A Preliminary Results

We introduce the following lemmas from prior literature.

Lemma 3 ((Ravikumar et al., 2011, Lemma 1))**.** *Consider a zero-mean random vector $x = (x_1, \dots, x_p)$ with covariance Σ, such that each $x_i/\sqrt{\Sigma_{i,i}}$ is sub-Gaussian with parameter σ^2. Given n i.i.d. samples, the associated sample covariance H satisfies the tail bound*

$$\mathbb{P}\left\{|H_{i,j} - \Sigma_{i,j}| > t\right\} \leq 4 \exp\left(-\frac{nt^2}{128(1+4\sigma^2)^2 \Sigma_{dmax}^2}\right), \tag{11}$$

for all $t \in (0, 8(1+4\sigma^2)\Sigma_{dmax})$.

Lemma 4 ((Hsu et al., 2012, Theorem 1))**.** *Consider a zero-mean sub-Gaussian random vector $x = (x_1, \dots, x_p)$ with parameter σ^2. Then for all $t > 0$,*

$$\mathbb{P}\left\{\|x\|^2 > \sigma^2 \cdot \left(p + 2\sqrt{pt} + 2pt\right)\right\} \leq \mathrm{e}^{-t}. \tag{12}$$

B Technical Lemmas

We introduce the necessary lemmas used in our analysis and provide the proofs.

Lemma 5. *For every feature $i \in [p]$, its empirical error ratio $\hat{\tau}_i$ fulfills*

$$\mathbb{P}\left\{|\tau_i - \hat{\tau}_i| \leq \frac{1}{2}|\tau_i|\right\} \geq 1 - 4\exp\left(-\frac{n\Sigma_{\Pi(i),\Pi(i)}^2}{3200(1+4\sigma^2)^2\Sigma_{dmax}^2}\right) - 4\exp\left(-\frac{n\Sigma_{i,\Pi(i)}^2}{3200(1+4\sigma^2)^2\Sigma_{dmax}^2}\right).$$

Proof. Setting $t = \frac{1}{5}\Sigma_{\Pi(i),\Pi(i)}$ in Lemma 3, we have $\left|H_{\Pi(i),\Pi(i)} - \Sigma_{\Pi(i),\Pi(i)}\right| \leq \frac{1}{5}\Sigma_{\Pi(i),\Pi(i)}$ with probability at least $1 - 4\exp\left(-\frac{n\Sigma_{\Pi(i),\Pi(i)}^2}{3200(1+4\sigma^2)^2\Sigma_{\mathrm{dmax}}^2}\right)$ for all i. Similarly, setting $t = \frac{1}{5}\Sigma_{i,\Pi(i)}$ in Lemma 3, we have $\left|H_{i,\Pi(i)} - \Sigma_{i,\Pi(i)}\right| \leq \frac{1}{2}\Sigma_{i,\Pi(i)}$ with probability at least $1 - 4\exp\left(-\frac{n\Sigma_{i,\Pi(i)}^2}{3200(1+4\sigma^2)^2\Sigma_{\mathrm{dmax}}^2}\right)$ for all i.

Using a union bound, with probability at least $1 - 4\exp\left(-\frac{n\Sigma_{\Pi(i),\Pi(i)}^2}{3200(1+4\sigma^2)^2\Sigma_{\mathrm{dmax}}^2}\right) - 4\exp\left(-\frac{n\Sigma_{i,\Pi(i)}^2}{3200(1+4\sigma^2)^2\Sigma_{\mathrm{dmax}}^2}\right)$, we have

$$\begin{aligned}
|\tau_i - \hat{\tau}_i| &= \left|\frac{H_{i,\Pi(i)}}{H_{\Pi(i),\Pi(i)}} - \frac{\Sigma_{i,\Pi(i)}}{\Sigma_{\Pi(i),\Pi(i)}}\right| \\
&\leq \left|\frac{\Sigma_{i,\Pi(i)} + \frac{1}{5}\Sigma_{i,\Pi(i)}}{\Sigma_{\Pi(i),\Pi(i)} - \frac{1}{5}\Sigma_{\Pi(i),\Pi(i)}} - \frac{\Sigma_{i,\Pi(i)}}{\Sigma_{\Pi(i),\Pi(i)}}\right| \\
&= \left|\frac{\Sigma_{i,\Pi(i)}}{2\Sigma_{\Pi(i),\Pi(i)}}\right| \\
&= \frac{1}{2}|\tau_i|.
\end{aligned}$$

