\mathbb{P}_0 \wedge \bar{\mathbb{P}}\|
\]

where \(\bar{\mathbb{P}} = p^{-1} \sum_{m=1}^{p} \mathbb{P}_m\) and \(\mathbb{P}_m\) is the distribution function of \(N_p(0, \Omega_m^{-1})\) with observation \(X_n\). Note that \(r_{\text{min}} = \tau(\log p/n)^{1/2}\). We only need to show that \(\|\mathbb{P}_0 \wedge \bar{\mathbb{P}}\| \geq c\) for some constant \(c > 0\). By the same argument with Cai et al. (2010) (page 2129), it suffices to show that

\[
\int \left(\frac{p^{-1} \sum_{m=1}^{p} f_m}{f_0}\right)^2 d\mu - 1 \rightarrow 0, \quad (S2.16)
\]
as \(n \to \infty\) where \(f_m\) is the density function of \(\mathbb{P}_m\) with respect to a \(\sigma\)-finite measure \(\mu\). Note that

\[
\int \left(\frac{p^{-1} \sum_{m=1}^{p} f_m}{f_0}\right)^2 d\mu - 1 = \frac{1}{p^2} \sum_{m=1}^{p} \int \frac{f_m^2}{f_0} d\mu + \frac{1}{p^2} \sum_{m \neq j} \int \frac{f_m f_j}{f_0} d\mu - 1
\]

and \(\int f_m f_j / f_0 d\mu = 1\) for any \(m \neq j\). Also note that

\[
\int \frac{f_m^2}{f_0} d\mu = (1 + b)^{n/2} \left(1 - \frac{b}{1 + 2b}\right)^{n/2} \leq e^{nb^2/(1+2b)} \leq e^{nb^2} = e^{\tau^2 \log p}
\]

where \(b = \tau(\log p/n)^{1/2}\). Thus, (S2.16) holds for some small \(\tau > 0\). It completes the proof for the case of polynomially decreasing \(\gamma(k)\).

For the case of exponentially decreasing \(\gamma(k) = Ce^{-\beta k}\), consider \(k = \)
min(\log n, p) for \( U_{11} \) instead of \( k = \min(n^{1/(2\alpha+1)}, p) \). Then, similar arguments for the lower bounds of \( U_{11} \) and \( U_{12} \) give the desired result.

For the exact banding \( \gamma(k) \), consider \( U_{11} \) with \( k = k_0 \) and \( a = (\log p/n)^{1/2} \), then it completes the proof.

**Proof of Theorem 3.** We follow closely the line of a proof in Cai and Zhou (2012). Consider the polynomially decreasing case, \( \gamma(k) = Ck^{-\alpha} \), first. Two parameter classes are considered depending on the relation between \( p \) and \( n \). For \( \exp(n^{1/(2\alpha+2)}) \geq p \) case, we show that

\[
\inf_{\hat{\Omega}_n} \sup_{\Omega_{0,n} \in G_{11}} \mathbb{E}_{0,n} \| \hat{\Omega}_n - \Omega_{0,n} \|_{\infty} \gtrsim \min \left( \frac{n^{-\alpha/(2\alpha+2)}}{\frac{p}{n^{1/2}}} \right), \tag{S2.17}
\]

and for \( \exp(n^{1/(2\alpha+2)}) \leq p \) case, we show that

\[
\inf_{\hat{\Omega}_n} \sup_{\Omega_{0,n} \in G_{12}} \mathbb{E}_{0,n} \| \hat{\Omega}_n - \Omega_{0,n} \|_{\infty} \gtrsim \left( \frac{\log p}{n} \right)^{\alpha/(2\alpha+1)}, \tag{S2.18}
\]

for some \( G_{11} \cup G_{12} \subset U(\epsilon_0, \gamma) \).

Consider \( \exp(n^{1/(2\alpha+2)}) \geq p \) case first. Define a class of precision matrices

\[
G_{11} = \left\{ \Omega(\theta) \in \mathbb{R}^{p \times p} : \Omega(\theta) = (I_p - A(\theta))^T(I_p - A(\theta)), \right. \\
A(\theta) = -\tau a \sum_{s=2}^{k} \theta_{s-1} G_s, \theta = (\theta_s) \in \{0, 1\}^{k-1} \}
\]

where \( G_s = (I(i = s, j = 1)) \) is a \( p \times p \) matrix and \( a = n^{-1/2} \) and \( k = \min(n^{1/(2\alpha+2)}, p) \). It is easy to show that \( G_{11} \subset U(\epsilon_0, \gamma) \) for some small
constant $\tau > 0$ and all sufficiently large $n$.

We use the Assouad’s lemma,

$$\inf_{\hat{\Omega}} \sup_{\Omega(\theta) \in \mathcal{G}} 2\mathbb{E}_\theta ||\hat{\Omega}_n - \Omega(\theta)||_\infty \geq \min_{H(\theta, \theta') \geq 1} \frac{||\Omega(\theta) - \Omega(\theta')||_\infty}{H(\theta, \theta')} \frac{k - 1}{2} \min_{H(\theta, \theta') = 1} ||P_\theta \wedge P_{\theta'}||.$$ 

It is easy to see that

$$\min_{H(\theta, \theta') \geq 1} \frac{||\Omega(\theta) - \Omega(\theta')||_\infty}{H(\theta, \theta')} \geq \tau a.$$

To show $\min_{H(\theta, \theta') = 1} ||P_\theta \wedge P_{\theta'}|| \geq c$ for some $c > 0$, it suffices to prove that $||P_\theta - P_{\theta'}||_1 \leq 1$. Note that

$$||P_\theta - P_{\theta'}||_1^2 \leq 2K(P_{\theta'} | P_\theta) \leq Cn||\Omega(\theta')^{-1/2}D_1\Omega(\theta')^{-1/2}||_F^2$$

for some constant $C > 0$ where $D_1 = \Omega(\theta) - \Omega(\theta')$. By the same argument used in the proof of Theorem 1, one can show that $||P_\theta - P_{\theta'}||_1^2 \leq C'n(\tau a)^2$ for some constant $C' > 0$, and it is smaller than 1 for some small constant $\tau > 0$. Thus, we have proved the (S2.17) part.

Now consider $\exp(n^{1/(2\alpha + 2)}) \leq p$ case. To show (S2.18) part, define a class of precision matrices

$$\mathcal{G}_{12} = \left\{ \Omega_m \in \mathbb{R}^{p \times p} : \Omega_m = (I_p - A_m)^T(I_p - A_m), A_m = -\tau B_m \left( \log \frac{p}{nk} \right)^{1/2}, 1 \leq m \leq m_* \right\}$$

where $B_m = (I(m+1 \leq i \leq m+k-1, j = m))$ is a $p \times p$ matrix, $m_* = p/k - 1$ and $k = (n/\log p)^{1/(2\alpha + 1)}$. Without loss of generality, we assume that $p$ can
be divided by \( k \). By the definition of \( \mathcal{G}_{12} \), tedious calculations yield that 
\( \mathcal{G}_{12} \subset \mathcal{U}(\epsilon_0, \gamma) \).

Let \( \Omega_0 = I_p \) and \( \mathbb{P}_m \) be the distribution function of \( N(0, \Omega_{-1}^{-1}) \) with observation \( X_n \). It is easy to check that for any \( 0 \leq m \neq m' \leq m_* \),
\[
\| \Omega_m - \Omega_{m'} \|_{\infty} \geq \tau \left( \frac{k \log p}{n} \right)^{1/2} = \tau \left( \frac{\log p}{n} \right)^{\alpha/(2\alpha+1)}
\]
by the definition of \( \mathcal{G}_{12} \) and \( k \). Since \( k^2 \leq p \), for any \( 1 \leq m \leq m_* \),
\[
K(\mathbb{P}_m \mid \mathbb{P}_0) \leq C n \| \Omega_{m'}^{-1/2} D_1 \Omega_{m'}^{-1/2} \|_F^2 \leq C' \tau^2 \log p \leq c \log m_*
\]
for some constants \( C, C' > 0 \), \( 0 < c < 1/8 \) and small \( \tau > 0 \), which implies that for any \( 1 \leq m \leq m_* \),
\[
\frac{1}{m_*} \sum_{m=1}^{m_*} K(\mathbb{P}_m \mid \mathbb{P}_0) \leq c \log m_*
\]
for some \( 0 < c < 1/8 \), so we can use Fano’s lemma,
\[
\inf_{\Omega_n} \sup_{\Omega_m \in \mathcal{G}_{12}} \mathbb{E}_m \| \widehat{\Omega}_n - \Omega_m \|_{\infty} \geq \min_{0 \leq m \neq m' \leq m_*} \frac{\| \Omega_m - \Omega_{m'} \|_{\infty}}{4} \frac{m_*^{1/2}}{1 + m_*^{1/2}} \left( 1 - 2c - \left( \frac{2c}{\log m_*} \right)^{1/2} \right).
\]

It completes the proof. For more details about Fano’s lemma, see Tsybakov (2008).

For the case of exponentially decreasing \( \gamma(k) = Ce^{-\beta k} \), consider \( k = \min([\log n \log p]^{1/2}, p) \) for \( \mathcal{G}_{11} \) instead of \( k = \min(n^{1/(2\alpha+2)}, p) \). Then, similar
arguments for the lower bound of $G_{11}$ give the desired result.

