Supplementary Material for
Nonparametric two-sample tests of high dimensional mean vectors via random integration

1. PROOFS OF MAIN RESULTS

Proof of Theorem 1. Let $(\mathbf{X}_1, \mathbf{Y}_1)$ be independent copies of $(\mathbf{X}, \mathbf{Y})$.

$$
\begin{aligned}
\mathrm{RID}_w(\mathbf{X}, \mathbf{Y}) &= \int E^2\left[\boldsymbol{\delta}^\top(\mathbf{X}-\mathbf{Y})\right] w(\boldsymbol{\delta}) d\boldsymbol{\delta} \\
&= \int E\left[\boldsymbol{\delta}^\top(\mathbf{X}-\mathbf{Y})\right] E\left[\boldsymbol{\delta}^\top(\mathbf{X}_1-\mathbf{Y}_1)\right] w(\boldsymbol{\delta}) d\boldsymbol{\delta} \\
&= \int E\left\{\left[\boldsymbol{\delta}^\top(\mathbf{X}-\mathbf{Y})\right]\left[\boldsymbol{\delta}^\top(\mathbf{X}_1-\mathbf{Y}_1)\right]\right\} w(\boldsymbol{\delta}) d\boldsymbol{\delta}.
\end{aligned}
$$

$\mathrm{RID}_w(\mathbf{X}, \mathbf{Y})$ may be evaluated easily for certain properly chosen w. Next, we assume that $w(\boldsymbol{\delta}) = w(\delta_1)\times\cdots\times w(\delta_p)$, and $w(\delta_i)$ is a density function with the mean α_i and the variance β_i^2 for $i = 1, \cdots, p$. Then, by Fubini's theorem, we have

$$
\begin{aligned}
\mathrm{RID}_{\boldsymbol{\theta}}(\mathbf{X}, \mathbf{Y}) &= \int E\left\{\left[\boldsymbol{\delta}^\top(\mathbf{X}-\mathbf{Y})\right]\left[\boldsymbol{\delta}^\top(\mathbf{X}_1-\mathbf{Y}_1)\right]\right\} w(\boldsymbol{\delta}) d\boldsymbol{\delta} \\
&= E\left\{\int \left[\boldsymbol{\delta}^\top(\mathbf{X}-\mathbf{Y})\right]\left[\boldsymbol{\delta}^\top(\mathbf{X}_1-\mathbf{Y}_1)\right] w(\boldsymbol{\delta}) d\boldsymbol{\delta}\right\} \\
&= \sum_{i=1}^{p}(\mu_{1i}-\mu_{2i})^2(\beta_i^2+\alpha_i^2) + \sum_{i\neq j}\alpha_i\alpha_j(\mu_{1i}-\mu_{2i})(\mu_{1j}-\mu_{2j}) \\
&= (\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^\top B(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2) + \left[(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^\top\mathbf{a}\right]^2.
\end{aligned}
$$

where $\boldsymbol{\theta} = (\alpha_1, \cdots, \alpha_p, \beta_1, \cdots, \beta_p)^\top$, $\mathbf{a} = (\alpha_1, \alpha_2, \cdots, \alpha_p)^\top$, and

$$
B = \begin{pmatrix} \beta_1^2 & 0 & \cdots & 0 \\ 0 & \beta_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_p^2 \end{pmatrix}.
$$

This completes the proof of Theorem 1. □

Proof of Theorem 2. Recall that

$$
\mathrm{RID}_{\boldsymbol{\theta},m,n} = \mathrm{RID}^1_{\boldsymbol{\theta},m} + \mathrm{RID}^2_{\boldsymbol{\theta},n} - 2\mathrm{RID}^3_{\boldsymbol{\theta},m,n},
$$

where

$$\begin{aligned}
\mathrm{RID}^1_{\boldsymbol{\theta},m} &= \frac{1}{C_m^2}\sum_{1\le i<j\le m}\mathbf{X}_i^\top W_{\boldsymbol{\theta}}\mathbf{X}_j,\\
\mathrm{RID}^2_{\boldsymbol{\theta},n} &= \frac{1}{C_n^2}\sum_{1\le i<j\le n}\mathbf{Y}_i^\top W_{\boldsymbol{\theta}}\mathbf{Y}_j,\\
\mathrm{RID}^3_{\boldsymbol{\theta},m,n} &= \frac{1}{C_m^1 C_n^1}\sum_{i=1}^m\sum_{j=1}^n\mathbf{X}_i^\top W_{\boldsymbol{\theta}}\mathbf{Y}_j.
\end{aligned}$$