□

Lemma 6. *The minimum eigenvalue of the sample covariance matrix follows*

$$\lambda_{min}(H_{S,S}) \geq \beta/2,$$

with probability at least $1 - 2\exp(-\min\{\frac{n\beta^2}{256\sigma^4\Sigma_{dmax}^2}, \frac{n\beta}{16\sigma^2\Sigma_{dmax}}\})$.

Proof. Using a variational characterization of eigenvalues, we have

$$\begin{aligned}
\lambda_{\min}(H_{S,S}) &= \min_{\|u\|=1} u^\top H_{S,S} u \\
&= \min_{\|u\|=1} u^\top \Sigma_{S,S} u + u^\top (H_{S,S} - \Sigma_{S,S}) u \\
&\geq \beta + \min_{\|u\|=1} \frac{1}{n} \sum_{k=1}^{n} u^\top (X_{k,S}^\top X_{k,S} - \mathbb{E}\left[X_{k,S}^\top X_{k,S}\right]) u \\
&\geq \beta - \max_{\|u\|=1} \left| \frac{1}{n} \sum_{k=1}^{n} (X_{k,S} u)^2 - \mathbb{E}\left[(X_{k,S} u)^2\right] \right| .
\end{aligned}$$

Note that for any u with $\|u\| \leq 1$, $X_{k,S}u$ is sub-Gaussian with parameter at most $\sigma^2 \Sigma_{\text{dmax}}$. It follows that $(X_{k,S}u)^2 - \mathbb{E}\left[(X_{k,S}u)^2\right]$ is sub-exponential with parameter $(32\sigma^4 \Sigma_{\text{dmax}}^2, 4\sigma^2 \Sigma_{\text{dmax}})$. Applying the sub-exponential tail bound leads to

$$\mathbb{P}\left\{ \left| \frac{1}{n} \sum_{k=1}^{n} (X_{k,S}u)^2 - \mathbb{E}\left[(X_{k,S}u)^2\right] \right| \geq t \right\} \leq 2 \exp\left(- \min\{ \frac{nt^2}{64\sigma^4 \Sigma_{\text{dmax}}^2}, \frac{nt}{8\sigma^2 \Sigma_{\text{dmax}}} \} \right) . \tag{13}$$

Setting $t = \beta/2$ leads to the result. ☐

Lemma 7. *The probability of* $\left|\left|\left| \hat{H}_{S^c,S} - H_{S^c,S} \right|\right|\right|_\infty \geq \frac{\beta\gamma}{24\sqrt{s}}$, *is bounded above in the order of* $O\left(s(p-s)\exp\left(-\frac{\beta^2\gamma^2 n}{s^3}\right)\right)$. *The probability of* $\left|\left|\left| \hat{H}_{S,S} - H_{S,S} \right|\right|\right|_\infty \geq \frac{\beta\gamma}{48(1-\gamma/2)\sqrt{s}}$, *is bounded above in the order of* $O\left(s^2 \exp\left(-\frac{\beta^2\gamma^2 n}{s^3(1-\gamma/2)^2}\right)\right)$.

Proof. Here we bound $\left|\left|\left| \hat{H}_{S^c,S} - H_{S^c,S} \right|\right|\right|_\infty$ and $\left|\left|\left| \hat{H}_{S,S} - H_{S,S} \right|\right|\right|_\infty$. We first consider $\left|\left|\left| \hat{H}_{S^c,S} - H_{S^c,S} \right|\right|\right|_\infty$. Note that for all $i \in S^c$, $j \in S$, we have

$$\begin{aligned}
\hat{H}_{i,j} &= \frac{1}{n} \sum_{k=1}^{n} \hat{X}_{k,i} \hat{X}_{k,j} \\
&= \frac{1}{n} \sum_{k=1}^{n} \left(M_{k,i} X_{k,i} + (1 - M_{k,i}) \hat{\tau}_i X_{k,\Pi(i)} \right) \left(M_{k,j} X_{k,j} + (1 - M_{k,j}) \hat{\tau}_j X_{k,\Pi(j)} \right) .
\end{aligned}$$