For the exact banding $\gamma(k)$, consider $G_{11}$ with $k = k_0$ and $a = (\log p/n)^{1/2}$, then it completes the proof. \hfill \Box

### S2.3 Proof of the P-loss Convergence Rates: Theorem 2 and Theorem 4

Lemma 1–5 are used to prove the main theorems.

**Lemma 1.** Let $X_1, \ldots, X_n \overset{iid}{\sim} N_p(0, \Omega_{0,n}^{-1})$ with $\Omega_{0,n} \in U(\epsilon_0, \gamma)$ defined at (2.8),

\[
N_{1n} = \left\{ X_n : \max_j \| \widehat{\text{Var}}(X_{1,(j-k):j}) \| \leq C_1 \right\},
\]

\[
N_{2n} = \left\{ X_n : \max_j \| \widehat{\text{Var}}^{-1}(X_{1,(j-k):j}) \| \leq C_2 \right\},
\]

\[
N_{3n} = \left\{ X_n : \max_j \| \widehat{\text{Var}}(X_{1,(j-k):j}) - \text{Var}(X_{1,(j-k):j}) \| \leq \left( C_3(k + \log(n \vee p))/n \right)^{1/2} \right\},
\]

\[
N_{4n} = \left\{ X_n : \max_j \| \widehat{\text{Var}}^{-1}(X_{1,(j-k):j}) - \text{Var}^{-1}(X_{1,(j-k):j}) \| \leq \left( C_4(k + \log(n \vee p))/n \right)^{1/2} \right\},
\]

where $C_1 = \epsilon_0^{-1}(2 + ((k + 1)/n)^{1/2})^2$, $C_2 = 4\epsilon_0^{-1}(1 - ((k + 1)/n)^{1/2})^{-2}$, $C_3 = C_3 C_2^2 \epsilon_0^{-2}$ and $N_n = \bigcap_{j=1}^{4} N_{jn}$. If $k + \log p = o(n)$ and $1 \leq k \leq p - 1$, then for any large constant $C_3$, there exists a positive constant $C_5$ such that

\[
P_{0n}(X_n \in N_n^c) \leq 6p e^{-n(1-((k+1)/n)^{1/2})^2/8 + 4 \times 5^k e^{-C_3 C_5 \epsilon_0^2 \log(n \vee p) + k}},
\]

for all sufficiently large $n$. Here, $C_5$ does not depend on $C_3$. 

Proof of Lemma 1. We will show that for any large constant \( C_3 \),

\[
\mathbb{P}_0(n \mathbb{N}_{c_1}^c) \leq 2pe^{-n/2}, \\
\mathbb{P}_0(n \mathbb{N}_{c_2}^c) \leq 2pe^{-n(1-((k+1)/n)^{1/2})^2/8}, \\
\mathbb{P}_0(n \mathbb{N}_{c_3}^c) \leq 2 \times 5^k \exp\{-C_3C_5c_0^2(k+\log(n\vee p))\}, \\
\mathbb{P}_0(n \mathbb{N}_{c_4}^c) \leq 2 \times 5^k \exp\{-C_3C_5c_0^2(k+\log(n\vee p))\} + 2pe^{-n(1-\sqrt{(k+1)/n})^2/8},
\]

for some positive constants \( C_4 \) and \( C_5 \). The inequalities (S2.19) and (S2.20) follow from Lemma B.7 in Lee and Lee (2018). Note that for any large constant \( C_3 > 0 \),

\[
\mathbb{P}_0(n \mathbb{N}_{c_3}^c) \leq p5^{k+1} \left( \exp\{-C_3C_6c_0^2(k+\log(n\vee p))\} + \exp\{-C_3^{1/2}C_7c_0\{n(k+\log(n\vee p))\}^{1/2}\} \right),
\]

for all sufficiently large \( n \) and some positive constants \( C_6 \) and \( C_7 \) by Lemma B.6 in Lee and Lee (2018). If we take \( C_5 = C_6/2 \), the right hand side (RHS) of (S2.23) is bounded by \( 2 \times 5^k \exp\{-C_3C_5c_0^2(k+\log(n\vee p))\} \) for any constant \( C_3 > 0 \) and all sufficiently large \( n \) because \( k + \log(n \vee p) = o(n) \). Similarly,

\[
\mathbb{P}_0(n \mathbb{N}_{c_4}^c) \leq 2 \times 5^k \exp\{-C_3C_5c_0^2(k+\log(n\vee p))\} + 2pe^{-n(1-((k+1)/n)^{1/2})^2/8}
\]
for $C_4 = C_3^2 \epsilon_0^{-2}$ and all sufficiently large $n$. Since the inequalities (S2.21) and (S2.22) also hold, this completes the proof.

Lemma 2. Consider model $X_1, \ldots, X_n \overset{iid}{\sim} N_p(0, \Omega_{0,n}^{-1})$ with $\Omega_{0,n} \in U(\epsilon_0, \gamma)$ defined at (2.8) and $\sum_{m=1}^{\infty} \gamma(m) < \infty$. Denote $\hat{\Omega}_{nk} = (I_p - \hat{A}_{nk})^T \hat{D}_{nk}^{-1} (I_p - \hat{A}_{nk})$, $\hat{D}_{nk} = \text{diag}(\hat{d}_{jk})$ and $\hat{A}_{nk} = (\hat{a}_{jl}^{(k)})$ for $1 \leq k \leq p - 1$, where $(\hat{a}_{j,j-k}^{(k)}, \ldots, \hat{a}_{j,j-1}^{(k)})^T = \hat{a}_{j}^{(k)}$, and $\hat{a}_{j}^{(k)} = 0$ if $1 \leq j \leq l \leq p$ or $|j - l| > k$. $\hat{a}_{j}^{(k)}$ and $\hat{d}_{jk}$ are defined at (2.6). If $k^{3/2}(k + \log(n \vee p)) = O(n)$, then

$$\mathbb{E}_0 n \left[ \|\hat{\Omega}_{nk} - \Omega_{0,n}\| I(X_n \in N_n) \right] \lesssim k^{3/4} \left[ \left( \frac{k + \log(n \vee p)}{n} \right)^{1/2} + \gamma(k) \right],$$

and if $k(k + \log(n \vee p)) = O(n)$, then

$$\mathbb{E}_0 n \left[ \|\hat{\Omega}_{nk} - \Omega_{0,n}\|_{\infty} I(X_n \in N_n) \right] \lesssim k \left[ \left( \frac{k + \log(n \vee p)}{n} \right)^{1/2} + \gamma(k) \right],$$

where the set $N_n$ is defined at Lemma 1.

Proof of Lemma 2. Let

$$A_{0,nk} = (a_{0,jl}^{(k)}) \quad \text{and} \quad D_{0,nk} = \text{diag}(d_{0,jk}), \quad (S2.24)$$

where $(a_{0,j,j-k}^{(k)}, \ldots, a_{0,j,j-1}^{(k)})^T = a_{0,j}^{(k)}$, and $a_{0,jl}^{(k)} = 0$ if $1 \leq j \leq l \leq p$ or
$|j - l| > k$. Define $\Omega_{0,nk} = (I_p - A_{0,nk})^T D_{0,nk}^{-1} (I_p - A_{0,nk})$. Note that

$$
\mathbb{E}_{0n} \left[ \|\hat{\Omega}_{nk} - \Omega_{0,n}\| I(X_n \in N_n) \right] \\
\leq \mathbb{E}_{0n} \left[ \|\hat{\Omega}_{nk} - \Omega_{0,nk}\| I(X_n \in N_n) \right] + \|\Omega_{0,nk} - \Omega_{0,n}\|
$$

(S2.25)

by the triangle inequality (See page 223 of Bickel and Levina (2008)). Also note that

$$
\|I_p - A_{0,nk}\|_{\infty} \leq 1 + \|A_{0,nk} - A_{0,n}\|_{\infty} + \|A_{0,n}\|_{\infty}
$$

$$
\leq 1 + C(k^{1/2} \gamma(k) + 1),
$$

$$
\|I_p - A_{0,nk}\|_1 \leq 1 + \|A_{0,nk} - A_{0,n}\|_{\infty} + \|A_{0,n}\|_1
$$

$$
\leq 1 + Ck \gamma(k) + \sum_{m=1}^{\infty} \gamma(m),
$$

for some constant $C > 0$ by Lemma 10, and $\|D_{0,nk}^{-1}\| \leq \max_j \var^{-1}(X_{1,(j-k):j}) \leq$
$\epsilon_0^{-1}$ using the similar argument to (S3.42). If we show that, on $(X_n \in N_n)$,

$$\| \hat{A}_{nk} - A_{0,nk} \|_{\infty} \lesssim k^{1/2} \left( \frac{k + \log(n \vee p)}{n} \right)^{1/2},$$

(S2.26)

$$\| \hat{A}_{nk} - A_{0,nk} \|_1 \lesssim k \left( \frac{k + \log(n \vee p)}{n} \right)^{1/2},$$

(S2.27)

$$\| \hat{D}_{nk}^{-1} - D_{0,nk}^{-1} \|_{\infty} \lesssim \left( \frac{k + \log(n \vee p)}{n} \right)^{1/2},$$

(S2.28)

$$\| \Omega_{nk} - \Omega_{0,n} \| \lesssim k^{3/4} \gamma(k)$$

and

$$\| \Omega_{nk} - \Omega_{0,n} \| \lesssim \| \Omega_{nk} - \Omega_{0,n} \|_{\infty} \lesssim k \gamma(k),$$

the proof is completed by (S2.25).

