Supplementary material for Weighted Conditional Network Testing for Multiple High-Dimensional Correlated Data Sets

Takwon Kim[1], Inyoung Kim[2,*], Ki-Ahm Lee[1,3]

[1] *School of Mathematics, Statistics and Data Science, Sungshin Women's University, Seoul, Republic of Korea*

[2] *Department of Statistics, Virginia Tech University, Blacksburg, VA 24061, USA*

[3] *Department of Mathematical Science, Seoul National University, Seoul, Republic of Korea*

Supplementary Material

S1 Proofs of the Main Theorems

In this section, we prove the main theorems of main document. Some technical lemmas for proving main theorems are placed in Section S2.

Condition 1 (Regularity of $\mathbf{\Omega}^d$)**.** There exists $C_0 > 0$ such that $C_0^{-1} \le \lambda_{\min}(\mathbf{\Omega}^d) \le \lambda_{\max}(\mathbf{\Omega}^d) \le C_0$, where $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ are the maximum and minimum eigenvalues of symmetric matrix A, respectively. Also assume that there exists $0 < \eta_d < 1$ such that $\max_{1\le i,j\le p_1} \left| \frac{\omega^d_{i,j}}{\sqrt{\omega^d_{i,i}\omega^d_{j,j}}} \right| \le \eta_d < 1$.

Condition 2 (Sparsity of $\mathbf{\Omega}^d$)**.** There exists $\tau > 0$ such that $|A_\tau| = O(p_1^{1/16})$, where $A_\tau = \left\{ (i,j) : \left| \frac{\omega^d_{i,j}}{\sqrt{\omega^d_{i,i}\omega^d_{j,j}}} \right| \ge (\log p_1)^{-2-\tau}, 1 \le i < j \le p_1,\ d = 1 \text{ or } 2 \right\}$

Condition 3 (Relationship among n_1, n_2, and p_1)**.** Assume $\log p_1 = o(n^{1/5})$ and $n = \min(n_1, n_2)$.

Remark on Condition 3. Xia et al. (2018) and our work appear similar in that both consider subnetworks. However, a significant distinction lies in the conditions related to the size of the subnetwork. Their focus is on identifying which subnetworks differ when the number of subnetworks ($\frac{p}{p_1}$) is very large. In contrast, our model aims to identify which edges differ within a subnetwork when the subnetwork itself is large, leading us to assume a relationship between p_1 and n. Therefore, the primary difference between Xia et al. (2018) and our work lies in whether the focus is on differences between subnetworks or within a subnetwork.

Condition 4 (Convergence rate of $\hat{\boldsymbol{\beta}}_i^d$)**.** The estimator $\hat{\beta}_i^d$ satisfies

$$\max_{1\le i\le p_1} \|\hat{\boldsymbol{\beta}}_i^d - \boldsymbol{\beta}_i^d\|_1 = o_p\{(\log p_1)^{-1}\} \quad \text{and} \quad \max_{1\le i\le p_1} \|\hat{\boldsymbol{\beta}}_i^d - \boldsymbol{\beta}_i^d\|_2 = o_p\{(n_d \log p_1)^{-1/4}\}.$$

According to (Liu, 2013, Proposition 4.2), Condition 4 is guaranteed under Condition 1, Condition 3, and $\max_{1\le i\le p_1}\sum_{j=1}^{p_1+p_2-1} I(\beta_{j,i}^d \neq 0) = o\left(\frac{\lambda_{\min}(\Sigma)\sqrt{n_d}}{\sqrt{\log p}^3}\right)$ by solving the following optimization problem:

$$\hat{\boldsymbol{\beta}}_i^d = \left(\mathbf{D}_{-i,-i}^d\right)^{-\frac{1}{2}} \underset{\mathbf{u}\in\mathbb{R}^{p_1+p_2-1}}{\arg\min} \left\{\lambda_{n_d,i}^d\|\mathbf{u}\|_1 + \frac{1}{2n_d}\left|\left(\mathbf{X}_{\cdot,-i}^d - \bar{\mathbf{X}}_{\cdot,-i}^d\right)\left(\mathbf{D}_{-i,-i}^d\right)^{-\frac{1}{2}}\mathbf{u} - \left(\mathbf{X}_{\cdot,i}^d - \bar{\mathbf{X}}_{\cdot,i}^d\right)\right|_2^2\right\}, \tag{S1.1}$$

where $\mathbf{D}^d = \text{diag}(\hat{\boldsymbol{\Sigma}}^d)$, $\hat{\boldsymbol{\Sigma}}^d = (\hat{\sigma}_{i,j}^d)_{1\le i,j\le p_1} = \left(\frac{1}{n_d}\sum_{k=1}^{n_d}(\mathbf{X}_{k,i}^d - \bar{\mathbf{X}}_{\cdot,i}^d)(\mathbf{X}_{k,j}^d - \bar{\mathbf{X}}_{\cdot,j}^d)\right)_{1\le i,j\le p_1}$, and $\lambda_{n_d,i}^d = \kappa_d\sqrt{(\hat{\sigma}_{i,i}^d \log p)/n_d}$ for $\kappa_d > 2$.

S1.1 Proof of Theorem 1

Proof of Theorem 1. We divide the index set $A := \{(i,j) : 1 \le i \le j \le p_1\}$ into two parts:

$$\text{Large } \left|\omega^d_{i,j}/\sqrt{\omega^d_{i,i}\omega^d_{j,j}}\right| \text{ elements: } A_\tau := \left\{(i,j) : \left|\frac{\omega^d_{i,j}}{\sqrt{\omega^d_{i,i}\omega^d_{j,j}}}\right| \ge (\log p_1)^{-1-\tau}\right\}$$

$$\text{Not large } \left|\omega^d_{i,j}/\sqrt{\omega^d_{i,i}\omega^d_{j,j}}\right| \text{ elements: } A \setminus A_\tau.$$

Let $y_{p_1} = 4\log p_1 - \log\log p_1 + t$ for fixed $t \in \mathbb{R}$. Then, by Lemma 11,

$$P\left(\max_{(i,j)\in A_\tau} |\tilde{\Delta}_{i,j}|^2 \ge y_{p_1}\right) = o(1).$$

Therefore, we can ignore A_τ parts:

$$\begin{aligned}
&P(\tilde{M}_n \ge y_{p_1})\\
&\quad = P(\max_{(i,j)\in A\setminus A_\tau} |\tilde{\Delta}_{i,j}|^2 \ge y_{p_1}) + P(\{\max_{(i,j)\in A\setminus A_\tau} |\tilde{\Delta}_{i,j}|^2 < y_{p_1}\} \cap \{\max_{(i,j)\in A_\tau} |\tilde{\Delta}_{i,j}|^2 \ge y_{p_1}\})\\
&\quad = P(\max_{(i,j)\in A\setminus A_\tau} |\tilde{\Delta}_{i,j}|^2 \ge y_{p_1}) + o(1),
\end{aligned}$$

For any $\varepsilon > 0$, we decompose $P(\tilde{M}_n \ge y_{p_1})$ into two parts:

$$\begin{aligned}
P(\tilde{M}_n \ge y_{p_1}) \le{}& P(\max_{(i,j)\in A\setminus A_\tau} |V_{i,j}| \ge \sqrt{y_{p_1} - \varepsilon})\\
&+ P(\max_{(i,j)\in A\setminus A_\tau} |\tilde{\Delta}_{i,j} - V_{i,j}| \ge \sqrt{y_{p_1}} - \sqrt{y_{p_1} - \varepsilon}) + o(1),
\end{aligned}$$

where $V_{i,j}$ is defined in Lemma 11. From Lemma 11, we derive

$$\begin{aligned}
P(\max_{(i,j)\in A\setminus A_\tau} |\tilde{\Delta}_{i,j} - V_{i,j}| \geq \sqrt{y_{p_1}} - \sqrt{y_{p_1} - \varepsilon}) &\leq P(\max_{(i,j)\in A\setminus A_\tau} |\tilde{\Delta}_{i,j} - V_{i,j}| \geq \frac{\varepsilon}{2}(y_{p_1})^{-1/2}) \\
&\leq P(\max_{(i,j)\in A\setminus A_\tau} |\tilde{\Delta}_{i,j} - V_{i,j}| \geq \frac{\varepsilon}{4}(\log p_1)^{-1/2}) \\
&= o(1).
\end{aligned}$$

Therefore, we obtain

$$P(\tilde{M}_n \geq y_{p_1}) \leq P(\max_{(i,j)\in A\setminus A_\tau} |V_{i,j}| \geq (y_{p_1} - \varepsilon)^{1/2}) + o(1).$$

and the opposite inequality can be showed similarly. Thus, for any $\varepsilon > 0$,

$$P(\max_{(i,j)\in A\setminus A_\tau} |V_{i,j}|^2 \geq y_{p_1} + \varepsilon) + o(1) \leq P(\tilde{M}_n \geq y_{p_1}) \leq P(\max_{(i,j)\in A\setminus A_\tau} |V_{i,j}|^2 \geq y_{p_1} - \varepsilon) + o(1).$$

Then, by Lemma 12,

$$P(\max_{1\leq m\leq q} |\hat{V}_m|^2 \geq y_{p_1} + 2\varepsilon) + o(1) \leq P(M_n \geq y_{p_1}) \leq P(\max_{1\leq m\leq q} |\hat{V}_m|^2 \geq y_{p_1} - 2\varepsilon) + o(1).$$

By Lemma 13, for any positive integer N

$$\begin{aligned}
\sum_{d=1}^{2N} \frac{(-1)^{d-1}}{d!} \left((8\pi)^{-1/2} e^{-(t+2\varepsilon)/2}\right)^d + o(1) &\leq P(\tilde{M}_n \geq y_{p_1}) \\
&\leq \sum_{d=1}^{2N-1} \frac{(-1)^{d-1}}{d!} \left((8\pi)^{-1/2} e^{-(t-2\varepsilon)/2}\right)^d + o(1).
\end{aligned}$$

By letting n and p_1 go to ∞, we obtain

$$\sum_{d=1}^{2N} \frac{(-1)^{d-1}}{d!} \left((8\pi)^{-1/2} e^{-(t+2\varepsilon)/2}\right)^d \leq \liminf P(M_n \geq y_{p_1})$$
$$\sum_{d=1}^{2N-1} \frac{(-1)^{d-1}}{d!} \left((8\pi)^{-1/2} e^{-(t-2\varepsilon)/2}\right)^d \geq \limsup P(M_n \geq y_{p_1}).$$

By letting N go ∞, we also obtain

$$1 - \exp\left(-(8\pi)^{-1/2} e^{-(t+2\varepsilon)/2}\right) \leq \liminf P(\tilde{M}_n \geq y_{p_1})$$
$$\leq \limsup P(\tilde{M}_n \geq y_{p_1}) \leq 1 - \exp\left(-(8\pi)^{-1/2} e^{-(t-2\varepsilon)/2}\right).$$

Finally, by letting ε go to 0, we complete the proof.

$$\lim_{n,p_1 \to \infty} P(\tilde{M}_n \geq y_{p_1}) = 1 - \exp\left(-(8\pi)^{-1/2} e^{-t/2}\right).$$

□

S1.2 Proof of Theorem 2

This proof is to demonstrate how the power of the test based on our weighted statistic converges to 1 as p_1 and n approach infinity in Theorem 2.

Proof of Theorem 2. In Theorem 1, we put $t = \frac{1}{2} \log \log p_1$ to

$$P(\tilde{M}_n - 4 \log p_1 + \log \log p_1 \leq t) \to \Phi_M(t) \quad \text{as } n, p_1 \to \infty,$$

then we obtain

$$\begin{aligned}
&\lim_{n,p_1\to\infty} P(\tilde{M}_n > 4\log p_1 - \frac{1}{2}\log\log p_1) \\
&= \lim_{n,p_1\to\infty} P(\tilde{M}_n - 4\log p_1 + \log\log p_1 > \frac{1}{2}\log\log p_1) \\
&= 1 - \lim_{p_1\to\infty} \Phi_M\left(\frac{1}{2}\log\log p_1\right) \\
&= 1 - \lim_{p_1\to\infty} \exp(-(8\pi)^{-1/2}(\log p_1)^{-1/4}) = 0.
\end{aligned} \tag{S1.2}$$

From the following inequality

$$A^2 \le 2(A-B)^2 + 2B^2 \quad \text{for any } A, B \in \mathbb{R},$$

we have

$$\max_{1\le i\le j\le p_1} \frac{1}{\hat{\theta}^1_{i,j} + \hat{\theta}^2_{i,j}} \left(\frac{\omega^1_{i,j}}{\sqrt{\omega^1_{i,i}\omega^1_{j,j}}} - \frac{\omega^2_{i,j}}{\sqrt{\omega^2_{i,i}\omega^2_{j,j}}} \right)^2 \le 2\tilde{M}_n + 2M_n. \tag{S1.3}$$

Thus, for $(\mathbf{\Omega}^1, \mathbf{\Omega}^2) \in \mathcal{U}(4)$, by (S1.3)

$$\begin{aligned}
1 &= P\left(16\log p_1 \le \max_{1\le i\le j\le p_1} \frac{1}{\hat{\theta}^1_{i,j} + \hat{\theta}^2_{i,j}} \left(\frac{\omega^1_{i,j}}{\sqrt{\omega^1_{i,i}\omega^1_{j,j}}} - \frac{\omega^2_{i,j}}{\sqrt{\omega^2_{i,i}\omega^2_{j,j}}} \right)^2\right) \\
&\le P(8\log p_1 - M_n \le \tilde{M}_n).
\end{aligned}$$

Decomposing a set $\{8\log p_1 - M_n \le \tilde{M}_n\}$ into

$$\begin{aligned}
&\{8\log p_1 - M_n \le \tilde{M}_n \le 4\log p_1 - \frac{1}{2}\log\log p_1\} \\
&\quad \bigcup \left(\{8\log p_1 - M_n \le \tilde{M}_n\} \cap \{\tilde{M}_n > 4\log p_1 - \frac{1}{2}\log\log p_1\} \right),
\end{aligned}$$

we have

$$\begin{aligned}
1 &= \lim_{n,p_1\to\infty} P(8\log p_1 - M_n \le \tilde{M}_n) \\
&= \lim_{n,p_1\to\infty} P(8\log p_1 - M_n \le \tilde{M}_n \le 4\log p_1 - \frac{1}{2}\log\log p_1) \\
&\quad + \lim_{n,p_1\to\infty} P(\{8\log p_1 - M_n \le \tilde{M}_n\} \cap \{\tilde{M}_n > 4\log p_1 - \frac{1}{2}\log\log p_1\}) \\
&= \lim_{n,p_1\to\infty} P(M_n \ge 4\log p_1 + \frac{1}{2}\log\log p_1) + 0.
\end{aligned}$$

The last inequality is from (S1.2). Thus, we obtain

$$\lim_{n,p\to\infty} P(M_n \ge 4\log p + \frac{1}{2}\log\log p) = 1,$$

where the convergence is uniform for $(\mathbf{\Omega}^1, \mathbf{\Omega}^2) \in \mathcal{U}(4)$, which implies

$$\begin{aligned}
\inf_{(\mathbf{\Omega}^1,\mathbf{\Omega}^2)\in\mathcal{U}(4)} P(\Psi_\alpha = 1) &= \inf_{(\mathbf{\Omega}^1,\mathbf{\Omega}^2)\in\mathcal{U}(4)} P(M_n \ge q_\alpha + 4\log p_1 - \log\log p_1) \\
&\ge \inf_{(\mathbf{\Omega}^1,\mathbf{\Omega}^2)\in\mathcal{U}(4)} P(M_n \ge 4\log p_1 + \frac{1}{2}\log\log p_1) \to 1
\end{aligned}$$

as $n, p_1 \to \infty$. Therefore, we obtain our goal. □

Remark on Theorem 2. Theorem 2 guarantees that the convergence of the power even when the dimension p_1 is sufficiently larger than the sample sizes n_1 and n_2, with $\log p_1 = o(n^{1/5})$ where $n = \min\{n_1, n_2\}$. Thus we need the scale assumption $\log p_1 = o(n^{1/5})$ for p_1 and n in Condition 3.

S1.3 Proof of Theorem 3

In fact, the probability measure P depends on $(\mathbf{\Omega}^1, \mathbf{\Omega}^2)$, so $P_{(\mathbf{\Omega}^1,\mathbf{\Omega}^2)}$ denote the probability measure P given $(\mathbf{\Omega}^1, \mathbf{\Omega}^2)$.

Proof of Theorem 3. To investigate the infimum of $\sup_{T_\alpha \in \mathcal{T}_\alpha} P_{(\mathbf{\Omega}^1, \mathbf{\Omega}^2)}(T_\alpha = 1)$, we should consider a case of $(\mathbf{\Omega}^1, \mathbf{\Omega}^2)$ with a small distance. Our strategy is finding a finite set of precision matrices $\mathcal{S}$ such that

$$\frac{1}{|\mathcal{S}|} \sum_{\Omega \in \mathcal{S}} \sup_{T_\alpha \in \mathcal{T}_\alpha} P_{(\mathbf{\Omega}, \mathbf{I}_{p_1+p_2})}(T_\alpha = 1) \leq \alpha + o(1)$$

Without loss of generality, we assume $n_1 \geq n_2$. Define $\mathbf{\Omega}_m = (\omega_{m,i,j}) \in \mathbb{R}^{(p_1+p_2)\times(p_1+p_2)}$ as follows,

$$\omega_{m,i,j} = \begin{cases} 1 & \text{if } i = j \\ \rho := \sqrt{\frac{n_1+n_2}{n_2(n_1 - c_0^2 \log p_1)}} c_0 (\log p_1)^{1/2} & \text{if } (i,j) = (m,1) \text{ or } (1,m) \\ 0 & \text{otherwise,} \end{cases}$$

where c_0 is to be determined later. From the assumption of Theorem 3, $\log p_1 = o(n_2)$, we know that $0 < \rho < 1$ is well-defined for sufficiently large n and p_1. Then if $(\mathbf{\Omega}^1, \mathbf{\Omega}^2) = (\mathbf{\Omega}_m, \mathbf{I}_{p_1+p_2})$,

$$\frac{1}{(\theta^1_{i,j} + \theta^2_{i,j})^{1/2}} \left| \frac{\omega^1_{i,j}}{\sqrt{\omega^1_{i,i}\omega^1_{j,j}}} - \frac{\omega^2_{i,j}}{\sqrt{\omega^2_{i,i}\omega^2_{j,j}}} \right| = \begin{cases} c_0 (\log p_1)^{1/2} & \text{if } (i,j) = (m,1) \text{ or } (1,m) \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathbf{I}_p$ is a $p \times p$ identity matrix. Thus, $(\mathbf{\Omega}_m, \mathbf{I}_{p_1+p_2}) \in \mathcal{U}(c_0)$ for sufficiently large n and p_1. Since $\mathcal{T}_\alpha$ is a set of all tests with size α,

$$P_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})}(T_\alpha = 1) \leq \alpha$$

for any $T_\alpha \in \mathcal{T}_\alpha$, which implies

$$\{\{T_\alpha = 1\} : T_\alpha \in \mathcal{T}_\alpha\} \subset \{A : P_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})}(A) \leq \alpha\}. \tag{S1.4}$$

Let $\mathcal{S}$ by $\{\Omega_m : m = 2, 3, \cdots, p_1\}$ and $dP_{\mathcal{S}} = \frac{1}{p_1-1}\sum_{m=2}^{p_1} dP_{(\boldsymbol{\Omega}_m, \mathbf{I}_{p_1+p_2})}$. Then, by (S1.4),

$$\begin{aligned}
\sup_{T_\alpha \in \mathcal{T}_\alpha} P_{\mathcal{S}}(T_\alpha = 1) &\leq \sup_{\{A: P_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})}(A) \leq \alpha\}} P_{\mathcal{S}}(A) \\
&= \sup_{\{A: P_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})}(A) \leq \alpha\}} |P_{\mathcal{S}}(A) - P_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})}(A)| + \alpha \\
&\leq \|P_{\mathcal{S}} - P_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})}\| + \alpha = \int |f-1| dP_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})} + \alpha,
\end{aligned}$$

where $\|\mu\|$ denotes the total variation norm of μ and $f = \frac{dP_{\mathcal{S}}}{dP_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})}}$ is a Radon-Nikodym derivative. Since

$$\begin{aligned}
\left(\int |f-1| dP_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})}\right)^2 &\leq \int (f-1)^2 dP_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})} \\
&= \int |f|^2 dP_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})} - 2\int f dP_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})} + 1 \\
&= \int |f|^2 dP_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})} - 1
\end{aligned}$$

then it is sufficient to show that

$$\int |f|^2 dP_{(\mathbf{I}_{p_1+p_2}, \mathbf{I}_{p_1+p_2})} = 1 + o(1).$$

By definition of $P_{(\mathbf{\Omega}_m,\mathbf{I}_{p_1+p_2})}$ and $P_{(\mathbf{I}_{p_1+p_2},\mathbf{I}_{p_1+p_2})}$,

$$dP_{(\mathbf{I}_{p_1+p_2},\mathbf{I}_{p_1+p_2})} = \sqrt{2\pi}^{-(p_1+p_2)(n_1+n_2)} \prod_{l=1}^{n_1} e^{-|\mathbf{y}_l^1|^2/2} d\mathbf{y}_l^1 \prod_{l=1}^{n_2} e^{-|\mathbf{y}_l^2|^2/2} d\mathbf{y}_l^2$$

$$dP_{(\mathbf{\Omega}_m,\mathbf{I}_{p_1+p_2})} = \sqrt{2\pi}^{-(p_1+p_2)(n_1+n_2)} \prod_{l=1}^{n_1} |\mathbf{\Omega}_m|^{1/2} \exp\left(-\frac{(\mathbf{y}_l^1)^T \Omega_m \mathbf{y}_l^1}{2}\right) d\mathbf{y}_l^1 \prod_{l=1}^{n_2} e^{-|\mathbf{y}_l^2|^2/2} d\mathbf{y}_l^2$$

so we can compute f directly,

$$\begin{aligned} f = \frac{dP_{\mathcal{S}}}{dP_{(\mathbf{I}_{p_1+p_2},\mathbf{I}_{p_1+p_2})}} &= \frac{1}{p_1-1}\sum_{m=1}^{p_1} \frac{dP_{(\mathbf{\Omega}_m,\mathbf{I}_{p_1+p_2})}}{dP_{(\mathbf{I}_{p_1+p_2},\mathbf{I}_{p_1+p_2})}} \\ &= \frac{1}{p_1-1}\sum_{m=1}^{p_1}\prod_{l=1}^{n_1} |\Omega_m|^{1/2} \exp\left(-\frac{(\mathbf{y}_l^1)^T(\mathbf{\Omega}_m - \mathbf{I}_{p_1+p_2})\mathbf{y}_l^1}{2}\right) \\ &= \frac{(1-\rho^2)^{n_1/2}}{p_1-1}\sum_{m=1}^{p_1}\prod_{l=1}^{n_1} \exp\left(-\frac{(\mathbf{y}_l^1)^T(\mathbf{\Omega}_m - \mathbf{I}_{p_1+p_2})\mathbf{y}_l^1}{2}\right) \end{aligned}$$

Using this, we can obtain

$$\begin{aligned} &\int |f|^2 dP_{(\mathbf{I}_{p_1+p_2},\mathbf{I}_{p_1+p_2})} \\ &= \frac{(1-\rho^2)^{n_1}}{(p_1-1)^2}\sum_{m,m'}\prod_{l=1}^{n_1}\int \exp\left(-\frac{(\mathbf{y}_l^1)^T(\mathbf{\Omega}_m + \mathbf{\Omega}_{m'} - 2\mathbf{I}_{p_1+p_2})\mathbf{y}_l^1}{2}\right) dP_{(\mathbf{I}_{p_1+p_2},\mathbf{I}_{p_1+p_2})} \\ &= \frac{(1-\rho^2)^{n_1}}{(p_1-1)^2}\sum_{m,m'}\prod_{l=1}^{n_1} |\mathbf{\Omega}_m + \mathbf{\Omega}_{m'} - \mathbf{I}_{p_1+p_2}|^{-1/2} \\ &= \frac{(1-\rho^2)^{n_1}}{(p_1-1)^2}\sum_{m=2}^{p_1}\left(1-4\rho^2\right)^{-n_1/2} + \frac{(1-\rho^2)^{n_1}}{(p_1-1)^2}\sum_{m\neq m'}(1-2\rho^2)^{-n_1/2} \\ &= \frac{(1-\rho^2)^{2n_1}}{(1-4\rho^2)^{n_1/2}}\frac{(1-\rho^2)^{-n_1}}{p_1-1} + \frac{(1-\rho^2)^{n_1}}{(1-2\rho^2)^{n_1/2}}\frac{p_1-2}{p_1-1} \end{aligned}$$

Finally, we will show that $\frac{(1-\rho^2)^{-n_1}}{p_1-1} = o(1)$ and $\frac{(1-a\rho^2)^{n_1/a}}{(1-2a\rho^2)^{n_1/(2a)}} = 1 + o(1)$ for any $a > 0$. Since $\rho = o(1)$, for sufficiently large n_1, n_2, p_1, we know $0 < \rho < \sqrt{1-\frac{1}{e}}$, which implies

$$\frac{1}{1-\rho^2} \leq \exp\left(\frac{e}{e-1}\rho^2\right). \tag{S1.5}$$

From (S1.5), we obtain

$$\begin{aligned}\frac{(1-\rho^2)^{-n_1}}{p_1-1} &= \frac{1}{p_1-1}\exp\left(\frac{en_1(n_1+n_2)}{(e-1)n_2(n_1-c_0^2\log p_1)}c_0^2\log p_1\right)\\ &\leq \frac{p_1}{p_1-1}p_1^{c_0^2\frac{Ke}{e-1}(1+o(1))-1},\end{aligned}$$

where $K>0$ is a constant for comparable condition between n_1 and n_2: $K^{-1}<\frac{n_1}{n_2}<K$. Thus, if we take c_0 sufficiently small then

$$\frac{(1-\rho^2)^{-n_1}}{p_1-1} = o(1) \quad \text{for sufficiently large } n_1, n_2, p_1.$$

Now, we will show $\frac{(1-a\rho^2)^{n_1/a}}{(1-2a\rho^2)^{n_1/(2a)}} = 1+o(1)$ for any $a>0$. We have

$$\begin{aligned}\frac{(1-a\rho^2)^{n_1/a}}{(1-2a\rho^2)^{n_1/(2a)}} &= \left(\frac{(1-a\rho^2)^2}{1-2a\rho^2}\right)^{n_1/(2a)} = \left(1+\frac{a^2\rho^4}{1-2a\rho^2}\right)^{n_1/(2a)}\\ &= \left(1+\frac{a^2\rho^4}{1-2a\rho^2}\right)^{\frac{1-2a\rho^2}{a^2\rho^4}\cdot\frac{an_1\rho^4}{2(1-2a\rho^2)}}.\end{aligned} \tag{S1.6}$$

In (S1.6),

$$\frac{a^2\rho^4}{1-2a\rho^2} = o(1) \quad \text{and} \quad \frac{an_1\rho^4}{2(1-2a\rho^2)},$$

which implies

$$\frac{(1-a\rho^2)^{n_1/a}}{(1-2a\rho^2)^{n_1/(2a)}} \to e^0 = 1$$

as $n, p_1 \to \infty$. Thus, $\frac{(1-a\rho^2)^{n_1/a}}{(1-2a\rho^2)^{n_1/(2a)}} = 1+o(1)$, then we also obtain

$$\frac{(1-\rho^2)^{n_1}}{(1-2\rho^2)^{n_1/2}} = 1+o(1)$$

$$\frac{(1-\rho^2)^{2n_1}}{(1-4\rho^2)^{n_1/2}} = \left(\frac{(1-\rho^2)^{n_1}}{(1-2\rho^2)^{n_1/2}}\cdot\frac{(1-2\rho^2)^{n_1/2}}{(1-4\rho^2)^{n_1/4}}\right)^2 = 1+o(1).$$

Therefore, if we take c_0 sufficiently small then

$$\frac{1}{p_1}\sum_{m=1}^{p_1}\sup_{T_\alpha\in\mathcal{T}_\alpha} P_{(\boldsymbol{\Omega}_m,\mathbf{I}_{p_1+p_2})}(T_\alpha=1)\leq \alpha+o(1)\leq 1-\gamma \quad \text{for sufficiently large } n, p_1.$$

It implies that there exists $1\leq \hat{m}\leq p_1$ such that

$$\sup_{T_\alpha\in\mathcal{T}_\alpha} P_{(\boldsymbol{\Omega}_{\hat{m}},\mathbf{I}_{p_1+p_2})}(T_\alpha=1)\leq 1-\gamma \quad \text{for sufficiently large } n, p_1.$$

Therefore,

$$\inf_{(\boldsymbol{\Omega}^1,\boldsymbol{\Omega}^2)\in\mathcal{U}(c_0)}\sup_{T_\alpha\in\mathcal{T}_\alpha} P_{(\boldsymbol{\Omega}^1,\boldsymbol{\Omega}^2)}(T_\alpha=1)\leq 1-\gamma \quad \text{for sufficiently large } n, p_1.$$

□

S1.4 Proof of Theorem 4

This proof is to demonstrate the FDP and FDR of the test based on our weighted statistic converge to $\frac{\alpha q_0}{q}$ as p_1 and n approach infinity in Theorem 4.