It is very easy to show that

$$\begin{aligned}
E(\mathrm{RID}_{\boldsymbol{\theta},m,n}) &= E(\mathrm{RID}^1_{\boldsymbol{\theta},m}) + E(\mathrm{RID}^2_{\boldsymbol{\theta},n}) - 2E(\mathrm{RID}^3_{\boldsymbol{\theta},m,n})\\
&= \boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2^\top W_{\boldsymbol{\theta}}\boldsymbol{\mu}_2 - 2\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}\boldsymbol{\mu}_2\\
&= \mathrm{RID}_{\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y}),
\end{aligned}$$

and

$$\begin{aligned}
Var(\mathrm{RID}_{\boldsymbol{\theta},m,n}) &= Var(\mathrm{RID}^1_{\boldsymbol{\theta},m}) + Var(\mathrm{RID}^2_{\boldsymbol{\theta},n}) + 4Var(\mathrm{RID}^3_{\boldsymbol{\theta},m,n})\\
&\quad -4Cov(\mathrm{RID}^1_{\boldsymbol{\theta},m},\mathrm{RID}^3_{\boldsymbol{\theta},m,n}) - 4Cov(\mathrm{RID}^2_{\boldsymbol{\theta},n},\mathrm{RID}^3_{\boldsymbol{\theta},m,n}).
\end{aligned}$$

Since

$$\begin{aligned}
Var(\mathrm{RID}^1_{\boldsymbol{\theta},m}) &= \frac{1}{(C_m^2)^2}\Big[C_m^2\left\{tr\{(W_{\boldsymbol{\theta}}\Sigma_1)^2\} + 2\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\boldsymbol{\mu}_1\right\}\\
&\quad +2C_m^2(m-2)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\boldsymbol{\mu}_1\Big]\\
&= \frac{tr\{(W_{\boldsymbol{\theta}}\Sigma_1)^2\}}{C_m^2} + \frac{4\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\boldsymbol{\mu}_1}{m}. \qquad \text{(A.1)}
\end{aligned}$$

Similarly, we have

$$Var(\mathrm{RID}^2_{\boldsymbol{\theta},n}) = \frac{tr\{(W_{\boldsymbol{\theta}}\Sigma_2)^2\}}{C_n^2} + \frac{4\boldsymbol{\mu}_2^\top W_{\boldsymbol{\theta}}\Sigma_2 W_{\boldsymbol{\theta}}\boldsymbol{\mu}_2}{n}, \qquad \text{(A.2)}$$

$$\begin{aligned}
Var(\mathrm{RID}^3_{\boldsymbol{\theta},m,n}) &= \frac{tr(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2)}{C_m^1 C_n^1} + \frac{\boldsymbol{\mu}_2^\top W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\boldsymbol{\mu}_2}{m}\\
&\quad + \frac{\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}\Sigma_2 W_{\boldsymbol{\theta}}\boldsymbol{\mu}_1}{n}, \qquad \text{(A.3)}
\end{aligned}$$

$$Cov(\mathrm{RID}^1_{\boldsymbol{\theta},m}, \mathrm{RID}^3_{\boldsymbol{\theta},m,n}) = \frac{2\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\boldsymbol{\mu}_2}{m}, \tag{A.4}$$

and

$$Cov(\mathrm{RID}^2_{\boldsymbol{\theta},n}, \mathrm{RID}^3_{\boldsymbol{\theta},m,n}) = \frac{2\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}\Sigma_2 W_{\boldsymbol{\theta}}\boldsymbol{\mu}_2}{n}. \tag{A.5}$$