To show (S2.26), note that

$$\| \hat{A}_{nk} - A_{0,nk} \|_{\infty} = \max_j \| \hat{a}_{j}^{(k)} - a_{0,j}^{(k)} \|_1$$

$$\leq k^{1/2} \max_j \| \hat{a}_{j}^{(k)} - a_{0,j}^{(k)} \|_2$$

$$= k^{1/2} \max_j \left\{ \Var^{-1}(X_{1,(j-k):(j-1)}) \Cov(X_{1,(j-k):(j-1)}, X_{1j}) 
- \Var^{-1}(X_{1,(j-k):(j-1)}) \Cov(X_{1,(j-k):(j-1)}, X_{1j}) \right\}_2$$

$$\leq k^{1/2} \left\{ \max_j \| \Var^{-1}(X_{1,(j-k):(j-1)}) \Cov(X_{1,(j-k):(j-1)}, X_{1j}) - \Cov(X_{1,(j-k):(j-1)}, X_{1j}) \|_2 
+ \max_j \| \Var^{-1}(X_{1,(j-k):(j-1)}) - \Var^{-1}(X_{1,(j-k):(j-1)}) \Cov(X_{1,(j-k):(j-1)}, X_{1j}) \|_2 \right\}.$$
The first part of the last line can be bounded above by
\[
k^{1/2} \max_j \left\| \Var^{-1}(X_{1,(j-k):(j-1)}) \left( \widehat{\Cov}(X_{1,(j-k):(j-1)}, X_{1j}) - \Cov(X_{1,(j-k):(j-1)}, X_{1j}) \right) \right\|_2 \leq k^{1/2} \max_j \left\| \Var^{-1}(X_{1,(j-k):(j-1)}) \left\| \widehat{\Cov}(X_{1,(j-k):(j-1)}, X_{1j}) - \Cov(X_{1,(j-k):(j-1)}, X_{1j}) \right\|_2 \right\|
\]
\[
\leq k^{1/2} \max_j \left\| \Var^{-1}(X_{1,(j-k):(j-1)}) \left\| \Var(X_{1,(j-k):j}) - \Var(X_{1,(j-k):j}) \right\|_2 \right\|
\]
\[
\lesssim k^{1/2} \left( \frac{k + \log(n \lor p)}{n} \right)^{1/2} \text{ on } (X_n \in N_n).
\]

The first inequality holds by the definition of the spectral norm, and the second inequality holds because the spectral norm of a matrix is larger than a \(\ell_2\) norm of any columns. The second part can be bounded similarly
\[
k^{1/2} \max_j \left\| \left( \Var^{-1}(X_{1,(j-k):(j-1)}) - \Var^{-1}(X_{1,(j-k):(j-1)}) \right) \Cov(X_{1,(j-k):(j-1)}, X_{1j}) \right\|_2 \leq k^{1/2} \max_j \left\| \Var^{-1}(X_{1,(j-k):(j-1)}) \right\| \left\| \Var(X_{1,(j-k):j}) \right\|
\]
\[
\lesssim k^{1/2} \left( \frac{k + \log(n \lor p)}{n} \right)^{1/2} \text{ on } (X_n \in N_n).
\]

By similar arguments, we can show that the inequality (S2.27) holds:
\[
\| \hat{A}_{nk} - A_{0,nk} \|_1 \leq k \max_j \| \hat{a}_{j}^{(k)} - a_{0,j}^{(k)} \|_{\text{max}} \leq k \max_j \| \hat{a}_{j}^{(k)} - a_{0,j}^{(k)} \|_2 \lesssim k \left( \frac{k + \log(n \lor p)}{n} \right)^{1/2} \text{ on } (X_n \in N_n).
\]

To show (S2.28), note that
\[
\| \hat{D}_{nk}^{-1} - D_{0,nk}^{-1} \|_\infty \leq \| \hat{D}_{nk}^{-1} \|_\infty \| D_{0,nk}^{-1} \|_\infty \| \hat{D}_{nk} - D_{0,nk} \|_\infty
\]
and \(\|\hat{D}_{nk}^{-1}\|_\infty \cdot \|D_{0,nk}^{-1}\|_\infty \leq \max_j \|\hat{\text{Var}}_{X_1,(j-k):j}\| \cdot \epsilon_0^{-1} \leq C_2 \epsilon_0^{-1}\) on \((X_n \in N_n)\) for \(C_2 > 0\) used in set \(N_{2n}\), by the similar argument to (S3.42). The rest part is easily bounded above as follows:

\[
\|\hat{D}_{nk} - D_{0,nk}\|_\infty = \max_j |\hat{d}_{jk} - d_{0,jk}|
\]

\[
\leq \max_j |\hat{\text{Var}}(X_{1j}) - \text{Var}(X_{1j})|
\]

\[
+ \max_j |\hat{\text{Cov}}(X_{1j}, X_{1,(j-k):(j-1)}) \hat{a}_j^{(k)} - \text{Cov}(X_{1j}, X_{1,(j-k):(j-1)}) a_{0,j}^{(k)}|
\]

\[
\leq \max_j |\hat{\text{Var}}(X_{1j}) - \text{Var}(X_{1j})| + \max_j |\hat{\text{Cov}}(X_{1j}, X_{1,(j-k):(j-1)}) (\hat{a}_j^{(k)} - a_{0,j}^{(k)})|
\]

\[
+ \max_j \left| \left( \hat{\text{Cov}}(X_{1j}, X_{1,(j-k):(j-1)}) - \text{Cov}(X_{1j}, X_{1,(j-k):(j-1)}) \right) a_{0,j}^{(k)} \right|
\]

\[
\lesssim \left( \frac{k + \log(n \lor p)}{n} \right)^{1/2} \quad \text{on } (X_n \in N_n).
\]

Hence, by (S2.25), we have shown that

\[
\mathbb{E}_0 \left[ \|\hat{\Omega}_{nk} - \Omega_{0,nk}\|_{I(X_n \in N_n)} \right] \lesssim k^{3/4} \left( \frac{k + \log(n \lor p)}{n} \right)^{1/2} + \|\Omega_{0,nk} - \Omega_{0,n}\|
\]

when \(k^{3/2}(k + \log(n \lor p)) = O(n)\), and

\[
\mathbb{E}_0 \left[ \|\hat{\Omega}_{nk} - \Omega_{0,nk}\|_{\infty} I(X_n \in N_n) \right] \lesssim k \left( \frac{k + \log(n \lor p)}{n} \right)^{1/2} + \|\Omega_{0,nk} - \Omega_{0,n}\|_{\infty}
\]

when \(k(k + \log(n \lor p)) = O(n)\). The conditions \(k^{3/2}(k + \log(n \lor p)) = O(n)\)

and \(k(k + \log(n \lor p)) = O(n)\) are required due to the term

\[
\mathbb{E}_0 \left[ \|D_{0,nk}^{-1}\| \cdot \|\hat{A}_{nk}^T - A_{0,nk}^T\| \cdot \|\hat{A}_{nk} - A_{0,nk}\| I(X_n \in N_n) \right]
\]

in (S2.25).
If we show that $\|\Omega_{0,nk} - \Omega_{0,n}\| \lesssim k^{3/4} \gamma(k)$ and $\|\Omega_{0,nk} - \Omega_{0,n}\|_\infty \lesssim k \gamma(k)$, this completes the proof. By Lemma 10, we have $\|A_{0,nk} - A_{0,n}\|_\infty \lesssim k^{1/2} \gamma(k)$ and $\|A_{0,nk} - A_{0,n}\|_1 \lesssim k \gamma(k)$. Note that

$$\|D_{0,nk} - D_{0,n}\|_\infty = \max_j \left| a_{0,j}^{(k)T} \text{Var}(X_{1,(j-k):(j-1)}) a_{0,j}^{(k)} - a_{0,j}^{T} \text{Var}(X_{1,1:(j-1)}) a_{0,j} \right|$$

$$= \max_j \left| (0, a_{0,j}^{(k)T}) - a_{0,j}^{T} \text{Var}(X_{1,1:(j-1)}) \left( \begin{pmatrix} 0 \\ a_{0,j}^{(k)} \end{pmatrix} + a_{0,j} \right) \right|$$

$$\leq \|A_{0,nk} - A_{0,n}\|_\infty \max_j \left( \|a_{0,j}^{(k)}\|_2 + \|a_{0,j}\|_2 \right) \|\text{Var}(X_{1,1:(j-1)})\|$$

$$\lesssim k^{1/2} \gamma(k).$$

Thus, it is easy to show that $\|\Omega_{0,nk} - \Omega_{0,n}\| \lesssim k^{3/4} \gamma(k)$ and $\|\Omega_{0,nk} - \Omega_{0,n}\|_\infty \lesssim k \gamma(k)$ by the triangle inequality in (S2.25). \(\square\)