It follows that

$$\begin{aligned}
\left| \hat{H}_{i,j} - H_{i,j} \right| \leq \Big| \frac{1}{n} \sum_{k=1}^{n} & (M_{k,i} M_{k,j} - 1) X_{k,i} X_{k,j} \\
& + (1 - M_{k,i})(1 - M_{k,j}) \hat{\tau}_i \hat{\tau}_j X_{k,\Pi(i)} X_{k,\Pi(j)} \\
& + M_{k,i}(1 - M_{k,j}) \hat{\tau}_j X_{k,i} X_{k,\Pi(j)} \\
& + (1 - M_{k,i}) M_{k,j} \hat{\tau}_i X_{k,\Pi(i)} X_{k,j} \Big| .
\end{aligned}$$

Since $M_{k,i}, M_{k,j}$ is either 0 or 1, we can upper bound the terms above by

$$\left| \hat{H}_{i,j} - H_{i,j} \right| \leq \frac{(1 + \hat{\tau}_{\max})^2}{n} \left| \sum_{k=1}^{n} X_{k,i} X_{k,j} \right| . \tag{14}$$

Using Lemma 3, we obtain

$$\mathbb{P}\left\{ \left| \frac{1}{n} \sum_{k=1}^{n} X_{k,i} X_{k,j} - \Sigma_{i,j} \right| \geq t \right\} \leq 4 \exp\left(- \frac{nt^2}{128(1 + 4\sigma_X^2)^2 \Sigma_{\text{dmax}}^2} \right) . \tag{15}$$

Then with probability at least $1 - 4\exp\left(-\frac{nt^2}{128(1+4\sigma_X^2)^2\Sigma_{\text{dmax}}^2}\right)$, it follows that

$$\left|\hat{H}_{i,j} - H_{i,j}\right| \le (1 + \hat{\tau}_{\max})^2(\Sigma_{i,j} + t)\,.$$

Setting $t = \Sigma_{i,j}$, we obtain that with high probability,

$$\left|\hat{H}_{i,j} - H_{i,j}\right| \le 2(1 + \hat{\tau}_{\max})^2\Sigma_{i,j}\,.$$

Using Lemma 5, with high probability we have

$$\left|\hat{H}_{i,j} - H_{i,j}\right| \le 2(1 + \hat{\tau}_{\max})^2\Sigma_{i,j}\,.$$

Now we consider the infinity norm bound. By using a union bound, we obtain

$$\begin{aligned}
\mathbb{P}\left\{\left|\!\left|\!\left|\hat{H}_{S^c,S} - H_{S^c,S}\right|\!\right|\!\right|_\infty \ge t\right\} &\le s(p-s)\mathbb{P}\left\{\left|\hat{H}_{i,j} - H_{i,j}\right| \ge \frac{t}{s}\right\} \\
&\le s(p-s)\mathbb{P}\left\{2(1+\hat{\tau}_{\max})^2\left|\frac{1}{n}\sum_{k=1}^{n} X_{k,i}X_{k,j}\right| \ge \frac{t}{s}\right\} \\
&= s(p-s)\mathbb{P}\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_{k,i}X_{k,j}\right| \ge \frac{t}{2(1+\hat{\tau}_{\max})^2 s}\right\} \\
&\le 2s(p-s)\exp\left(-\frac{nt^2}{16(1+\hat{\tau}_{\max})^4 s^2}\right)\,,
\end{aligned}$$

where the last inequality follows a sub-Exponential tail bound.