Proof of theorem 4. Let $\widehat{\mathrm{FDP}}(t)=\dfrac{(1-\Phi(t))(p_1^2-p_1)}{\max(N(t),1)}$ then

$$\lim_{t\nearrow(4\log p_1-\log\log p_1+\log\log\log p_1)^{1/2}}\widehat{\mathrm{FDP}}(t)\leq(\log\log p_1)^{-1/2}<\alpha,$$

$$\lim_{t\searrow 0}\widehat{\mathrm{FDP}}(t)=1>\alpha \quad \text{a.s.},$$

for sufficiently large p_1 and $0<\alpha<1$. Moreover, for $0<t<P_{\alpha/2}$,

$$\widehat{\mathrm{FDP}}(t)\geq 2(1-\Phi(t))>2(1-\Phi(P_{\alpha/2}))=\alpha,$$

where $P_{\alpha/2}$ is a value with $\big(1-\Phi(P_{\alpha/2})\big)=\alpha/2$. Thus,

$$0 < P_{\alpha/2} \le \hat{t}_0 < (4\log p_1 - \log\log p_1 + \log\log\log p_1)^{1/2} < 2(\log p_1)^{1/2} \quad \text{a.s..} \tag{S1.7}$$

Moreover, since $\widehat{\mathrm{FDP}}$ is continuous at almost everywhere points and increases at the discontinuous points, so

$$\frac{(1-\Phi(\hat{t}_0))(p_1^2-p_1)}{\max(N(\hat{t}_0),1)} = \alpha$$

for sufficiently large p_1 and $0<\alpha<1$. Then,

$$\begin{aligned}
\left|\frac{\mathrm{FDP}(\hat{t}_0)}{\alpha q_0/q}-1\right| &= \left|\frac{N_0(\hat{t}_0)}{\max(N(\hat{t}_0),1)}\cdot\frac{q}{\alpha q_0}-1\right| = \left|\frac{N_0(\hat{t}_0)}{2(1-\Phi(\hat{t}_0))q_0}-1\right| \\
&= \left|\frac{\sum_{(i,j)\in\mathcal{I}_0}\left[I(|\Delta_{i,j}|\ge\hat{t}_0)-2(1-\Phi(\hat{t}_0))\right]}{2(1-\Phi(\hat{t}_0))q_0}\right| \\
&\le \left|\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau}\left[I(|\Delta_{i,j}|\ge\hat{t}_0)-2(1-\Phi(\hat{t}_0))\right]}{2(1-\Phi(\hat{t}_0))q_0}\right| \\
&\quad + \left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\left[I(|\Delta_{i,j}|\ge\hat{t}_0)-I(|V_{i,j}|\ge\hat{t}_0)\right]}{2(1-\Phi(\hat{t}_0))q_0}\right| \\
&\quad + \left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\left[P(|V_{i,j}|\ge\hat{t}_0)-2(1-\Phi(\hat{t}_0))\right]}{2(1-\Phi(\hat{t}_0))q_0}\right| \\
&\quad + \left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\left[I(|V_{i,j}|\ge\hat{t}_0)-P(|V_{i,j}|\ge\hat{t}_0)\right]}{2(1-\Phi(\hat{t}_0))q_0}\right| =: B_1+B_2+B_3+B_4
\end{aligned}$$

Using Lemmas 14-17, we can show that each B_i is $o_p\{1\}$. Therefore, we obtain

$$\left|\frac{\mathrm{FDP}(\hat{t}_0)}{\alpha q_0/q}-1\right| = o_p\{1\}.$$

□

Remark on Theorem 4. According to Theorem 4, the FDP and FDR converge even if the dimension p_1 is much larger than the sample sizes n_1 and n_2, as long as $p_1 \le cn^r$ for some constants $c > 0$ and $r > 0$, where $n = \min\{n_1, n_2\}$. Unlike Theorem 2, this result ensures convergence even in cases of unbalanced sample sizes. That is, the convergence of the FDP and FDR can be expected even when $n_1 \gg n_2$ or $n_2 \gg n_1$.

S2 Technical Lemmas

In this section, we collect some technical lemmas and their proofs for proving main theorems. Lemmas 1-10 are about convergence rates of some quantities. The following Lemma is a particular case of (Cai and Liu, 2011, Lemma 1).

Lemma 1. *Let $\xi_1, \cdots, \xi_n$ be independent random variables with mean 0. Suppose that there exists $D_n > 0$ such that*

$$\sum_{k=1}^{n} E[\xi_k^2 e^{|\xi_k|}] \leq D_n^2.$$

Then, for $0 < x < D_n$,

$$P\left(\left|\sum_{k=1}^{n} \xi_k\right| \geq 2D_n x\right) \leq 2e^{-x^2}.$$

Proof. For any $t > 0$,

$$P\left(\sum_{k=1}^{n} \xi_k - 2D_n x \geq 0\right) \leq E\left[e^{t(\sum_{k=1}^{n} \xi_k - 2D_n x)}\right] = e^{-2tD_n x} \prod_{k=1}^{n} E\left[e^{t\xi_k}\right].$$

By using an inequality $e^s \leq 1 + s + s^2 e^{|s|}$, we obtain

$$P\left(\sum_{k=1}^{n} \xi_k - 2D_n x \geq 0\right) \leq e^{-2tD_n x} \prod_{k=1}^{n} \left(1 + tE[\xi_k] + t^2 E\left[\xi_k^2 e^{t|\xi_k|}\right]\right).$$

From $E[\xi_k] = 0$,

$$P\left(\sum_{k=1}^{n} \xi_k - 2D_n x \geq 0\right) \leq e^{-2tD_n x} \prod_{k=1}^{n} \left(1 + t^2 E\left[\xi_k^2 e^{t|\xi_k|}\right]\right) \leq e^{-2tD_n x} \prod_{k=1}^{n} \exp\left(t^2 E\left[\xi_k^2 e^{t|\xi_k|}\right]\right).$$

By taking $t = x/D_n \leq 1$, we obtain

$$P\left(\sum_{k=1}^{n} \xi_k \geq 2D_n x\right) \leq e^{-2x^2} \exp\left(\frac{x^2}{D_n^2} \sum_{k=1}^{n} E\left[\xi_k^2 e^{|\xi_k|}\right]\right) \leq e^{-x^2}.$$

And replacing ξ_k with $-\xi_k$ makes we obtain the opposite inequality.

$$P\left(-\sum_{k=1}^{n} \xi_k \geq 2D_n x\right) \leq e^{-x^2}.$$

Therefore,

$$P\left(\left|\sum_{k=1}^{n} \xi_k\right| \geq 2D_n x\right) \leq P\left(\sum_{k=1}^{n} \xi_k \geq 2D_n x\right) + P\left(-\sum_{k=1}^{n} \xi_k \geq 2D_n x\right) \leq 2e^{-x^2}.$$

□

Lemma 2. *Let $\{W_k\}_{k=1,\cdots,n} \subset \mathbb{R}$ be a set of identical independent observations such that $E\Big[W_k\Big] = \mu$, $\mathrm{Var}(W_k) = \sigma^2 < \infty$, and $E\Big[(W_k - \mu)^2 e^{|W_k - \mu|/(\log n)}\Big] < C$ for $k = 1, \cdots, n$ and for sufficiently large n. Then, there exists a function $f_W = f_W(\sigma, n) = o(1)$ such that*

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} W_k - \mu\right| \geq (2\sigma + f_W)\sqrt{M}(\log p/n)^{1/2}\right) \leq 2p^{-M}$$

for any $M > 0$ and for sufficiently large n and p with $\log p = o(n^{1/5})$.

Proof. Let $\xi_k = (\log n)^{-1}(W_k - \mu)$ then

$$\begin{aligned}\sum_{k=1}^{n} E\left[\xi_k^2 e^{|\xi_k|}\right] =& \sum_{k=1}^{n} E\left[\xi_k^2 e^{|\xi_k|} I_{\{|\xi_k|\leq(\log n)^{-1/2}\}}\right] + \sum_{k=1}^{n} E\left[\xi_k^2 e^{|\xi_k|} I_{\{|\xi_k|>(\log n)^{-1/2}\}}\right]\\ \leq& nE\left[\frac{(W_1-\mu)^2}{(\log n)^2} e^{\frac{1}{\sqrt{\log n}}}\right] + nE\left[\xi_1^2 e^{|\xi_1|} I_{\{|\xi_1|>(\log n)^{-1/2}\}}\right]\\ \leq& \frac{n}{(\log n)^2} e^{\frac{1}{\sqrt{\log n}}} \mathrm{Var}(W_k) + \frac{n}{(\log n)^2} E\left[(W_1-\mu)^2 e^{|W_1-\mu|/(\log n)} I_{\{|W_1-\mu|>(\log n)^{1/2}\}}\right]\\ =& \frac{n}{(\log n)^2}\left(\sigma^2 + \sigma^2(e^{\frac{1}{\sqrt{\log n}}} - 1) + E\left[(W_1-\mu)^2 e^{|W_1-\mu|/(\log n)} I_{\{|W_1-\mu|>(\log n)^{1/2}\}}\right]\right)\end{aligned}$$

Let $f_W := \sigma(e^{\frac{1}{\sqrt{\log n}}} - 1) + \sigma^{-1} E\left[(W_1-\mu)^2 e^{|W_1-\mu|/(\log n)} I_{\{|W_1-\mu|>(\log n)^{1/2}\}}\right]$ and $D_n = \frac{\sqrt{n}}{\log n}(\sigma + f_W/2)$ then by the dominated convergence theorem

$$f_W = o(1) \quad \text{as } n \to \infty$$

and

$$\sum_{k=1}^{n_d} E\left[\xi_k^2 e^{|\xi_k|}\right] \leq \frac{n}{(\log n)^2}(\sigma^2 + \sigma f_W) \leq D_n^2.$$

Let $x = \sqrt{M \log p}$, then for sufficiently large n and p with $n^{1/5} = o(\log p)$,

$$0 < x < D_n$$

and by Lemma 1,

$$P\left((\log n)^{-1}\left|\sum_{k=1}^{n}(W_k - \mu)\right| \geq 2\frac{\sqrt{n}}{\log n}(\sigma + f_W/2)\sqrt{M\log p}\right) \leq 2e^{-x^2} = 2p^{-M}$$

Thus,

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} W_k - \mu\right| \geq (2\sigma + f_W)\sqrt{M}(\log p/n)^{1/2}\right) \leq 2p^{-M}$$

□

Lemma 3. *Let* $\{W_k\}_{k=1,\cdots,n} \subset \mathbb{R}$ *be a set of identical independent observations such that* $E[W_k] = \dfrac{n-1}{n}\mu$, $\mathrm{Var}(W_k) = \sigma^2 < \infty$, *and* $E\left[(W_k - \frac{n-1}{n}\mu)^2 e^{|W_k - \frac{n-1}{n}\mu|/(\log n)}\right] < C$ *for* $k = 1, \cdots, n$ *and for sufficiently large* n*. Then, there exists a function* $f_W = f_W(\sigma, n)$ *such that*

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} W_k - \mu\right| \geq (2\sigma + f_W)\sqrt{M}(\log p/n)^{1/2}\right) \leq 2p^{-M}$$

for any $M > 0$ *and for sufficiently large* n *and* p *with* $\log p = o(n^{1/5})$.

Proof. Let $\xi_k = (\log n)^{-1}\left(W_k - \frac{n-1}{n}\mu\right)$ and decompose $\xi_k^2 e^{|\xi_k|}$ into two parts:

$$\xi_k^2 e^{|\xi_k|} = \xi_k^2 e^{|\xi_k|} I_{\{|\xi_k| \leq (\log n)^{-1/2}\}} + \xi_k^2 e^{|\xi_k|} I_{\{|\xi_k| > (\log n)^{-1/2}\}}. \tag{S2.8}$$

Then, from the first part of (S2.8), we derive

$$\begin{aligned}\sum_{k=1}^{n} E\left[\xi_k^2 e^{|\xi_k|} I_{\{|\xi_k| \leq (\log n)^{-1/2}\}}\right] &\leq \sum_{k=1}^{n} E\left[\frac{(W_k - \frac{n-1}{n}\mu)^2}{(\log n)^2} e^{\frac{1}{\sqrt{\log n}}}\right] \\ &\leq \sum_{k=1}^{n} E\left[\frac{(W_k - \frac{n-1}{n}\mu)^2}{(\log n)^2} e^{\frac{1}{\sqrt{\log n}}}\right] \\ &\leq \frac{n}{(\log n)^2} e^{\frac{1}{\sqrt{\log n}}} \mathrm{Var}(W_k) \leq \frac{n}{(\log n)^2}\left(\sigma^2 + \sigma^2(e^{\frac{1}{\sqrt{\log n}}} - 1)\right),\end{aligned}$$

and from the second part of (S2.8), we derive

$$\begin{aligned}&\sum_{k=1}^{n} E\left[\xi_k^2 e^{|\xi_k|} I_{\{|\xi_k| > (\log n)^{-1/2}\}}\right] \\ &\quad = nE\left[\xi_1^2 e^{|\xi_1|} I_{\{|\xi_1| > (\log n)^{-1/2}\}}\right] \\ &\quad = \frac{n}{(\log n)^2} E\left[(W_1 - \frac{n-1}{n}\mu)^2 e^{|W_1 - \frac{n-1}{n}\mu|/(\log n)} I_{\{|W_1 - \frac{n-1}{n}\mu| > (\log n)^{1/2}\}}\right].\end{aligned}$$

Let $f_W := 2\sigma(e^{\frac{1}{\sqrt{\log n}}} - 1) + 2\sigma^{-1}E\left[(W_1 - \frac{n-1}{n}\mu)^2 e^{|W_1 - \frac{n-1}{n}\mu|/(\log n)} I_{\{|W_1 - \frac{n-1}{n}\mu| > (\log n)^{1/2}\}}\right]$ and $D_n = \frac{\sqrt{n}}{\log n}(\sigma + f_W/2 - n^{-1/2}|\mu|/2)$ then by the dominated convergence theorem

$$\frac{2\sigma}{\sqrt{\log n}} \leq f_W = o(1) \quad \text{as } n \to \infty. \tag{S2.9}$$

Also,

$$\begin{aligned}\sum_{k=1}^{n_d} E\left[\xi_k^2 e^{|\xi_k|}\right] &= \sum_{k=1}^{n} E\left[\xi_k^2 e^{|\xi_k|} I_{\{|\xi_k| \leq (\log n)^{-1/2}\}}\right] + \sum_{k=1}^{n} E\left[\xi_k^2 e^{|\xi_k|} I_{\{|\xi_k| > (\log n)^{-1/2}\}}\right] \\ &\leq \frac{n}{(\log n)^2}(\sigma^2 + \sigma f_W/2) \leq \frac{n}{(\log n)^2}(\sigma + f_W/4)^2 \leq D_n^2\end{aligned}$$

for sufficiently large n. The last inequality is derived from

$$\frac{|\mu|}{2\sqrt{n}} \leq \frac{\sigma}{2\sqrt{\log n}} \leq \frac{f_W}{4}, \quad \text{by (S2.9)}$$

for sufficiently large n. By Lemma 1, we obtain

$$P\left((\log n)^{-1}\left|\sum_{k=1}^{n}(W_k - \frac{n-1}{n}\mu)\right| \geq 2\frac{\sqrt{n}}{\log n}(\sigma + f_W/2 - n^{-1/2}|\mu|/2)\sqrt{M\log p}\right) \leq 2e^{-x^2}$$

for $0 < x < D_n$. Let $x = \sqrt{M \log p}$, then for sufficiently large n and p with $n^{1/5} = o(\log p)$, we have

$$0 < x < D_n.$$

Thus,

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} W_k - \frac{n-1}{n}\mu\right| \geq (2\sigma + f_W - n^{-1/2}|\mu|)\sqrt{M}(\log p/n)^{1/2}\right) \leq 2p^{-M}$$

Therefore,

$$\begin{aligned}
&P\left(\left|\frac{1}{n}\sum_{k=1}^{n}W_k-\mu\right|\geq(2\sigma+f_W)\sqrt{M}(\log p/n)^{1/2}\right)\\
&\quad\leq P\left(\left|\frac{1}{n}\sum_{k=1}^{n}W_k-\frac{n-1}{n}\mu\right|\geq(2\sigma+f_W)\sqrt{M}(\log p/n)^{1/2}-\frac{1}{n}|\mu|\right)\\
&\quad\leq P\left(\left|\frac{1}{n}\sum_{k=1}^{n}W_k-\frac{n-1}{n}\mu\right|\geq(2\sigma+f_W-|\mu|(Mn\log p)^{-1/2})\sqrt{M}(\log p/n)^{1/2}\right).
\end{aligned}$$

If $p>e^{\frac{1}{M}}$, then

$$\begin{aligned}
&P\left(\left|\frac{1}{n}\sum_{k=1}^{n}W_k-\mu\right|\geq(2\sigma+f_W)\sqrt{M}(\log p/n)^{1/2}\right)\\
&\quad\leq P\left(\left|\frac{1}{n}\sum_{k=1}^{n}W_k-\frac{n-1}{n}\mu\right|\geq(2\sigma+f_W-n^{-1/2}|\mu|)\sqrt{M}(\log p/n)^{1/2}\right)\leq 2p^{-M}.
\end{aligned}$$

□

Lemma 4. *Assume Condition 1. For any $d=1,2,\ 1\leq i,j,i'\leq p_1,\ i'\neq i$, and $k=1,\cdots,n_d$ and for sufficiently large n_d,*

$$\varepsilon^d_{k,i},\ \ \varepsilon^d_{k,i}\varepsilon^d_{k,j},\ \ (\varepsilon^d_{k,i})^2/r^d_{i,i},\ \ \tilde{\varepsilon}^d_{k,i}\tilde{\mathbf{X}}^d_{k,i'},\ \ \text{and}\ \tilde{\varepsilon}^d_{k,i}(\tilde{\mathbf{X}}^d_{k,-i})^T\boldsymbol{\beta}^d_i$$

satisfies assumptions in Lemma 2, and

$$\tilde{\mathbf{X}}^d_{k,i}\tilde{\mathbf{X}}^d_{k,j},\ \ \tilde{\varepsilon}^d_{k,i}\tilde{\varepsilon}^d_{k,j},\ \ \text{and}\ \tilde{\varepsilon}^d_{k,i}\tilde{\varepsilon}^d_{k,j}/r^d_{i,j}$$

satisfies assumptions in Lemma 3, and where $\tilde{\varepsilon}^d_{k,i}=\varepsilon^d_{k,i}-\bar{\varepsilon}^d_{\cdot,i}$, $\tilde{\mathbf{X}}^d_{k,i}=\mathbf{X}^d_{k,i}-\bar{\mathbf{X}}^d_{\cdot,i}$, and $\tilde{\mathbf{X}}^d_{k,-i}=\mathbf{X}^d_{k,-i}-\bar{\mathbf{X}}^d_{\cdot,-i}$.

Proof. From the definition of $\varepsilon^d_{k,i}$ and its normality, we obtain the results for sufficeintly large

---

n_d. ∎

Lemma 5. *Assume Conditions 1 and 3. There exists $f = f(C_0, n, K)$ such that for any $M > 0$, $1 \le i, j, i' \le p_1$, $i' \ne i$, and sufficiently large n and p_1,*

$$P\left(\left|\bar{\varepsilon}^d_{\cdot,i}\right| \ge (2\sqrt{3}C_0 + f)\sqrt{M}(\log p_1/n_d)^{1/2}\right) \le 2p_1^{-M}$$

$$P\left(\left|U^d_{i,j}\right| \ge (2\sqrt{3}C_0 + f)\sqrt{M}(\log p_1/n_d)^{1/2}\right) \le 2p_1^{-M}$$

$$P\left(\left|\frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}\tilde{\mathbf{X}}^d_{k,i'}\right| \ge (2C_0 + f)\sqrt{M}(\log p_1/n_d)^{1/2}\right) \le 2p_1^{-M}$$

$$P\left(\left|\frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}(\tilde{\mathbf{X}}^d_{k,i'})^T\boldsymbol{\beta}^d_i\right| \ge (2\sqrt{3}C_0 + f)\sqrt{M}(\log p_1/n_d)^{1/2}\right) \le 2p_1^{-M}$$

Moreover, for any $M > 0$ and sufficiently large n and p_1,

$$P\left(\max_i\left|\bar{\varepsilon}^d_{\cdot,i}\right| \ge (2\sqrt{3}C_0 + f)\sqrt{M+1}(\log p_1/n_d)^{1/2}\right) = O(p_1^{-M})$$

$$P\left(\max_{i,j}\left|U^d_{i,j}\right| \ge (2\sqrt{3}C_0 + f)\sqrt{M+2}(\log p_1/n_d)^{1/2}\right) = O(p_1^{-M})$$

$$P\left(\max_{i,i'}\left|\frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}\tilde{\mathbf{X}}^d_{k,i'}\right| \ge (2\sqrt{3}C_0 + f)\sqrt{M+2}(\log p_1/n_d)^{1/2}\right) = O(p_1^{-M})$$

$$P\left(\max_i\left|\frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}(\tilde{\mathbf{X}}^d_{k,i'})^T\boldsymbol{\beta}^d_i\right| \ge (2\sqrt{3}C_0 + f)\sqrt{M+1}(\log p_1/n_d)^{1/2}\right) = O(p_1^{-M})$$

Proof. By Lemma 2 and Lemma 4, this results are derived directly. ∎

Lemma 6. *Assume Conditions 1 and 3. There exists $f = f(C_0, n, K)$ such that for any $M > 0$, $1 \le i, j \le p_1$, and sufficiently large n and p_1,*

$$P\left(\left|\hat{\sigma}^d_{i,j} - \sigma^d_{i,j}\right| \ge (2\sqrt{3}C_0 + f)\sqrt{M}(\log p_1/n_d)^{1/2}\right) \le 2p_1^{-M}$$

$$P\left(\left|\tilde{R}^d_{i,j} - r^d_{i,j}\right| \ge (2\sqrt{3}C_0 + f)\sqrt{M}(\log p_1/n_d)^{1/2}\right) \le 2p_1^{-M}$$

where $\hat{\boldsymbol{\Sigma}}^d = (\hat{\sigma}^d_{i,j})_{1\le i,j\le p_1+p_2} = \frac{1}{n_d}\sum_{k=1}^{n_d}(\mathbf{X}^d_k - \bar{\mathbf{X}}^d)(\mathbf{X}^d_k - \bar{\mathbf{X}}^d)^T$ *and* $\tilde{R}^d_{i,j} = \frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}\tilde{\varepsilon}^d_{k,j}$.

Moreover, for any $M > 0$ and sufficiently large n and p_1,

$$P\left(\max_{i,j}\left|\hat{\sigma}^d_{i,j} - \sigma^d_{i,j}\right| \geq (2\sqrt{3}C_0 + f)\sqrt{M+2}(\log p_1/n_d)^{1/2}\right) = O(p_1^{-M})$$
$$P\left(\max_{i,j}\left|\tilde{R}^d_{i,j} - r^d_{i,j}\right| \geq (2\sqrt{3}C_0 + f)\sqrt{M+2}(\log p_1/n_d)^{1/2}\right) = O(p_1^{-M})$$

Proof. By Lemma 3 and Lemma 4, the results are derived directly. □

Lemma 7. *Assume Conditions 1, 3, and 4, then we have*

$$\tilde{r}^d_{i,j} = \tilde{R}^d_{i,j} - \tilde{r}^d_{i,i}(\hat{\beta}^d_{i,j} - \beta^d_{i,j}) - \tilde{r}^d_{j,j}(\hat{\beta}^d_{j-1,i} - \beta^d_{j-1,i}) + o_p\{(n_d \log p_1)^{-1/2}\}$$

uniformly in $1 \leq i < j \leq p_1$ and $d = 1, 2$, where $\tilde{R}^d_{i,j} = \frac{1}{n_d}\sum_{k=1}^{n_d} \tilde{\varepsilon}^d_{k,i}\tilde{\varepsilon}^d_{k,j}$.

$$\tilde{r}^d_{i,i} = \tilde{R}^d_{i,i} + o_p\{(n_d \log p_1)^{-1/2}\}$$

uniformly in $1 \leq i \leq p_1$ and $d = 1, 2$.

Proof. We prove the lemma in 2 steps.

Step 1. For $1 \leq i \leq j \leq p_1$, the following equality holds uniformly,

$$\begin{aligned}\tilde{r}^d_{i,j} =& \tilde{R}^d_{i,j} - \frac{1}{n_d}\sum_{k=1}^{n_d} \tilde{\varepsilon}^d_{k,i}\tilde{\mathbf{X}}^d_{k,i}(\hat{\beta}^d_{i,j} - \beta^d_{i,j})I(i \neq j) \\ & - \frac{1}{n_d}\sum_{k=1}^{n_d} \tilde{\varepsilon}^d_{k,j}\tilde{\mathbf{X}}^d_{k,j}(\hat{\beta}^d_{j-1,i} - \beta^d_{j-1,i})I(i \neq j) + o_p\{(n_d \log p_1)^{-1/2}\}.\end{aligned}$$

Proof of step 1. By definition of $\hat{\varepsilon}^d_{k,i} = \tilde{\mathbf{X}}^d_{k,i} - \tilde{\mathbf{X}}^d_{k,-i}\hat{\boldsymbol{\beta}}^d_i$ and $\tilde{\varepsilon}^d_{k,i} = \tilde{\mathbf{X}}^d_{k,i} - \tilde{\mathbf{X}}^d_{k,-i}\boldsymbol{\beta}^d_i$,

$$\begin{aligned}\frac{1}{n_d}\sum_{k=1}^{n_d}\hat{\varepsilon}^d_{k,i}\hat{\varepsilon}^d_{k,j} =& \frac{1}{n_d}\sum_{k=1}^{n_d}\left(\tilde{\varepsilon}^d_{k,i} - \tilde{\mathbf{X}}^d_{k,-i}(\hat{\boldsymbol{\beta}}^d_i - \boldsymbol{\beta}^d_i)\right)\left(\tilde{\varepsilon}^d_{k,j} - \tilde{\mathbf{X}}^d_{k,-j}(\hat{\boldsymbol{\beta}}^d_j - \boldsymbol{\beta}^d_j)\right)\\ =& \frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}\tilde{\varepsilon}^d_{k,j} - \frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}\tilde{\mathbf{X}}^d_{k,-j}(\hat{\boldsymbol{\beta}}^d_j - \boldsymbol{\beta}^d_j)\\ &- \frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,j}\tilde{\mathbf{X}}^d_{k,-i}(\hat{\boldsymbol{\beta}}^d_i - \boldsymbol{\beta}^d_i) + (\hat{\boldsymbol{\beta}}^d_i - \boldsymbol{\beta}^d_i)^T\hat{\boldsymbol{\Sigma}}^d_{-i,-j}(\hat{\boldsymbol{\beta}}^d_j - \boldsymbol{\beta}^d_j)\\ &=: \frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}\tilde{\varepsilon}^d_{k,j} - a_1 - a_2 + a_3,\end{aligned}$$

where $\hat{\boldsymbol{\Sigma}}^d = (\hat{\sigma}^d_{i,j})_{1\le i,j\le p_1+p_2} = \frac{1}{n_d}\sum_{k=1}^{n_d}(\mathbf{X}^d_k - \bar{\mathbf{X}}^d)(\mathbf{X}^d_k - \bar{\mathbf{X}}^d)^T$. Now we will bound the last three terms separately. By Conditions 1 and 4 and Lemma 6,

$$\begin{aligned}|a_3| \le& \left|(\hat{\boldsymbol{\beta}}^d_i - \boldsymbol{\beta}^d_i)^T(\hat{\boldsymbol{\Sigma}}^d_{-i,-j} - \boldsymbol{\Sigma}^d_{-i,-j})(\hat{\boldsymbol{\beta}}^d_j - \boldsymbol{\beta}^d_j)\right| + \left|(\hat{\boldsymbol{\beta}}^d_i - \boldsymbol{\beta}^d_i)^T\boldsymbol{\Sigma}^d_{-i,-j}(\hat{\boldsymbol{\beta}}^d_j - \boldsymbol{\beta}^d_j)\right|\\ =& o_p\{(\log p_1)^{-1}\}O_p\{(\log p_1/n_d)^{1/2}\}o_p\{(\log p_1)^{-1}\} + o_p\{(n_d\log p_1)^{-1/2}\}\\ =& o_p\{(n_d\log p_1)^{-1/2}\}\end{aligned}$$

uniformly in $1 \le i \le j \le p_1$. For a_1, by Condition 4 and Lemma 5,

$$\begin{aligned}\left|a_1 - \frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}\tilde{\mathbf{X}}^d_{k,i}(\hat{\beta}^d_{i,j} - \beta^d_{i,j})I(i \ne j)\right| \le& \max_i\max_{i'\ne i,j}\left|\frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,i}\tilde{\mathbf{X}}^d_{k,i'}\right| o_p\{(\log p_1)^{-1}\}\\ =& o_p\{(n_d\log p_1)^{-1/2}\}\end{aligned}$$

uniformly in $1 \le i \le j \le p_1$. Since a_2 has a similar form to a_1, so

$$a_2 = o_p\{(n_d\log p_1)^{-1/2}\} + \frac{1}{n_d}\sum_{k=1}^{n_d}\tilde{\varepsilon}^d_{k,j}\tilde{\mathbf{X}}^d_{k,j}(\hat{\beta}^d_{j-1,i} - \beta^d_{j-1,i})I(i \ne j)$$

□

Step 2. For $1 \le i < j \le p_1$, the following equality holds

$$\tilde{r}^d_{i,j} = \tilde{R}^d_{i,j} - \tilde{r}^d_{i,i}(\hat{\beta}^d_{i,j} - \beta^d_{i,j}) - \tilde{r}^d_{j,j}(\hat{\beta}^d_{j-1,i} - \beta^d_{j-1,i}) + o_p\{(n_d \log p_1)^{-1/2}\}$$

Proof of step 2. Since

$$\begin{aligned}\tilde{R}^d_{i,i,d} =& \frac{1}{n_d}\sum_{k=1}^{n_d} \tilde{\varepsilon}^d_{k,i}\tilde{\varepsilon}^d_{k,i} = \frac{1}{n_d}\sum_{k=1}^{n_d} \tilde{\varepsilon}^d_{k,i}(\tilde{\mathbf{X}}^d_{k,i} - \tilde{\mathbf{X}}^d_{k,-i}\boldsymbol{\beta}^d_i) \\ =& \frac{1}{n_d}\sum_{k=1}^{n_d} \tilde{\varepsilon}^d_{k,i}\tilde{\mathbf{X}}^d_{k,i} + O_p\{(\log p_1/n_d)^{1/2}\}.\end{aligned}$$

The last equality is from Lemma 5. Thus by step 1 for $i = j$,

$$\begin{aligned}\frac{1}{n_d}\sum_{k=1}^{n_d} \tilde{\varepsilon}^d_{k,i}\tilde{X}^d_{k,i} =& \tilde{R}^d_{i,i} + O_p\{(\log p_1/n_d)^{1/2}\} \\ =& \tilde{r}^d_{i,i} + o_p\{(n_d \log p_1)^{-1/2}\} + O_p\{(\log p_1/n_d)^{1/2}\} \\ =& \tilde{r}^d_{i,i} + O_p\{(\log p_1/n_d)^{1/2}\}\end{aligned}$$

uniformly in $1 \le i \le p_1$. Therefore,

$$\begin{aligned}\tilde{r}^d_{i,j} =& \tilde{R}^d_{i,j} - \tilde{r}^d_{i,i}(\hat{\beta}^d_{i,j} - \beta^d_{i,j}) - \tilde{r}^d_{j,j}(\hat{\beta}^d_{j-1,i} - \beta^d_{j-1,i}) \\ &+ O_p\{(\log p_1/n_d)^{1/2}\}o_p\{(\log p_1)^{-1}\} + o_p\{(n_d \log p_1)^{-1/2}\} \\ =& \tilde{R}^d_{i,j} - \tilde{r}^d_{i,i}(\hat{\beta}^d_{i,j} - \beta^d_{i,j}) - \tilde{r}^d_{j,j}(\hat{\beta}^d_{j-1,i} - \beta^d_{j-1,i}) + o_p\{(n_d \log p_1)^{-1/2}\}\end{aligned}$$

□

□

Lemma 8. *Assume Conditions 1, 3, and 4, then we have*

$$\hat{r}^d_{i,j} - (\omega^d_{i,i}\tilde{R}^d_{i,i} + \omega^d_{j,j}\tilde{R}^d_{j,j} - 1)r^d_{i,j} + U^d_{i,j} = o_p\{(n_d \log p_1)^{-1/2}\}$$

uniformly in $1 \le i < j \le p_1$ and $d = 1, 2$. Also, we have

$$\hat{r}^d_{i,i} = r^d_{i,i} + U^d_{i,i} + o_p\{(n_d \log p_1)^{-1/2}\} \tag{S2.10}$$

uniformly in $1 \le i \le p_1$ and $d = 1, 2$.