**Lemma 3.** Consider model $X_1,\ldots,X_n \overset{iid}{\sim} N_p(0, \Omega_{0,n}^{-1})$ and the k-BC prior. Assume that $\Omega_{0,n} \in \mathcal{U}(\epsilon_0, \gamma)$ defined at (2.8) and $\sum_{m=1}^{\infty} \gamma(m) < \infty$. Let

$$\pi(d_j \mid X_n) = IG \left( d_j \mid \frac{n_j}{2}, \frac{n}{2} \hat{d}_{jk}, d_j \leq M \right),$$

$$\bar{\pi}(d_j \mid X_n) = IG \left( d_j \mid \frac{n_j}{2}, \frac{n}{2} \hat{\bar{d}}_{jk} \right),$$

for $j = 1,\ldots,p$, where $\hat{d}_{jk}$ defined at (2.6). If $M \geq 9\epsilon_0^{-1}$, $\nu_0 = o(n)$, $k + \log p = o(n)$ and $1 \leq k \leq p - 1$, then on $(X_n \in N_n)$,

$$\pi(A_n, D_n \mid X_n) = \pi(d_1 \mid X_n) \prod_{j=2}^{p} \pi(a_j \mid d_j, X_n) \pi(d_j \mid X_n)$$

$$\leq \bar{\pi}(d_1 \mid X_n) \prod_{j=2}^{p} \pi(a_j \mid d_j, X_n) \bar{\pi}(d_j \mid X_n)$$

(S2.29)

for all sufficiently large $n$, where the set $N_n$ is defined at Lemma 1.
Proof of Lemma 3. We have

\[
\pi(d_j \mid X_n) = \frac{IG\left(d_j \mid n_j/2, n\hat{d}_{jk}/2\right) I(d_j \leq M)}{\int_0^M IG\left(d_j' \mid n_j/2, n\hat{d}_{jk}/2\right) dd_j'}
\]

for \( j = 1, \ldots, p \). To show (S2.29), it suffices to prove, on \((X_n \in N_n)\),

\[
\left[ \min_j \tilde{\pi}(d_j \leq M \mid X_n) \right]^{-p} \leq C
\]

for some constant \( C > 0 \). Note that on \((X_n \in N_n)\), \( C^{-1} \leq \hat{d}_{jk}^{-1} \leq C_2 \) and

\[
\tilde{\pi}(d_j \leq M \mid X_n) = \tilde{\pi}(M^{-1} \leq d_j^{-1} \mid X_n)
\]

\[= \tilde{\pi}\left(M^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1} \leq d_j^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1} \mid X_n\right)\]

\[= 1 - \tilde{\pi}\left(d_j^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1} < M^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1} \mid X_n\right) .\]

By page 29 of Boucheron et al. (2013), if \( X \) is a sub-gamma random variable with variance factor \( \nu \) and scale parameter \( c \),

\[
\max \left[ P(X > (2\nu t)\sqrt{2} + ct), P(X < -(2\nu t)\sqrt{2} - ct) \right] \leq e^{-t} \quad (S2.30)
\]

for all \( t > 0 \). Since a centered Gamma\((a, b)\) random variable is a sub-gamma random variable with \( \nu = a/b^2 \) and \( c = 1/b \), applying \( t = nt' \) with \( t' = (M - 2C_1)^2/(8M)^2 < 1 \) to the inequality (S2.30),

\[
e^{-nt'} \geq \tilde{\pi}\left( d_j^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1} < -2\left(\frac{n_j}{n}\right)^{1/2} \hat{d}_{jk}^{-1} (t')^{1/2} - 2\hat{d}_{jk}^{-1} t' \mid X_n\right)\]

\[\geq \tilde{\pi}\left( d_j^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1} < -2\hat{d}_{jk}^{-1} (t')^{1/2} \mid X_n\right)\]

\[\geq \tilde{\pi}\left( d_j^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1} < M^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1} \mid X_n\right) .\]
because \( M \geq 9e_0^{-1} > 2C_1 \) for all sufficiently large \( n \) and \( \nu_0 = o(n) \). Thus, for some constant \( C > 0 \), on \( (X_n \in N_n) \),

\[
\tilde{\pi}(d_j \leq M \mid X_n) \geq 1 - e^{-Cn}, \quad \text{(S2.31)}
\]

and

\[
\left[ \min_j \tilde{\pi}(d_j \leq M \mid X_n) \right]^{-p} \leq (1 - e^{-Cn})^{-p} = (1 - e^{-Cn})^{-e^{Cn} \times p/e^{Cn}} \leq (C')^p/e^{Cn} \to 1
\]

as \( n \to \infty \) for some constant \( C' > 0 \).

**Lemma 4.** Consider the model \( X_1, \ldots, X_n \overset{iid}{\sim} N_p(0, \Omega_0^{-1}) \) and the \( k \)-BC prior. Assume that \( \Omega_{0,n} \in \mathcal{U}(\epsilon_0, \gamma) \) defined at (2.8) and \( \sum_{m=1}^{\infty} \gamma(m) < \infty \).

If \( M \geq 9e_0^{-1}, \nu_0 = o(n), k + \log p = o(n) \) and \( 1 \leq k \leq p-1 \), then

\[
E^\pi \left( \| A_n - \hat{A}_{nk} \|_\infty^2 \mid X_n \right) \leq Ck \left( \frac{k + \log p}{n} \right) \quad \text{on } (X_n \in N_n),
\]

\[
E^\pi \left( \| A_n - \hat{A}_{nk} \|_1^2 \mid X_n \right) \leq Ck \left( \frac{k + \log p}{n} \right) \quad \text{on } (X_n \in N_n),
\]

for some constant \( C > 0 \) and all sufficiently large \( n \), where \( \hat{A}_{nk} \) is defined at Lemma 2.

**Proof of Lemma 4.** Let \( E^\pi(\cdot \mid X_n) \) denote the expectation with respect to \( \tilde{\pi}(d_1 \mid X_n) \prod_{j=2}^{p} \pi(a_j \mid d_j, X_n) \tilde{\pi}(d_j \mid X_n) \) in Lemma 3. Note that on \( (X_n \in \)
\[ N_n, \]

\[
\mathbb{E}^\pi \left( \| A_n - \hat{A}_{nk} \|_\infty^2 \mid X_n \right) \\
\leq k \mathbb{E}^\pi \left( \max_j \| a_j - \hat{a}_j^{(k)} \|_2^2 \mid X_n \right) \\
\leq k \mathbb{E}^\pi \left( \max_j \frac{d_j}{n} \left\| \text{Var}^{-1}(X_{1,(j-k):(j-1)}) \left( \frac{n}{d_j} \right)^{1/2} \text{Var}^{1/2}(X_{1,(j-k):(j-1)}) (a_j - \hat{a}_j^{(k)}) \right\|_2^2 \mid X_n \right) \\
\leq \frac{kMC_2}{n} \mathbb{E}^\pi \left( \max_j \left\| \left( \frac{n}{d_j} \right)^{1/2} \text{Var}^{1/2}(X_{1,(j-k):(j-1)}) (a_j - \hat{a}_j^{(k)}) \right\|_2^2 \mid X_n \right) \\
\leq \frac{k}{n} \mathbb{E}^\pi \left( \max_j \left\| \left( \frac{n}{d_j} \right)^{1/2} \text{Var}^{1/2}(X_{1,(j-k):(j-1)}) (a_j - \hat{a}_j^{(k)}) \right\|_2^2 \mid X_n \right) \\
= \frac{k}{n} \mathbb{E} \left( \max_j \chi_{jk}^2 \right)
\]

by Lemma 3. \( \chi_{jk}^2 \) is a chi-square random variable with \( k_j = \min(j - 1, k) \) degree of freedom. By the maximal inequality for chi-square random variables (Example 2.7 in Boucheron et al. (2013)),

\[
\mathbb{E} \left( \max_j \chi_{jk}^2 \right) = k_j + \mathbb{E} \left( \max_j \chi_{jk}^2 - k_j \right) \\
\leq C (k + \log p)
\]

for some constant \( C > 0 \). Thus, we have

\[
\mathbb{E}^\pi \left( \| A_n - \hat{A}_{nk} \|_\infty^2 \mid X_n \right) \leq C k \left( \frac{k + \log p}{n} \right)
\]
on \( (X_n \in N_n) \), for some constant \( C > 0 \).

Let \( a_{c_j} = (a_{j+1,j}, \ldots, a_{\min(j+k,p),j})^T \) be the nonzero column vector of \( A_n \).

Since the posterior distributions for \( a_{c_j} \)'s are the independent multivariate normal distributions with finite variances whose rate is \( 1/n \) on \( (X_n \in N_n) \),
it is easy to show that

\[
\mathbb{E}^\pi \left( \|A_n - \hat{A}_{nk}\|_1^2 \mid X_n \right) \leq Ck \left( \frac{k + \log p}{n} \right)
\]

on \((X_n \in N_n)\), for some constant \(C > 0\) using similar arguments.