Similarly, for the other infinity norm, we have

$$\mathbb{P}\left\{\left|\!\left|\!\left|\hat{H}_{S,S} - H_{S,S}\right|\!\right|\!\right|_\infty \ge t\right\} \le 2s^2\exp\left(-\frac{nt^2}{16(1+\hat{\tau}_{\max})^4 s^2}\right)\,. \tag{16}$$

□

C Proof of Lemma 1

Proof. Setting $t = \frac{1}{2}\Sigma_{\text{dmin}}$ in Lemma 3, we have $|H_{i,i} - \Sigma_{i,i}| \le \frac{1}{2}\Sigma_{\text{dmin}}$ with probability at least $1 - 4\exp\left(-\frac{n\Sigma_{\text{dmin}}^2}{512(1+4\sigma_X^2)^2\Sigma_{\text{dmax}}^2}\right)$. With the same probability and some algebra, we have

$$\begin{aligned}
\frac{H_{i,i}}{\Sigma_{i,i}} &\le \frac{\Sigma_{i,i} + \frac{1}{2}\Sigma_{\text{dmin}}}{\Sigma_{i,i}} \le \frac{\Sigma_{\text{dmin}} + \frac{1}{2}\Sigma_{\text{dmin}}}{\Sigma_{\text{dmin}}} \le \frac{3}{2}\,, \\
\frac{H_{i,i}}{\Sigma_{i,i}} &\ge \frac{\Sigma_{i,i} - \frac{1}{2}\Sigma_{\text{dmin}}}{\Sigma_{i,i}} \ge \frac{\Sigma_{\text{dmin}} - \frac{1}{2}\Sigma_{\text{dmin}}}{\Sigma_{\text{dmin}}} \ge \frac{1}{2}\,.
\end{aligned}$$

□

D Proof of Lemma 2

Proof. We first look at the concentration properties of $|\!|\!|\Sigma_{S^c,S} - H_{S^c,S}|\!|\!|_\infty$ and $|\!|\!|\Sigma_{S,S} - H_{S,S}|\!|\!|_\infty$. By using Lemma 3 and a union bound, we obtain

$$\mathbb{P}\left\{|\!|\!|\Sigma_{S^c,S} - H_{S^c,S}|\!|\!|_\infty \ge t\right\} \le s(p-s)\mathbb{P}\left\{|\Sigma_{i,j} - H_{i,j}| \ge \frac{t}{s}\right\}$$

$$\leq 4s(p-s)\exp\left(-\frac{nt^2}{128s^2(1+4\sigma^2)^2\Sigma_{\mathrm{dmax}}^2}\right).$$

Setting $t = \frac{\beta\gamma}{6\sqrt{s}}$, we obtain that

$$\mathbb{P}\left\{ |||\Sigma_{S^c,S} - H_{S^c,S}|||_\infty \geq \frac{\beta\gamma}{6\sqrt{s}} \right\} \leq 4s(p-s)\exp\left(-\frac{\beta^2\gamma^2 n}{4608s^3(1+4\sigma^2)^2\Sigma_{\mathrm{dmax}}^2}\right). \tag{17}$$

Similarly, for the other infinity norm, we have

$$\begin{aligned}\mathbb{P}\left\{ |||\Sigma_{S,S} - H_{S,S}|||_\infty \geq t \right\} &\leq s^2\mathbb{P}\left\{ |\Sigma_{i,j} - H_{i,j}| \geq \frac{t}{s} \right\} \\ &\leq 4s^2\exp\left(-\frac{nt^2}{128s^2(1+4\sigma^2)^2\Sigma_{\mathrm{dmax}}^2}\right).\end{aligned}$$

Setting $t = \frac{\beta\gamma}{12(1-\gamma)\sqrt{s}}$, we obtain that

$$\mathbb{P}\left\{ |||\Sigma_{S,S} - H_{S,S}|||_\infty \geq \frac{\beta\gamma}{12(1-\gamma)\sqrt{s}} \right\} \leq 4s^2\exp\left(-\frac{\beta^2\gamma^2 n}{18432s^3(1-\gamma)^2(1+4\sigma^2)^2\Sigma_{\mathrm{dmax}}^2}\right). \tag{18}$$

Now we proceed with the main bound. Note that

$$H_{S^c,S}H_{S,S}^{-1} = HT_1 + HT_2 + HT_3 + HT_4\,, \tag{19}$$

where

$$HT_1 = \Sigma_{S^c,S}\Sigma_{S,S}^{-1}\,, \tag{20}$$

$$HT_2 = (H_{S^c,S} - \Sigma_{S^c,S})\Sigma_{S,S}^{-1}\,, \tag{21}$$

$$HT_3 = \Sigma_{S^c,S}(H_{S,S}^{-1} - \Sigma_{S,S}^{-1})\,, \tag{22}$$

$$HT_4 = (H_{S^c,S} - \Sigma_{S^c,S})(H_{S,S}^{-1} - \Sigma_{S,S}^{-1})\,. \tag{23}$$

