Supplementary material

We will use $C_1, C_2, \ldots$ as generic positive constants whose values may change from line to line.

Technical results of Section 2.1

In this section of the appendix we collect proofs of Section 2.1 and some additional technical results.

Some deviation results

Let $\hat{\mathbf{\Sigma}}^\tau = (\hat{\sigma}_{ab}^\tau)$ and $\mathbf{\Sigma}^\tau = (\sigma_{ab}^\tau)$. To bound the element-wise deviation of the weighted sample covariance matrix $\hat{\mathbf{\Sigma}}^\tau$ from the population covariance matrix Σ^τ, we use the following decomposition

$$|\sum_i w_i^\tau x_a^i x_b^i - \sigma_{ab}^\tau| \leq |\hat{\sigma}_{ab}^\tau - \mathbb{E}\hat{\sigma}_{ab}^\tau| + |\mathbb{E}\hat{\sigma}_{ab}^\tau - \sigma_{ab}^\tau|. \quad (22)$$

Standard treatment of the expectation integrals gives us that $|\mathbb{E}\hat{\sigma}_{ab}^\tau - \sigma_{ab}^\tau| = \mathcal{O}(h)$, see for example Tsybakov (2009). The following Lemma characterizes the first term in Equation (22).

Lemma 10. *Let $\tau \in [0,1]$ be a fixed time point. Assume that Σ^τ satisfies the assumptions* **S** *and* **C** *and the kernel function satisfies the assumption* **K**. *Let $\{\mathbf{x}^i\}$ be an independent sample according to the model (1). Then*

$$\mathbb{P}[|\hat{\sigma}_{ab}^\tau - \mathbb{E}\hat{\sigma}_{ab}^\tau| > \epsilon) \leq C_1 \exp(-C_2 nh\epsilon^2), \quad |\epsilon| \leq \delta, \quad (23)$$

where C_1, C_2 and δ depend only on $\Lambda_{\max}$ and M_K.

Proof. The argument is quite standard. We use some ideas presented in Bickel and Levina (2004). Let us define $\tilde{x}_a^i = \frac{x_a^i}{\sqrt{\sigma_{aa}^i}}$ and $\tilde{x}_b^i = \frac{x_b^i}{\sqrt{\sigma_{bb}^i}}$. Note that $\tilde{x}_a^i, \tilde{x}_b^i \sim \mathcal{N}(0,1)$ and $\mathrm{Corr}(\tilde{x}_a^i, \tilde{x}_b^i) = \rho_{ab}^i$, where

$$\rho_{ab}^i = \frac{\sigma_{ab}^i}{\sqrt{\sigma_{aa}^i \sigma_{bb}^i}}.$$

Now we have

$$\begin{aligned}
&\mathbb{P}[|\hat{\sigma}_{ab}^\tau - \mathbb{E}\hat{\sigma}_{ab}^\tau| > \epsilon] \\
&= \mathbb{P}[|\sum_i \frac{2}{nh} K_h(i-\tau)(x_a^i x_b^i - \sigma_{ab}^i)| > \epsilon] \\
&= \mathbb{P}[|\sum_i \frac{2}{nh} K_h(i-\tau)\sqrt{\sigma_{aa}^i \sigma_{bb}^i}(\tilde{x}_a^i \tilde{x}_b^i - \rho_{ab}^i)| > \epsilon].
\end{aligned}$$

A simple calculation gives that

$$\begin{aligned}
\tilde{x}_a^i \tilde{x}_b^i - \rho_{ab}^i = \frac{1}{4}\big((\tilde{x}_a^i + \tilde{x}_b^i)^2 - 2(1+\rho_{ab}^i) \\
- (\tilde{x}_a^i - \tilde{x}_b^i)^2 - 2(1-\rho_{ab}^i)\big),
\end{aligned}$$

which combined with the equation above and union bound gives

$$\begin{aligned}
&\mathbb{P}[|\hat{\sigma}_{ab}^\tau - \mathbb{E}\hat{\sigma}_{ab}^\tau| > \epsilon] \\
&\leq 2\mathbb{P}\left[M_K \Lambda_{\max} \sum_i \frac{4}{nh}((Z^i)^2 - 1) \geq \epsilon\right],
\end{aligned} \quad (24)$$

where Z^i are independent $\mathcal{N}(0,1)$. The lemma follows from the standard results on the large deviation of χ^2 random variables. □

The bandwidth parameter needs to be chosen to balance the bias and variance in (22). If the bandwidth is chosen as $h = \mathcal{O}(n^{-1/3})$, the following result is straight forward.