Proof. By using Lemma 7,

$$\begin{aligned}
\hat{r}^d_{i,j} =& -(\tilde{r}^d_{i,j} + \tilde{r}^d_{i,i}\hat{\beta}^d_{i,j} + \tilde{r}^d_{j,j}\hat{\beta}^d_{j-1,i}) \\
=& -(\tilde{R}^d_{i,j} + \tilde{r}^d_{i,i}\beta^d_{i,j} + \tilde{r}^d_{j,j}\beta^d_{j-1,i}) + o_p\{(n_d \log p_1)^{-1/2}\} \\
=& -(\tilde{R}^d_{i,j} + \tilde{R}^d_{i,i}\beta^d_{i,j} + \tilde{R}^d_{j,j}\beta^d_{j-1,i}) + o_p\{(n_d \log p_1)^{-1/2}\} \\
=& -\left(\tilde{R}^d_{i,j} + \tilde{R}^d_{i,i}\frac{\omega^d_{i,j}}{\omega^d_{j,j}} + \tilde{R}^d_{j,j}\frac{\omega^d_{i,j}}{\omega^d_{i,i}}\right) + o_p\{(n_d \log p_1)^{-1/2}\} \\
=&(-1 + \tilde{R}^d_{i,i}\omega^d_{i,i} + \tilde{R}^d_{j,j}\omega^d_{j,j})r^d_{i,j} + r^d_{i,j} - \tilde{R}^d_{i,j} + o_p\{(n_d \log p_1)^{-1/2}\}.
\end{aligned}$$

From the definitions of $r^d_{i,j}$, $\tilde{R}^d_{i,j}$, and $U^d_{i,j}$,

$$\begin{aligned}
\hat{r}^d_{i,j} - (\omega^d_{i,i}\tilde{R}^d_{i,i} + \omega^d_{j,j}\tilde{R}^d_{j,j} - 1)r^d_{i,j} =& E[\varepsilon^d_{k,i}\varepsilon^d_{k,j}] - \left(\frac{1}{n_d}\sum_{k=1}^{n_d}\varepsilon^d_{k,i}\varepsilon^d_{k,j} - \bar{\varepsilon}^d_{\cdot,i}\bar{\varepsilon}^d_{\cdot,j}\right) + o_p\{(n_d \log p_1)^{-1/2}\} \\
=& -U^d_{i,j} + \bar{\varepsilon}^d_{\cdot,i}\bar{\varepsilon}^d_{\cdot,j} + o_p\{(n_d \log p_1)^{-1/2}\}.
\end{aligned}$$

By Lemma 5,

$$\hat{r}^d_{i,j} - (\omega^d_{i,i}\tilde{R}^d_{i,i} + \omega^d_{j,j}\tilde{R}^d_{j,j} - 1)r^d_{i,j} + U^d_{i,j} = \Big(O_p\{(\log p_1/n_d)^{1/2}\}\Big)^2 + o_p\{(n_d \log p_1)^{-1/2}\}$$
$$= o_p\{(n_d \log p_1)^{-1/2}\}.$$

Also, we can obtain (S2.10) by the similar way. □

Lemma 9. *Assume Conditions 1, 3, and 4, then we have*

$$\hat{\omega}^{d,w}_{i,j} - \tilde{U}^d_{i,j} = \frac{(\omega^d_{i,i}U_{i,i} + \omega^d_{j,j}U_{j,j})r^d_{i,j}}{2\sqrt{r^d_{i,i}r^d_{j,j}}} + o_p\{(n_d \log p_1)^{-1/2}\}$$

uniformly in $1 \leq i < j \leq p_1$ *and* $d = 1, 2$. *In particular, for* $\tilde{A} \subset A$, *we have*

$$\max_{(i,j)\in\tilde{A}} |\hat{\omega}^{d,w}_{i,j} - \tilde{U}^d_{i,j}| = O_p\{(\log p_1/n_d)^{1/2}\} \max_{(i,j)\in\tilde{A}} |r^d_{i,j}| + o_p\{(n_d \log p_1)^{-1/2}\}. \tag{S2.11}$$

Remark on Lemma 9. Using equation (S2.11), we can approximate the estimator $\hat{\omega}^d_{i,j}$ of $\frac{\omega^d_{i,j}}{\sqrt{\hat{\omega}^d_{i,i}\hat{\omega}^d_{j,j}}}$ by the random variable $\tilde{U}^d_{i,j}$, derived from the multivariate normal distribution $(\varepsilon^d_{k,i})_{i=1,\cdots,n_d}$. However, while this approximation requires a sufficient convergence rate, equation (S2.11) only provides a relatively slow convergence rate of $O_p\{(\log p_1/n_d)^{1/2}\}$. This necessitates Condition 2, which indicates that the set A_τ of indices (i,j), where $|r^d_{i,j}|$ is not small, has a negligible size. As a result, for indices in A^c_τ, we have $\max_{(i,j)\in A^c_\tau} |r^d_{i,j}| \leq (\log p_1)^{-1-\tau}$. By using the equation (S2.11), we can achieve a faster convergence rate for $\hat{\omega}^d_{i,j}$, specifically $o_p\{(n_d \log p_1)^{-1/2}\}$, when considering the indices in A^c_τ. For the remaining indices in A_τ, we can disregard the set because A_τ is a small set. Therefore, using (S2.11), $\hat{\omega}^d_{i,j}$ can be approximated by the random variable $\tilde{U}^d_{i,j}$ with an appropriate convergence rate. Since the distribution of $\tilde{U}^d_{i,j}$ is much easier to compute compared to that of $\hat{\omega}^d_{i,j}$, we can derive the asymptotic distribution of $\hat{\omega}_{i,j}^d$ through (S2.11).

Proof. By Lemma 6, we have $U_{i,i}^d = O_p\{(\log p_1/n_d)^{1/2}\}$, then

$$\begin{aligned}
\sqrt{r_{i,i}^d r_{j,j}^d} - \sqrt{(r_{i,i}^d + U_{i,i}^d)(r_{j,j}^d + U_{j,j}^d)} &= \frac{-\left(r_{j,j}^d U_{i,i}^d + r_{i,i}^d U_{j,j}^d + U_{i,i}^d U_{j,j}^d\right)}{\sqrt{r_{i,i}^d r_{j,j}^d} + \sqrt{(r_{i,i}^d + U_{i,i}^d)(r_{j,j}^d + U_{j,j}^d)}} \\
&= \frac{-\left(r_{j,j}^d U_{i,i}^d + r_{i,i}^d U_{j,j}^d + U_{i,i}^d U_{j,j}^d\right)}{2\sqrt{r_{i,i}^d r_{j,j}^d}} + O_p\{\log p_1/n_d\} \\
&= \frac{-\left(r_{j,j}^d U_{i,i}^d + r_{i,i}^d U_{j,j}^d\right)}{2\sqrt{r_{i,i}^d r_{j,j}^d}} + O_p\{\log p_1/n_d\}.
\end{aligned}$$

Also, we know

$$\hat{r}_{i,i}^d = r_{i,i}^d + U_{i,i}^d + o_p\{(n_d \log p_1)^{-1/2}\}$$

from Lemma and 8, so we obtain

$$\begin{aligned}
\frac{1}{\sqrt{\hat{r}_{i,i}^d \hat{r}_{j,j}^d}} - \frac{1}{\sqrt{r_{i,i}^d r_{j,j}^d}} &= \frac{1}{\sqrt{(r_{i,i}^d + U_{i,i}^d)(r_{j,j}^d + U_{j,j}^d)}} - \frac{1}{\sqrt{r_{i,i}^d r_{j,j}^d}} + o_p\{(n_d \log p_1)^{-1/2}\} \\
&= \frac{\sqrt{r_{i,i}^d r_{j,j}^d} - \sqrt{(r_{i,i}^d + U_{i,i}^d)(r_{j,j}^d + U_{j,j}^d)}}{\sqrt{r_{i,i}^d r_{j,j}^d (r_{i,i}^d + U_{i,i}^d)(r_{j,j}^d + U_{j,j}^d)}} + o_p\{(n_d \log p_1)^{-1/2}\} \\
&= \frac{-\left(r_{j,j}^d U_{i,i}^d + r_{i,i}^d U_{j,j}^d\right)}{2 r_{i,i}^d r_{j,j}^d \sqrt{(r_{i,i}^d + U_{i,i}^d)(r_{j,j}^d + U_{j,j}^d)}} + o_p\{(n_d \log p_1)^{-1/2}\} \\
&= \frac{-\left(\omega_{i,i}^d U_{i,i}^d + \omega_{j,j}^d U_{j,j}^d\right)}{2\sqrt{r_{i,i}^d r_{j,j}^d}} + o_p\{(n_d \log p_1)^{-1/2}\}.
\end{aligned}$$

From Lemma 8, we also have

$$\begin{aligned}
\hat{r}_{i,j}^d &= (\omega_{i,i}^d \tilde{R}_{i,i} + \omega_{j,j}^d \tilde{R}_{j,j} - 1) r_{i,j}^d - U_{i,j}^d + o_p\{(n_d \log p_1)^{-1/2}\} \\
&= \left(\omega_{i,i}^d \left(r_{i,i}^d + U_{i,i}^d\right) + \omega_{j,j}^d \left(r_{j,j}^d + U_{j,j}^d\right) - 1\right) r_{i,j}^d - U_{i,j}^d + o_p\{(n_d \log p_1)^{-1/2}\} \\
&= \left(\omega_{i,i}^d U_{i,i}^d + \omega_{j,j}^d U_{j,j}^d + 1\right) r_{i,j}^d - U_{i,j}^d + o_p\{(n_d \log p_1)^{-1/2}\}.
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
\hat{\omega}_{i,j}^{d,w} - \tilde{U}_{i,j}^{d} &= \frac{\hat{r}_{i,j}^{d}}{\sqrt{\hat{r}_{i,i}^{d}\hat{r}_{j,j}^{d}}} - \frac{r_{i,j}^{d} - U_{i,j}^{d}}{\sqrt{r_{i,i}^{d}r_{j,j}^{d}}} \\
&= \frac{r_{i,j}^{d} - U_{i,j}^{d} + (\omega_{i,i}^{d}U_{i,i} + \omega_{j,j}^{d}U_{j,j})r_{i,j}^{d}}{\sqrt{\hat{r}_{i,i}^{d}\hat{r}_{j,j}^{d}}} - \frac{r_{i,j}^{d} - U_{i,j}^{d}}{\sqrt{r_{i,i}^{d}r_{j,j}^{d}}} + o_p\{(n_d \log p_1)^{-1/2}\} \\
&= \frac{r_{i,j}^{d}}{\sqrt{\hat{r}_{i,i}^{d}\hat{r}_{j,j}^{d}}} - \frac{r_{i,j}^{d}}{\sqrt{r_{i,i}^{d}r_{j,j}^{d}}} + \frac{(\omega_{i,i}^{d}U_{i,i} + \omega_{j,j}^{d}U_{j,j})r_{i,j}^{d}}{\sqrt{r_{i,i}^{d}r_{j,j}^{d}}} \\
&\quad + \left(-U_{i,j}^{d} + (\omega_{i,i}^{d}U_{i,i} + \omega_{j,j}^{d}U_{j,j})r_{i,j}^{d}\right)\left(\frac{1}{\sqrt{\hat{r}_{i,i}^{d}\hat{r}_{j,j}^{d}}} - \frac{1}{\sqrt{r_{i,i}^{d}r_{j,j}^{d}}}\right) \\
&= \frac{(\omega_{i,i}^{d}U_{i,i} + \omega_{j,j}^{d}U_{j,j})r_{i,j}^{d}}{2\sqrt{r_{i,i}^{d}r_{j,j}^{d}}} + o_p\{(n_d \log p_1)^{-1/2}\}.
\end{aligned}$$

□

Lemma 10. *Under Condition 1, for any* $1 \le i, j \le p_1$,

$$n_d \theta_{i,j}^{d} \ge 1. \tag{S2.12}$$

Assume Conditions 1, 3, and 4, then we have

$$\max_{1 \le i \le j \le p_1} \left| \hat{\theta}_{i,j}^{d} - \theta_{i,j}^{d} \right| = O_p\{(\log p_1 / n_d^3)^{1/2}\} \tag{S2.13}$$

uniformly in $1 \le i < j \le p_1$ *and* $d = 1, 2$.

Proof. For any $1 \le i, j \le p_1$,

$$n_d \theta_{i,j}^{d} = \frac{r_{i,i}^{d} r_{j,j}^{d} + (r_{i,j}^{d})^2}{r_{i,i}^{d} r_{j,j}^{d}} \ge 1 \tag{S2.14}$$

And by Lemmas 6 and 7, $\hat{r}_{i,i}^{d} = r_{i,i}^{d} + O_p\{(\log p_1/n_d)^{1/2}\}$ uniformly in $1 \le i \le p_1$ and by

Condition 4,

$$
\begin{aligned}
\left|\hat{\theta}^d_{i,j}-\theta^d_{i,j}\right| &= \left|\frac{1+(\hat{\beta}^d_{i,j})^2\hat{r}^d_{i,i}/\hat{r}^d_{j,j}}{n_d}-\frac{1+(\beta^d_{i,j})^2 r^d_{i,i}/r^d_{j,j}}{n_d}\right| \\
&= \frac{1}{n_d}\left|(\hat{\beta}^d_{i,j})^2\frac{\hat{r}^d_{i,i}}{\hat{r}^d_{j,j}}-(\beta^d_{i,j})^2\frac{r^d_{i,i}}{r^d_{j,j}}\right| = O_p\{(\log p_1/n_d^3)^{1/2}\}
\end{aligned} \tag{S2.15}
$$

uniformly in $1 \le i < j \le p_1$. □

Lemma 11. *In the setting of Theorem 1, let* $A = \{(i,j) : 1 \le i \le j \le p\}$ *and*

$$
\begin{aligned}
V_{i,j} &:= \frac{(\tilde{U}^1_{i,j}-\tilde{U}^2_{i,j})-(\omega^1_{i,j}/\sqrt{\omega^1_{i,i}\omega^1_{j,j}}-\omega^2_{i,j}/\sqrt{\omega^2_{i,i}\omega^2_{j,j}})}{(\theta^1_{i,j}+\theta^2_{i,j})^{1/2}} \\
\tilde{\Delta}_{i,j} &:= \frac{(\hat{\omega}^{1,w}_{i,j}-\hat{\omega}^{2,w}_{i,j})-(\omega^1_{i,j}/\sqrt{\omega^1_{i,i}\omega^1_{j,j}}-\omega^2_{i,j}/\sqrt{\omega^2_{i,i}\omega^2_{j,j}})}{(\hat{\theta}^1_{i,j}+\hat{\theta}^2_{i,j})^{1/2}}
\end{aligned}
$$

then under Conditions 1, 2, 3, and 4, we obtain

$$
\begin{aligned}
\max_{(i,j)\in A\setminus A_\tau}|\tilde{\Delta}_{i,j}-V_{i,j}| =& o_p\{(\log p_1)^{-1/2}\}, \\
P\left(\max_{(i,j)\in A_\tau}|\tilde{\Delta}_{i,j}|^2 \ge 3.9\log p_1\right) =& o(1),
\end{aligned}
$$

for any $t > 0$.

Proof. First, we will show the first equality. Without loss of generality, we may assume $n_1 \ge n_2$ in the proof. Since $n_d\theta^d_{i,j} \ge 1$ and $n_d \max_{(i,j)\in A}|\hat{\theta}^d_{i,j}-\theta^d_{i,j}| = O_p\{(\log p_1/n_d)^{1/2}\}$ from and Lemma 10,

$$
\begin{aligned}
&\max_{(i,j)\in A}\left|\frac{1}{(\hat{\theta}^1_{i,j}+\hat{\theta}^2_{i,j})^{1/2}}-\frac{1}{(\theta^1_{i,j}+\theta^2_{i,j})^{1/2}}\right| \\
&= \sqrt{n_2}\max_{(i,j)\in A}\left|\frac{1}{\left(\frac{n_2}{n_1}\left(n_1\hat{\theta}^1_{i,j}\right)+n_2\hat{\theta}^2_{i,j}\right)^{1/2}}-\frac{1}{\left(\frac{n_2}{n_1}\left(n_1\theta^1_{i,j}\right)+n_2\theta^2_{i,j}\right)^{1/2}}\right|.
\end{aligned}
$$

Since $\frac{n_2}{n_1}\left(n_1\theta^1_{i,j}\right)+n_2\theta^2_{i,j}\geq 1$, we obtain

$$\max_{(i,j)\in A}\left|\frac{1}{(\hat{\theta}^1_{i,j}+\hat{\theta}^2_{i,j})^{1/2}}-\frac{1}{(\theta^1_{i,j}+\theta^2_{i,j})^{1/2}}\right|=\sqrt{n_2}\left(O_p\{(\log p_1)^{1/2}n_2/n_1^2\}+O_p\{(\log p_1)^{1/2}/n_2\}\right)$$
$$=O_p\{(\log p_1/n_2)^{1/2}\}.$$

Thus, since $\hat{\omega}^{d,w}_{i,j}-\frac{\omega^1_{i,j}}{\sqrt{\omega^1_{i,i}\omega^1_{j,j}}}=O_p\{(\log p_1/n_d)^{1/2}\}$, we have

$$\tilde{\Delta}_{i,j}-V_{i,j}=\frac{\left(\hat{\omega}^{1,w}_{i,j}-\hat{\omega}^{2,w}_{i,j}\right)-\left(\frac{\omega^1_{i,j}}{\sqrt{\omega^1_{i,i}\omega^1_{j,j}}}-\frac{\omega^2_{i,j}}{\sqrt{\omega^2_{i,i}\omega^2_{j,j}}}\right)}{(\hat{\theta}^1_{i,j}+\hat{\theta}^2_{i,j})^{1/2}}-\frac{(\tilde{U}^1_{i,j}-\tilde{U}^2_{i,j})-\left(\frac{\omega^1_{i,j}}{\sqrt{\omega^1_{i,i}\omega^1_{j,j}}}-\frac{\omega^2_{i,j}}{\sqrt{\omega^2_{i,i}\omega^2_{j,j}}}\right)}{(\theta^1_{i,j}+\theta^2_{i,j})^{1/2}}$$
$$=\frac{\left(\hat{\omega}^{1,w}_{i,j}-\hat{\omega}^{2,w}_{i,j}\right)-(\tilde{U}^1_{i,j}-\tilde{U}^2_{i,j})}{(\theta^1_{i,j}+\theta^2_{i,j})^{1/2}}+O_p\{\log p_1/n_2\}.$$

From Lemma 9, we obtain

$$\left|\tilde{\Delta}_{i,j}-V_{i,j}\right|=\left|\frac{\left(\hat{\omega}^{1,w}_{i,j}-\tilde{U}^1_{i,j}\right)-\left(\hat{\omega}^{2,w}_{i,j}-\tilde{U}^2_{i,j}\right)}{(\theta^1_{i,j}+\theta^2_{i,j})^{1/2}}\right|+O_p\{\log p_1/n_2\}$$
$$\leq\frac{1}{\sqrt{2}}\sum_{d=1,2}\left(\left|\frac{\frac{1}{n_d}\sum_{k=1}^{n_d}\frac{\varepsilon^d_{k,i}}{\sqrt{r^d_{i,i}}}\frac{\varepsilon^d_{k,i}}{\sqrt{r^d_{i,i}}}}{\sqrt{\hat{\theta}^d_{i,i}}}\right|+\left|\frac{\frac{1}{n_d}\sum_{k=1}^{n_d}\frac{\varepsilon^d_{k,j}}{\sqrt{r^d_{j,j}}}\frac{\varepsilon^d_{k,j}}{\sqrt{r^d_{j,j}}}}{\sqrt{\hat{\theta}^d_{j,j}}}\right|\right)+O_p\{\log p_1/n_2\}\tag{S2.16}$$

and

$$\max_{(i,j)\in A_\tau}|\tilde{\Delta}_{i,j}-V_{i,j}|=O_p\{(\log p_1)^{1/2}\}\max_{(i,j)\in A_\tau}|r^d_{i,j}|+O_p\{\log p_1/\sqrt{n_2}\}.\tag{S2.17}$$

We have

$$|r^d_{i,j}|\leq\left|\frac{\omega^d_{i,j}}{\sqrt{\omega^d_{i,i}\omega^d_{j,j}}}\right|\sqrt{r^d_{i,i}r^d_{j,j}}=o((\log p_1)^{-1})\tag{S2.18}$$

for $(i,j)\in A\setminus A_\tau$. Therefore, combining (S2.17) and (S2.18), we obtain the first equality.

Next, we will show the second equality. By (S2.16) and Lemma 18, we have for sufficiently any $\varepsilon_0 > 0$,

$$
\begin{aligned}
&P\left(\max_{(i,j)\in A_\tau}\left|\tilde{\Delta}_{i,j}-V_{i,j}\right| \geq (\sqrt{2}+\varepsilon_0)(\log p_1)^{1/2}\right) \\
&\leq 2\sum_{d=1,2} P\left(\sum_{(i,j)\in A_\tau}\left|\frac{\frac{1}{n_d}\sum_{k=1}^{n_d}\frac{\varepsilon_{k,i}^d}{\sqrt{r_{i,i}^d}}\frac{\varepsilon_{k,j}^d}{\sqrt{r_{j,j}^d}}}{\sqrt{\hat{\theta}_{i,j}^d}}\right| \geq \frac{(\log p_1)^{1/2}}{2}\right) + o(1) \\
&= C|A_\tau|\left(1-\Phi\left(\frac{(\log p_1)^{1/2}}{2\sqrt{2}}\right)\right) + o(1) = C|A_\tau|\frac{p_1^{-1/16}}{(\log p_1)^{1/4}} + o(1).
\end{aligned}
$$

Also, by Lemma 18, we obtain

$$
\begin{aligned}
P\left(\max_{(i,j)\in A_\tau}|V_{i,j}| \geq \frac{(\log p_1)^{1/2}}{2}\right) &\leq C|A_\tau|\left(1-\Phi\left(\frac{(\log p_1)^{1/2}}{2\sqrt{2}}\right)\right) + o(1) \\
&= C|A_\tau|\frac{p_1^{-1/16}}{(\log p_1)^{1/4}} + o(1).
\end{aligned}
$$

Therefore, for sufficiently large n and p_1 and for sufficiently small $\varepsilon_0 > 0$, we obtain

$$
\begin{aligned}
&P\left(\max_{(i,j)\in A_\tau}|\tilde{\Delta}_{i,j}|^2 \geq 3.9\log p_1\right) \\
&= P\left(\max_{(i,j)\in A_\tau}|V_{i,j}| \geq 0.5(\log p_1)^{1/2}\right) + P\left(\max_{(i,j)\in A_\tau}\left|\tilde{\Delta}_{i,j}-V_{i,j}\right| \geq (\sqrt{2}+\varepsilon_0)(\log p_1)^{1/2}\right) \\
&= C|A_\tau|\frac{p_1^{-1/16}}{(\log p_1)^{1/4}} + o(1).
\end{aligned}
$$

Since $|A_\tau| = O(p_1^{1/16})$, we can conclude

$$
P\left(\max_{(i,j)\in A_\tau}|\tilde{\Delta}_{i,j}|^2 \geq 3.9\log p_1\right) = o(1).
$$

□

Lemma 12. *In the setting of Theorem 1, let* $A\backslash A_\tau = \{(i_m, j_m) : m = 1, \cdots, q\}$ *with* $q = |A\backslash A_\tau|$

and

$$\eta_{m,d} = \frac{\mathrm{Var}(\varepsilon^d_{k,i_m}\varepsilon^d_{k,j_m})}{r^d_{i_m,i_m} r^d_{j_m,j_m}}$$

$$Z_{k,m} = \begin{cases} \frac{\varepsilon^1_{k,i_m}\varepsilon^1_{k,j_m} - E[\varepsilon^1_{k,i_m}\varepsilon^1_{k,j_m}]}{\sqrt{r^1_{i_m,i_m} r^1_{j_m,j_m}}} & \textit{for } 1 \le k \le n_1 \\ -\frac{\varepsilon^2_{k-n_1,i_m}\varepsilon^2_{k-n_1,j_m} - E[\varepsilon^2_{k-n_1,i_m}\varepsilon^2_{k-n_1,j_m}]}{\sqrt{r^2_{i_m,i_m} r^2_{j_m,j_m}}} & \textit{for } n_1+1 \le k \le n_1+n_2 \end{cases}$$

$$\hat{Z}_{k,m} = Z_{m,k} I(|Z_{k,m}| \le \tau_n) - E[Z_{k,m} I(|Z_{k,m}| \le \tau_n)], \quad \textit{where } \tau_n = 4\log(p_1 + n_1 + n_2)$$

then using the above $\eta_{m,d}$, $Z_{k,m}$, and $\hat{Z}_{k,m}$, we can represent V_{i_m,j_m} again and find approximation of V_{i_m,j_m}. That is

$$V_m := \frac{\frac{1}{n_1}\sum_{k=1}^{n_1} Z_{k,m} + \frac{1}{n_2}\sum_{k=n_1+1}^{n_1+n_2} Z_{k,m}}{(\eta_{m,1}/n_1 + \eta_{m,2}/n_2)^{1/2}} = V_{i_m,j_m}$$

$$\textit{and} \quad \hat{V}_m := \frac{\frac{1}{n_1}\sum_{k=1}^{n_1} \hat{Z}_{k,m} + \frac{1}{n_2}\sum_{k=n_1+1}^{n_1+n_2} \hat{Z}_{k,m}}{(\eta_{m,1}/n_1 + \eta_{m,2}/n_2)^{1/2}}$$

then

$$\max_{1\le m\le q} |V_m - \hat{V}_m| = O_p\{p_1^{-3}\}.$$

Proof. By the definition of $\eta_{m,d}$, we know that $\eta_{m,d} = 1 + \frac{(r^d_{i_m,j_m})^2}{r^d_{i_m,i_m} r^d_{j_m,j_m}} \ge 1$. Thus, we obtain

$$\begin{aligned} |V_m - \hat{V}_m| &\le \frac{1}{n_1}\sum_{k=1}^{n_1} \frac{|Z_{k,m} - \hat{Z}_{k,m}|}{(\eta_{m,2}/n_2 + \eta_{m,1}/n_1)^{1/2}} + \frac{1}{n_2}\sum_{k=n_1+1}^{n_1+n_2} \frac{|Z_{k,m} - \hat{Z}_{k,m}|}{(\eta_{m,2}/n_2 + \eta_{m,1}/n_1)^{1/2}} \\ &\le \frac{1}{n_2}\sum_{k=1}^{n_1+n_2} \frac{|Z_{k,m}| I(|Z_{k,m}| > \tau_n) + |E[Z_{k,m} I(|Z_{k,m}| \le \tau_n)]|}{(1/n_2 + 1/n_1)^{1/2}}. \end{aligned}$$

From $E[Z_{k,m}] = 0$, we know $E[Z_{k,m} I(|Z_{k,m}| \le \tau_n)] = -E[Z_{k,m} I(|Z_{k,m}| > \tau_n)]$, so $\max_{1\le m\le q} \max_{1\le k\le n_1+n_2} |Z_{k,m}| \le \tau_n$

implies

$$\max_{1\le m\le q}|V_m-\hat V_m| \le \sum_{k=1}^{n_1+n_2}\frac{1}{\sqrt{n_2}}|E[Z_{k,m}I(|Z_{k,m}|>\tau_n)]|$$
$$\le \frac{2n_1}{\sqrt{n_2}}\max_{1\le k\le n_1+n_2}E[|Z_{k,m}|\exp(|Z_{k,m}|-4\log(p_1+n_1+n_2))]$$
$$=\frac{2n_1}{\sqrt{n_2}(p_1+n_1+n_2)^4}E[|Z_{k,m}|\exp(|Z_{k,m}|)]\le Cp_1^{-3}$$

because the nomality of ε^d_{k,i_m} implies the uniformly boundedness of $E\left[|Z_{k,m}|\exp\left(|Z_{k,m}|\right)\right]$.