By (A.1)-(A.5), we have

$$\begin{aligned} Var(\mathrm{RID}_{\boldsymbol{\theta},m,n}) &= \frac{tr\{(W_{\boldsymbol{\theta}}\Sigma_1)^2\}}{C_m^2} + \frac{tr\{(W_{\boldsymbol{\theta}}\Sigma_2)^2\}}{C_n^2} + \frac{4tr(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2)}{C_m^1 C_n^1} \\ &\quad + \frac{4(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^\top W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{m} + \frac{4(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^\top W_{\boldsymbol{\theta}}\Sigma_2 W_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{n}. \end{aligned}$$

Denote

$$\sigma^2_{m,n} = \frac{tr\{(W_{\boldsymbol{\theta}}\Sigma_1)^2\}}{C_m^2} + \frac{tr\{(W_{\boldsymbol{\theta}}\Sigma_2)^2\}}{C_n^2} + \frac{4tr(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2)}{C_m^1 C_n^1}.$$

Therefore, under $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, we have $Var(\mathrm{RID}_{\boldsymbol{\theta},m,n}) = \sigma^2_{m,n}$. By condition E3, under $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, we have $Var(\mathrm{RID}_{\boldsymbol{\theta},m,n}) = \sigma^2_{m,n}(1 + o(1))$.

Since

$$\begin{aligned} \mathrm{RID}^1_{\boldsymbol{\theta},m} &= \frac{1}{C_m^2}\sum_{1\le i<j\le m}(\mathbf{X}_i - \boldsymbol{\mu}_1 + \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}(\mathbf{X}_j - \boldsymbol{\mu}_1 + \boldsymbol{\mu}_1), \\ \mathrm{RID}^2_{\boldsymbol{\theta},n} &= \frac{1}{C_n^2}\sum_{1\le i<j\le n}(\mathbf{Y}_i - \boldsymbol{\mu}_2 + \boldsymbol{\mu}_2)^\top W_{\boldsymbol{\theta}}(\mathbf{Y}_j - \boldsymbol{\mu}_2 + \boldsymbol{\mu}_2), \\ \mathrm{RID}^3_{\boldsymbol{\theta},m,n} &= \frac{1}{C_m^1 C_n^1}\sum_{i=1}^m\sum_{j=1}^n(\mathbf{X}_i - \boldsymbol{\mu}_1 + \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}(\mathbf{Y}_j - \boldsymbol{\mu}_2 + \boldsymbol{\mu}_2), \end{aligned}$$

then, we have

$$\mathrm{RID}_{\boldsymbol{\theta},m,n} = \widehat{\mathrm{RID}}_{\boldsymbol{\theta},m,n} + \widetilde{\mathrm{RID}}_{\boldsymbol{\theta},m,n},$$

where

$$\begin{aligned} \widehat{\mathrm{RID}}_{\boldsymbol{\theta},m,n} &= \frac{1}{C_m^2}\sum_{1\le i<j\le m}(\mathbf{X}_i - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}(\mathbf{X}_j - \boldsymbol{\mu}_1) \\ &\quad + \frac{1}{C_n^2}\sum_{1\le i<j\le n}(\mathbf{Y}_i - \boldsymbol{\mu}_2)^\top W_{\boldsymbol{\theta}}(\mathbf{Y}_j - \boldsymbol{\mu}_2) \\ &\quad - 2\frac{1}{C_m^1 C_n^1}\sum_{i=1}^m\sum_{j=1}^n(\mathbf{X}_i - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}(\mathbf{Y}_j - \boldsymbol{\mu}_2), \end{aligned}$$

and

$$\begin{aligned}\widetilde{\mathrm{RID}}_{\boldsymbol{\theta},m,n} &= \frac{2}{m}\sum_{i=1}^{m}(\mathbf{X}_i-\boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2) \\ &\quad +\frac{2}{n}\sum_{j=1}^{n}(\mathbf{Y}_j-\boldsymbol{\mu}_2)^\top W_{\boldsymbol{\theta}}(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)+\mathrm{RID}_{\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y}).\end{aligned}$$