**Lemma 5.** Consider the model \(X_1, \ldots, X_n \overset{iid}{\sim} N_p(0, \Omega_{0,n}^{-1})\) and the \(k\)-BC prior. Assume that \(\Omega_{0,n} \in \mathcal{U}(\epsilon_0, \gamma)\) defined at (2.8) and \(\sum_{m=1}^{\infty} \gamma(m) < \infty\). If \(M \geq 9\epsilon_0^{-1}, \nu_0 = o(n), k + \log p = o(n), 1 \leq k \leq p - 1\) and \(k^2 = O(n \log p)\), then

\[
\mathbb{E}^\pi \left( \|D_n^{-1} - \hat{D}_{nk}^{-1}\|_\infty \mid X_n \right) \leq C \left( \frac{\log p}{n} \right)^{1/2} \text{ on } (X_n \in N_n)
\]

for some constant \(C > 0\) and all sufficiently large \(n\), where \(\hat{D}_{nk}\) is defined at Lemma 2.

**Proof of Lemma 5.** By Lemma 3, on \((X_n \in N_n)\),

\[
\mathbb{E}^\pi \left( \|D_n^{-1} - \hat{D}_{nk}^{-1}\|_\infty \mid X_n \right) \leq C \mathbb{E}^\tilde{\pi} \left( \|D_n^{-1} - \hat{D}_{nk}^{-1}\|_\infty \mid X_n \right)
\]

for some constant \(C > 0\). It is easy to show that

\[
\mathbb{E}^\tilde{\pi} \left( \|D_n^{-1} - \hat{D}_{nk}^{-1}\|_\infty \mid X_n \right) \leq \mathbb{E}^\tilde{\pi} \left( \max_j |d_j^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1}| \mid X_n \right) + \max_j \left| \frac{n - n_j}{n} \hat{d}_{jk}^{-1} \right|
\]

\[
\leq \frac{1}{\lambda} \log \exp \mathbb{E}^\tilde{\pi} \left( \lambda \max_j |d_j^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1}| \mid X_n \right) \frac{2k}{n} C_2
\]

\[
\leq \frac{1}{\lambda} \log \left[ \frac{p}{\lambda} \max_j \mathbb{E}^\tilde{\pi} \left( e^{\lambda |d_j^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1}|} \mid X_n \right) \right] + \frac{2k}{n} C_2
\]
for any \( \lambda > 0 \), on \( (X_n \in N_n) \). Let \( \lambda < n\hat{d}_{jk}/2 \). Note that the upper bound for the moment generating function of \( |d_j^{-1} - n_j\hat{d}_{jk}^{-1}/n| \) is

\[
\mathbb{E}^\pi \left( e^{\lambda |d_j^{-1} - n_j\hat{d}_{jk}^{-1}|} \mid X_n \right) = \int_0^\infty e^{\lambda |d_j^{-1} - n_j\hat{d}_{jk}^{-1}|} \Gamma \left( d_j^{-1} \mid \frac{n_j}{2}, \frac{n}{2}\hat{d}_{jk} \right) dd_j^{-1} \\
\leq \int_0^{n_j\hat{d}_{jk}/n} e^{\lambda |n_j\hat{d}_{jk}^{-1} - d_j^{-1}|} \Gamma \left( d_j^{-1} \mid \frac{n_j}{2}, \frac{n}{2}\hat{d}_{jk} \right) dd_j^{-1} \\
+ \mathbb{E}^\pi \left( e^{\lambda (d_j^{-1} - n_j\hat{d}_{jk}^{-1})} \mid X_n \right) \\
\leq e^{\lambda n_j\hat{d}_{jk}/n} \int_0^\infty e^{-\lambda d_j^{-1}} \Gamma \left( d_j^{-1} \mid \frac{n_j}{2}, \frac{n}{2}\hat{d}_{jk} \right) dd_j^{-1} \\
+ \exp \left( \frac{n_j\lambda^2}{n\hat{d}_{jk}(n\hat{d}_{jk} - 2\lambda)} \right) \\
\leq e^{\lambda n_j\hat{d}_{jk}/n} \left( \frac{n\hat{d}_{jk}}{n\hat{d}_{jk} + 2\lambda} \right)^{n_j/2} + \exp \left( \frac{n_j\lambda^2}{n\hat{d}_{jk}(n\hat{d}_{jk} - 2\lambda)} \right) .
\]

The second inequality follow from page 28 of Boucheron et al. (2013). Since \( \lambda < n\hat{d}_{jk}/2 \),

\[
e^{\lambda n_j\hat{d}_{jk}/(n\hat{d}_{jk} + 2\lambda)} \left( \frac{n\hat{d}_{jk}}{n\hat{d}_{jk} + 2\lambda} \right)^{n_j/2} = e^{\lambda n_j/(n\hat{d}_{jk})} \left( 1 + \frac{2\lambda}{n\hat{d}_{jk}} \right)^{-n_j/2} \\
\leq \left( 1 + \frac{2\lambda}{n\hat{d}_{jk}} \right)^{\lambda n_j/(2n\hat{d}_{jk})} \\
= \left( 1 + \frac{2\lambda}{n\hat{d}_{jk}} \right)^{n\hat{d}_{jk}/(2\lambda) \lambda^2 n_j/(n^2\hat{d}_{jk}^2)} \\
\leq \exp \left( \frac{\lambda^2 n_j}{n^2\hat{d}_{jk}^2} \right) ,
\]
where the first inequality follows from Lemma 7. Thus, on \((X_n \in N_n)\),

\[
\mathbb{E}^\pi \left( \|D_n^{-1} - \hat{D}_{nk}^{-1}\|_\infty \mid X_n \right) \leq \frac{1}{\lambda} \log \left[ p \max_j \mathbb{E}^\pi \left( e^{\lambda i d_j^{-1} - \frac{n_j}{n} \hat{d}_{jk}^{-1}} \mid X_n \right) \right] + \frac{2k}{n} C_2 \\
\leq \frac{\log p}{\lambda} + \frac{1}{\lambda} \max_j \left[ \exp \left( \frac{\lambda^2 n_j}{n^2 d_{jk}^2} \right) + \exp \left( \frac{n_j \lambda^2}{n d_{jk} (n d_{jk} - 2 \lambda)} \right) \right] \\
+ \frac{2k}{n} C_2 \\
\leq \frac{\log p}{\lambda} + \frac{2 \log 2}{\lambda} + \max_j \left( \frac{\lambda n_j}{n^2 d_{jk}^2} + \frac{n_j \lambda}{n d_{jk} (n d_{jk} - 2 \lambda)} \right) + \frac{2k}{n} C_2 \\
\leq \frac{\log p}{\lambda} + \frac{2 \log 2}{\lambda} + \frac{\lambda C_2}{n} + \frac{\lambda C_2}{(n C_2^{-1} - 2 \lambda)} + \frac{2k}{n} C_2 \\
\leq C \left( \frac{\log p}{n} \right)^{1/2}
\]

for some constant \(C > 0\) if we choose \(\lambda \asymp (n \log p)^{1/2}\). \( \square \)

**Proof of Theorem 2.** Note that

\[
\mathbb{E}_{0n} \mathbb{E}^\pi \left( \|\Omega_n - \Omega_{0,n}\| \mid X_n \right) \\
\leq \mathbb{E}_{0n} \mathbb{E}^\pi \left( \|\Omega_n - \Omega_{0,n}\| \mid X_n \right) I(X_n \in N_n) \quad \text{(S2.32)} \\
+ \mathbb{E}_{0n} \mathbb{E}^\pi \left( \|\Omega_n - \Omega_{0,n}\| \mid X_n \right) I(X_n \in N_n^c) \quad \text{(S2.33)}
\]

where the set \(N_n\) is defined at Lemma 1. The term (S2.33) is bounded
above by
\[
\mathbb{E}_0 \left[ (\mathbb{E}^\pi(\|\Omega_n\| | X_n) + \|\Omega_{0,n}\|) I(X_n \in N_n^c) \right] \\
\leq \mathbb{E}_0 \left[ (\mathbb{E}^\pi(\|I_p - A_n\|_1 + \|I_p - A_n\|_\infty D_n^{-1} | X_n) + \|\Omega_{0,n}\|) I(X_n \in N_n^c) \right] \\
\leq \left\{ \mathbb{E}_0 \left[ \mathbb{E}^\pi(\|I_p - A_n\|_1 + \|I_p - A_n\|_\infty D_n^{-1} | X_n) \right]^2 \right\}^{1/2} P_0(X_n \in N_n^c)^{1/2} \\
+ \|\Omega_{0,n}\| \mathbb{P}_0(X_n \in N_n^c) \\
\leq p^\kappa P_0(X_n \in N_n^c)^{1/2} + \|\Omega_{0,n}\| \mathbb{P}_0(X_n \in N_n^c) \\
\leq (p^\kappa + C) \left( 6p e^{-n((k+1)/n)^{1/2}) / 4 + 4 \times 5^k e^{-C_3 C_5 \kappa \log(n \sqrt{p})} \right)^{1/2} \\
\leq n^{-1}
\]

for all sufficiently large \( n \) and some positive constants \( \kappa, \kappa_3 \) and \( \kappa_5 \). The fourth inequality follows from Lemmas 1 and 8. The third inequality holds because
\[
\left[ \mathbb{E}^\pi(\|I_p - A_n\|_1 + \|I_p - A_n\|_\infty D_n^{-1} | X_n) \right]^2 \\
\leq \left[ \mathbb{E}^\pi(p^3 \max_{j,l} \|I_p - A_n\|_\infty \cdot \max_j \|D_n^{-1} \|_{\max} | X_n) \right]^2 \\
\leq p^6 \left[ \mathbb{E}^\pi \left( (1 + \sum_{j,l} a_{jl})^2 \cdot \sum_j d_j^{-1} | X_n) \right)^2 \\
\leq 4p^6 \left( \sum_j \mathbb{E}^\pi \left( d_j^{-1} | X_n) + \mathbb{E}^\pi \left( (\sum_{j,l} a_{jl})^2 \cdot \sum_j d_j^{-1} | X_n) \right)^2 \\
\leq 4p^6 \left( p \max_j \mathbb{E}^\pi \left( d_j^{-1} | X_n) + p^5 \mathbb{E}^\pi \left( \max_{j,j'} d_{j,l}^2 d_{j',l}^{-1} | X_n) \right)^2 \\
\leq 4p^6 \left( p \max_j \mathbb{E}^\pi \left( d_j^{-1} | X_n) + p^8 \max_{j,j',l} \mathbb{E}^\pi \left( \max_j a_{jl}^2 d_{j,l}^{-1} | X_n) \right)^2 \\
\leq 4p^6 \left( p \max_j \sum_{j,l} \hat{a}_{jl}^{-1} + p^8 \max_{j,j',l} \left( (\hat{a}_{jl}^{(k)})^2 + M[(n \text{Var}(X_{1,(j-k):(j-1)})^{-1})_{l-j+k+1,l-j+k+1}] \right) \right) n_{j,l}^{-1} d_{j,l}^{-1} \right)^2,
\]