Thus,

$$P\left(\max_{1\le m\le q}|V_m-\hat V_m|\ge Cp_1^{-3}\right)$$
$$\le P\left(\max_{1\le m\le q}\max_{1\le k\le n_1+n_2}|Z_{k,m}|>\tau_n\right)$$
$$\le (n_1+n_2)q\max_{1\le m\le q}\max_{1\le k\le n_1+n_2}P\left(|Z_{k,m}|>4\log(p_1+n_1+n_2)\right)$$
$$\le (n_1+n_2)q\max_{1\le m\le q}\max_{1\le k\le n_1+n_2}E\left[\exp\left(|Z_{k,m}|-4\log(p_1+n_1+n_2)\right)\right]$$
$$=\frac{(n_1+n_2)q}{(p_1+n_1+n_2)^4}\max_{1\le m\le q}\max_{1\le k\le n_1+n_2}E\left[\exp\left(|Z_{k,m}|\right)\right]=O(p_1^{-1})$$

because the nomality of ε^d_{k,i_m} implies the uniformly boundedness of $E\left[\exp\left(|Z_{k,m}|\right)\right]$. □

Lemma 13. *In the setting of Theorem 1, let $y_{p_1}=4\log p_1-\log\log p_1+t$ then for any positive integer N, the following inequalities hold.*

$$\sum_{d=1}^{2N}\frac{(-1)^{d-1}}{d!}\left((8\pi)^{-1/2}e^{-t/2}\right)^d+o(1)\le P(\max_{1\le m\le q}|\hat V_m|^2\ge y_{p_1})$$
$$\le\sum_{d=1}^{2N-1}\frac{(-1)^{d-1}}{d!}\left((8\pi)^{-1/2}e^{-t/2}\right)^d+o(1),$$

where $\hat V_m$ and q are defined in Lemma 12.

Proof. Without loss of generality, we may assume $n_1\ge n_2$.

By Bonferroni inequality,

$$
\begin{aligned}
&\sum_{d=1}^{2N}(-1)^{d-1}\sum_{1\le m_1<\cdots<m_d\le q}P\left(\bigcap_{j=1}^{d}\{|\hat{V}_{m_j}|^2\ge y_{p_1}\}\right)\\
&\quad\le P(\max_{1\le m\le q}|\hat{V}_m|^2\ge y_{p_1})\\
&\quad\le\sum_{d=1}^{2N-1}(-1)^{d-1}\sum_{1\le m_1<\cdots<m_d\le q}P\left(\bigcap_{j=1}^{d}\{|\hat{V}_{m_j}|^2\ge y_{p_1}\}\right).
\end{aligned}
$$

For fixed $d>0$, let $\boldsymbol{\xi}_k=(\tilde{Z}_{k,m_1},\cdots,\tilde{Z}_{k,m_d})^T$ and

$$
\tilde{Z}_{k,m}=\begin{cases}\frac{\hat{Z}_{k,m}}{(\eta_{m,1}+n_1\eta_{m,2}/n_2)^{1/2}} & \text{for } 1\le k\le n_1\\[2ex] \frac{\hat{Z}_{k,m}}{(n_2\eta_{m,1}/n_1+\eta_{m,2})^{1/2}} & \text{for } n_1+1\le k\le n_1+n_2,\end{cases}
$$

then

$$
\left\{\left|\sum_{k=1}^{n_1+n_2}\boldsymbol{\xi}_k\right|_{\min}\ge(n_1y_{p_1})^{1/2}\right\}=\bigcap_{j=1}^{d}\{|\hat{V}_{m_j}|\ge y_{p_1}^{1/2}\}
$$

and $\boldsymbol{\xi}_k$ satisfies the conditions of (Zaitsev, 1987, Theorem 1.1) with for any $u,v\in\mathbb{R}^d$ and $m=3,4,\cdots,$

$$
\begin{aligned}
E[(\boldsymbol{\xi}_k\cdot v)^2(\boldsymbol{\xi}_k\cdot u)^{m-2}]&\le\left(C\sqrt{d}\tau_n\right)^{m-2}\|u\|^{m-2}E[(\boldsymbol{\xi}_k\cdot v)^2]\\
&\le\frac{m!}{2}\left(C\sqrt{d}\tau_n\right)^{m-2}\|u\|^{m-2}E[(\boldsymbol{\xi}_k\cdot v)^2]
\end{aligned}
$$

Thus, there exist $c_1,c_2,c_3>0$ such that for any $\lambda>0$,

$$
\begin{aligned}
&P\left(|\tilde{\mathbf{N}}_d|_{\min}\ge(n_1y_{p_1})^{1/2}+\lambda\right)-c_1d^{5/2}\exp\left(-\frac{\lambda}{c_2Cd^3\tau_n}\right)\\
\le&P\left(\left|\sum_{l=1}^{n_1+n_2}\boldsymbol{\xi}_l\right|_{\min}\ge(n_1y_{p_1})^{1/2}\right)\\
\le&P\left(|\tilde{\mathbf{N}}_d|_{\min}\ge(n_1y_{p_1})^{1/2}-\lambda\right)+c_1d^{5/2}\exp\left(-\frac{\lambda}{c_2Cd^3\tau_n}\right),
\end{aligned}
$$

where $\tilde{\mathbf{N}}_d \sim N(0, n_1\text{Cov}(\boldsymbol{\xi}_1) + n_2\text{Cov}(\boldsymbol{\xi}_{n_1+1}))$. Let $\mathbf{N}_d = n_1^{-1/2}\tilde{\mathbf{N}}_d$ and $\lambda = \sqrt{n_1}\varepsilon_n(\log p_1)^{-1/2}$ for ε_n determined later, then $\mathbf{N}_d \sim N(0, \text{Cov}(\boldsymbol{\xi}_1) + n_2\text{Cov}(\boldsymbol{\xi}_{n_1+1})/n_1)$ and

$$
\begin{aligned}
&P\left(|\mathbf{N}_d|_{\min} \geq y_{p_1}^{1/2} + \varepsilon_n(\log p_1)^{-1/2}\right) - c_1 d^{5/2}\exp\left(-\frac{\sqrt{n_1}\varepsilon_n}{c_2 C d^3 \tau_n (\log p_1)^{1/2}}\right) \\
\leq &P\left(\left|\sum_{l=1}^{n_1+n_2}\boldsymbol{\xi}_l\right|_{\min} \geq (n_1 y_{p_1})^{1/2}\right) \\
\leq &P\left(|\mathbf{N}_d|_{\min} \geq y_{p_1}^{1/2} - \varepsilon_n(\log p_1)^{-1/2}\right) + c_1 d^{5/2}\exp\left(-\frac{\sqrt{n_1}\varepsilon_n}{c_2 C d^3 \tau_n (\log p_1)^{1/2}}\right)
\end{aligned}
$$

Set $\varepsilon_n = (\log p_1)^{1/2} n_1^{-1/10}$, then $\varepsilon_n \to 0$ and

$$
\begin{aligned}
&c_1 d^{5/2}\exp\left(-\frac{\sqrt{n_1}\varepsilon_n}{c_2 C d^3 \tau_n (\log p_1)^{1/2}}\right) \\
&\quad = c_1 d^{5/2}\exp\left(-\frac{n_1^{\frac{2}{5}}}{16 c_2 C d^3 \log(p_1 + n_1 + n_2)}\right) \to 0, \quad \text{as } n, p_1 \to \infty.
\end{aligned}
$$

By (Cai et al., 2013, Lemma 5), we obtain

$$
\sum_{1 \leq m_1 < \cdots < m_d \leq q} P(|\mathbf{N}|_{\min} \geq y_{p_1}^{1/2} \pm \varepsilon_n(\log p)^{-1/2}) = \frac{1}{d!}\left((8\pi)^{-1/2}e^{-t/2}\right)^d (1 + o(1)),
$$

so we can conclude the proof by combining with Bonferroni inequaltiy. □

Lemma 14. *Under the same conditions as in Theorem 4, we have*

$$
\left|\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau}\left[I(|\Delta_{i,j}| \geq \hat{t}_0) - 2(1 - \Phi(\hat{t}_0))\right]}{2(1 - \Phi(\hat{t}_0))q_0}\right| = o_p\{1\}.
$$

Proof. Let

$$
\begin{aligned}
B_1 &:= \left|\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau}\left[I(|\Delta_{i,j}| \geq \hat{t}_0) - 2(1 - \Phi(\hat{t}_0))\right]}{2(1 - \Phi(\hat{t}_0))q_0}\right| \\
&\leq \left|\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau} I(|\Delta_{i,j}| \geq \hat{t}_0))}{2(1 - \Phi(\hat{t}_0))q_0}\right| + \frac{|\mathcal{I}_0 \cap A_\tau|}{q_0}.
\end{aligned}
$$

Then, we have for any $\varepsilon > 0$

$$P\left(\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau} I(|\Delta_{i,j}| \geq \hat{t}_0)}{2(1-\Phi(\hat{t}_0))q_0} > \varepsilon\right)$$
$$= P\left(\left\{\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau} I(|\Delta_{i,j}| \geq \hat{t}_0)}{2(1-\Phi(\hat{t}_0))q_0} > \varepsilon\right\} \cap \{\hat{t}_0 \geq (3.9\log p_1)^{1/2}\}\right) \tag{S2.19}$$
$$+ P\left(\left\{\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau} I(|\Delta_{i,j}| \geq \hat{t}_0)}{2(1-\Phi(\hat{t}_0))q_0} > \varepsilon\right\} \cap \{\hat{t}_0 < (3.9\log p_1)^{1/2}\}\right).$$

By Lemma 11, we have

$$P\left(\left\{\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau} I(|\Delta_{i,j}| \geq \hat{t}_0)}{2(1-\Phi(\hat{t}_0))q_0} > \varepsilon\right\} \cap \{\hat{t}_0 \geq (3.9\log p_1)^{1/2}\}\right)$$
$$\leq P\left(\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau} I(|\Delta_{i,j}| \geq (3.9\log p_1)^{1/2})}{2(1-\Phi(\hat{t}_0))q_0} > 0\right) \tag{S2.20}$$
$$= P\left(\max_{(i,j)\in\mathcal{I}_0\cap A_\tau} |\Delta_{i,j}| > (3.9\log p_1)^{1/2}\right) = o(1).$$

Also, by the assumption of Theorem 4 ($|\mathcal{I}_0 \cap A_\tau| = o(p_1^\nu)$ for any $\nu > 0$), we have

$$P\left(\left\{\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau} I(|\Delta_{i,j}| \geq \hat{t}_0)}{2(1-\Phi(\hat{t}_0))q_0} > \varepsilon\right\} \cap \{\hat{t}_0 < (3.9\log p_1)^{1/2}\}\right)$$
$$\leq P\left(\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau} I(|\Delta_{i,j}| \geq \hat{t}_0) > 2\varepsilon(1-\Phi((3.9\log p_1)^{1/2}))q_0\right) \tag{S2.21}$$
$$\leq P\left(|\mathcal{I}_0\cap A_\tau| > C\frac{p_1^{2-1.95}}{\sqrt{\log p_1}}\right) = o(1).$$

Thus, by combining (S2.19), (S2.20), and (S2.21), we obtain

$$\frac{\sum_{(i,j)\in\mathcal{I}_0\cap A_\tau} I(|\Delta_{i,j}| \geq \hat{t}_0)}{2(1-\Phi(\hat{t}_0))q_0} = o_p\{1\},$$

which implies $B_1 = o_p\{1\}$. $\square$

Lemma 15. *Under the same conditions as in Theorem 4, we assume*

$$0 < \hat{t}_0 < (4\log p_1 - \log\log p_1 + \log\log\log p_1)^{1/2} \quad a.s..$$

Then, we have

$$\left| \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\left[I(|\Delta_{i,j}| \geq \hat{t}_0) - I(|V_{i,j}| \geq \hat{t}_0)\right]}{2(1-\Phi(\hat{t}_0))q_0} \right| = o_p\{1\}.$$

Proof. Let

$$\begin{aligned} B_2 &:= \left| \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\left[I(|\Delta_{i,j}| \geq \hat{t}_0) - I(|V_{i,j}| \geq \hat{t}_0)\right]}{2(1-\Phi(\hat{t}_0))q_0} \right| \\ &= \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\left[I(|\Delta_{i,j}| \geq \hat{t}_0 > |V_{i,j}|) + I(|V_{i,j}| \geq \hat{t}_0 > |\Delta_{i,j}|)\right]}{2(1-\Phi(\hat{t}_0))q_0}. \end{aligned}$$

For $(i,j) \in \mathcal{I}_0 \setminus A_\tau$ and for any $\varepsilon > 0$, we have

$$\begin{aligned} & I(|\Delta_{i,j}| \geq \hat{t}_0 > |V_{i,j}|) \\ &= I(|\Delta_{i,j}| \geq \hat{t}_0 > |V_{i,j}|)I(|\Delta_{i,j} - V_{i,j}| < \varepsilon(\log p_1)^{-1/2}) \\ &\quad + I(|\Delta_{i,j}| \geq \hat{t}_0 > |V_{i,j}|)I(|\Delta_{i,j} - V_{i,j}| \geq \varepsilon(\log p_1)^{-1/2}) \\ &\leq I(|V_{i,j}| + \varepsilon(\log p_1)^{-1/2} \geq \hat{t}_0 > |V_{i,j}|) + I(|\Delta_{i,j} - V_{i,j}| \geq \varepsilon(\log p_1)^{-1/2}) \end{aligned} \tag{S2.22}$$

and

$$\begin{aligned} & I(|V_{i,j}| \geq \hat{t}_0 > |\Delta_{i,j}|) \\ &\leq I(|V_{i,j}| \geq \hat{t}_0 > |V_{i,j}| - \varepsilon(\log p_1)^{-1/2}) + I(|\Delta_{i,j} - V_{i,j}| \geq \varepsilon(\log p_1)^{-1/2}). \end{aligned} \tag{S2.23}$$

By Lemma 11, we have

$$\begin{aligned} &P\left(\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} I(|\Delta_{i,j}-V_{i,j}|\geq\varepsilon(\log p_1)^{-1/2})}{2(1-\Phi(\hat{t}_0))q_0}>0\right) \\ &\quad = P\left(\max_{(i,j)\in\mathcal{I}_0\setminus A_\tau}|\Delta_{i,j}-V_{i,j}|\geq\varepsilon(\log p_1)^{-1/2}\right)=o(1). \end{aligned} \tag{S2.24}$$

Combining (S2.22), (S2.23), and (S2.24), we have

$$\begin{aligned} B_2 &\leq \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} I(t_0+\varepsilon(\log p_1)^{-1/2}>|V_{i,j}|>\hat{t}_0-\varepsilon(\log p_1)^{-1/2})}{2(1-\Phi(\hat{t}_0))q_0}+o_p\{1\} \\ &=: B_{21}+o_p\{1\}. \end{aligned}$$

We will decompose B_{21} into M parts as follows:

$$B_{21}=\sum_{m=1}^{M}\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\frac{I(t_0+\varepsilon(\log p_1)^{-1/2}>|V_{i,j}|>\hat{t}_0-\varepsilon(\log p_1)^{-1/2})}{2(1-\Phi(\hat{t}_0))q_0}I(J_m), \tag{S2.25}$$

where

$$J_m=\{v_{m-1}\leq\hat{t}_0<v_m\},$$

$M=\lceil v_M/v_1\rceil$, $v_1=(\log p_1)^{-1/2}(\log\log\log\log p_1)^{-1}$, and

$$v_m=\begin{cases} mv_1 & \text{if } 0\leq m<M \\ \sqrt{4\log p_1-\log\log p_1+\log\log\log p_1} & \text{if } m=M, \end{cases}$$

and $\lceil x\rceil$ is the smallest integer greater than or equal to x. Using (S2.25), we obtain

$$\begin{aligned} E[B_{21}] &\leq \sum_{m=1}^{M}\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\frac{P(v_m+\varepsilon(\log p_1)^{-1/2}>|V_{i,j}|>v_{m-1}-\varepsilon(\log p_1)^{-1/2})}{2(1-\Phi(v_m))q_0}P(J_m) \\ &\leq C\sum_{m=1}^{M}\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\frac{P(v_m+\varepsilon(\log p_1)^{-1/2}>|V_{i,j}|>v_{m-1}-\varepsilon(\log p_1)^{-1/2})}{\phi(v_m)q_0/v_m}P(J_m). \end{aligned}$$

By (PETROV, 1975, Theorem 5 in Chapter VIII), we obtain

$$
\begin{aligned}
&P(v_m + \varepsilon(\log p_1)^{-1/2} > |V_{i,j}| > v_{m-1} - \varepsilon(\log p_1)^{-1/2}) \\
&\quad \le 2\Phi(v_m + \varepsilon(\log p_1)^{-1/2}) - 2\Phi(v_{m-1} - \varepsilon(\log p_1)^{-1/2}) \\
&\qquad + O\left(\frac{1}{\sqrt{n}} e^{-\frac{\left[v_{m-1}-\varepsilon(\log p_1)^{-1/2}\right]_+^2}{2}}\right) \\
&\quad \le \frac{4\varepsilon\phi\left(\left[v_{m-1} - \varepsilon(\log p_1)^{-1/2}\right]_+\right)}{(\log p_1)^{1/2}} + O\left(\frac{1}{\sqrt{n}} e^{-\frac{\left[v_{m-1}-\varepsilon(\log p_1)^{-1/2}\right]_+^2}{2}}\right)
\end{aligned}
$$

uniformly in $(i,j) \in \mathcal{I}_0 \setminus A_\tau$ and $m = 1, \cdots, M$. Then, we have

$$
\begin{aligned}
E[B_{21}] \le C\varepsilon \sum_{m=1}^{M} &\frac{\phi\left(\left[v_{m-1} - \varepsilon(\log p_1)^{-1/2}\right]_+\right)}{(\log p_1)^{1/2}\phi(v_m)} v_m P(J_m) \\
&+ C\sum_{m=1}^{M} \frac{v_m}{\phi(v_m)} O\left(\frac{1}{\sqrt{n}} e^{-\frac{\left[v_{m-1}-\varepsilon(\log p_1)^{-1/2}\right]_+^2}{2}}\right) P(J_m).
\end{aligned}
$$

We note that

$$
\begin{aligned}
&\frac{\phi\left(\left[v_{m-1} - \varepsilon(\log p_1)^{-1/2}\right]_+\right)}{\phi(v_m)} \\
&\quad \le e^{-\frac{1}{2}\left((v_m - v_1) - \varepsilon(\log p_1)^{-1/2}\right)_+^2 + \frac{1}{2}v_m^2} \\
&\quad = e^{v_1 v_m + \varepsilon v_m(\log p_1)^{-1/2} - \frac{v_1^2}{2} - \varepsilon v_1(\log p_1)^{-1/2} - \frac{\varepsilon^2}{2}(\log p_1)^{-1}} I(v_{m-1} > \varepsilon(\log p_1)^{-1/2}) \\
&\qquad + e^{v_m^2/2} I(v_{m-1} < \varepsilon(\log p_1)^{-1/2}) \\
&\quad \le e^{\left(v_1(\log p_1)^{1/2} + 2\varepsilon\right)v_M(\log p_1)^{-1/2}} + e^{\frac{(\varepsilon(\log p_1)^{-1/2} + v_1)^2}{2}}
\end{aligned}
$$

Since $v_m \le v_M \le 2(\log p_1)^{1/2}$ and $v_1(\log p_1)^{1/2} = o(1)$, we have

$$
\max_{1\le m\le M} \frac{\phi\left(\left[v_{m-1} - \varepsilon(\log p_1)^{-1/2}\right]_+\right)}{(\log p_1)^{1/2}\phi(v_m)} v_m < C.
$$

Similarily, we have

$$\frac{v_m e^{-\frac{\left[v_{m-1}-\varepsilon(\log p_1)^{-1/2}\right]_+^2}{2}}}{\sqrt{n}\phi(v_m)} \leq \frac{C(\log p_1)^{1/2}}{\sqrt{n}}.$$

Thus, we obtain

$$E[B_{21}] \leq \sum_{m=1}^{M} \left(C\varepsilon + O\left(\frac{(\log p_1)^{1/2}}{\sqrt{n}} \right) \right) P(J_M) = C\varepsilon + O\left(\frac{(\log p_1)^{1/2}}{\sqrt{n}} \right).$$

Therefore, we obtain

$$\lim_{n,p_1\to\infty} E[B_2] \leq C\varepsilon$$

for any $\varepsilon > 0$, which implies

$$\lim_{n,p_1\to\infty} E[B_2] = 0.$$

Then,

$$\lim_{n,p_1\to\infty} P(|B_2| > \varepsilon) \leq \lim_{n,p_1\to\infty} \frac{E[|B_2|]}{\varepsilon} = \lim_{n,p_1\to\infty} \frac{E[B_2]}{\varepsilon} = 0.$$

Therefore, $B_2 = o_p\{1\}$. □

Lemma 16. *Under the same conditions as in Theorem 4, we assume*

$$0 < \hat{t}_0 < (4\log p_1 - \log\log p_1 + \log\log\log p_1)^{1/2} \quad a.s..$$

Then, we have

$$\left| \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} \left[P(|V_{i,j}| \geq \hat{t}_0) - 2(1-\Phi(\hat{t}_0)) \right]}{2(1-\Phi(\hat{t}_0))q_0} \right| = o_p\{1\}.$$

Proof. Let

$$B_3 := \left| \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} \left[P(|V_{i,j}| \geq \hat{t}_0) - 2(1-\Phi(\hat{t}_0)) \right]}{2(1-\Phi(\hat{t}_0))q_0} \right|.$$

Then, by (PETROV, 1975, Theorem 5 in Chapter VIII), we obtain

$$B_3 \leq C\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} e^{-\hat{t}_0^2/2}}{2\sqrt{n}(1-\Phi(\hat{t}_0))q_0} \leq C\frac{e^{-\hat{t}_0^2/2}}{2\sqrt{n}(1-\Phi(\hat{t}_0))}.$$

As in Lemma 15, we also decompose B_3 into M parts as follows:

$$B_3 \leq C\sum_{m=1}^{M}\frac{e^{-\hat{t}_0^2/2}}{2\sqrt{n}(1-\Phi(\hat{t}_0))}I(J_m) \leq C\sum_{m=1}^{M}\frac{e^{-v_{m-1}^2/2}}{2\sqrt{n}(1-\Phi(v_m))}I(J_m).$$

Since

$$\frac{e^{-v_{m-1}^2/2}}{1-\Phi(v_m)} \leq Cv_m e^{-\frac{1}{2}\left(v_{m-1}^2-v_m^2\right)} \leq Cv_m e^{-\frac{1}{2}\left((v_m-v_1)^2-v_m^2\right)} \leq C(\log p_1)^{1/2}e^{v_1v_m}$$

and $v_1v_m = o(1)$, we obtain

$$B_3 \leq C\sum_{m=1}^{M}\frac{(\log p_1)^{1/2}}{\sqrt{n}}I(J_m) = C\frac{(\log p_1)^{1/2}}{\sqrt{n}}(1+o_p\{1\}) = o_p\{1\}.$$

□

Lemma 17. *Under the same conditions as in Theorem 4, we assume*

$$0 < P_{\alpha/2} \leq < \hat{t}_0 < (4\log p_1 - \log\log p_1 + \log\log\log p_1)^{1/2} \quad a.s., \tag{S2.26}$$

where $P_{\alpha/2}$ is a value with $\left(1-\Phi(P_{\alpha/2})\right) = \alpha/2$. Then, we have

$$\left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\left[I(|V_{i,j}|\geq\hat{t}_0) - P(|V_{i,j}|\geq\hat{t}_0)\right]}{2(1-\Phi(\hat{t}_0))q_0}\right| = o_p\{1\}.$$

Proof. Let

$$B_4 := \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}\left[I(|V_{i,j}|\geq \hat{t}_0) - P(|V_{i,j}|\geq \hat{t}_0)\right]}{2(1-\Phi(\hat{t}_0))q_0}.$$

Using the M partition $\{J_m : m = 1, \cdots, M\}$, we have

$$\begin{aligned}\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} I(|V_{i,j}|\geq v_m)}{2(1-\Phi(v_{m-1}))q_0} I(J_m) \quad &\leq \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} I(|V_{i,j}|\geq \hat{t}_0)}{2(1-\Phi(\hat{t}_0))q_0} I(J_m) \\ &\leq \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} I(|V_{i,j}|\geq v_{m-1})}{2(1-\Phi(v_{m-2}))q_0} \frac{1-\Phi(v_{m-2})}{1-\Phi(v_m)} I(J_m)\end{aligned}$$

for $m = 2, \cdots, M$, and we know

$$I(J_1), I\left(\bigcap_{m=2}^{M} J_m^c\right) = o_p\{1\}$$

by (S2.26) for sufficiently large p_1. Thus, we have

$$\begin{aligned}&\sum_{m=1}^{M} \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}[I(|V_{i,j}|\geq v_m) - P(|V_{i,j}|\geq v_m)]}{2(1-\Phi(v_{m-1}))q_0} I(J_m) \\ &\quad - \sum_{m=1}^{M} \sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} \frac{P(v_{m-1}\leq |V_{i,j}| < v_m)}{2(1-\Phi(v_{m-1}))q_0} I(J_m) + o_p\{1\} \\ &\leq B_4 \\ &\leq \sum_{m=1}^{M} \frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}[I(|V_{i,j}|\geq v_{m-1}) - P(|V_{i,j}|\geq v_{m-1})]}{2(1-\Phi(v_{m-2}))q_0} \frac{1-\Phi(v_{m-2})}{1-\Phi(v_m)} I(J_m) \\ &\quad - \sum_{m=1}^{M} \sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} \frac{P(v_{m-2}\leq |V_{i,j}| < v_{m-1})}{2(1-\Phi(v_{m-2}))q_0} \frac{1-\Phi(v_{m-2})}{1-\Phi(v_m)} I(J_m) + o_p\{1\}.\end{aligned}$$

Hence, for $B_4 = o_p\{1\}$, it is sufficient to show

$$\max_{1\leq m\leq M} \sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau} \frac{P(v_{m-1}\leq |V|_{i,j} < v_m)}{2(1-\Phi(v_{m-1}))q_0} = o_p\{1\}, \tag{S2.27}$$

$$\max_{2\leq m\leq M} \frac{1-\Phi(v_{m-2})}{1-\Phi(v_m)} = 1 + o(1), \tag{S2.28}$$

and

$$\max_{1\le m\le M}\left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}[I(|V_{i,j}|\ge v_m)-P(|V_{i,j}|\ge v_m)]}{2(1-\Phi(v_{m-1}))q_0}\right| = o_p\{1\}. \tag{S2.29}$$