It is straightforward to show that

$$\begin{aligned} E(\widehat{\mathrm{RID}}_{\boldsymbol{\theta},m,n}) &= 0, \quad E(\widetilde{\mathrm{RID}}_{\boldsymbol{\theta},m,n}) = \mathrm{RID}_{\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y}), \\ Var(\widetilde{\mathrm{RID}}_{\boldsymbol{\theta},m,n}) &= \frac{4(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^\top W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{m}+\frac{4(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^\top W_{\boldsymbol{\theta}}\Sigma_2 W_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{n}.\end{aligned}$$

Under Condition E3, we have

$$Var\left(\frac{\widetilde{\mathrm{RID}}_{\boldsymbol{\theta},m,n}-\mathrm{RID}_{\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{\sigma_{m,n}}\right)=o(1).$$

Therefore, we can obtain

$$\begin{aligned}\frac{\mathrm{RID}_{\boldsymbol{\theta},m,n}-\mathrm{RID}_{\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{\sigma_{m,n}} &= \frac{\widehat{\mathrm{RID}}_{\boldsymbol{\theta},m,n}}{\sigma_{m,n}}+\frac{\widetilde{\mathrm{RID}}_{\boldsymbol{\theta},m,n}-\mathrm{RID}_{\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{\sigma_{m,n}} \\ &= \frac{\widehat{\mathrm{RID}}_{\boldsymbol{\theta},m,n}}{\sigma_{m,n}}+o_p(1).\end{aligned}$$

Thus, in order to show Theorem 2, it is sufficient to show that

$$\frac{\widehat{\mathrm{RID}}_{\boldsymbol{\theta},m,n}}{\sigma_{m,n}}\xrightarrow{\mathscr{D}}\mathcal{N}(0,1).$$

To obtain the asymptotic normality of $\widehat{\mathrm{RID}}_{\boldsymbol{\theta},m,n}$, we can assume without loss of generality that $\boldsymbol{\mu}_1=\boldsymbol{\mu}_2=0$. Let $\mathbf{E}_i=W_{\boldsymbol{\theta}}^{1/2}\mathbf{X}_i$ for $i=1,\cdots,m$, and $\mathbf{E}_{j+m}=W_{\boldsymbol{\theta}}^{1/2}\mathbf{Y}_j$ for $j=1,\cdots,n$, and for $i\neq j$

$$\rho_{ij}=\begin{cases}(C_m^2)^{-1}\mathbf{E}_i^\top\mathbf{E}_j, & \text{if } i,j\in\{1,2,\cdots,m\},\\ -2(mn)^{-1}\mathbf{E}_i^\top\mathbf{E}_j, & \text{if } i\in\{1,2,\cdots,m\} \text{ and } j\in\{m+1,2,\cdots,m+n\}.\\ (C_n^2)^{-1}\mathbf{E}_i^\top\mathbf{E}_j, & \text{if } i,j\in\{m+1,2,\cdots,m+n\}.\end{cases}$$

Denote $R_j = \sum_{i=1}^{j-1} \rho_{ij}$ for $j = 2, \cdots, m+n$, and $S_k = \sum_{j=2}^{k} R_j$, and $\mathscr{F}_k = \sigma\{\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_k\}$ which is the σ-algebra generated by $\{\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_k\}$. Then, we have