S2. PROOFS
whose expectation is bounded above by $p^c$ for some constant $c > 0$ by Lemma 6 and its proof, where the fifth and sixth inequalities follow from Lemma 3.

We decompose the term (S2.32) as follows:

$$
E_n^0 \left[ E^\pi \left( \| \Omega - \Omega_{0,n} \| \mid X_n \right) I(X_n \in N_n) \right] 
\leq E_n^0 \left[ E^\pi \left( \| \Omega - \hat{\Omega}_{nk} \| \mid X_n \right) I(X_n \in N_n) \right] 
+ E_n^0 \left[ \| \hat{\Omega}_{nk} - \Omega_{0,n} \| I(X_n \in N_n) \right],
$$

(S2.34)

(S2.35)

where $\hat{\Omega}_{nk}$ is defined at Lemma 2. By Lemma 2, the upper bound for (S2.35) is $Ck^{3/4}[(k + \log(n \lor p))/n]^{1/2} + \gamma(k)$ for some constant $C > 0$ because we assume that $k^{3/2}(k + \log(n \lor p)) = O(n)$. Note that the term (S2.34) can be decomposed as (S2.25) and

$$\| I_p - \hat{A}_{nk} \|_1 \leq \| I_p - A_{0,nk} \|_1 + \| \hat{A}_{nk} - A_{0,nk} \|_1$$

$$\| I_p - \hat{A}_{nk} \|_\infty \leq \| I_p - A_{0,nk} \|_\infty + \| \hat{A}_{nk} - A_{0,nk} \|_\infty$$

$$\| I_p - \hat{A}_{nk} \| \leq \| I_p - A_{0,nk} \| + \| \hat{A}_{nk} - A_{0,nk} \|$$

and $\| \hat{D}_{nk}^{-1} \| \leq C_2$ on $(X_n \in N_n)$ for some constant $C > 0$. By Lemma 4 and Lemma 5, it is easy to show that the upper bound for (S2.34) is...
$C k^{1/2}((k + \log(n \lor p))/n)^{1/2}$ for some constant $C > 0$ because we assume that $k^{3/2}(k + \log(n \lor p)) = O(n)$.

**Proof of Theorem 4.** We can use the same arguments used in the proof of Theorem 2. It suffices to prove that

$$
\|I_p - \hat{A}_{nk}\|_1 \lesssim k^{1/2} \text{ on } (X_n \in N_n).
$$

It trivially holds because we assume that $k(k + \log(n \lor p)) = O(n)$.

**S2.4 Proof of Corollary 1**

Lemma 6 is used to prove Corollary 1.

**Lemma 6.** Consider the model $X_1, \ldots, X_n \overset{iid}{\sim} N_p(0, \Omega_{0,n}^{-1})$ and $\Omega_{0,n} \in \mathcal{U}(\epsilon_0, \gamma)$ defined at (2.8). If $k = o(n)$, then for given positive integer $m$,

$$
\mathbb{E}_0n(\hat{d}_{jk}^{-m}) \lesssim (k + 1)^{m+1},
$$

$$
\mathbb{E}_0n((\hat{a}_{ji}^{(k)})^m) \lesssim (k + 1)^{2m+1},
$$

where $\hat{d}_{jk}$ and $\hat{a}_{ji}^{(k)}$ be defined at (2.6).

**Proof.** Note that

$$
\mathbb{E}_0n(\hat{d}_{jk}^{-m}) \leq \mathbb{E}_0n\left\|\text{Var}^{-1}(X_{1,(j-k);j})\right\|^m
$$

$$
\leq \mathbb{E}_0n\left[\text{tr}\left(\text{Var}^{-1}(X_{1,(j-k);j})\right)^m\right]
$$

$$
\leq (k + 1)^m \sum_{l=1}^{k+1} \mathbb{E}_0n\left[\text{Var}^{-1}(X_{1,(j-k);j}(l))\right]^m
$$
KYOUNGJAЕ LEE AND JAEYONG LEE

where for any $p \times p$ matrix $A$, $A_{(i)}$ is the $(i, i)$ component of $A$. Also note that $[\widehat{\text{Var}}^{-1}(X_{1,(j-k):j})]_{(l)}$ is a inverse-gamma distribution $IG((n - k)/2, n[\text{Var}^{-1}(X_{1,(j-k):j})]_{(l)}/2)$ because diagonal elements of an inverse-Wishart matrix are inverse-gamma random variables (Huang and Wand, 2013).

Since $\Omega_{0,n} \in \mathcal{U}(\epsilon_0, \gamma)$,

$$(k + 1)^m \sum_l \mathbb{E}_0 \left[ [\widehat{\text{Var}}^{-1}(X_{1,(j-k):j})]_{(l)} \right]^m \leq (k + 1)^{m+1} \left( \frac{n\epsilon_0^{-1}}{n - k - 2m} \right)^m \approx (k + 1)^{m+1}.$$  

Similarly,

$$\mathbb{E}_0(\hat{a}_{ji}^{(k)})^m \leq \mathbb{E}_0 \left[ \|\widehat{\text{Var}}^{-1}(X_{1,(j-k):j})\|_2^m \|\widehat{\text{Var}}(X_{1,(j-k):j})\|_2^m \right] \leq \mathbb{E}_0 \left\{ \left[ \text{tr} (\widehat{\text{Var}}^{-1}(X_{1,(j-k):j})) \right]^m \left[ \text{tr} (\widehat{\text{Var}}(X_{1,(j-k):j})) \right]^m \right\} \leq \left\{ \mathbb{E}_0 \left[ \text{tr} (\widehat{\text{Var}}^{-1}(X_{1,(j-k):j})) \right]^{2m} \mathbb{E}_0 \left[ \text{tr} (\widehat{\text{Var}}(X_{1,(j-k):j})) \right]^{2m} \right\}^{1/2} \approx (k + 1)^{2m+1}$$

because diagonal elements of a Wishart matrix are gamma random variables (Rao, 2009), i.e. $[\widehat{\text{Var}}(X_{1,(j-k):j})]_{(l)} \sim \text{Gamma}(n/2, n[\text{Var}(X_{1,(j-k):j})]_{(l)}/2)$.

Proof of Corollary 1. Since

$$\mathbb{E}_0 \|\hat{\Omega}_{nk}^{LL} - \Omega_{0,n}\| \leq \mathbb{E}_0 \|\hat{\Omega}_n - \Omega_{n,n}\| + \mathbb{E}_0 \|\hat{\Omega}_n - \Omega_{nk}\| \leq \mathbb{E}_0 \|\hat{\Omega}_n - \Omega_{n,n}\| + \mathbb{E}_0 \|\hat{\Omega}_n - \hat{\Omega}_{nk}\|.$$
it suffices to prove

\[ E_{0n} \| E^\pi(\Omega_n \mid X_n) - \hat{\Omega}_{nk}^{LL} \|_\infty \leq \frac{Ck^2}{n} \leq k^{3/4} \left[ \left( \frac{k + \log(n \lor p)}{n} \right)^{1/2} + \gamma(k) \right] \]

for some constant \( C > 0 \) because of the assumption \( k(k + \log(n \lor p)) = O(n) \).