From Assumption 2, we know that $|||HT_1|||_\infty \leq 1-\gamma$. For HT_2, we have

$$\begin{aligned}|||HT_2|||_\infty &= \left|\left|\left|(H_{S^c,S} - \Sigma_{S^c,S})\Sigma_{S,S}^{-1}\right|\right|\right|_\infty \\ &\leq |||H_{S^c,S} - \Sigma_{S^c,S}|||_\infty \left|\left|\left|\Sigma_{S,S}^{-1}\right|\right|\right|_\infty \\ &\leq \sqrt{s}\cdot|||H_{S^c,S} - \Sigma_{S^c,S}|||_\infty \left|\left|\left|\Sigma_{S,S}^{-1}\right|\right|\right| \\ &\leq \frac{\sqrt{s}}{\beta}\cdot|||H_{S^c,S} - \Sigma_{S^c,S}|||_\infty \\ &\leq \frac{\gamma}{6}\,,\end{aligned}$$

where the last inequality follows from (17) . For HT_3, we have

$$\begin{aligned}|||HT_3|||_\infty &= \left|\left|\left|\Sigma_{S^c,S}(H_{S,S}^{-1} - \Sigma_{S,S}^{-1})\right|\right|\right|_\infty \\ &\leq \left|\left|\left|\Sigma_{S^c,S}\Sigma_{S,S}^{-1}(\Sigma_{S,S} - H_{S,S})H_{S,S}^{-1}\right|\right|\right|_\infty \\ &\leq \left|\left|\left|\Sigma_{S^c,S}\Sigma_{S,S}^{-1}\right|\right|\right|_\infty |||\Sigma_{S,S} - H_{S,S}|||_\infty \left|\left|\left|H_{S,S}^{-1}\right|\right|\right|_\infty \\ &\leq \frac{2(1-\gamma)\sqrt{s}}{\beta}\cdot|||\Sigma_{S,S} - H_{S,S}|||_\infty \\ &\leq \frac{\gamma}{6}\,,\end{aligned}$$

where the last inequality follows from (18) . For HT_4, we have

$$
\begin{aligned}
|||HT_4|||_\infty &= \left|\left|\left|(H_{S^c,S} - \Sigma_{S^c,S})(H_{S,S}^{-1} - \Sigma_{S,S}^{-1})\right|\right|\right|_\infty \\
&\le |||H_{S^c,S} - \Sigma_{S^c,S}|||_\infty \left|\left|\left|H_{S,S}^{-1} - \Sigma_{S,S}^{-1}\right|\right|\right|_\infty \\
&\le |||H_{S^c,S} - \Sigma_{S^c,S}|||_\infty \left|\left|\left|H_{S,S}^{-1}(\Sigma_{S,S} - H_{S,S})\Sigma_{S,S}^{-1}\right|\right|\right|_\infty \\
&\le |||H_{S^c,S} - \Sigma_{S^c,S}|||_\infty \left|\left|\left|H_{S,S}^{-1}\right|\right|\right|_\infty |||\Sigma_{S,S} - H_{S,S}|||_\infty \left|\left|\left|\Sigma_{S,S}^{-1}\right|\right|\right|_\infty \\
&\le s \cdot |||H_{S^c,S} - \Sigma_{S^c,S}|||_\infty \left|\left|\left|H_{S,S}^{-1}\right|\right|\right| |||\Sigma_{S,S} - H_{S,S}|||_\infty \left|\left|\left|\Sigma_{S,S}^{-1}\right|\right|\right| \\
&\le \frac{2s}{\beta^2} \cdot |||H_{S^c,S} - \Sigma_{S^c,S}|||_\infty |||\Sigma_{S,S} - H_{S,S}|||_\infty \\
&\le \frac{\gamma}{6} ,
\end{aligned}
$$

where the last inequality follows from (17) and (18), given that $\gamma < \frac{6}{7}$. This completes the proof. ☐