As we showed $B_{21}=o_p\{1\}$ in the proof of Lemma 15, we also can show (S2.27). Since

$$\begin{aligned}\left|\max_{2\le m\le M}\frac{\Phi(v_m)-\Phi(v_{m-2})}{1-\Phi(v_m)}\right| &\le C\frac{2v_1\phi(v_{m-2})}{\phi(v_m)/v_m}\le Cv_1v_Me^{-\frac{v_{m-2}^2-v_m^2}{2}}\\ &\le Cv_1v_Me^{-\frac{(v_m-2v_1)^2-v_m^2}{2}}\\ &\le Cv_1v_Me^{2v_1v_m}=o(1)\quad(\because v_1v_M=O(\log\log\log\log p_1)^{-1})),\end{aligned}$$

we have (S2.28). To show (S2.29), we use the following inequality,

$$\begin{aligned}&P\left(\max_{0\le m\le M}\left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}[I(|V_{i,j}|\ge v_m)-P(|V_{i,j}|\ge v_m)]}{2(1-\Phi(v_{m-1}))q_0}\right|\ge\varepsilon\right)\\ &\le\sum_{m=0}^{M}P\left(\left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}[I(|V_{i,j}|\ge v_m)-P(|V_{i,j}|\ge v_m)]}{2(1-\Phi(v_{m-1}))q_0}\right|\ge\varepsilon\right)\end{aligned}$$

From (S2.28), we also know that

$$\max_{2\le m\le M}\frac{1-\Phi(v_{m-2})}{1-\Phi(v_m)}=1+o(1),$$

which implies

$$\max_{1\le m\le M}\frac{1-\Phi(v_{m-1})}{1-\Phi(v_m)}\ge C>0$$

for sufficiently large p_1. Therefore, we have

$$\begin{aligned}&P\left(\left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}[I(|V_{i,j}|\ge v_m)-P(|V_{i,j}|\ge v_m)]}{2(1-\Phi(v_{m-1}))q_0}\right|\ge\varepsilon\right)\\ &\le P\left(\left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}[I(|V_{i,j}|\ge v_m)-P(|V_{i,j}|\ge v_m)]}{2(1-\Phi(v_m))q_0}\right|\ge C\varepsilon\right).\end{aligned}$$

By (Xia et al., 2015, Lemma 4), we obtain

$$\sum_{1\le m\le M} P\left(\left|\frac{\sum_{(i,j)\in\mathcal{I}_0\setminus A_\tau}[I(|V_{i,j}|\ge v_m)-P(|V_{i,j}|\ge v_m)]}{2(1-\Phi(v_m))q_0}\right|\ge C\varepsilon\right)=o(1)$$

Therefore, we complete the proof. □

Lemma 18. *Let* $\mathbf{X}_k\in\mathbb{R}^p\sim N(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)$ *for* $k=1,\cdots,n_1$ *and* $\mathbf{Y}_k\in\mathbb{R}^p\sim N(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)$ *for* $k=1,\cdots,n_2$. *Define*

$$\tilde{\boldsymbol{\Sigma}}_1=(\tilde\sigma_{i,j,1})=\frac{1}{n_1}\sum_{k=1}^{n_1}(\mathbf{X}_k-\boldsymbol{\mu}_1)(\mathbf{X}_k-\boldsymbol{\mu}_1)^T,\quad \tilde{\boldsymbol{\Sigma}}_2=(\tilde\sigma_{i,j,2})=\frac{1}{n_2}\sum_{k=1}^{n_2}(\mathbf{Y}_k-\boldsymbol{\mu}_2)(\mathbf{Y}_k-\boldsymbol{\mu}_2)^T.$$

Then, for some constant $C>0$, $\tilde\sigma_{i,j,1}-\tilde\sigma_{i,j,2}$ *satisfies the large deviation bound*

$$\begin{aligned}&P\left(\max_{(i,j)\in\mathcal{S}}\frac{(\tilde\sigma_{i,j,1}-\tilde\sigma_{i,j,2}-\sigma_{i,j,1}-\sigma_{i,j,2})^2}{\mathrm{Var}((\mathbf{X}_{k,i}-\mu_{1,i})(\mathbf{X}_{k,j}-\mu_{1,j}))/n_1+\mathrm{Var}((\mathbf{Y}_{k,i}-\mu_{2,i})(\mathbf{Y}_{k,j}-\mu_{2,j}))/n_2}\ge x^2\right)\\&\le C|\mathcal{S}|(1-\Phi(x/\sqrt{2}))+o(1)\end{aligned}$$

uniformly for $0\le x\le\sqrt{n_2}$ *and any subset* $\mathcal{S}\subset\{(i,j):1\le i\le j\le p\}$ *where* Φ *is the cumulative distribution function of the standard normal.*

Proof. Assume $n_1\ge n_2$, and define

$$Z_{k,i,j}=\begin{cases}\dfrac{n_2}{n_1}\big((X_{k,i}-\mu_{1,i})(X_{k,j}-\mu_{1,j})-\sigma_{i,j,1}\big) & \text{for } 1\le k\le n_1\\[2ex] -\big((Y_{k-n_1,i}-\mu_{2,i})(Y_{k-n_1,j}-\mu_{2,j})-\sigma_{i,j,2}\big) & \text{for } n_1+1\le k\le n_1+n_2\end{cases}$$

Using the nomalities of X_k and Y_k, we can apply Lemma 2 to

$$(X_{k,i}-\mu_{1,i})(X_{k,j}-\mu_{1,j}),\ (Y_{k,i}-\mu_{2,i})(Y_{k,j}-\mu_{2,j}),\ \big((X_{k,i}-\mu_{1,i})(X_{k,j}-\mu_{1,j})-\sigma_{i,j,1}\big)^2,$$

and $\big((Y_{k,i}-\mu_{2,i})(Y_{k,j}-\mu_{2,j})-\sigma_{i,j,2}\big)^2$

as in Lemma 5, so we obtain

$$\begin{aligned}
P\left(\max_{i,j}|\tilde{\sigma}_{i,j,d}-\sigma_{i,j,d}|\geq C(\log p/n_d)^{1/2}\right)&=o(1),\\
P\left(\max_{i,j}\left|\frac{1}{n_1}\sum_{k=1}^{n_1}(Z_{k,i,j})^2-\frac{n_2^2}{n_1^2}\mathrm{Var}(X_{k,i}X_{k,j})\right|\geq C\left(\frac{n_2}{n_1}\right)^2(\log p/n_1)^{1/2}\right)&=o(1),\\
P\left(\max_{i,j}\left|\frac{1}{n_2}\sum_{k=n_1+1}^{n_1+n_2}(Z_{k,i,j})^2-\mathrm{Var}(Y_{k,i}Y_{k,j})\right|\geq C(\log p/n_1)^{1/2}\right)&=o(1).
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
&P\left(\max_{i,j}\frac{\sum_{k=1}^{n_1+n_2}(Z_{k,i,j})^2}{n_2^2\mathrm{Var}(X_{k,i}X_{k,j})/n_1+n_2\mathrm{Var}(Y_{k,i}Y_{k,j})}\geq 2\right)\\
&\leq P\left(\max_{i,j}\frac{\sum_{k=1}^{n_1}(Z_{k,i,j})^2}{n_2^2\mathrm{Var}(X_{k,i}X_{k,j})/n_1}\geq 2\right)+P\left(\max_{i,j}\frac{\sum_{k=n_1+1}^{n_1+n_2}(Z_{k,i,j})^2}{n_2\mathrm{Var}(Y_{k,i}Y_{k,j})}\geq 2\right)\\
&\leq P\left(\max_{i,j}\left(\sum_{k=1}^{n_1}(Z_{k,i,j})^2-\frac{n_2^2}{n_1}\mathrm{Var}(X_{k,i}X_{k,j})\right)\geq\frac{n_2^2}{n_1}H_{\min}\right)\\
&\quad+P\left(\max_{i,j}\left(\sum_{k=n_1+1}^{n_1+n_2}(Z_{k,i,j})^2-n_2\mathrm{Var}(Y_{k,ik}Y_{k,j})\right)\geq n_2H_{\min}\right)\\
&\leq P\left(\max_{i,j}\left(\frac{1}{n_1}\sum_{k=1}^{n_1}(Z_{k,i,j})^2-\frac{n_2^2}{n_1^2}\mathrm{Var}(X_{k,i}X_{k,j})\right)\geq\frac{n_2^2}{n_1^2}H_{\min}\right)\\
&\quad+P\left(\max_{i,j}\left(\frac{1}{n_2}\sum_{k=n_1+1}^{n_1+n_2}(Z_{k,i,j})^2-\mathrm{Var}(Y_{k,i}Y_{k,j})\right)\geq H_{\min}\right)=o(1),
\end{aligned}$$

where $H_{\min}:=\min\left(\min_{i,j}\mathrm{Var}(X_{k,i}X_{k,j})+\min_{i,j}\mathrm{Var}(Y_{k,i}Y_{k,j})\right)>0$. By (Jing et al., 2003, Theorem 1), we obtain

$$\max_{(i,j)\in\mathcal{S}} P\left(\frac{\left(\sum_{k=1}^{n_1+n_2} Z_{k,i,j}\right)^2}{\sum_{k=1}^{n_1+n_2}(Z_{k,i,j})^2} \geq x^2\right) \leq C|\mathcal{S}|(1-\Phi(x))$$

for $0 \leq x \leq \sqrt{n_2}$. Therefore,

$$\begin{aligned}
&P\left(\max_{(i,j)\in\mathcal{S}} \frac{(\tilde{\sigma}_{i,j,1}-\tilde{\sigma}_{i,j,2}-\sigma_{i,j,1}-\sigma_{i,j,2})^2}{\mathrm{Var}((\mathbf{X}_{k,i}-\mu_{1,i})(\mathbf{X}_{k,j}-\mu_{1,j}))/n_1+\mathrm{Var}((\mathbf{Y}_{k,i}-\mu_{2,i})(\mathbf{Y}_{k,j}-\mu_{2,j}))/n_2} \geq x^2\right)\\
&\leq P\left(\max_{(i,j)\in\mathcal{S}} \frac{\left(\sum_{k=1}^{n_1+n_2} Z_{i,j,k}\right)^2}{\sum_{k=1}^{n_1+n_2}(Z_{k,i,j})^2} \times \max_{i,j} \frac{\sum_{k=1}^{n_1+n_2}(Z_{k,i,j})^2}{n_2^2\mathrm{Var}(X_{k,i}X_{k,j})/n_1+n_2\mathrm{Var}(Y_{k,i}Y_{k,j})} \geq x^2\right)\\
&\leq P\left(\max_{(i,j)\in\mathcal{S}} \frac{\left(\sum_{k=1}^{n_1+n_2} Z_{k,i,j}\right)^2}{\sum_{k=1}^{n_1+n_2}(Z_{k,i,j})^2} \geq x^2/2\right)\\
&\quad + P\left(\max_{i,j} \frac{\sum_{k=1}^{n_1+n_2}(Z_{k,i,j})^2}{n_2^2\mathrm{Var}(X_{k,i}X_{k,j})/n_1+n_2\mathrm{Var}(Y_{k,i}Y_{k,j})} \geq 2\right)\\
&= C|\mathcal{S}|(1-\Phi(x/\sqrt{2}))+o(1).
\end{aligned}$$

□

S3 Algorithms for Weighted Conditional Testing

Our algorithms consist of (1) the estimation of the precision matrix, (2) the global testing procedure, and (3) the multiple testing procedure. The algorithms explained in this section are described in Algorithms 1-3. In addition, the computational plots when the number of nodes $p = p_1 = p_2$ and sample sizes $n_1 = n_2 = 50, 100$ are displayed in Figures 13 and 14.

In Figures 13 and 14 in the supplementary materials, it can be observed that the WCT model requires approximately 1.5 to 4 times more computation time compared to the NCT model. That is because WCT uses $(p_1+p_2)\times(p_1+p_2)$ network (given network + network given other network), but NCT uses only $p_1 \times p_1$ network (network given other network). In these testing procedures, the computation times are more influenced by the network dimensions than by sample size.

S3.1 Precision Matrix Estimation

Step 1) Using lasso, we estimate the regression coefficients $\boldsymbol{\beta}_i \in \mathbb{R}^{p_1+p_2-1}$ and residual $\hat{\varepsilon}_{k,i}$ of $\mathbf{X}_{k,i}$ to $\mathbf{X}_{k,-i}$ described in (S1.1) with the smoothing parameter $\lambda_{n,i} = \kappa\sqrt{(\hat{\sigma}_{i,i}\log p)/n}$.

Step 2) Using the estimator of $\boldsymbol{\beta}_i$, we then estimate $\omega^d_{i,j}$, $\theta^{d,w}_{i,j}$, and $\Delta_{i,j}$.

S3.2 Global Weighted Conditional Testing

Step 1) From the precision matrix estimation with $\kappa = 2$, compute $\Delta_{i,j}$ and $M_n = \max_{i \le i \le j \le p_1}(\Delta_{i,j})^2$.

Step 2) If $M_n \ge q_\alpha + 4\log p_1 - \log\log p_1$, where $q_\alpha = -\log(8\pi) - 2\log|\log(1-\alpha)|$ and a significance level of α, then reject $H_0 : \max_{1\le i\le j\le p_1}\left|\frac{\omega^1_{i,j}}{\sqrt{\omega^1_{i,i}\omega^1_{j,j}}} - \frac{\omega^2_{i,j}}{\sqrt{\omega^2_{i,i}\omega^2_{j,j}}}\right| = 0$.

S3.3 Multiple Weighted Conditional Testing

Step 1: Repeat the precision matrix estimation for $\kappa = 1/20, 2/20, \cdots, 39/20, 40/20$, and collect $\Delta_{i,j}$s for each κ. Then, select $\hat{\kappa}_0$ to minimize $\sum_{l=1}^{10}\left(\frac{\sum_{1\le i<j\le p_1} I\left\{|\Delta_{i,j}|\ge\Phi^{-1}(1-l[1-\Phi\{(\log p_1)^{1/2}\}]/10)\right\}}{lp_1(p_1-1)[1-\Phi\{(\log p_1)^{1/2}\}]/10} - 1\right)^2$. With $\hat{\kappa}_0$, compute $\Delta_{i,j}$.

Step 2: Start t from 0 and continue increasing by 0.001 until $\frac{[1-\Phi(t)]({p_1}^2-p_1)}{\max(N(t),1)} \le \alpha$ or $t > 2(\log p_1)^{1/2}$, where $N(t)$ defined in Section 3.2 of main document.

Step 3: For each $1 \le i < j \le p_1$, if $|\Delta_{i,j}| \ge \hat{t}_0$, then reject $H_{0,i,j} : \frac{\omega^1_{i,j}}{\sqrt{\omega^1_{i,i}\omega^1_{j,j}}} = \frac{\omega^2_{i,j}}{\sqrt{\omega^2_{i,i}\omega^2_{j,j}}}$.

S4 Settings of Simulation Studies

In Section 6 of the main document, we generate four types of sparse precision matrices, $\Omega^{(m)}$, $m = 1, \ldots, 4$ for the simulation studies. These simple examples of $\Omega^{(m)}$ with $p_1 = p_2 = 10$ and

$m = 1, \ldots, 4$ are displayed in Figure 1.

We conducted simulations for various scenarios as follows: 1) same sample size and same dimension of network, 2) different sample sizes and same dimension of network, 3) same sample size and different dimensions of network, and 4) different diagonal elements. We describe scenarios 2), 3), and 4) in this section. In the second scenario (Tables 3-6), we set $n_1 = 100$, $n_2 = 50, 75, 100$, and $p_1 = p_2 = 50$. For each combination of n_1, n_2, p_1, p_2, and m, we conduct 1000 simulations and study the performance of our approach based on the difference between n_1 and n_2. In the third scenario (Tables 7-10), we set $n = n_1 = n_2 = 100$ and $p_1 = 100$, $p_2 = 30, 50, 100$. For each combination of n, p_1, p_2, and m, we conduct 1000 simulations and study the performance of our approach based on the difference between p_1 and p_2. In the fourth scenario (Tables 11-14), we set $n = n_1 = n_2 = 100$ and $\tilde{p} = p_1 = p_2 = 50$. However, for each m, we consider $\mathbf{\Omega}^1$ and $\mathbf{\Omega}^2$ with

$$\mathbf{\Omega}^d = \frac{1}{1+\delta}\mathbf{D}_d^{1/2}(\mathbf{\Omega}^{(m)} + \delta\mathbf{I})\mathbf{D}_d^{1/2},$$

where $\mathbf{D}_d$ is a diagonal matrix whose diagonal elements are uniformly distributed between 0.5 and 2.5 independently and $\delta = |\lambda_{\min}(\mathbf{\Omega}^{(m)})| + 0.05$. We note that we set $\mathbf{D}_1 = \mathbf{D}_2$ in other scenarios. For each combination of n, $\tilde{p}$, and m, we conduct 1000 simulations and study the performance of our approach based on the difference between $\mathbf{D}_1$ and $\mathbf{D}_2$.

Using Tables 3 and 11 of the supplementary materials, we can also see the influence of adjusting diagonal elements. Adjusting the diagonal elements is crucial for analyzing pathway-based genetics data by controlling the effects of individual genes (or the diagonal elements of the precision matrices). Also, the unbalanced sample sizes between the two groups magnify the effects of individual genes. Therefore, by adjusting the diagonal elements, we can also reduce the impact of unbalanced sample sizes.

In Table 11 of the supplementary materials, we consider the case where the effects of individual genes (or the diagonal elements of the precision matrices) are different. As we can see in Table 11, WCT, the only test that adjusts the diagonal elements, achieves an empirical size similar to the significance level 5%, whereas the other tests exhibit much higher empirical sizes (above 98%). This demonstrates that adjusting the diagonal elements is crucial for analyzing pathway data by controlling the effects of individual genes.

Additionally, Table 3 of the supplementary materials reveals that the effectiveness of adjusting the diagonal elements becomes more pronounced as the sample size becomes unbalanced. For instance, in the case of $\Omega^{(2)}$ with the same given networks in Table 3, when $(n_1, n_2) = (100, 100)$, WCT, NWCT, and NCT have similar empirical sizes of 2.2%, 3.5%, and 4.2%, respectively. However, when $(n_1, n_2) = (100, 50)$, WCT achieves an empirical size of 1.4%, while the other two methods show empirical sizes exceeding 25%. This indicates that adjusting the diagonal elements is useful not only when the diagonal elements of the precision matrices differ but also when sample sizes are unbalanced.

As observed in the empirical size, Table 4 of the supplementary materials also illustrates the influence of adjusting diagonal elements. For example, in the case of $\Omega^{(4)}$ with the same given networks in Table 4, when $(n_1, n_2) = (100, 100)$, the empirical power of WCT, NWCT, and NCT exceeds 98% (specifically, 99.7%, 98.1%, and 98.7%, respectively). However, when $(n_1, n_2) = (100, 50)$, WCT still achieves an empirical power of 99.6%, while the empirical power of the other two methods falls below 90%.

S5 Testing Results on Breast Cancer Genetic Pathways

We apply our weighted conditional testing (WCT) and non-weighted conditional test (NWCT) to breast cancer genetic pathway data in Section 7. The human breast cancer data set consists of 22283 gene expression measurements from 176 White patients and 102 non-White patients. We collected 25 pathways whose ID, name, and the number of genes were listed in Table 15. We compare our approaches (WCT, NCT) with the nonconditional test (NCT) by Xia et al. (2015). We display selected results in Figures 2-12 and Tables 17, 18.

We found that some genes in Table 16 of the supplementary materials are associated with breast cancers. Functional polymorphisms of FAS and FASL gene were associated with the risk of breast cancer in the examined population (Hashemi et al., 2013). We also found that functional polymorphisms in the death pathway genes FAS and FASL significantly contribute to the occurrence of breast cancer (Wang et al., 2012). Reimer et al. (2002) reports that FASL:FAS ratio may be useful not only as a prognostic factor but also as a predictive factor for projecting response to the antioestrogen tamoxifen. The results strongly support a correlation between FASL:FAS ratio greater than 1 and lack of efficacy of tamoxifen in hormone receptor positive patients. A tumor-associated splice-isoform of MAP2K7 drives dedifferentiation in MBNL1-low cancers via JNK activation (Ray et al., 2020). Inhibition of the MAP2K7-JNK pathway with 5Z-7-oxozeaenol induces apoptosis in T-cell acute lymphoblastic leukemia (Chen et al., 2021).

We found some biological results of genes in Table 18 (WCT result). Zhao et al. (2022) report that RFC2 is a prognosis biomarker correlated with the immune signature in diffuse lower-grade gliomas. An essential function of a modified PSMA3, the immunoproteasome, is the processing of class I MHC peptides (GeneCards; The Human Gene Database). Saidy et al.

(2021) reports that PKA expression and PP1 expression are of significant interest in cancer as they are involved in a wide array of cellular processes, and these data indicate PKA and PP1 may play an important role in patient outcomes. CENPA, an essential centromere protein, is a prognostic marker for relapse in estrogen receptor-positive breast cancer (McGovern et al., 2012). PSMA7 - NUP133; these genes are both protein coding genes. Diseases associated with PSMA7 include Hepatitis and Hepatitis C Virus. Diseases associated with NUP133 include Galloway-Mowat Syndrome 8 and Nephrotic Syndrome, Type 18 (GeneCards; The Human Gene Database; https://www.genecards.org/). PMF1-BGLAP is a Protein Coding gene. Diseases associated with PMF1-BGLAP include Phototoxic Dermatitis and Cerebral Arteriopathy, Autosomal Dominant, With Subcortical Infarcts And Leukoencephalopathy, Type 1 (GeneCards; The Human Gene Database; https://www.genecards.org/).

Bibliography

Cai, T. and W. Liu (2011). Adaptive thresholding for sparse covariance matrix estimation. *J. Amer. Statist. Assoc. 106*(494), 672–684.

Cai, T., W. Liu, and Y. Xia (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J. Amer. Statist. Assoc. 108*(501), 265–277.

Chen, T., W. Du, J. Junco, and et al. (2021). Inhibition of the map2k7-jnk pathway with 5z-7-oxozeaenol induces apoptosis in t-cell acute lymphoblastic leukemia. *Oncotarget. 12(18)*, 1787–1801.

Hashemi, M., A. Fazaeli, S. Ghavami, and et al. (2013). Functional polymorphisms of fas and fasl gene and risk of breast cancer - pilot study of 134 cases. *PLoS One 8*.

Jing, B.-Y., Q.-M. Shao, and Q. Wang (2003). Self-normalized cramér-type large deviations for independent random variables. *The Annals of probability 31*(4), 2167–2215.

Liu, W. (2013). Gaussian graphical model estimation with false discovery rate control. *Ann. Statist. 41*(6), 2948–2978.

McGovern, S., Y. Qi, L. Pusztai, W. Symmans, and T. Buchholz (2012). Centromere protein-a, an essential centromere protein, is a prognostic marker for relapse in estrogen receptor-positive breast cancer. *Breast Cancer Res. 14(3)*, R72.

PETROV, V. (1975). Sums of independent random variables.

Ray, D., Y. C. Yun, M. Idris, S. Cheng, A. Boot, T. B. H. Iain, S. G. Rozen, P. Tan, and D. M. Epstein (2020). A tumor-associated splice-isoform of ¡i¿map2k7¡/i¿ drives dedifferentiation in mbnl1-low cancers via jnk activation. *Proceedings of the National Academy of Sciences 117*(28), 16391–16400.

Reimer, T., D. Koczan, H. Müller, K. Friese, H.-J. Thiesen, and B. Gerber (2002). Tumour fas ligand:fas ratio greater than 1 is an independent marker of relative resistance to tamoxifen therapy in hormone receptor positive breast cancer. *Breast Cancer Res. 4(5)*, R9.

Saidy, B., S. Kotecha, A. Butler, and et al. (2021). Pp1, pka and darpp-32 in breast cancer: A retrospective assessment of protein and mrna expression. *J Cell Mol Med. 25(11)*, 5015–5024.

Wang, W., Z. Zheng, W. Yu, H. Lin, B. Cui, and F. Cao (2012). Polymorphisms of the fas and fasl genes and risk of breast cancer. *Oncol Lett. 3(3)*, 625–628.

Xia, Y., T. Cai, and T. Cai (2015). Testing differential network with applications to the detection of gene-gene interaction. *Biometrika 102*, 247–266.

Zaitsev, A. Y. (1987). On the gaussian approximation of convolutions under multidimensional analogues of sn bernstein's inequality conditions. *Probability theory and related fields 74*(4), 535–566.

Zhao, X., Y. Wang, J. Li, and et al (2022). Rfc2: a prognosis biomarker correlated with the immune signature in diffuse lower-grade gliomas. *Sci Rep. 12(1)*, 3122.

Table 1: Empirical FDR (%) and the power of testing the equality of two precision matrices (Ω_1, Ω_2) when same given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ are generated by $\boldsymbol{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = \tilde{p} = 50, 100, 200$, the number of nodes of the given network $p_2 = \tilde{p} = 50, 100, 200$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.1$ and 0.2. For $m = 1, 2, 3, 4$, $\boldsymbol{\Omega}^{(1)}$ is a pentadiagonal matrix, $\boldsymbol{\Omega}^{(2)}$ is a scale-free network, $\boldsymbol{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\boldsymbol{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\boldsymbol{\Omega}^{(1)}$			$\boldsymbol{\Omega}^{(2)}$			$\boldsymbol{\Omega}^{(3)}$			$\boldsymbol{\Omega}^{(4)}$		
Same given networks case $((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2}))$												
$\tilde{p}$	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
	empirical FDR with $\alpha = 0.1$											
50	10.0	12.1	10.6	11.4	13.2	14.1	9.3	9.7	10.4	12.5	13.0	13.5
100	10.8	12.8	13.4	10.5	11.2	10.9	13.1	13.1	13.2	10.8	10.7	11.6
200	10.5	9.9	10.4	11.3	10.1	10.4	11.6	10.9	10.9	10.9	10.7	9.9
	empirical FDR with $\alpha = 0.2$											
50	23.5	23.0	23.0	21.7	24.6	24.5	20.2	21.8	21.3	24.0	26.4	25.3
100	23.3	22.8	24.3	19.6	20.8	20.5	22.6	24.1	25.2	20.6	22.7	23.3
200	21.6	19.1	19.8	21.3	20.4	20.0	22.9	21.8	23.6	20.4	19.3	20.0
	Power with $\alpha = 0.1$											
50	44.3	34.8	38.2	46.0	37.4	38.9	47.9	38.0	40.4	46.9	38.4	42.9
100	48.7	38.6	43.1	50.4	41.4	41.8	50.4	40.5	44.6	47.7	39.2	42.6
200	50.9	39.7	41.8	50.5	40.4	41.3	53.9	43.4	44.7	48.3	37.9	38.6
	Power with $\alpha = 0.2$											
50	53.1	42.2	46.7	53.5	44.2	44.8	54.1	38.0	40.4	55.3	46.1	51.4
100	54.6	44.8	49.3	56.5	48.1	49.8	57.0	46.5	50.6	55.1	46.0	48.9
200	55.9	44.9	47.0	55.6	45.4	46.1	58.8	48.9	50.6	53.8	43.6	44.5

Table 2: Empirical FDR(%) and the power of testing the equality of two precision matrices (Ω_1, Ω_2) when different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ are generated by $\mathbf{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = \tilde{p} = 50, 100, 200$, the number of nodes of the given network $p_2 = \tilde{p} = 50, 100, 200$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.1$ and 0.2. For $m = 1, 2, 3, 4$, $\mathbf{\Omega}^{(1)}$ is a pentadiagonal matrix, $\mathbf{\Omega}^{(2)}$ is a scale-free network, $\mathbf{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\mathbf{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal element is 0.5.