Let \( \Omega_n = (\omega_{ij}) \) and \( \hat{\Omega}_{nk}^{LL} = (\hat{\omega}_{ij}^{LL}) \), then for \( i < j \leq i + k \),

\[
E_{0n} \left| \mathbb{E}^\pi(\omega_{ij} \mid X_n) - \hat{\omega}_{ij}^{LL} \right|
\leq E_{0n} \left| \mathbb{E}^\pi(d_{ij}^{-1}a_{ji} \mid X_n) - \frac{n_j}{n} \hat{d}_{jk}^{-1}a_{ji} \right| \quad (S2.36)
+ \sum_{l=j+1}^{i+k} E_{0n} \left| \mathbb{E}^\pi(d_l^{-1}a_{li}a_{lj} \mid X_n) - \frac{n_l}{n} \hat{d}_{lk}^{-1}a_{li}a_{lj} \right| \quad (S2.37)
\]

by (S2.4). The (S2.36) term can be decomposed by

\[
E_{0n} \left| \left( \mathbb{E}^\pi(d_{ij}^{-1}a_{ji} \mid X_n) - \frac{n_j}{n} \hat{d}_{jk}^{-1}a_{ji} \right) I(X_n \in N_n) \right| \quad (S2.38)
+ E_{0n} \left| \left( \mathbb{E}^\pi(d_l^{-1}a_{li}a_{lj} \mid X_n) - \frac{n_l}{n} \hat{d}_{lk}^{-1}a_{li}a_{lj} \right) I(X_n \in N_n^c) \right|. \quad (S2.39)
\]

To deal with the above terms, we need to compute the expectation of truncated distributions. When \( Y \) is a truncated gamma distribution \( Y \sim Gamma^{Tr}(\alpha, \beta, c_1 \leq Y \leq c_2) \), the expectation of \( Y \) is

\[
\mathbb{E}Y = \frac{\alpha}{\beta} \int_{c_1}^{c_2} \text{Gamma}(y \mid \alpha + 1, \beta)dy
\]

(Coffey and Muller, 2000). Thus, one can show that (S2.38) is bounded
above by
\[
\mathbb{E}_0 n \left| \frac{n_j}{n} \hat{d}_{jk}^{-1} \hat{a}_{ji}^{(k)} \left( \int_0^M \text{Gamma}(d_j^{-1} \mid \frac{n_j}{2} + 1, \frac{n}{2} \hat{d}_{jk}^{d_j^{-1}}) - 1 \right) I(X_n \in N_n) \right| 
\leq C_1 C_2^2 e^{-cn}
\]
for all sufficiently large \( n \) and some positive constant \( c \) by the same argument with (S2.31). On the other hand, (S2.39) is bounded above by
\[
C \left[ \mathbb{E}_0 n (\hat{d}_{jk}^{-2} (\hat{a}_{ji}^{(k)})^2) \right]^{1/2} \mathbb{P}_0 (X_n \in N_n^c) \leq (k + 1)^{7/2} \mathbb{P}_0 (X_n \in N_n^c) \leq \frac{1}{n^2}
\]
for some constant \( C > 0 \) and all sufficiently large \( n \) by Lemma 1, Lemma 6 and the choice of large \( C_3 \) in the set \( N_n \).

The (S2.37) can be decomposed by
\[
\sum_{l=j+1}^{i+k} \mathbb{E}_0 n \left| \left( \mathbb{E}^\pi (d_i^{-1} a_i a_j \mid X_n) - \frac{n_l}{n} \hat{d}_{lk}^{-1} \hat{a}_{li}^{(k)} \hat{a}_{lj}^{(k)} \right) I(X_n \in N_n) \right| (S2.40)
\]
\[
+ \sum_{l=j+1}^{i+k} \mathbb{E}_0 n \left| \left( \mathbb{E}^\pi (d_i^{-1} a_i a_j \mid X_n) - \frac{n_l}{n} \hat{d}_{lk}^{-1} \hat{a}_{li}^{(k)} \hat{a}_{lj}^{(k)} \right) I(X_n \in N_n^c) \right| (S2.41)
\]
Note that in (S2.40),
\[
\mathbb{E}^\pi (d_i^{-1} a_i a_j \mid X_n) = \mathbb{E}^\pi (d_i^{-1} \mathbb{E}^\pi (a_i a_j \mid d_i, X_n) \mid X_n)
\]
\[
= \mathbb{E}^\pi (d_i^{-1} \mathbb{E}^\pi (a_i \mid d_i, X_n) \mathbb{E}^\pi (a_j \mid d_i, X_n) \mid X_n)
\]
\[
+ \mathbb{E}^\pi (d_i^{-1} \text{Cov}^\pi (a_i, a_j \mid d_i, X_n) \mid X_n).
\]
If we prove that \( \sum_{l=j+1}^{i+k} E_{0n} |E^\pi(d_l^{-1}\text{Cov}^\pi(a_{ti}, a_{tj} \mid d_l, X_n) \mid X_n)I(X_n \in N_n)| \lesssim k/n \), (S2.40) is bounded above by \( Ck/n \) for some constant \( C > 0 \) by the similar arguments used in (S2.38). It is easy to show that

\[
\sum_{l=j+1}^{i+k} E_{0n} |E^\pi(d_l^{-1}\text{Cov}^\pi(a_{ti}, a_{tj} \mid d_l, X_n) \mid X_n)I(X_n \in N_n)|
\]

\[
\leq \sum_{l=j+1}^{i+k} E_{0n} \left[ E^\pi\left( d_l^{-1} |\text{Cov}^\pi(a_{ti}, a_{tj} \mid d_l, X_n)| \right) \mid X_n \right] I(X_n \in N_n)
\]

\[
\leq \sum_{l=j+1}^{i+k} E_{0n} \left( E^\pi\left( d_l^{-1} |\text{Var}^\pi(a_{ti} \mid d_l, X_n)\text{Var}^\pi(a_{tj} \mid d_l, X_n)|^{1/2} \right) \mid X_n \right) I(X_n \in N_n)
\]

\[
\lesssim \frac{k}{n}.
\]

Similar to (S2.39), (S2.41) is bounded above by \( C/n^2 \) for some constant \( C > 0 \). Thus, we have shown

\[
E_{0n} |E^\pi(\omega_{ij} \mid X_n) - \hat{\Omega}_{ij}^{LL}| \lesssim \frac{k}{n}
\]

for any \( i < j \leq i+k \). Since \( \omega_{ii} = d_i^{-1} + \sum_{l=i+1}^{i+k} d_l^{-1}a_{ti}^2 \) for \( i < p \) and \( \omega_{pp} = d_p^{-1} \),

\[
E_{0n} |E^\pi(\Omega_{ii} \mid X_n) - \hat{\Omega}_{ii}^{LL}| \lesssim \frac{k}{n}
\]

can be shown easily for \( 1 \leq i \leq p \) by similar arguments. Thus, it implies

\[
E_{0n}\|E^\pi(\Omega_n \mid X_n) - \hat{\Omega}_{nk}^{LL}\|_\infty \lesssim \frac{k^2}{n}.
\]
S3 Auxiliary results

Lemma 7. For any \(x, n > 0\),

\[ e^x \leq \left(1 + \frac{x}{n}\right)^{n+x/2}. \]

The proof can be obtained by a simple algebra.

Lemma 8. If we assume that \(\Omega_{0,n} \in U(\epsilon_0, \gamma)\) (defined at (2.8)) and \(\sum_{k=1}^{\infty} \gamma(k) < \infty\), then

\[ \|\Omega_{0,n}\|_{\infty} < C \]

for some \(C > 0\) not depending on \(p\).

Proof. Let \(\Omega_{0,n} = (I_p - A_{0,n})^T D_{0,n}^{-1} (I_p - A_{0,n})\) be the modified Cholesky decomposition of \(\Omega_{0,n}\). Since \(\|\Omega_{0,n}\|_{\infty} \leq \|I_p - A_{0,n}\|_1 \|D_{0,n}^{-1}\|_\infty \|I_p - A_{0,n}\|_\infty\)
and

\[ \|I_p - A_{0,n}\|_\infty \leq 1 + \|A_{0,n}\|_\infty \leq 1 + \gamma(1), \]

\[ \|D_{0,n}^{-1}\|_\infty = \max_j d_{0,j}^{-1} \]

\[ = \max_j \left\| \text{Var}^{1/2}(X_{1,1:j}) \left( \begin{array}{c} -a_{0,j} \\ 1 \end{array} \right) \right\|_2^{-2} \]

\[ \leq \max_j \lambda_{\min} (\text{Var}(X_{1,1:j}))^{-1} = \max_j \left\| \text{Var}^{-1}(X_{1,1:j}) \right\| \leq \epsilon_0^{-1}, \]

(S3.42)

we only need to prove \(\|A_{0,n}\|_1 \leq C\) for some \(C > 0\). By the definition of
\( U(\epsilon_0, \gamma) \), it is easy to show \( |a_{0,ij}| \leq \gamma(i-j) \) for all \( i > j \). Thus,
\[
\|A_{0,n}\|_1 = \max_j \sum_{i=j+1}^p |a_{0,ij}|
\]
\[
\leq \max_j \sum_{i=j+1}^p \gamma(i-j)
\]
\[
\leq \sum_{m=1}^\infty \gamma(m) < \infty.
\]

\[\square\]

**Lemma 9.** For any positive integers \( p_1 \) and \( p_2 \), let \( A_{11}, A_{12} \) and \( A_{22} \) be a \( p_1 \times p_1, p_1 \times p_2 \) and \( p_2 \times p_2 \) matrix,
\[
\|A_{12}\| \leq \left\| \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{pmatrix} \right\|,
\]
where \( \|\cdot\| \) is the matrix \( L_2 \) norm.