Lemma 11. *Under the assumptions* **K**, **S** *and* **C**, *if the bandwidth parameter satisfies $h = \mathcal{O}(n^{-1/3})$, then*

$$\mathbb{P}[\max_{a,b} |\hat{\sigma}_{ab}^\tau - \sigma_{ab}^\tau| > \epsilon) \leq C_1 \exp(-C_2 n^{2/3}\epsilon^2 + \log p),$$

where C_1 and C_2 are constants depending only on M_K, M_Σ and $\Lambda_{\max}$.

Proof. The lemma follows from (22) by applying the union bound. □

Next, we directly apply Lemma 5 and Lemma 6 from Ravikumar et al. (2008) to obtain bounds on the deviation term $\mathbf{\Delta} = \hat{\mathbf{\Omega}}^\tau - \mathbf{\Omega}^\tau$ and the remainder term $R(\mathbf{\Delta})$.

Lemma 12. *Assume that the conditions of Theorem 1 are satisfied. There exist constants $C_1, C_2 > 0$ depending only on $\Lambda_{\max}$, M_∞, M_Σ, M_K, $M_\mathcal{I}$ and α such that with probability at least $1 - C_1\exp(-C_2 \log p)$, the following two statements hold:*

1. *There exists some $M_\Delta > 0$ depending on $\Lambda_{\max}$, M_∞, M_Σ, M_K, $M_\mathcal{I}$ and α such that $||\mathbf{\Delta}||_\infty \leq M_\Delta n^{-1/3}\sqrt{\log p}$.*
2. *Furthermore, element-wise maximum of the remainder term $R(\mathbf{\Delta})$ can be bounded $||R(\mathbf{\Delta})||_\infty \leq \frac{\alpha\lambda}{8}$.*

Proof. We perform the analysis on the event $\mathcal{A}$ defined in (27). Under the assumption of the lemma, we have that $n > Cd^3(\log p)^{3/2}$ and on the event $\mathcal{A}$,

$$||\hat{\mathbf{\Sigma}}^\tau - \mathbf{\Sigma}^\tau||_\infty + \lambda \leq M_\Delta \lambda \leq \frac{M_\Delta}{d}. \quad (25)$$

This implies that under the conditions of Lemma 6 and Lemma 5 in Ravikumar et al. (2008) are satisfied and we apply them to conclude the statement of the lemma. □

The following lemma gives us deviation of the minimum eigenvalue of the weighted empirical covariance matrix from the population quantity.

Lemma 13. *Let $\tau \in [0,1]$ be a fixed time point. Assume that Σ^τ satisfies the assumptions* **S** *and* **C** *and the kernel function satisfies the assumption* **K**. *Let $\{\mathbf{x}^i\}$ be an independent sample according to the model (1). Then*

$$\mathbb{P}[|\Lambda_{\min}(\hat{\mathbf{\Sigma}}^\tau_{NN}) - \Lambda_{\min}(\mathbf{\Sigma}^\tau_{NN})| > \epsilon) \leq C_1 \exp(-C_2 \frac{nh}{|N|^2}\epsilon^2 + C_3 \log|N|), \quad (26)$$

where C_1, C_2 and C_3 are constants that depend only on $\Lambda_{\max}$, M_Σ and M_K.

Proof. Using perturbation theory results (see for example Stewart and Sun (1990)), we have that

$$|\Lambda_{\min}(\hat{\mathbf{\Sigma}}^\tau_{NN}) - \Lambda_{\min}(\mathbf{\Sigma}^\tau_{NN})| \leq ||\hat{\mathbf{\Sigma}}^\tau_{NN} - \mathbf{\Sigma}^\tau_{NN}||_F \leq |N| \max_{a\in N, b\in N} |\hat{\sigma}^\tau_{ab} - \sigma^\tau_{ab}|.$$