	$\mathbf{\Omega}^{(1)}$			$\mathbf{\Omega}^{(2)}$			$\mathbf{\Omega}^{(3)}$			$\mathbf{\Omega}^{(4)}$		
Different given networks case $((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2}))$												
$\tilde{p}$	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
	empirical FDR with $\alpha = 0.1$											
50	10.0	10.6	10.1	10.2	12.1	11.0	11.1	12.1	13.1	9.9	11.1	9.8
100	12.1	12.7	12.6	9.1	10.8	10.5	9.3	11.1	9.7	11.2	11.0	11.3
200	8.7	9.0	9.0	10.3	9.5	9.3	10.2	10.6	10.5	9.2	8.2	7.7
	empirical FDR with $\alpha = 0.2$											
50	18.5	18.9	19.0	21.8	22.1	23.1	20.4	21.7	23.6	20.5	20.7	21.6
100	20.9	21.8	22.4	19.2	18.8	20.0	20.5	21.1	22.1	19.7	19.9	20.4
200	19.0	17.8	18.1	21.4	20.6	19.3	20.4	20.6	19.2	16.7	18.6	16.6
	Power with $\alpha = 0.1$											
50	39.0	32.8	36.2	38.6	30.9	32.3	37.3	29.9	30.2	36.0	29.7	33.1
100	43.9	36.1	37.9	41.9	33.4	34.2	43.9	35.3	37.7	41.4	32.6	33.4
200	44.3	38.9	40.0	42.6	30.8	31.1	44.5	33.4	33.4	44.3	32.9	33.9
	Power with $\alpha = 0.2$											
50	46.1	40.2	43.4	46.5	39.6	41.3	45.1	37.6	39.4	43.6	37.2	41.6
100	50.4	42.7	44.3	47.9	39.3	39.8	50.8	42.6	44.8	47.7	38.8	40.1
200	49.9	38.9	40.0	47.9	36.9	36.5	49.7	39.9	40.3	48.7	38.7	39.1

Table 3: Empirical sizes (%) for testing the equality of two precision matrices $(\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ when same or different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ are generated under H_0 using $\mathbf{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = 50$, the number of nodes of the given network $p_2 = 50$, sample size $n_1 = 100$, $n_2 = 50, 75, 100$, and significance level $\alpha = 0.05$. For $m = 1, 2, 3, 4$, $\mathbf{\Omega}^{(1)}$ is a pentadiagonal matrix, $\mathbf{\Omega}^{(2)}$ is a scale-free network, $\mathbf{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\mathbf{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\mathbf{\Omega}^{(1)}$			$\mathbf{\Omega}^{(2)}$			$\mathbf{\Omega}^{(3)}$			$\mathbf{\Omega}^{(4)}$		
Same given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
n_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
50	2.5	2.9	6.5	1.4	33.6	25.5	1.8	7.7	8.8	2.8	3.2	7.0
75	2.3	2.6	4.3	2.4	6.3	6.1	2.3	4.1	5.9	3.2	3.3	5.8
100	3.5	3.1	5.7	2.2	3.5	4.2	2.2	3.7	4.6	2.4	2.8	3.6
different given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
n_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
50	6.0	11.0	20.0	3.2	8.7	14.4	4.7	10.7	21.0	4.5	9.5	16.7
75	6.2	6.8	21.8	3.7	5.8	15.1	5.1	5.7	20.7	5.4	6.4	17.6
100	5.5	4.7	21.1	3.5	4.1	13.4	6.2	5.3	21.1	4.5	4.5	18.0

Table 4: Empirical powers(%) for testing the equality of two precision matrices $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ when same or different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ are generated under H_1 using $\boldsymbol{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = 50$, the number of nodes of the given network $p_2 = 50$, sample size $n_1 = 100$, $n_2 = 50, 75, 100$, and significance level $\alpha = 0.05$. For $m = 1, 2, 3, 4$, $\boldsymbol{\Omega}^{(1)}$ is a pentadiagonal matrix, $\boldsymbol{\Omega}^{(2)}$ is a scale-free network, $\boldsymbol{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\boldsymbol{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\boldsymbol{\Omega}^{(1)}$			$\boldsymbol{\Omega}^{(2)}$			$\boldsymbol{\Omega}^{(3)}$			$\boldsymbol{\Omega}^{(4)}$		
Same given networks case $\big((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\big)$												
n_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
50	97.4	88.5	89.6	94.5	84.1	84.8	98.8	90.3	91.4	96.2	87.2	89.2
75	99.4	94.0	96.7	98.3	92.7	93.8	99.5	96.1	97.1	97.9	93.1	94.3
100	99.3	97.4	98.2	98.4	94.4	95.4	99.6	97.9	98.4	99.7	98.1	98.7
Different given networks case $\big((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\big)$												
n_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
50	87.0	60.7	70.0	82.2	56.6	63.7	90.1	63.3	72.7	85.3	59.9	68.2
75	95.9	82.6	88.8	91.4	75.9	81.4	97.2	85.0	88.4	95.4	80.9	87.2
100	97.6	89.7	94.2	96.1	88.5	91.0	98.1	91.6	94.1	97.4	90.2	92.9

Table 5: Empirical FDR (%) and the power of testing the equality of two precision matrices (Ω_1, Ω_2) when same given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ are generated by $\mathbf{\Omega}^{(m)}$, $m = 1, \dots, 4$, with the number of nodes of the network given other networks $p_1 = 50$, the number of nodes of the given network $p_2 = 50$, sample size $n_1 = 100$, $n_2 = 50, 75, 100$, and significance level $\alpha = 0.1$ and 0.2. For $m = 1, 2, 3, 4$, $\mathbf{\Omega}^{(1)}$ is a pentadiagonal matrix, $\mathbf{\Omega}^{(2)}$ is a scale-free network, $\mathbf{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\mathbf{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\mathbf{\Omega}^{(1)}$			$\mathbf{\Omega}^{(2)}$			$\mathbf{\Omega}^{(3)}$			$\mathbf{\Omega}^{(4)}$		
Same given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
n_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
	empirical FDR with $\alpha = 0.1$											
50	16.8	15.5	13.8	12.6	14.8	12.1	15.0	15.4	15.5	11.2	11.9	12.6
75	13.2	15.0	13.0	12.6	11.5	10.7	13.4	13.8	15.0	11.7	13.0	13.7
100	10.0	12.1	10.6	11.4	13.2	14.1	9.3	9.7	10.4	12.5	13.0	13.5
	empirical FDR with $\alpha = 0.2$											
50	26.4	27.1	27.3	25.0	25.3	23.7	23.5	24.5	25.2	22.4	22.6	23.7
75	24.7	25.9	27.3	24.5	25.2	24.2	23.7	25.6	26.2	24.3	25.8	26.1
100	23.5	23.0	23.0	21.7	24.6	24.5	20.2	21.8	21.3	24.0	26.4	25.3
	Power with $\alpha = 0.1$											
50	35.5	25.6	26.7	35.7	28.4	27.4	35.8	26.2	26.8	37.1	29.2	30.1
75	42.8	34.9	37.2	43.0	34.2	34.7	42.7	33.4	34.9	41.0	31.4	35.8
100	44.3	34.8	38.2	46.0	37.4	38.9	47.9	38.0	40.4	46.9	38.4	42.9
	Power with $\alpha = 0.2$											
50	42.9	34.0	36.7	44.1	35.3	34.5	43.7	34.7	36.9	43.4	35.8	37.9
75	50.8	42.7	45.9	49.3	42.5	42.7	49.5	42.4	44.2	48.5	40.6	44.8
100	53.1	42.2	46.7	53.5	44.2	44.8	54.1	38.0	40.4	55.3	46.1	51.4

Table 6: Empirical FDR(%) and the power of testing the equality of two precision matrices (Ω_1, Ω_2) when different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ are generated by $\mathbf{\Omega}^{(m)}$, $m = 1, \dots, 4$, with the number of nodes of the network given other networks $p_1 = 50$, the number of nodes of the given network $p_2 = 50$, sample size $n_1 = 100$, $n_2 = 50, 75, 100$, and significance level $\alpha = 0.1$ and 0.2. For $m = 1, 2, 3, 4$, $\mathbf{\Omega}^{(1)}$ is a pentadiagonal matrix, $\mathbf{\Omega}^{(2)}$ is a scale-free network, $\mathbf{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\mathbf{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal element is 0.5.

	$\mathbf{\Omega}^{(1)}$			$\mathbf{\Omega}^{(2)}$			$\mathbf{\Omega}^{(3)}$			$\mathbf{\Omega}^{(4)}$		
Different given networks case $((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2}))$												
n_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
	empirical FDR with $\alpha = 0.1$											
50	14.9	14.9	14.9	11.9	13.0	12.3	12.6	12.9	12.8	11.2	11.1	11.5
75	12.5	12.2	12.3	12.7	12.6	13.4	10.8	13.2	13.1	11.1	11.9	10.8
100	10.0	10.6	10.1	10.2	12.1	11.0	11.1	12.1	13.1	9.9	11.1	9.8
	empirical FDR with $\alpha = 0.2$											
50	25.4	27.3	25.8	26.7	25.3	24.5	24.1	24.5	26.5	21.3	21.7	24.3
75	22.1	22.0	22.8	22.8	23.1	25.8	21.5	22.1	23.2	21.5	21.5	22.1
100	18.5	18.9	19.0	21.8	22.1	23.1	20.4	21.7	23.6	20.5	20.7	21.6
	Power with $\alpha = 0.1$											
50	27.8	21.6	22.9	24.2	19.9	18.9	25.1	20.3	20.7	26.9	20.0	21.8
75	31.6	25.0	28.7	31.7	26.7	27.4	30.5	24.6	25.1	32.4	25.6	27.1
100	39.0	32.8	36.2	38.6	30.9	32.3	37.3	29.9	30.2	36.0	29.7	33.1
	Power with $\alpha = 0.2$											
50	33.9	27.9	29.0	31.3	26.6	25.8	31.4	26.3	27.9	33.5	26.4	28.9
75	40.6	33.5	36.8	39.1	34.0	35.2	37.4	32.3	32.4	40.6	33.5	36.0
100	46.1	40.2	43.4	46.5	39.6	41.3	45.1	37.6	39.4	43.6	37.2	41.6

Table 7: Empirical sizes (%) for testing the equality of two precision matrices $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ when same or different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ are generated under H_0 using $\boldsymbol{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = 100$, the number of nodes of the given network $p_2 = 30, 50, 100$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.05$. For $m = 1, 2, 3, 4$, $\boldsymbol{\Omega}^{(1)}$ is a pentadiagonal matrix, $\boldsymbol{\Omega}^{(2)}$ is a scale-free network, $\boldsymbol{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\boldsymbol{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\boldsymbol{\Omega}^{(1)}$			$\boldsymbol{\Omega}^{(2)}$			$\boldsymbol{\Omega}^{(3)}$			$\boldsymbol{\Omega}^{(4)}$		
Same given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
p_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
30	3.4	3.3	5.1	2.4	3.7	3.6	2.3	3.4	4.2	3.6	3.8	5.6
50	2.4	2.5	4.6	3.1	3.7	3.8	2.3	3.3	3.3	2.6	2.6	3.4
100	2.7	2.7	3.7	3.0	3.2	3.4	2.3	2.9	3.7	2.9	2.9	4.4
different given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
p_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
30	5.3	5.8	18.3	3.9	4.2	12.7	4.7	4.1	16.7	4.9	4.4	18.0
50	4.3	4.0	18.5	3.8	3.9	12.3	5.4	4.9	19.8	3.4	3.7	17.5
100	3.9	3.9	17.1	4.6	5.2	12.5	3.9	3.7	19.6	4.2	4.1	18.0

Table 8: Empirical powers(%) for testing the equality of two precision matrices $(\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ when same or different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ are generated under H_1 using $\mathbf{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = 100$, the number of nodes of the given network $p_2 = 30, 50, 100$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.05$. For $m = 1, 2, 3, 4$, $\mathbf{\Omega}^{(1)}$ is a pentadiagonal matrix, $\mathbf{\Omega}^{(2)}$ is a scale-free network, $\mathbf{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\mathbf{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\mathbf{\Omega}^{(1)}$			$\mathbf{\Omega}^{(2)}$			$\mathbf{\Omega}^{(3)}$			$\mathbf{\Omega}^{(4)}$		
Same given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
p_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
30	99.2	95.7	96.6	98.5	94.0	94.0	100.0	97.5	97.9	99.4	94.4	95.8
50	99.4	95.2	96.5	98.8	93.9	94.8	99.9	97.5	98.0	99.2	95.7	96.6
100	99.9	96.4	97.0	99.0	93.7	94.7	99.9	97.3	97.8	99.3	96.1	97.3
Different given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
p_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
30	95.2	88.3	90.7	96.3	85.2	88.1	98.7	90.8	93.9	95.3	83.1	86.5
50	97.2	88.4	91.9	94.7	82.7	85.0	99.3	91.2	93.8	95.6	86.6	89.6
100	97.1	88.7	92.0	95.6	85.9	89.6	98.4	87.9	93.5	96.7	85.5	90.1

Table 9: Empirical FDR (%) and the power of testing the equality of two precision matrices (Ω_1, Ω_2) when same given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ are generated by $\mathbf{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = 100$, the number of nodes of the given network $p_2 = 30, 50, 100$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.1$ and 0.2. For $m = 1, 2, 3, 4$, $\mathbf{\Omega}^{(1)}$ is a pentadiagonal matrix, $\mathbf{\Omega}^{(2)}$ is a scale-free network, $\mathbf{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\mathbf{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\mathbf{\Omega}^{(1)}$			$\mathbf{\Omega}^{(2)}$			$\mathbf{\Omega}^{(3)}$			$\mathbf{\Omega}^{(4)}$		
Same given networks case $((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2}))$												
p_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
	empirical FDR with $\alpha = 0.1$											
30	11.4	12.4	12.0	11.5	12.4	12.9	11.0	12.6	12.8	10.8	11.7	11.3
50	11.0	11.5	12.6	9.8	10.9	12.0	11.9	12.6	13.0	11.7	13.1	13.6
100	10.8	12.8	13.4	10.5	11.2	10.9	13.1	13.1	13.2	10.8	10.7	11.6
	empirical FDR with $\alpha = 0.2$											
30	20.6	22.9	23.1	21.3	22.4	23.2	21.8	23.8	23.0	20.9	21.1	22.5
50	22.1	23.1	24.5	20.2	20.9	21.9	20.7	21.7	24.0	23.6	24.2	24.8
100	23.3	22.8	24.3	19.6	20.8	20.5	22.6	24.1	25.2	20.6	22.7	23.3
	Power with $\alpha = 0.1$											
30	48.8	40.8	43.1	46.2	38.1	38.2	49.9	41.2	41.9	47.7	39.1	43.1
50	50.1	41.7	44.3	49.1	41.6	42.7	51.2	41.5	43.1	51.1	44.7	46.5
100	48.7	38.6	43.1	50.4	41.4	41.8	50.4	40.5	44.6	47.7	39.2	42.6
	Power with $\alpha = 0.2$											
30	54.3	47.2	49.1	52.8	44.8	45.0	57.0	49.1	50.0	53.3	46.2	48.6
50	56.5	48.5	51.6	54.8	48.7	49.3	56.9	48.3	50.0	57.9	51.2	53.0
100	54.6	44.8	49.3	56.5	48.1	49.8	57.0	46.5	50.6	55.1	46.0	48.9

Table 10: Empirical FDR(%) and the power of testing the equality of two precision matrices (Ω_1, Ω_2) when different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ are generated by $\boldsymbol{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = 100$, the number of nodes of the given network $p_2 = 30, 50, 100$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.1$ and 0.2. For $m = 1, 2, 3, 4$, $\boldsymbol{\Omega}^{(1)}$ is a pentadiagonal matrix, $\boldsymbol{\Omega}^{(2)}$ is a scale-free network, $\boldsymbol{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\boldsymbol{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal element is 0.5.

	$\boldsymbol{\Omega}^{(1)}$			$\boldsymbol{\Omega}^{(2)}$			$\boldsymbol{\Omega}^{(3)}$			$\boldsymbol{\Omega}^{(4)}$		
Different given networks case $((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2}))$												
p_2	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
	empirical FDR with $\alpha = 0.1$											
30	9.7	10.0	10.9	10.6	10.0	10.3	10.9	11.6	12.8	10.9	11.9	12.7
50	11.0	10.9	10.4	10.8	10.0	9.9	9.6	11.1	10.4	9.1	9.4	10.0
100	12.1	12.7	12.6	9.1	10.8	10.5	9.3	11.1	9.7	11.2	11.0	11.3
	empirical FDR with $\alpha = 0.2$											
30	21.0	21.4	21.8	21.5	21.1	21.5	21.7	23.5	22.7	23.7	23.9	24.0
50	21.2	22.5	22.3	21.1	21.5	22.9	21.1	19.3	19.5	19.0	20.1	19.4
100	20.9	21.8	22.4	19.2	18.8	20.0	20.5	21.1	22.1	19.7	19.9	20.4
	Power with $\alpha = 0.1$											
30	43.1	41.8	43.6	38.8	32.8	32.7	38.0	29.5	30.4	42.4	34.9	35.7
50	50.9	42.5	44.5	42.5	35.3	35.5	40.1	34.0	34.1	43.7	34.9	35.7
100	43.9	36.1	37.9	41.9	33.4	34.2	43.9	35.3	37.7	41.4	32.6	33.4
	Power with $\alpha = 0.2$											
30	50.3	41.8	43.6	46.2	39.5	39.0	45.6	36.5	37.4	49.8	41.4	43.0
50	50.9	42.5	44.5	48.7	42.1	42.5	47.6	40.0	40.5	50.4	41.1	42.5
100	50.4	42.7	44.3	47.9	39.3	39.8	50.8	42.6	44.8	47.7	38.8	40.1

Table 11: Empirical sizes (%) for testing the equality of two precision matrices $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ when same or different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ are generated under H_0 using $\boldsymbol{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = p$, the number of nodes of the given network $p_2 = p$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.05$. For $m = 1, 2, 3, 4$, $\boldsymbol{\Omega}^{(1)}$ is a pentadiagonal matrix, $\boldsymbol{\Omega}^{(2)}$ is a scale-free network, $\boldsymbol{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\boldsymbol{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\boldsymbol{\Omega}^{(1)}$			$\boldsymbol{\Omega}^{(2)}$			$\boldsymbol{\Omega}^{(3)}$			$\boldsymbol{\Omega}^{(4)}$		
Same given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
$\mathbf{D}_d$	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
same	3.5	3.1	5.7	2.2	3.5	4.2	2.2	3.7	4.6	2.4	2.8	3.6
different	2.6	100	100	2.5	100	100	2.4	100	100	1.9	99.9	100
different given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
$\mathbf{D}_d$	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
same	5.5	4.7	21.1	3.5	4.1	13.4	6.2	5.3	21.1	4.5	4.5	18.0
different	5.1	98.9	99.5	5.0	99.7	100	5.2	98.3	99.6	5.1	99.0	99.5

Table 12: Empirical powers(%) for testing the equality of two precision matrices $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ when same or different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ are generated under H_1 using $\boldsymbol{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = p$, the number of nodes of the given network $p_2 = p$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.05$. For $m = 1, 2, 3, 4$, $\boldsymbol{\Omega}^{(1)}$ is a pentadiagonal matrix, $\boldsymbol{\Omega}^{(2)}$ is a scale-free network, $\boldsymbol{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\boldsymbol{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\boldsymbol{\Omega}^{(1)}$			$\boldsymbol{\Omega}^{(2)}$			$\boldsymbol{\Omega}^{(3)}$			$\boldsymbol{\Omega}^{(4)}$		
Same given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
$\mathbf{D}_d$	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
same	99.3	97.4	98.2	98.4	94.4	95.4	99.6	97.9	98.4	99.7	98.1	98.7
different	99.5	100	100	99.3	100	100	99.6	100	100	99.7	100	100
Different given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
$\mathbf{D}_d$	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
same	97.6	89.7	94.2	96.1	88.5	91.0	98.1	91.6	94.1	97.4	90.2	92.9
different	97.3	99.7	99.6	95.8	99.8	99.7	98.9	99.6	99.8	97.1	99.5	99.6

Table 13: Empirical FDR (%) and the power of testing the equality of two precision matrices (Ω_1, Ω_2) when same given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ are generated by $\mathbf{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = p$, the number of nodes of the given network $p_2 = p$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.1$ and 0.2. For $m = 1, 2, 3, 4$, $\mathbf{\Omega}^{(1)}$ is a pentadiagonal matrix, $\mathbf{\Omega}^{(2)}$ is a scale-free network, $\mathbf{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\mathbf{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal elements is 0.5.

	$\mathbf{\Omega}^{(1)}$			$\mathbf{\Omega}^{(2)}$			$\mathbf{\Omega}^{(3)}$			$\mathbf{\Omega}^{(4)}$		
Same given networks case $((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) = (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2}))$												
$\mathbf{D}_d$	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
	empirical FDR with $\alpha = 0.1$											
same	10.0	12.1	10.6	11.4	13.2	14.1	9.3	9.7	10.4	12.5	13.0	13.5
different	10.3	11.2	11.3	10.7	11.2	10.3	12.5	14.0	11.8	12.2	12.8	13.3
	empirical FDR with $\alpha = 0.2$											
same	23.5	23.0	23.0	21.7	24.6	24.5	20.2	21.8	21.3	24.0	26.4	25.3
different	17.3	21.3	21.6	20.4	21.7	19.1	21.5	25.8	23.8	21.7	23.0	24.0
	Power with $\alpha = 0.1$											
same	44.3	34.8	38.2	46.0	37.4	38.9	47.9	38.0	40.4	46.9	38.4	42.9
different	46.9	32.7	38.8	47.1	36.4	35.7	47.3	36.3	38.5	47.1	34.2	39.5
	Power with $\alpha = 0.2$											
same	53.1	42.2	46.7	53.5	44.2	44.8	54.1	38.0	40.4	55.3	46.1	51.4
different	54.7	43.7	48.3	55.5	44.3	44.3	54.8	45.0	47.6	55.4	43.1	48.0

Table 14: Empirical FDR(%) and the power of testing the equality of two precision matrices (Ω_1, Ω_2) when different given networks using the weighted conditional test (WCT), nonweighted conditional test (NWCT), and nonconditional test (NCT); $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ are generated by $\boldsymbol{\Omega}^{(m)}$, $m = 1, \ldots, 4$, with the number of nodes of the network given other networks $p_1 = p$, the number of nodes of the given network $p_2 = p$, sample size $n_1 = n_2 = 100$, and significance level $\alpha = 0.1$ and 0.2. For $m = 1, 2, 3, 4$, $\boldsymbol{\Omega}^{(1)}$ is a pentadiagonal matrix, $\boldsymbol{\Omega}^{(2)}$ is a scale-free network, $\boldsymbol{\Omega}^{(3)}$ is a symmetric matrix whose upper off-diagonal elements are from *Binomial* distribution, and $\boldsymbol{\Omega}^{(4)}$ is a symmetric matrix whose randomly assigned off-diagonal element is 0.5.

	$\boldsymbol{\Omega}^{(1)}$			$\boldsymbol{\Omega}^{(2)}$			$\boldsymbol{\Omega}^{(3)}$			$\boldsymbol{\Omega}^{(4)}$		
Different given networks case $\left((\Omega^1_{I_1,I_2}, \Omega^1_{I_2,I_2}) \neq (\Omega^2_{I_1,I_2}, \Omega^2_{I_2,I_2})\right)$												
$\mathbf{D}_d$	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT	WCT	NWCT	NCT
	empirical FDR with $\alpha = 0.1$											
same	10.0	10.6	10.1	10.2	12.1	11.0	11.1	12.1	13.1	9.9	11.1	9.8
different	9.9	9.7	10.2	11.2	11.8	11.0	9.3	11.7	11.8	11.2	11.9	9.9
	empirical FDR with $\alpha = 0.2$											
same	18.5	18.9	19.0	21.8	22.1	23.1	20.4	21.7	23.6	20.5	20.7	21.6
different	18.3	20.1	19.3	19.9	19.6	19.8	19.7	20.3	23.6	20.9	21.4	20.8
	Power with $\alpha = 0.1$											
same	39.0	32.8	36.2	38.6	30.9	32.3	37.3	29.9	30.2	36.0	29.7	33.1
different	43.4	33.2	35.7	39.4	30.6	30.5	38.7	30.7	33.0	37.6	27.0	29.5
	Power with $\alpha = 0.2$											
same	46.1	40.2	43.4	46.5	39.6	41.3	45.1	37.6	39.4	43.6	37.2	41.6
different	50.9	43.3	46.7	47.7	40.2	41.2	47.0	40.3	42.9	46.2	34.3	38.6

Table 15: Pathway id, name, and the number of genes of the pathway in the breast cancer genetic pathways.

Pathway Id	Pathway Name	The Number of Genes
113	KEGG_FOCAL_ADHESION	195
137	KEGG_NEUROTROPHIN_SIGNALING_PATHWAY	118
141	KEGG_REGULATION_OF_ACTIN_CYTOSKELETON	197
142	KEGG_INSULIN_SIGNALING_PATHWAY	129
153	KEGG_ALZHEIMERS_DISEASE	145
672	REACTOME_SIGNALLING_BY_NGF	196
683	REACTOME_TCA_CYCLE_AND_RESPIRATORY _ELECTRON_TRANSPORT	106
711	REACTOME_ORC1_REMOVAL_FROM_CHROMATIN	63
717	REACTOME_SIGNALING_BY_ERBB2	91
750	REACTOME_SIGNALING_BY_THE_B_CELL _RECEPTOR_BCR	113
772	REACTOME_SIGNALING_BY_FGFR_IN_DISEASE	112
852	REACTOME_METABOLISM_OF_AMINO_ACIDS _AND_DERIVATIVES	176
857	REACTOME_BIOLOGICAL_OXIDATIONS	117
930	REACTOME_TRANSMEMBRANE_TRANSPORT_OF _SMALL_MOLECULES	338
943	REACTOME_SIGNALING_BY_PDGF	113
955	REACTOME_METABOLISM_OF_PROTEINS	366
956	REACTOME_DOWNSTREAM_SIGNAL_TRANSDUCTION	85
968	REACTOME_TRANSCRIPTION	177
1075	REACTOME_METABOLISM_OF_RNA	241
1111	REACTOME_METABOLISM_OF_LIPIDS_AND _LIPOPROTEINS	399
1206	REACTOME_DNA_REPLICATION	173
1219	REACTOME_HEMOSTASIS	415
1271	REACTOME_ADAPTIVE_IMMUNE_SYSTEM	453
1277	REACTOME_CLASS_I_MHC_MEDIATED_ANTIGEN _PROCESSING_PRESENTATION	203
1289	REACTOME_PLATELET_ACTIVATION_SIGNALING _AND_AGGREGATION	185

Table 16: List of significant edges under the nonconditional test (NCT), the non-weighted conditional test (NWCT), and the weighted conditional test (WCT) with FDR 0.2 for breast cancer genetic pathway with the testing pathway: 137, the conditioned pathway: 717

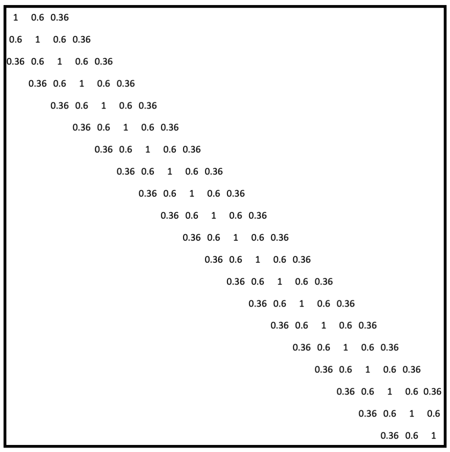
NCT	NWCT	WCT
	CALM1 - ATF4	CALM1 - ATF4
	ARHGDIA - GRB2	ARHGDIA - GRB2
	MAPK14 - HRAS	MAPK14 - HRAS
		PLCG1 - MAPK1
	RPS6KA4 - CAMK2G	RPS6KA4 - CAMK2G
		MAPK11 - YWHAB
		NFKB1 - GRB2
MAP2K7 - FASLG	MAP2K7 - FASLG	MAP2K7 - FASLG

Table 17: List of significant edges under the nonconditional test (NCT), the non-weighted conditional test (NWCT), and the weighted conditional test (WCT) with FDR 0.2 for breast cancer genetic pathway data with the testing pathway: 113, the conditioned pathway: 137

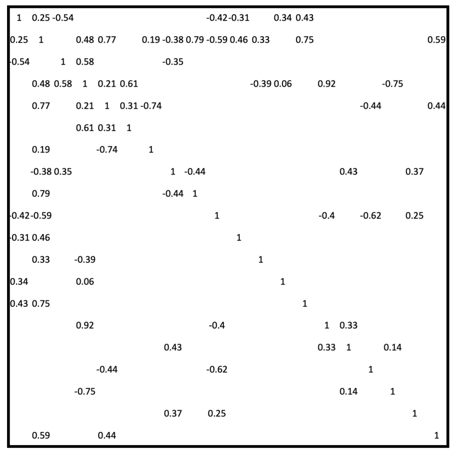
NCT	NWCT	WCT
COL1A2 - COL5A2		
VEGFB - ACTN2		
PIK3CA - LAMB2		

Table 18: List of significant edges under the nonconditional test (NCT), the non-weighted conditional test (NWCT), and the weighted conditional test (WCT) with FDR 0.2 for the breast cancer genetic pathway with the testing pathway: 1206, the conditioned pathway: 750

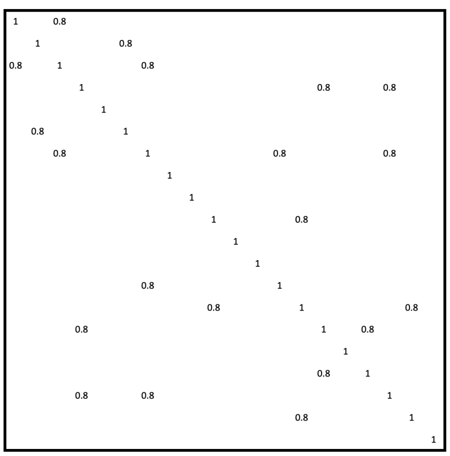
NCT	NWCT	WCT
	RFC2 - PSMA3	RFC2 - PSMA3
	PPP2R1A - PSMD7	
PPP2R1A - LIG1	PPP2R1A - LIG1	PPP2R1A - LIG1
	PPP1CC - PSMA1	PPP1CC - PSMA1
RPS27 - POLE2	RPS27 - POLE2	RPS27 - POLE2
		PSMD8 - PSMC3
PSMD2 - CLASP1		
	PSMD4 - CENPA	
	PSMA7 - NUP133	
		PSMA7 - PSME4
		PSMA7 - UBA52
PSMA5 - ZWINT	PSMA5 - ZWINT	PSMA5 - ZWINT
PSMA3 - STAG2	PSMA3 - STAG2	PSMA3 - STAG2
PSMA1 - CDK2	PSMA1 - CDK2	PSMA1 - CDK2
NUP133 - PSMB10	NUP133 - PSMB10	NUP133 - PSMB10
PLK1 - STAG1		
PLK1 - SPC25	PLK1 - SPC25	PLK1 - SPC25
CCNA2 - SEH1L	CCNA2 - SEH1L	CCNA2 - SEH1L
	PMF1///BGLAP - POLD3	PMF1///BGLAP - POLD3
		STAG2 - NUP85
CKAP5 - NUP107	CKAP5 - NUP107	CKAP5 - NUP107
		SKA1 - FBXO5



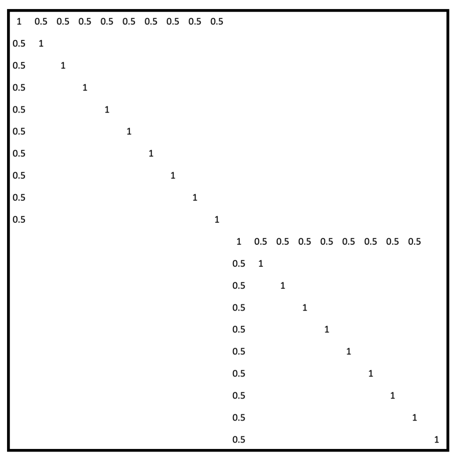
(a) An example of $\Omega^{(1)}$



(b) An example of $\Omega^{(2)}$



(c) An example of $\Omega^{(3)}$



(d) An example of $\Omega^{(4)}$

Figure 1: Examples of $\mathbf{\Omega^{(d)}}$ for $p_1 = p_2 = 10$, $d = 1, 2, 3, 4$ in the simulation setting. Shaded part is a network given other networks, and non-shaded part is a given network.