**Proof.** Note
\[
\left\| \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{pmatrix} \right\| = \sup_{\|x\|_2 = 1} \left\| \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{pmatrix} x \right\|_2
\]
\[
\geq \sup_{\|x_2\|_2 = 1} \left\| \begin{pmatrix} A_{12}x_2 \\ A_{22}x_2 \end{pmatrix} \right\|_2 \geq \sup_{\|x_2\|_2 = 1} \|A_{12}x_2\|_2 = \|A_{12}\|
\]
where \( x = (x_1^T, x_2^T)^T \) and \( x_1 \in \mathbb{R}^{p_1}, x_2 \in \mathbb{R}^{p_2} \). This completes the proof. \(\square\)
Lemma 10. If we assume that $\Omega_{0,n} \in \mathcal{U}(\epsilon_0, \gamma)$, which is defined at (2.8), then

$$ \|A_{0,nk} - A_{0,n}\|_\infty \leq Ck^{1/2}\gamma(k),$$

$$ \|A_{0,nk} - A_{0,n}\|_1 \leq Ck\gamma(k) $$

for some $C > 0$, where $A_{0,nk}$ is defined at (S2.24).

Proof of Lemma 10. We only consider $k < j - 1$ case because $A_{0,nk} = A_{0,n}$ trivially holds when $k \geq j - 1$. Note first that

$$ \|A_{0,nk} - A_{0,n}\|_\infty \leq \|A_{0,nk} - B_k(A_{0,n})\|_\infty + \|B_k(A_{0,n}) - A_{0,n}\|_\infty. $$

The second term is bounded above by $\gamma(k)$ by the definition of $\mathcal{U}(\epsilon_0, \gamma)$. Denote

$$ \text{Var}^{-1}(X_{1,1:(j-1)}) = \begin{pmatrix} \Omega_{11,j} & \Omega_{12,j} \\ \Omega_{21,j} & \Omega_{22,j} \end{pmatrix}, $$

$$ \text{Cov}(X_{1,1:(j-1)}, X_{1j}) = \begin{pmatrix} \Sigma_{1j} \\ \Sigma_{2j} \end{pmatrix}, $$

where $\Omega_{11,j}$ is a $(j-k-1) \times (j-k-1)$ matrix, $\Omega_{22,j}$ is a $k \times k$ matrix and $\Sigma_{2j} = \text{Cov}(X_{1,(j-k):(j-1)}, X_{1j})$ is a $k$-dimensional vector. Since $\max_j \|a_{0,j} - B_{k-1,j}(a_{0,j})\|_1 \leq \gamma(k)$ by assumption, it directly implies

$$ \max_j \|\Omega_{11,j} \Sigma_{1j} + \Omega_{12,j} \Sigma_{2j}\|_1 \leq \gamma(k). $$

(S3.43)
Also note that \( \text{Var}^{-1}(X_{1,(j-k):j-1}) = \Omega_{22,j} - \Omega_{21,j}\Omega_{11,j}^{-1}\Omega_{12,j} \) by the inversion of partitioned matrix. With this fact, we have the following upper bound for \( \|A_{0,nk} - B_k(A_{0,n})\|_\infty \),

\[
\|A_{0,nk} - B_k(A_{0,n})\|_\infty = \max_j \|a_{0,j}^{(k)} - B_{k-1,j}(a_{0,j})\|_1 \\
= \max_j \|\Omega_{21,j}\Sigma_{ij} + \Omega_{21,j}\Omega_{11,j}^{-1}\Omega_{12,j}\Sigma_j\|_1 \\
= \max_j \|\Omega_{21,j}\Omega_{11,j}^{-1}(\Omega_{11,j}\Sigma_{ij} + \Omega_{12,j}\Sigma_j)\|_1 \\
\leq \max_j \|\Omega_{21,j}\Omega_{11,j}^{-1}\|_1 \|\Omega_{11,j}\Sigma_{ij} + \Omega_{12,j}\Sigma_j\|_1 \\
\leq \max_j k^{1/2}\|\Omega_{21,j}\Omega_{11,j}^{-1}\| \|\Omega_{11,j}\Sigma_{ij} + \Omega_{12,j}\Sigma_j\|_1 \\
\leq \max_j k^{1/2}\|\Omega_{21,j}\|_1 \|\Omega_{11,j}^{-1}\| \cdot \gamma(k) \\
\leq \epsilon_0^{-2} k^{1/2} \gamma(k).
\]

The second inequality holds because \( \|A\|_1 \leq p_1^{1/2}\|A\| \) for any \( p_1 \times p_2 \) matrix \( A \) (Horn and Johnson, 1990). The third inequality follows from the Cauchy-Schwarz inequality and (S3.43). The last inequality holds because \( \|\Omega_{21,j}\| \leq \|\text{Var}^{-1}(X_{1,1:(j-1)})\| = \lambda_{\min}(\text{Var}(X_{1,1:(j-1)}))^{-1} \leq \lambda_{\min}(\Omega_{0,n})^{-1} \leq \epsilon_0^{-1} \) and \( \|\Omega_{11,j}^{-1}\| = \lambda_{\min}(\Omega_{11,j})^{-1} \leq \lambda_{\min}(\text{Var}^{-1}(X_{1,1:(j-1)}))^{-1} = \lambda_{\max}(\text{Var}(X_{1,1:(j-1)})) \leq \lambda_{\max}(\Omega_{0,n}) \leq \epsilon_0^{-1} \) by Lemma 9 and \( \Omega_{0,n} \in \mathcal{U}(\epsilon_0, \gamma) \). It proves the first part of Lemma 10.

To show the second argument of Lemma 10, note that

\[
\|A_{0,nk} - A_{0,n}\|_1 \leq \|A_{0,nk} - B_k(A_{0,n})\|_1 + \|B_k(A_{0,n}) - A_{0,n}\|_1.
\]
The first term is bounded above by

\[ \| A_{0,nk} - B_k(A_{0,n})\|_1 \leq k \max_j \| a_{0,j}^{(k)} - B_{k-1,j}(a_{0,j})\|_{\text{max}} \]

\[ \leq k \max_j \| a_{0,j}^{(k)} - B_{k-1,j}(a_{0,j})\| \]

\[ = k \max_j \| \Omega_{21,j}\Omega_{11,j}^{-1}(\Omega_{11,j}\Sigma_{1j} + \Omega_{12,j}\Sigma_{2j})\|_2 \]

\[ \leq k \max_j \| \Omega_{21,j}\Omega_{11,j}^{-1}\| \| \Omega_{11,j}\Sigma_{1j} + \Omega_{12,j}\Sigma_{2j}\|_2 \]

\[ \leq \epsilon_0^{-2}k\gamma(k) \]

by the similar arguments used in the previous paragraph. Also note that

\[ \| B_k(A_{0,n}) - A_{0,n}\|_1 = \sum_{i=j+k}^p |a_{0,ij}| \]

\[ \leq \sum_{i=j+k}^p \sum_{j'\leq 1} |a_{0,ij'}| \]

\[ \leq \sum_{i=j+k}^p \gamma(i - j) \]

\[ \leq \sum_{m=k}^\infty \gamma(m). \]

If we assume the polynomially decreasing \( \gamma(k) = Ck^{-\alpha} \), we have \( \sum_{m=k}^\infty \gamma(m) \leq C'k\gamma(k) \) for some constant \( C' > 0 \). If we assume the exact band or exponentially decreasing \( \gamma(k) = Ce^{-\beta k} \), it is easy to show that \( \sum_{m=k}^\infty \gamma(m) \leq C''\gamma(k) \) for some constant \( C'' > 0 \). Thus, \( \| A_{0,nk} - A_{0,n}\|_1 \) is bounded above by \( C''k\gamma(k) \) for some constant \( C'' > 0 \). \( \square \)
References

Assouad, P. (1983). Deux remarques sur l’estimation. *Comptes rendus des séances de l’Académie des sciences. Série 1 Mathématique* 296(23), 1021–1024.

Bickel, P. J. and E. Levina (2008). Regularized estimation of large covariance matrices. *The Annals of Statistics* 36(1), 199–227.

Boucheron, S., G. Lugosi, and P. Massart (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford.

Cai, T. T., C.-H. Zhang, and H. H. Zhou (2010). Optimal rates of convergence for covariance matrix estimation. *The Annals of Statistics* 38(4), 2118–2144.

Cai, T. T. and H. H. Zhou (2012). Minimax estimation of large covariance matrices under 1-norm. *Statistica Sinica* 22(4), 1319–1349.

Coffey, C. S. and K. E. Muller (2000). Properties of doubly-truncated gamma variables. *Communications in Statistics-Theory and Methods* 29(4), 851–857.

Horn, R. and C. Johnson (1990). *Matrix Analysis*. Cambridge University Press.

Huang, A. and M. P. Wand (2013). Simple marginally noninformative prior distributions for covariance matrices. *Bayesian Analysis* 8(2), 439–452.

LeCam, L. (1973). Convergence of estimates under dimensionality restrictions. *The Annals of Statistics* 1(1), 38–53.

Lee, K. and J. Lee (2018). Optimal Bayesian Minimax Rates for Unconstrained Large Covariance
Matrices. *Bayesian Analysis* 13(4), 1211–1229.

Liu, Y. and Z. Ren (2017, December). Minimax Estimation of Large Precision Matrices with Bandable Cholesky Factor. *ArXiv e-prints*.

Rao, C. (2009). *Linear Statistical Inference and its Applications*. Wiley Series in Probability and Statistics. Wiley.

Tsybakov, A. (2008). *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer New York.