But then using (22), Lemma 10 and the union bound, the result follows. □

Proof of Proposition 3

We will perform analysis on the event

$$\mathcal{A} = \left\{ ||\hat{\mathbf{\Sigma}}^\tau - \mathbf{\Sigma}^\tau||_\infty \leq \frac{\alpha\lambda}{8} \right\}. \quad (27)$$

Under the assumptions of the proposition, it follows from Lemma 11 that $\mathbb{P}[\mathcal{A}] \geq 1 - C_1\exp(-C_2 \log p)$. Also, under the assumptions of the proposition, Lemma 12 can be applied to conclude that $R(\mathbf{\Delta}) \leq \frac{\alpha\lambda}{8}$. Let $e_j \in \mathbb{R}^{|S|}$ be a unit vector with 1 at position j and zeros elsewhere. On the event $\mathcal{A}$, it holds that

$$\begin{aligned}
&\max_{1\leq j\leq |S|} |e_j'(\mathcal{I}_{SS})^{-1}[(\vec{\hat{\mathbf{\Sigma}}}^\tau - \vec{\mathbf{\Sigma}}^\tau) - \overrightarrow{R(\mathbf{\Delta})} + \lambda \overrightarrow{\text{sign}}(\mathbf{\Omega}^\tau)]_S| \\
&\leq |||(\mathcal{I}_{SS})^{-1}|||_{\infty,\infty}\big(||(\vec{\hat{\mathbf{\Sigma}}}^\tau - \vec{\mathbf{\Sigma}}^\tau)_S||_\infty + ||\overrightarrow{R(\mathbf{\Delta})}_S||_\infty \\
&\quad + \lambda||\overrightarrow{\text{sign}}(\mathbf{\Omega}^\tau_S)||_\infty\big) \\
&\quad (\text{ using the Hölder's inequality }) \\
&\leq M_{\mathcal{I}}\frac{4+\alpha}{4}\lambda \leq C\frac{\sqrt{\log p}}{n^{1/3}} \\
&< \omega_{\min} = M_\omega \frac{\sqrt{\log p}}{n^{1/3}},
\end{aligned}$$

for a sufficiently large constant M_ω.

Proof of Proposition 4

We will work on the event $\mathcal{A}$ defined in (27). Under the assumptions of the proposition, Lemma 12 gives $R(\mathbf{\Delta}) \leq \frac{\alpha\lambda}{8}$. Let $e_j \in \mathbb{R}^{p^2-|S|}$ be a unit vector with 1 at position j and zeros elsewhere. On the event $\mathcal{A}$, it holds that

$$\begin{aligned}
&\max_{1\leq j\leq (p^2-|S|)} \Big|e_j'(\mathcal{I}_{S^CS}(\mathcal{I}_{SS})^{-1}[(\vec{\mathbf{\Sigma}}^\tau - \vec{\hat{\mathbf{\Sigma}}}^\tau) + \overrightarrow{R(\mathbf{\Delta})}]_S + \\
&\qquad (\vec{\hat{\mathbf{\Sigma}}}^\tau - \vec{\mathbf{\Sigma}}^\tau)_{S^C} - \overrightarrow{R(\mathbf{\Delta})}_{S^C})\Big| \\
&\leq |||\mathcal{I}_{S^CS}(\mathcal{I}_{SS})^{-1}|||_{\infty,\infty}\Big(||\vec{\mathbf{\Sigma}}^\tau - \vec{\hat{\mathbf{\Sigma}}}^\tau||_\infty + ||\overrightarrow{R(\mathbf{\Delta})}||_\infty\Big) + \\
&\qquad ||\vec{\hat{\mathbf{\Sigma}}}^\tau - \vec{\mathbf{\Sigma}}^\tau||_\infty + ||\overrightarrow{R(\mathbf{\Delta})}||_\infty \\
&\leq (1-\alpha)\frac{\alpha\lambda}{4} + \frac{\alpha\lambda}{4} \\
&\leq \alpha\lambda,
\end{aligned}$$

which concludes the proof.

Technical results of Section 3

In this subsection, we provide a proof of Lemma 9.