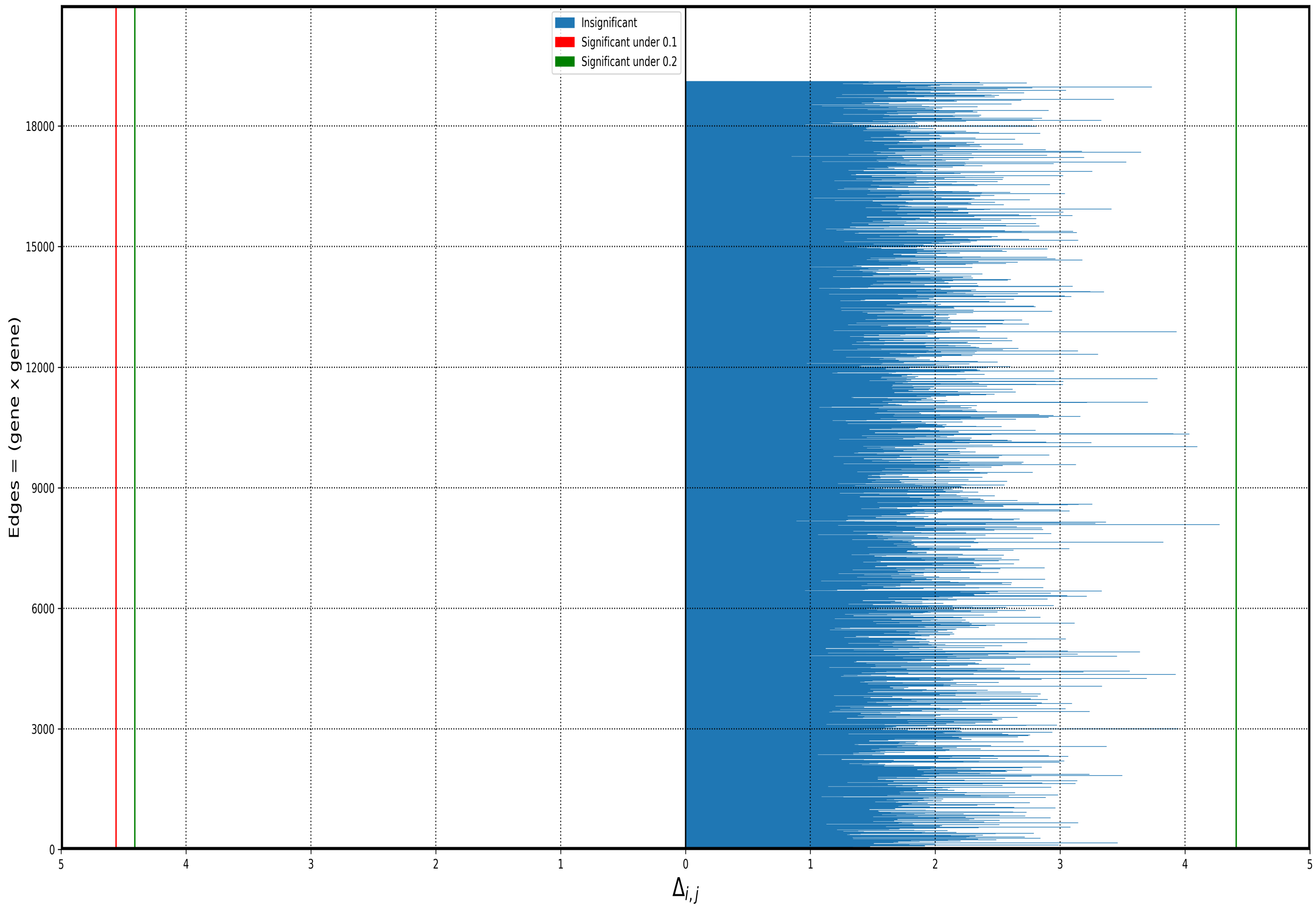


Figure 2: $|\Delta_{i,j}| = \dfrac{|\hat{\omega}^{1,w}_{i,j} - \hat{\omega}^{2,w}_{i,j}|}{(\hat{\theta}^{1}_{i,j} + \hat{\theta}^{2}_{i,j})^{1/2}}$ for breast cancer genetic pathway with the testing pathway: 113, the conditioned pathway: 137

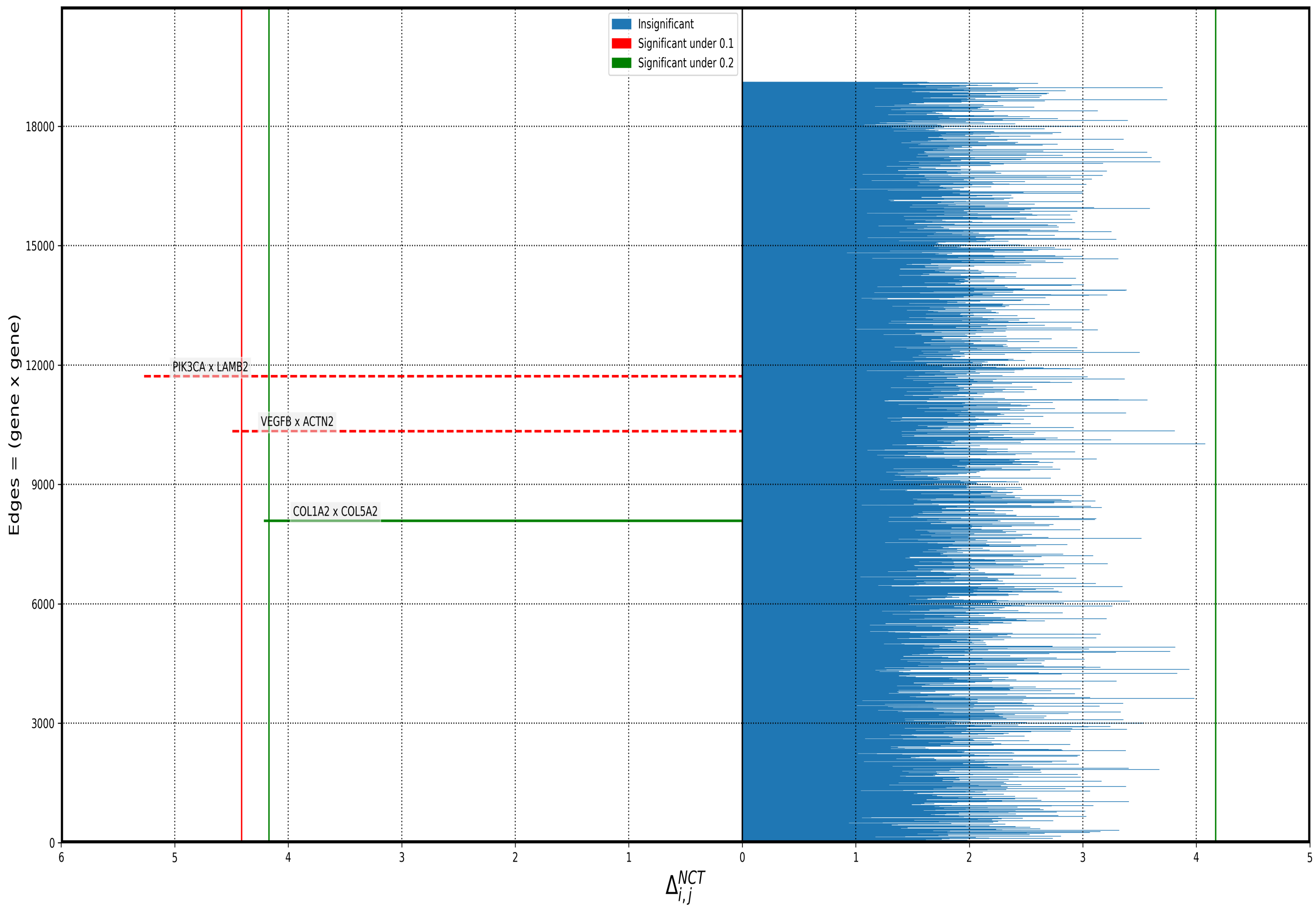


Figure 3: $|\Delta_{i,j}^{NCT}| = \dfrac{|\hat{\omega}_{i,j}^1 - \hat{\omega}_{i,j}^2|}{(\hat{\theta}_{i,j}^{1,NCT} + \hat{\theta}_{i,j}^{2,NCT})^{1/2}}$ for breast cancer genetic pathway with the testing pathway: 113, where $\hat{\theta}_{i,j}^{d,NCT} = \hat{\theta}_{i,j}^d / (\hat{r}_{i,i}^d \hat{r}_{j,j}^d)$

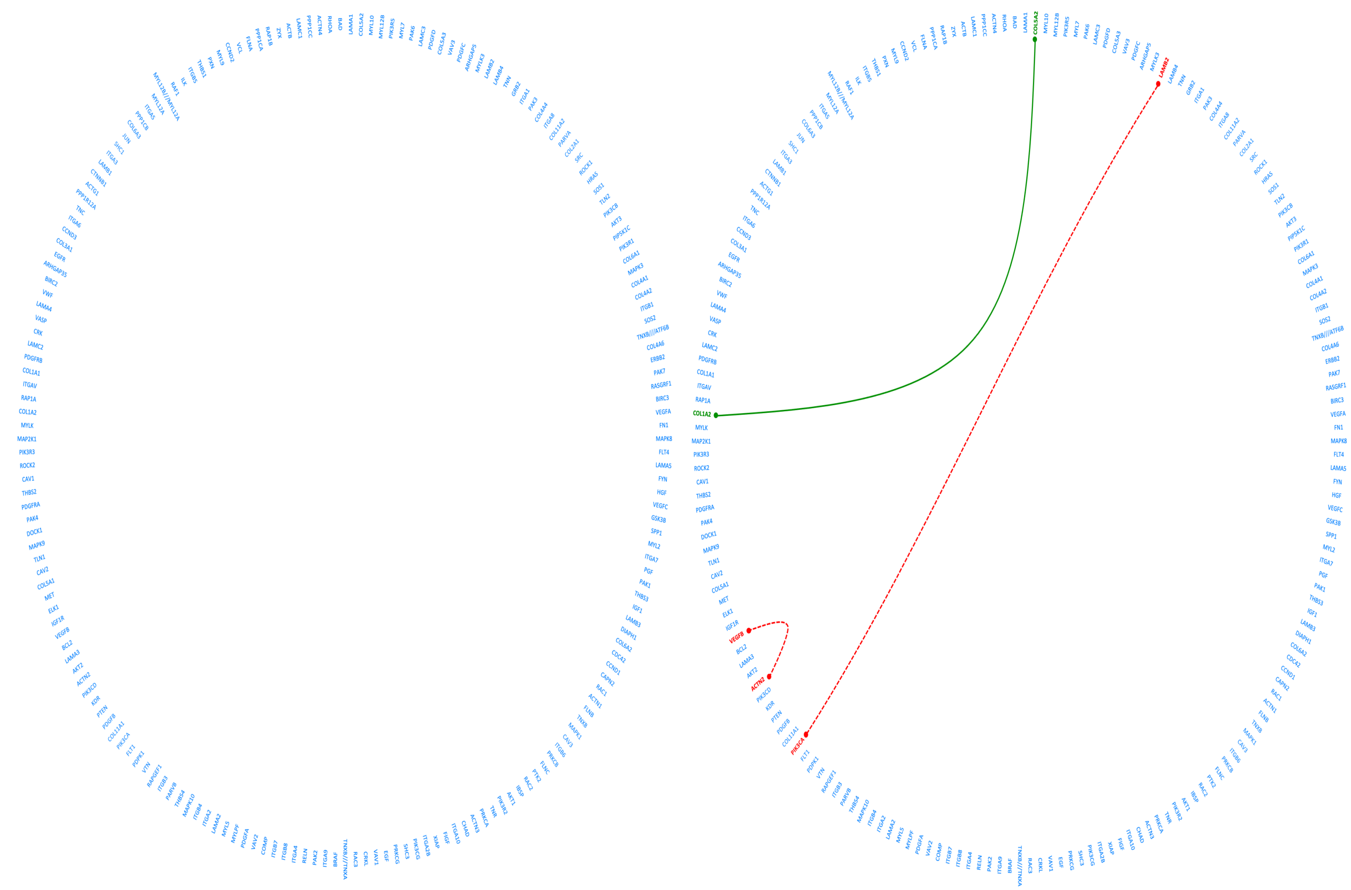


Figure 4: $\Delta_{i,j}$ for breast cancer genetic pathway with the testing pathway: 113, the conditional pathway: 137

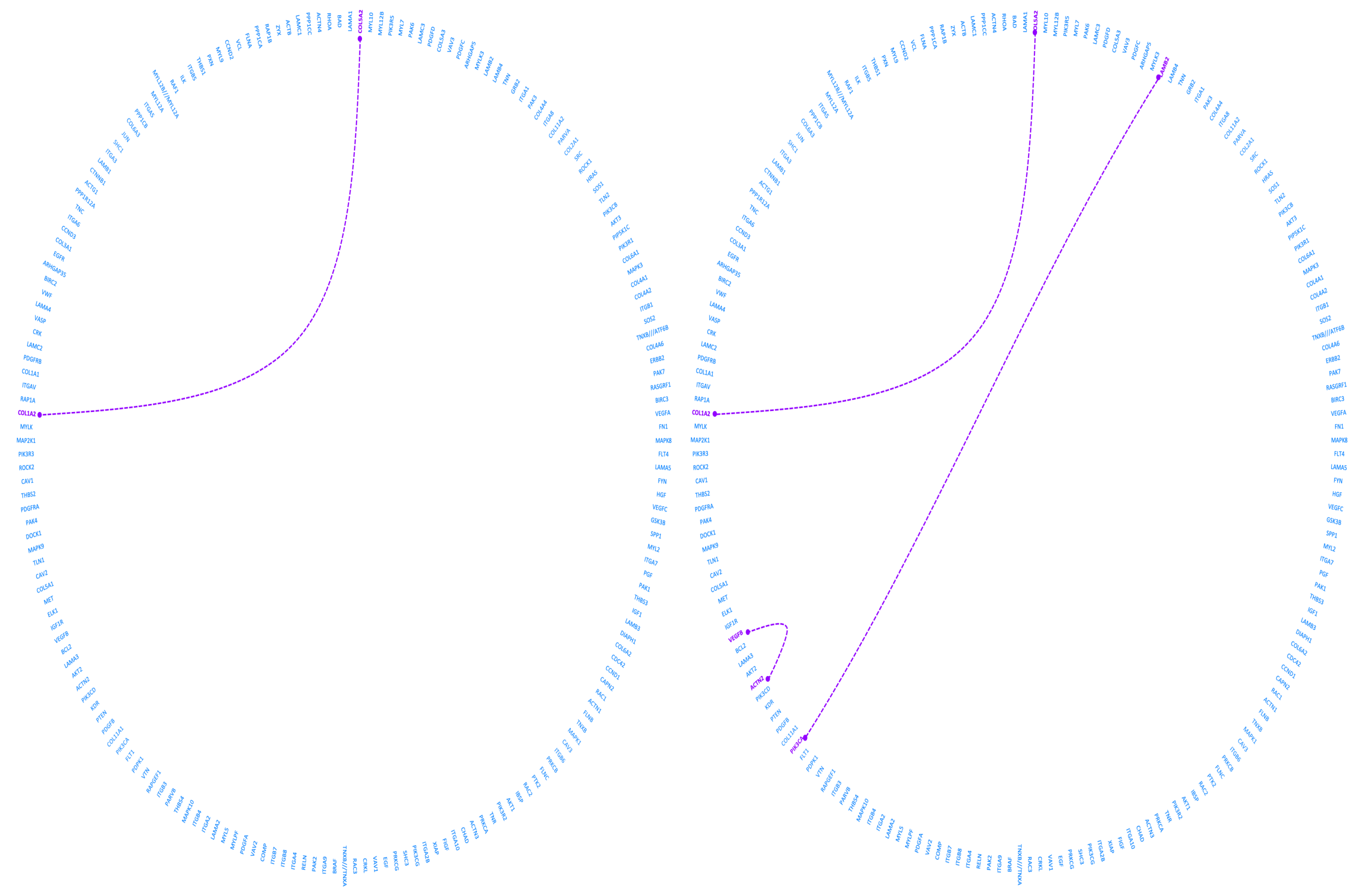


Figure 5: $\Delta_{i,j}$ for breast cancer genetic pathway with the testing pathway: 113, the conditional pathway: 137

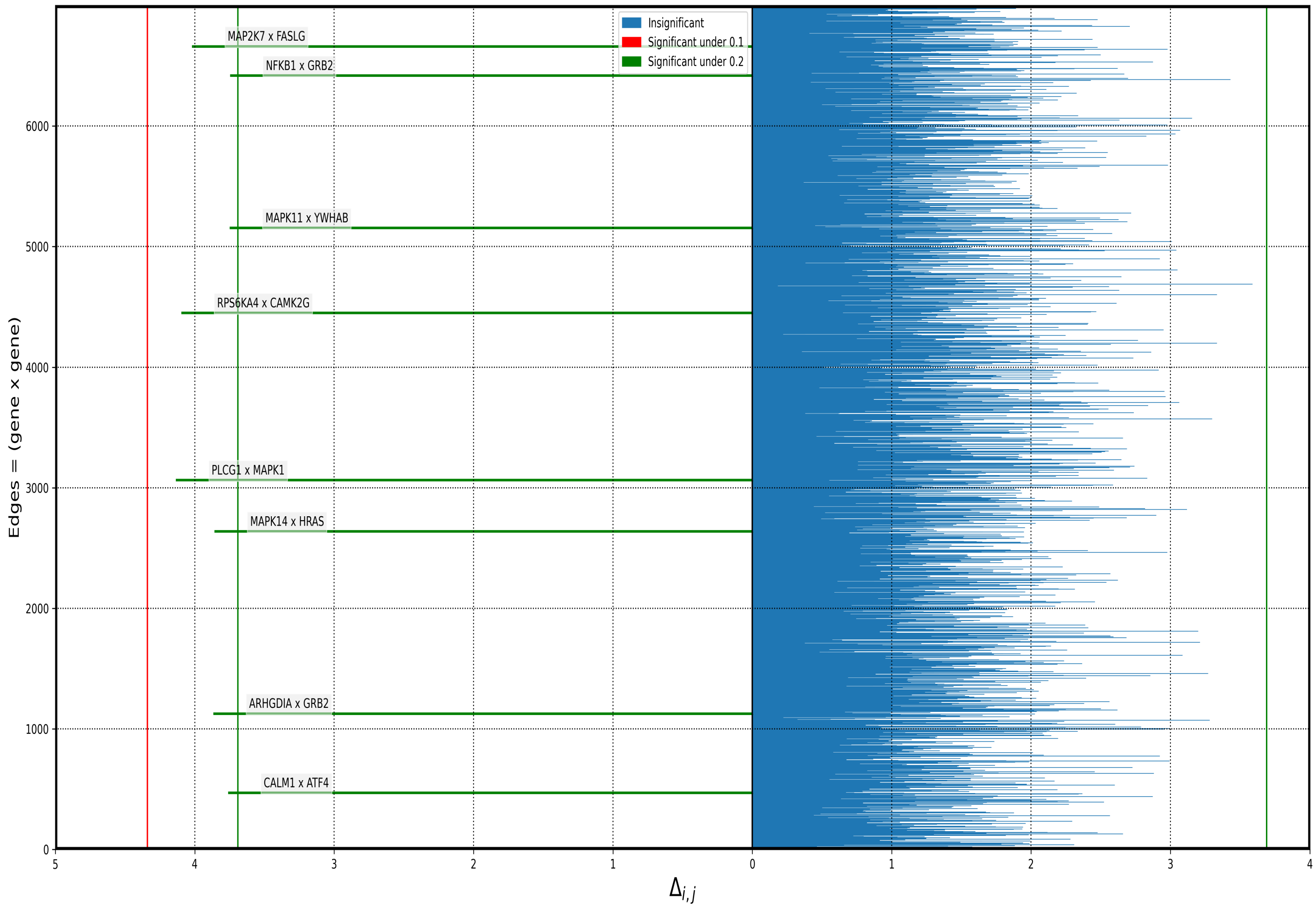


Figure 6: $|\Delta_{i,j}| = \dfrac{|\hat{\omega}^{1,w}_{i,j} - \hat{\omega}^{2,w}_{i,j}|}{(\hat{\theta}^1_{i,j} + \hat{\theta}^2_{i,j})^{1/2}}$ for breast cancer genetic pathway with the testing pathway: 137, the conditioned pathway: 717

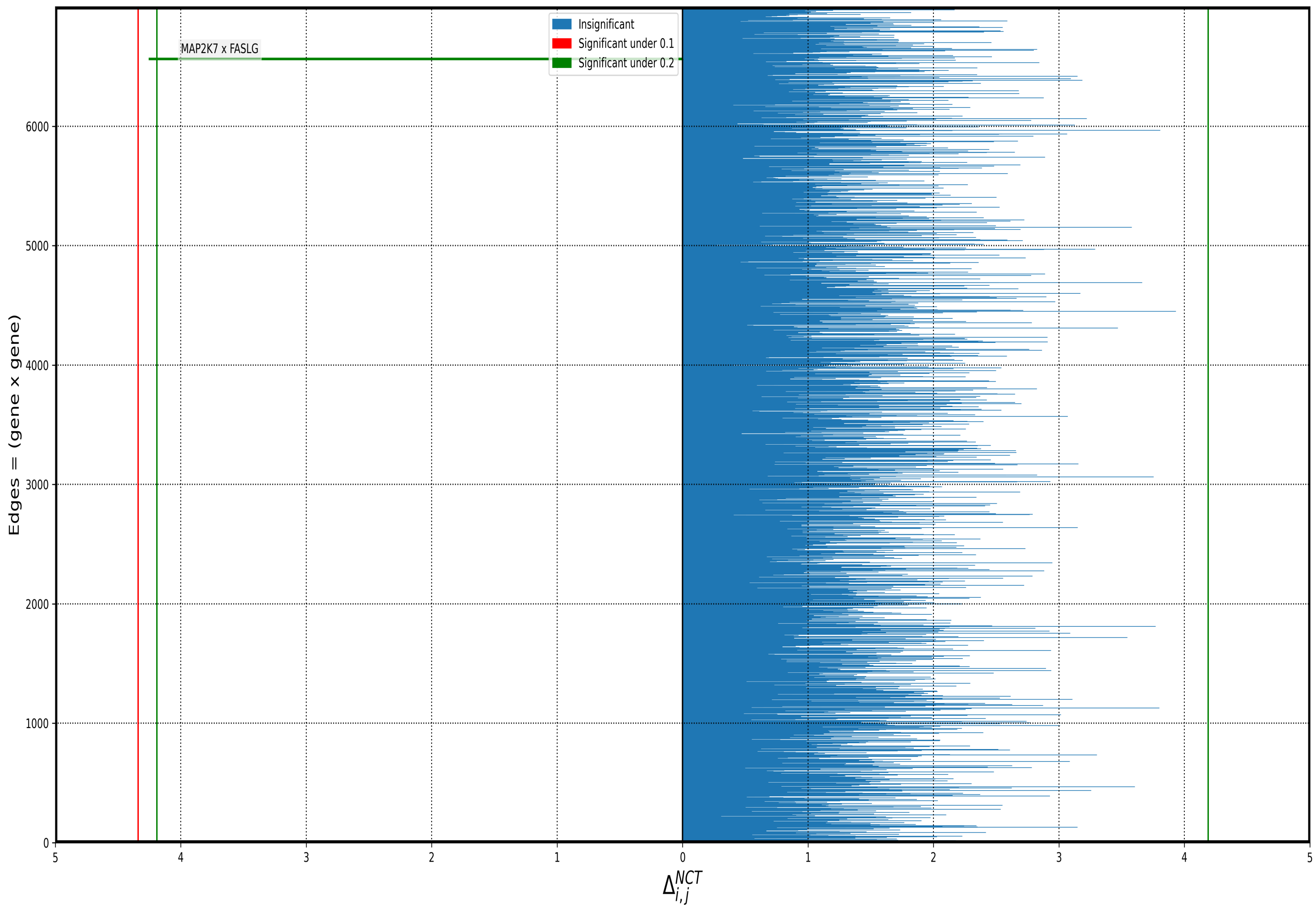


Figure 7: $|\Delta_{i,j}^{NCT}| = \dfrac{|\hat{\omega}_{i,j}^1 - \hat{\omega}_{i,j}^2|}{(\hat{\theta}_{i,j}^{1,NCT} + \hat{\theta}_{i,j}^{2,NCT})^{1/2}}$ for breast cancer genetic pathway with the testing pathway: 137, where $\hat{\theta}_{i,j}^{d,NCT} = \hat{\theta}_{i,j}^d / (\hat{r}_{i,i}^d \hat{r}_{j,j}^d)$

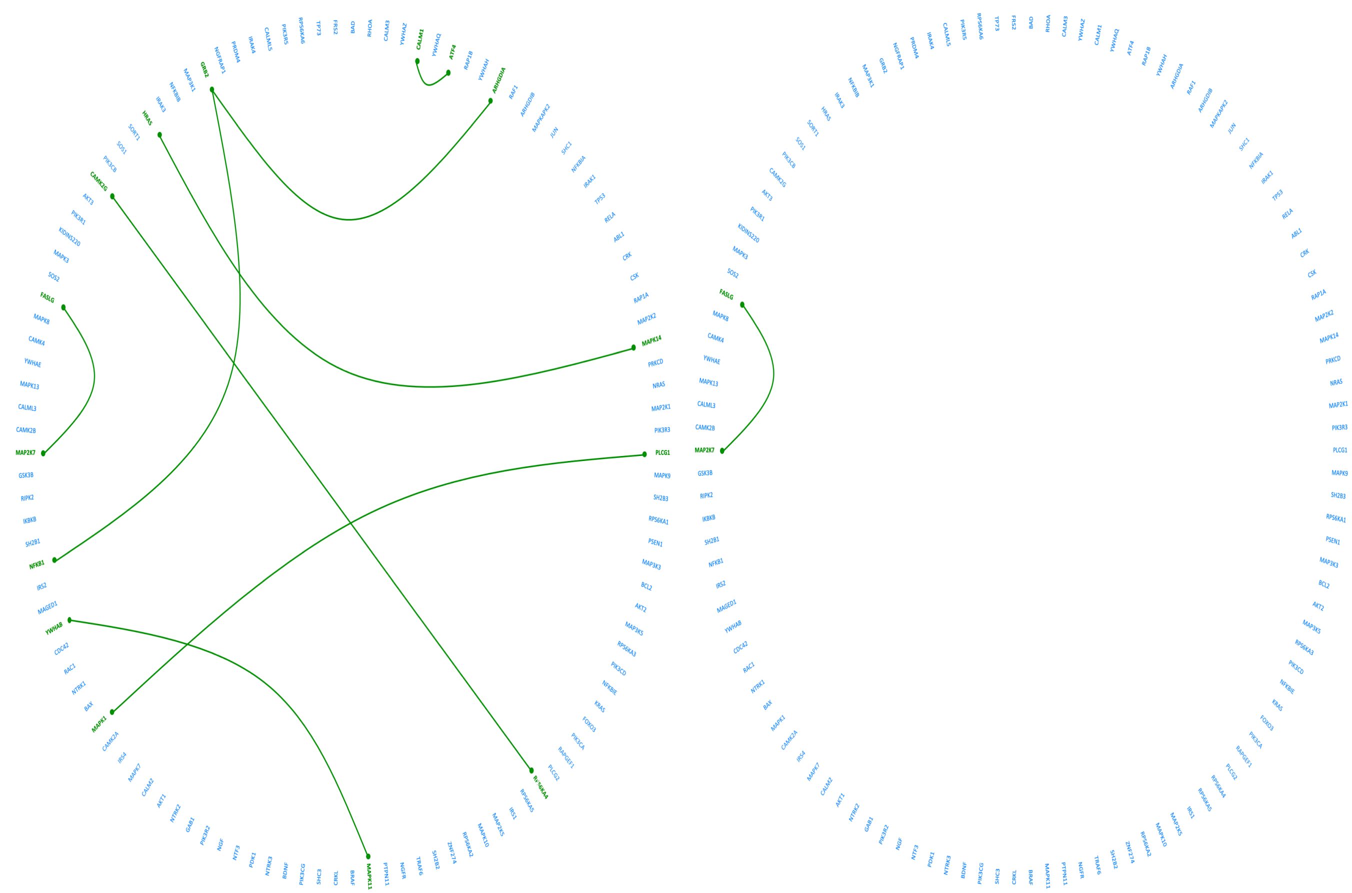


Figure 8: $\Delta_{i,j}$ for breast cancer genetic pathway with the testing pathway: 137, the conditional pathway: 717