E Proof of Theorem 1

Proof. Setting $t = \frac{1}{2}\Sigma_{\text{dmin}}$ in Lemma 3, we have $|H_{i,j} - \Sigma_{i,j}| \le \frac{1}{2}\Sigma_{\text{dmin}}$ with probability at least $1 - 4\exp\left(-\frac{n\Sigma_{\text{dmin}}^2}{512(1+4\sigma_X^2)^2\Sigma_{\text{dmax}}^2}\right)$ for all i and j. It follows that $|H_{i,j} + \Sigma_{i,j}| \le \frac{1}{2}\Sigma_{\text{dmin}} + 2|\Sigma_{i,j}|$, and $\left|H_{i,j}^2 - \Sigma_{i,j}^2\right| \le \frac{1}{4}\Sigma_{\text{dmin}}^2 + \Sigma_{\text{dmin}}|\Sigma_{i,j}|$. Dividing both sides by $\Sigma_{j,j}$, we obtain

$$
\left|\frac{H_{i,j}^2}{\Sigma_{j,j}} - \frac{\Sigma_{i,j}^2}{\Sigma_{j,j}}\right| \le \frac{\Sigma_{\text{dmin}}^2}{4\Sigma_{j,j}} + \frac{\Sigma_{\text{dmin}}}{\Sigma_{j,j}}|\Sigma_{i,j}| \le \frac{1}{4}\Sigma_{\text{dmin}} + |\Sigma_{j,j}| .
$$

Note that

$$
\begin{aligned}
\left|\frac{H_{i,j}^2}{\Sigma_{j,j}} - \frac{\Sigma_{i,j}^2}{\Sigma_{j,j}}\right| &= \left|\frac{H_{i,j}^2 H_{j,j}}{\Sigma_{j,j} H_{j,j}} - \frac{\Sigma_{i,j}^2}{\Sigma_{j,j}}\right| \\
&= \left|\frac{H_{j,j}}{\Sigma_{j,j}}\hat{\zeta}_{i,j} - \zeta_{i,j}\right| \le \frac{1}{4}\Sigma_{\text{dmin}} + |\Sigma_{j,j}| .
\end{aligned}
$$

Rewriting the last inequality, we have

$$
\begin{aligned}
&\frac{\Sigma_{j,j}}{H_{j,j}}\left(\zeta_{i,j} - \frac{1}{4}\Sigma_{\text{dmin}} - |\Sigma_{j,j}|\right) \\
&\le \hat{\zeta}_{i,j} \le \frac{\Sigma_{j,j}}{H_{j,j}}\left(\zeta_{i,j} + \frac{1}{4}\Sigma_{\text{dmin}} + |\Sigma_{j,j}|\right) .
\end{aligned}
$$

It follows from Lemma 1, that

$$
\begin{aligned}
&\frac{2}{3}\left(\zeta_{i,j} - \frac{1}{4}\Sigma_{\text{dmin}} - |\Sigma_{j,j}|\right) \\
&\le \hat{\zeta}_{i,j} \le 2\left(\zeta_{i,j} + \frac{1}{4}\Sigma_{\text{dmin}} + |\Sigma_{j,j}|\right) .
\end{aligned}
$$

As a result, to ensure that $\hat{\zeta}_{i,\Pi(i)} > \hat{\zeta}_{i,j}$, it is sufficient to ensure

$$
\begin{aligned}
\hat{\zeta}_{i,\Pi(i)} &\ge \frac{2}{3}\left(\zeta_{i,\Pi(i)} - \frac{1}{4}\Sigma_{\text{dmin}} - \left|\Sigma_{\Pi(i),\Pi(i)}\right|\right) \\
&> 2\left(\zeta_{i,j} + \frac{1}{4}\Sigma_{\text{dmin}} + |\Sigma_{j,j}|\right) \ge \hat{\zeta}_{i,j} ,
\end{aligned}
$$

and simplification leads to

$$\zeta_{i,\Pi(i)} - 3\zeta_{i,j} > \left|\Sigma_{\Pi(i),\Pi(i)}\right| + 3\left|\Sigma_{j,j}\right| + \Sigma_{\mathrm{dmin}}\,.$$

□