Proof of Lemma 9

Only a proof sketch is provided here. We analyze the event defined in (18) by splitting it into several terms. Observe that for $b \in N^C$, we can write

$$\begin{aligned}
x_b^i &= \mathbf{\Sigma}^\tau_{bN}(\mathbf{\Sigma}^\tau_{NN})^{-1}\mathbf{x}^i_N \\
&\quad + [\mathbf{\Sigma}^{t_i}_{bN}(\mathbf{\Sigma}^{t_i}_{NN})^{-1} - \mathbf{\Sigma}^\tau_{bN}(\mathbf{\Sigma}^\tau_{NN})^{-1}]'\mathbf{x}^i_N \\
&\quad + v_b^i
\end{aligned}$$

where $v_b^i \sim \mathcal{N}(0, (\sigma_b^i)^2)$ with $\sigma_b^i \leq 1$. Let us denote $\tilde{\mathbf{V}}_b \in \mathbb{R}^n$ the vector with components $\tilde{v}_b^i = \sqrt{w_i^\tau}v_b^i$. With this, we have the following decomposition of the components of the event $\mathcal{E}_4$. For all $b \in N^c$,

$$w_{b,1} = \mathbf{\Sigma}^\tau_{bN}(\mathbf{\Sigma}^\tau_{NN})^{-1}\lambda\,\text{sign}(\boldsymbol{\theta}^\tau_N),$$

$$w_{b,2} = \tilde{\mathbf{V}}_b'\Big[(\tilde{\mathbf{X}}_N(\hat{\mathbf{\Sigma}}_{NN})^{-1}\lambda\,\text{sign}(\boldsymbol{\theta}^\tau_N) + \Pi^\perp_{\tilde{\mathbf{X}}_N}(\mathbf{E}^1)\Big],$$

$$w_{b,3} = \tilde{\mathbf{V}}_b'\Pi^\perp_{\tilde{\mathbf{X}}_N}(\mathbf{E}^2)$$

and

$$w_{b,4} = \tilde{\mathbf{F}}_b'\Big[(\tilde{\mathbf{X}}_N(\hat{\mathbf{\Sigma}}_{NN})^{-1}\lambda\,\text{sign}(\boldsymbol{\theta}^\tau_N) + \Pi^\perp_{\tilde{\mathbf{X}}_N}(\mathbf{E}^1 + \mathbf{E}^2)\Big],$$

where $\Pi^\perp_{\tilde{\mathbf{X}}_N}$ is the projection operator defined as $\mathbf{I}_p - \tilde{\mathbf{X}}_N(\tilde{\mathbf{X}}_N'\tilde{\mathbf{X}}_N)^{-1}\tilde{\mathbf{X}}_N'$, $\mathbf{E}^1$ and $\mathbf{E}^2$ are defined in the proof of Lemma 8 and we have introduced $\tilde{\mathbf{F}}_b \in \mathbb{R}^n$ as the vector with components $\tilde{f}_b^i = \sqrt{w_i^\tau}[\mathbf{\Sigma}^{t_i}_{bN}(\mathbf{\Sigma}^{t_i}_{NN})^{-1} - \mathbf{\Sigma}^\tau_{bN}(\mathbf{\Sigma}^\tau_{NN})^{-1}]'\mathbf{x}^i_N$. The lemma will follow using the triangle inequality if we show that

$$\max_{b\in N^C} |w_{b,1}| + |w_{b,2}| + |w_{b,3}| + |w_{b,4}| \leq \lambda.$$

Under the assumptions of the lemma, it holds that $\max_{b\in N^C}|w_{b,1}| < (1-\gamma)\lambda$.

Next, we deal with the term $w_{b,2}$. We observe that conditioning on $\mathbf{X}_S$, we have that $w_{b,2}$ is normally distributed with variance that can be bounded combining results of Lemma 13 from the supplementary material with the proof of Lemma 4 in Wainwright (2009). Next, we use the Gaussian tail bound to conclude that $\max_{b \in N^C} |w_{b,2}| < \gamma\lambda/2$ with probability at least $1 - \exp(-C_2 nh(d \log p)^{-1})$.

An upper bound on the term $w_{b,3}$ is obtained as follows $w_{b,3} \leq ||\tilde{\mathbf{V}}_b||_2 ||\Pi^{\perp}_{\tilde{\mathbf{X}}_N}(\mathbf{E}^2)||_2$ and then observing that the term is asymptotically dominated by the term $w_{b,2}$. Using similar reasoning, we also have that $w_{b,4}$ is asymptotically smaller than $w_{b,2}$.

Combining all the upper bounds, we obtain the desired result.