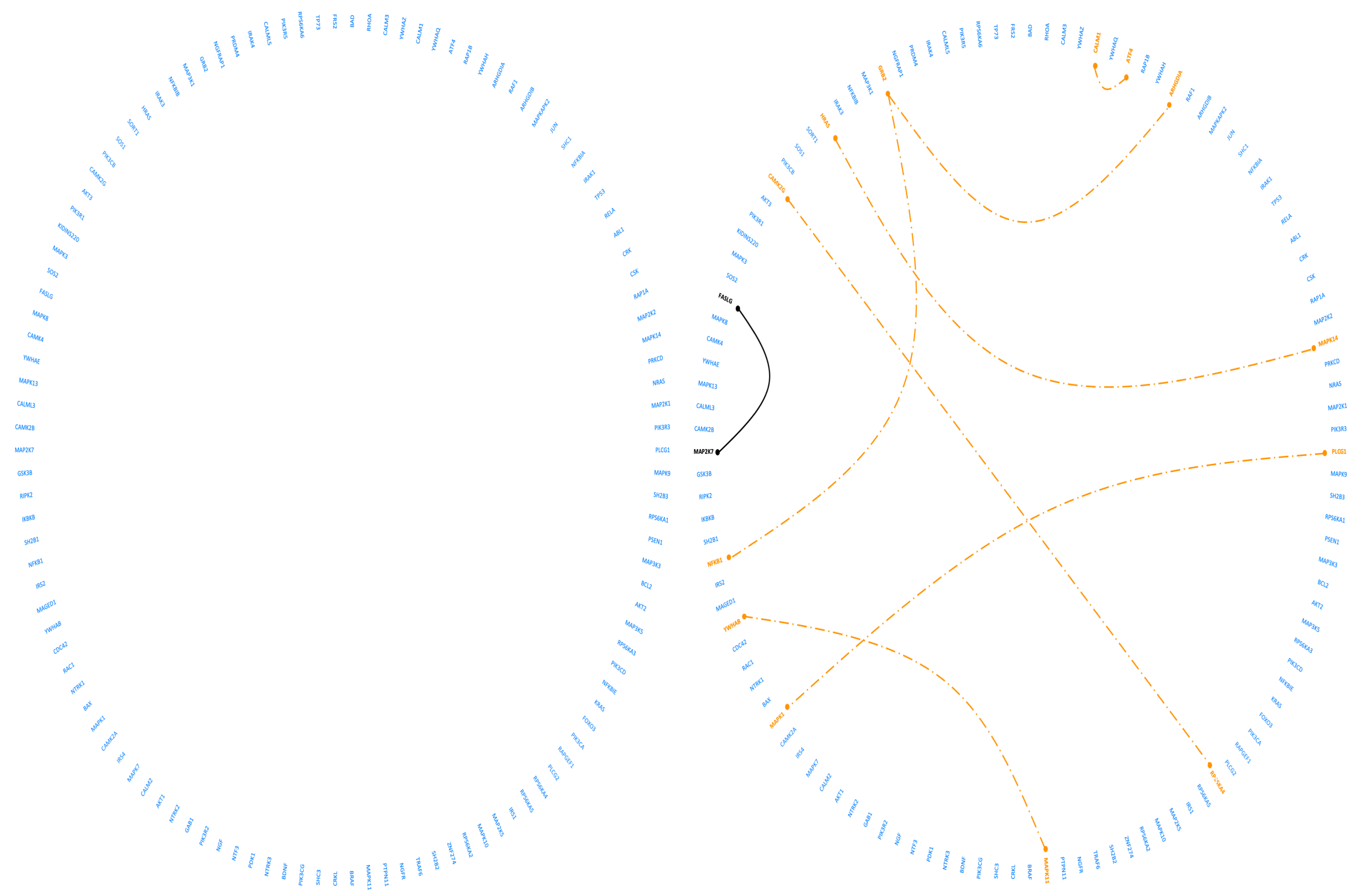


Figure 9: $\Delta_{i,j}$ for breast cancer genetic pathway with the testing pathway: 137, the conditional pathway: 717

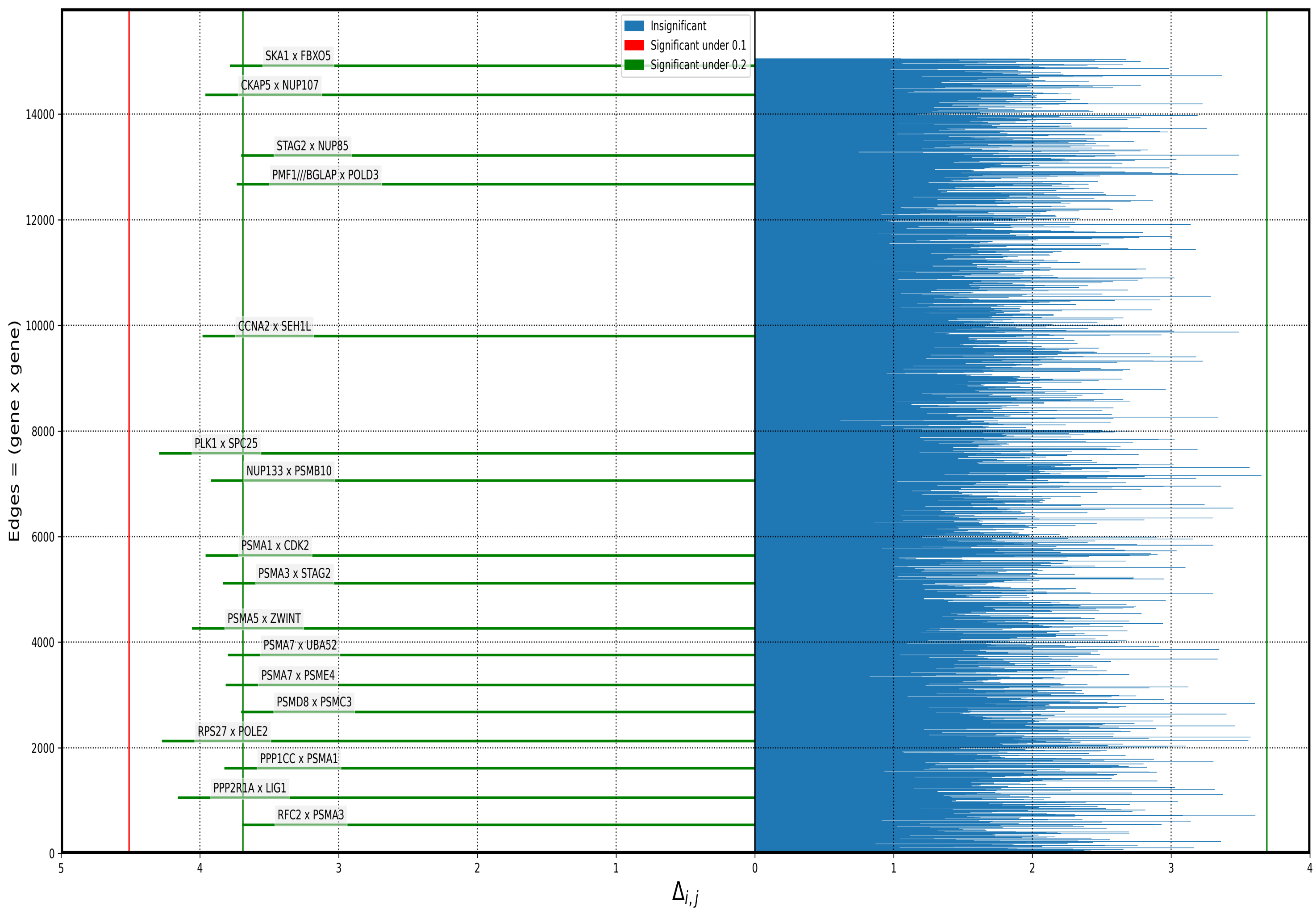


Figure 10: $|\Delta_{i,j}| = \dfrac{|\hat{\omega}^{1,w}_{i,j} - \hat{\omega}^{2,w}_{i,j}|}{(\hat{\theta}^1_{i,j} + \hat{\theta}^2_{i,j})^{1/2}}$ for breast cancer genetic pathway with the testing pathway: 1206, the conditional pathway: 750

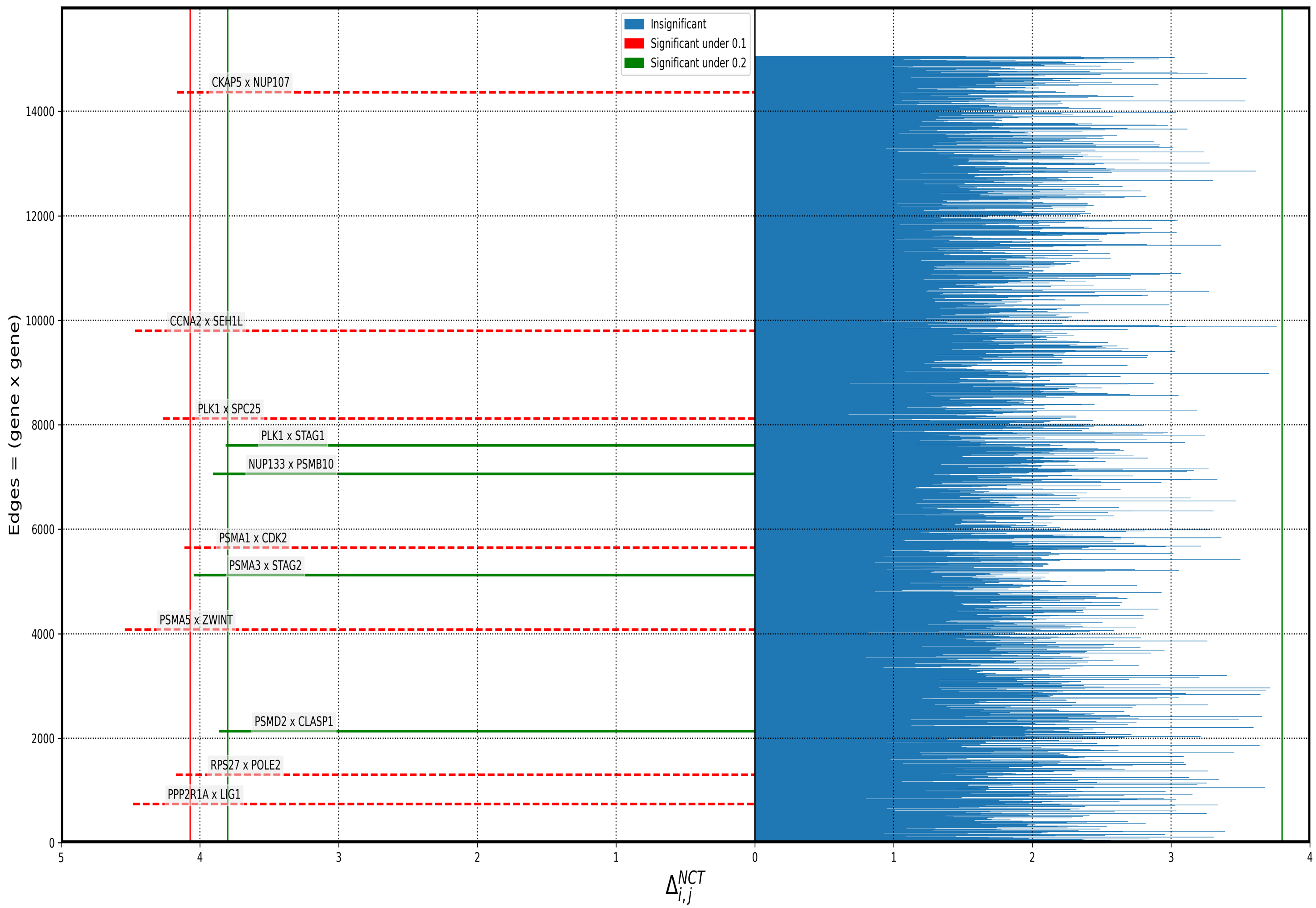


Figure 11: $|\Delta_{i,j}^{NCT}| = \dfrac{|\hat{\omega}_{i,j}^{1} - \hat{\omega}_{i,j}^{2}|}{(\hat{\theta}_{i,j}^{1,NCT} + \hat{\theta}_{i,j}^{2,NCT})^{1/2}}$ for breast cancer genetic pathway with the testing pathway: 1206, where $\hat{\theta}_{i,j}^{d,NCT} = \hat{\theta}_{i,j}^{d}/(\hat{r}_{i,i}^{d}\hat{r}_{j,j}^{d})$

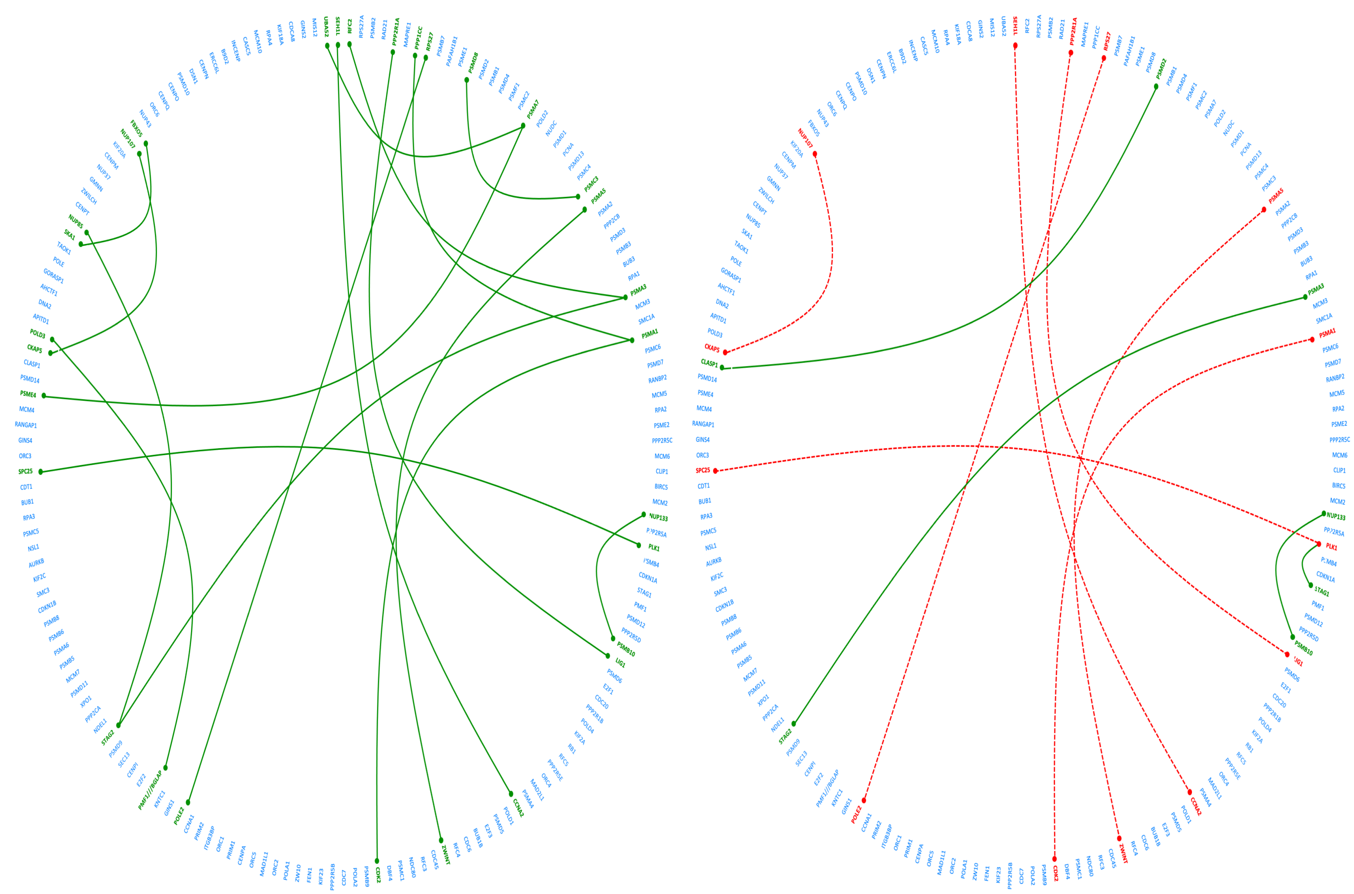


Figure 12: $\Delta_{i,j}$ for the breast cancer genetic pathway with the testing pathway: 1206, the conditional pathway: 750

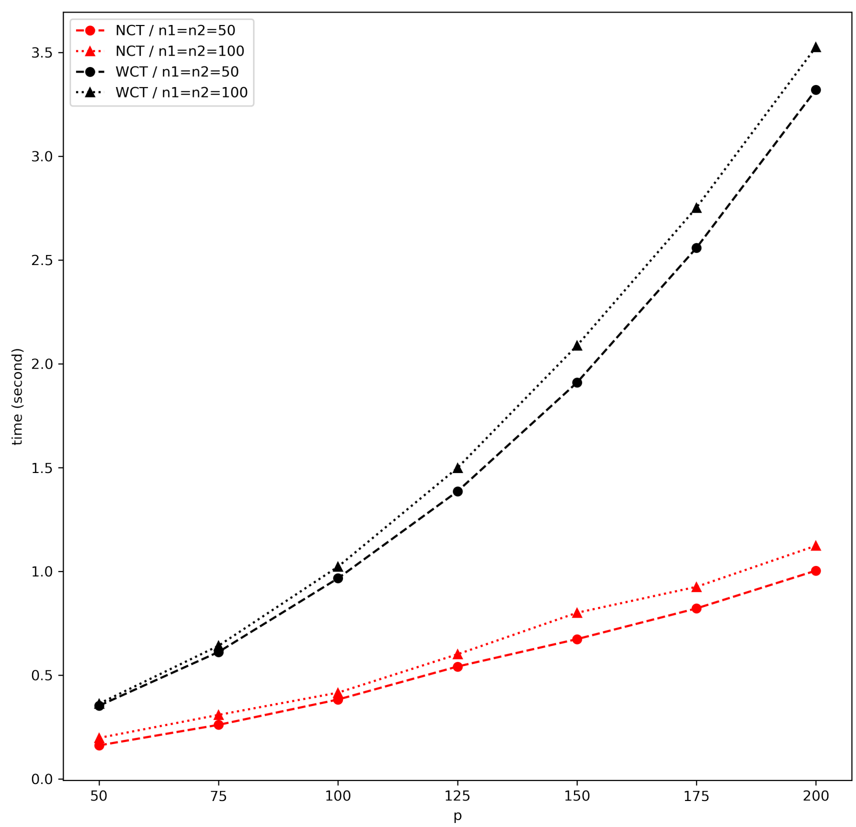


Figure 13: Computational time (second) for global test: the weighted conditional test (WCT) and nonconditional test (NCT) with the number of nodes of the test network $p_1 = p$, the number of nodes of the conditioned network $p_2 = p$, samples sizes $n_1 = n_2 = 50, 100$; black-dashed line with circle=WCT when $n_1 = n_2 = 50$; black-dotted line with triangle=WCT when $n_1 = n_2 = 100$; red-dashed line with circle=NCT when $n_1 = n_2 = 50$; red-dotted line with triangle=NCT when $n_1 = n_2 = 100$

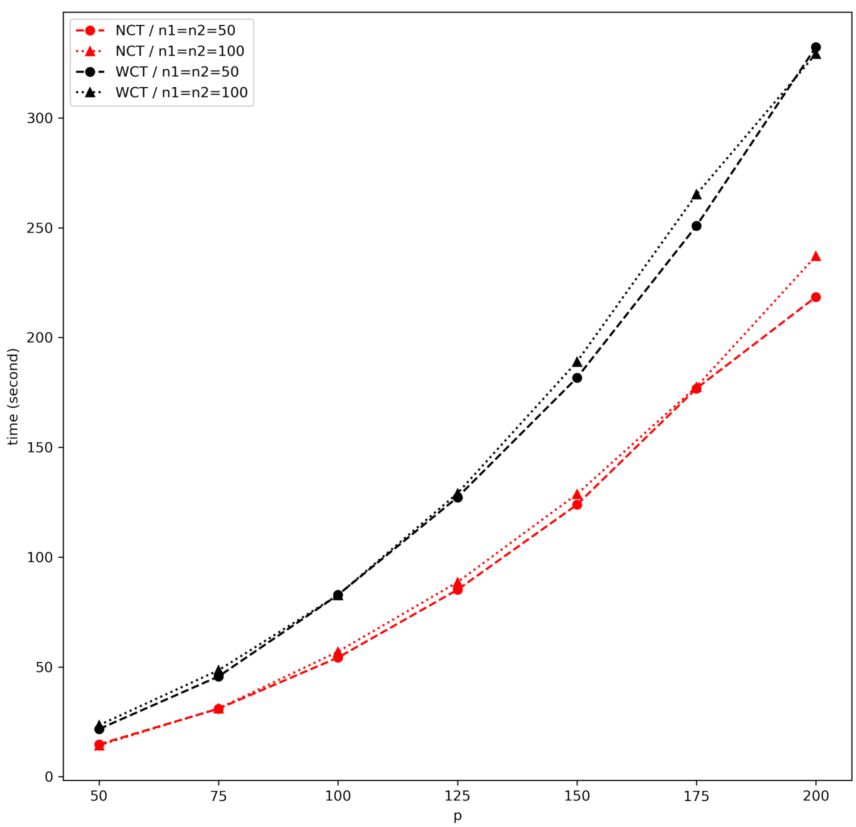


Figure 14: Computational time (second) for multiple test: the weighted conditional test (WCT) and nonconditional test (NCT) with the number of nodes of the test network $p_1 = p$, the number of nodes of the conditioned network $p_2 = p$, samples sizes $n_1 = n_2 = 50, 100$; black-dashed line with circle=WCT when $n_1 = n_2 = 50$; black-dotted line with triangle=WCT when $n_1 = n_2 = 100$; red-dashed line with circle=NCT when $n_1 = n_2 = 50$; red-dotted line with triangle=NCT when $n_1 = n_2 = 100$

Algorithm 1 : Precision Matrix Estimation

Require: $p_1, p_2, n_1, n_2 \geq 0$, $n_1 \times (p_1 + p_2)$ matrix $\mathbf{X}^1$, $n_2 \times (p_1 + p_2)$ matrix $\mathbf{X}^2$, and $0 < \kappa$

Ensure: $(\Delta_{i,j})$: Precision Matrix Estimator

$p \leftarrow p_1 + p_2$

$n \leftarrow \max(n_1, n_2)$

for $d = 1, 2$ **do**

$\bar{\mathbf{X}}^d \leftarrow \frac{1}{n_d} \sum_{k=1}^{n_d} \mathbf{X}^d_{k,\cdot}$

$\hat{\mathbf{\Sigma}}^d = (\hat{\sigma}^d_{i,j})_{1 \leq i,j \leq p_1} \leftarrow \frac{1}{n_d} \sum_{k=1}^{n_d} (\mathbf{X}^d_{k,\cdot} - \bar{\mathbf{X}}^d)(\mathbf{X}^d_{k,\cdot} - \bar{\mathbf{X}}^d)^T$

for $i = 1, \cdots, p_1$ **do**

$\mathbf{D} \leftarrow \mathrm{diag}\left(\hat{\mathbf{\Sigma}}^d_{-i,-i}\right)$

$\lambda^d_{n_d,i} \leftarrow 2\sqrt{(\hat{\sigma}^d_{i,i} \log p)/n_d}$

$\hat{\boldsymbol{\beta}}^d_i \leftarrow \mathbf{D}^{-1/2} \arg\min_{\mathbf{u} \in \mathbb{R}^{p_1+p_2-1}} \left\{ (2n_d)^{-1} \left| \left(\mathbf{X}^d_{\cdot,-i} - \bar{\mathbf{X}}^d_{\cdot,-i}\right) \mathbf{D}^{-1/2}\mathbf{u} - \left(\mathbf{X}^d_{\cdot,i} - \bar{\mathbf{X}}^d_{\cdot,i}\right) \right|_2^2 + \lambda^d_{n_d,i} |\mathbf{u}|_1 \right\}$

end for

for $i = 1, \cdots, p_1$ **do**

$\hat{\varepsilon}^d_{\cdot,i} \leftarrow \mathbf{X}^d_{\cdot,i} - \bar{\mathbf{X}}^d_i - (\mathbf{X}^d_{\cdot,-i} - \bar{\mathbf{X}}^d_{-i}) \cdot \hat{\boldsymbol{\beta}}^d_i$

end for

$\tilde{r}^d \leftarrow \frac{1}{n_d} \sum_{k=1}^{n_d} \hat{\varepsilon}^d_{k,\cdot} (\hat{\varepsilon}^d_{k,\cdot})^T$

for $i = 1, \cdots, p_1$ **do**

for $j = i+1, \cdots, p_1$ **do**

$\hat{r}^d_{i,j} \leftarrow -\tilde{r}^d_{i,j} - \tilde{r}^d_{i,i} \hat{\beta}^d_{i,j} - \tilde{r}^d_{j,j} \hat{\beta}^d_{j-1,i}$

$\hat{\omega}^{d,w}_{i,j} \leftarrow \frac{\hat{r}^d_{i,j}}{\sqrt{\tilde{r}^d_{i,i} \tilde{r}^d_{j,j}}}$

$\hat{\theta}^d_{i,j} \leftarrow \frac{1}{n_d} \left(1 + (\hat{\beta}^d_{i,j})^2 \frac{\hat{r}^d_{i,i}}{\hat{r}^d_{j,j}} \right)$

end for

end for

end for

for $i = 1, \cdots, p_1$ **do**

for $j = i + 1, \cdots, p_1$ **do**

$\Delta_{i,j} \leftarrow \frac{\left|\hat{\omega}_{i,j}^{1,w} - \hat{\omega}_{i,j}^{2,w}\right|}{\sqrt{\hat{\theta}_{i,j}^{1} + \hat{\theta}_{i,j}^{2}}}$

end for

end for

Algorithm 2 : The Global Test

Require: $p_1, p_2, n_1, n_2 \geq 0$, $(p_1 + p_2) \times n_1$ matrix $\mathbf{X}^1$, $(p_1 + p_2) \times n_2$ matrix $\mathbf{X}^2$, and $0 < \alpha < 1$

Ensure: Ψ_α: Whether the null hypothesis would be rejected in the global test.

$\mathbf{\Delta} \leftarrow \text{PrecisionMatrixEstimate}(p_1, p_2, n_1, n_2, \mathbf{X}^1, \mathbf{X}^2, 2)$

$M_n \leftarrow \max_{1 \leq i < j \leq p_1} (Z_{i,j})^2$

$\Phi_\alpha \leftarrow I\{M_n \geq -\log(8\pi) - 2\log|\log(1-\alpha)| + 4\log p_1 - \log\log p_1\}$

Algorithm 3 : The Multiple Test

Require: $p_1, p_2, n_1, n_2 \geq 0$, $(p_1 + p_2) \times n_1$ matrix $\mathbf{X}^1$, $(p_1 + p_2) \times n_2$ matrix $\mathbf{X}^2$, and $0 < \alpha < 1$

Ensure: N_α: Which null hypotheses would be rejected in the multiple test.

for $s = 1, \cdots, 40$ **do**

$\boldsymbol{\Delta}^s \leftarrow \text{PrecisionMatrixEstimate}(p_1, p_2, n_1, n_2, \mathbf{X}^1, \mathbf{X}^2, s/20)$

$E_s \leftarrow \sum_{l=1}^{10} \left(\frac{\sum_{1 \leq i < j \leq p_1} I\left\{\Delta^s_{i,j} \geq \Phi^{-1}(1 - l[1-\Phi\{(\log p_1)^{1/2}\}]/10)\right\}}{l p_1 (p_1 - 1)[1 - \Phi\{(\log p_1)^{1/2}\}]/10} \right)^2$ ▷ (Φ is the cumulate distribution of standard normal.)

end for

$\hat{\kappa} \leftarrow \arg\min_{s=1,2,\cdots,40} E_s$

for $T = 1, 2, \cdots, 200(\log p_1)^{1/2}$ **do**

$t \leftarrow T/100$

$N \leftarrow \left|\left\{(i,j) : 1 \leq i < j \leq p_1 \text{ and } Z^{\hat{\kappa}}_{i,j} \geq t\right\}\right|$

if $\frac{(1-\Phi(t))(p_1^2 - p_1)}{\max(N,1)} \leq \alpha$ **then**

Break

end if

end for

$N_\alpha \leftarrow \left\{(i,j) : 1 \leq i < j \leq p_1 \text{ and } Z^{\hat{\kappa}}_{i,j} \geq t\right\}$