$$\widehat{\mathrm{RID}}_{\boldsymbol{\theta},m,n} = \sum_{j=2}^{m+n} R_j.$$

Note

$$\begin{aligned} E(R_j^2|\mathscr{F}_{j-1}) &= E\left(\left(\sum_{i=1}^{j-1} \rho_{ij}\right)^2 \Big| \mathscr{F}_{j-1}\right) = E\left(\sum_{i_1,i_2=1}^{j-1} \rho_{i_1 j}\rho_{i_2 j}\Big|\mathscr{F}_{j-1}\right) \\ &= \sum_{i_1,i_2=1}^{j-1} E\left(\rho_{i_1 j}\rho_{i_2 j}\Big|\mathscr{F}_{j-1}\right) \\ &= \begin{cases} (C_m^2)^{-2} \sum_{i_1,i_2=1}^{j-1} \mathbf{E}_{i_i}^\top W_{\boldsymbol{\theta}}^{1/2} \Sigma_1 W_{\boldsymbol{\theta}}^{1/2} \mathbf{E}_{i_2}, & \text{if } j \leq m, \\ 4(mn)^{-2} \sum_{i_1,i_2=1}^{m} \mathbf{E}_{i_i}^\top W_{\boldsymbol{\theta}}^{1/2} \Sigma_2 W_{\boldsymbol{\theta}}^{1/2} \mathbf{E}_{i_2}, & \text{if } j = m+1. \\ 4(mn)^{-2} \sum_{i_1,i_2=1}^{m} \mathbf{E}_{i_i}^\top W_{\boldsymbol{\theta}}^{1/2} \Sigma_2 W_{\boldsymbol{\theta}}^{1/2} \mathbf{E}_{i_2} & \text{if } j \geq m+2 \\ +(C_n^2)^{-2} \sum_{i_1,i_2=m+1}^{j-1} W_{\boldsymbol{\theta}}^{1/2} \Sigma_2 W_{\boldsymbol{\theta}}^{1/2} \mathbf{E}_{i_2}. & \end{cases} \end{aligned}$$

Denote

$$\varsigma_{m,n} = \sum_{j=2}^{m+n} E(R_j^2|\mathscr{F}_{j-1}).$$

Then, we have

$$\begin{aligned} E(\varsigma_{m,n}) &= \frac{1}{(C_m^2)^2} \sum_{j=2}^{m} \sum_{i_1,i_2=1}^{j-1} E(\mathbf{E}_{i_i}^\top W_{\boldsymbol{\theta}}^{1/2} \Sigma_1 W_{\boldsymbol{\theta}}^{1/2} \mathbf{E}_{i_2}) \\ &\quad + \frac{4}{(mn)^2} \sum_{i_1,i_2=1}^{m} E(\mathbf{E}_{i_i}^\top W_{\boldsymbol{\theta}}^{1/2} \Sigma_2 W_{\boldsymbol{\theta}}^{1/2} \mathbf{E}_{i_2}) \\ &\quad + \sum_{j=m+2}^{m+n} \Big[\frac{4}{(mn)^2} \sum_{i_1,i_2=1}^{m} E(\mathbf{E}_{i_i}^\top W_{\boldsymbol{\theta}}^{1/2} \Sigma_2 W_{\boldsymbol{\theta}}^{1/2} \mathbf{E}_{i_2}) \\ &\quad + \frac{1}{(C_n^2)^2} \sum_{i_1,i_2=m+1}^{j-1} E(W_{\boldsymbol{\theta}}^{1/2} \Sigma_2 W_{\boldsymbol{\theta}}^{1/2} \mathbf{E}_{i_2}) \Big] \\ &= \frac{tr\{(W_{\boldsymbol{\theta}}\Sigma_1)^2\}}{C_m^2} + \frac{tr\{(W_{\boldsymbol{\theta}}\Sigma_2)^2\}}{C_n^2} + \frac{4tr(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2)}{C_m^1 C_n^1} \\ &= \sigma_{m,n}^2. \end{aligned} \tag{A.6}$$

and

$$
\begin{aligned}
E(\varsigma_{m,n}^2) &= E\Big\{\frac{1}{(C_m^2)^2}\sum_{j=2}^{m}\sum_{i_1,i_2=1}^{j-1}\mathbf{E}_{i_i}^{\top}W_{\boldsymbol{\theta}}^{1/2}\Sigma_1 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_{i_2} \\
&\quad +\frac{4}{(mn)^2}\sum_{i_1,i_2=1}^{m}\mathbf{E}_{i_i}^{\top}W_{\boldsymbol{\theta}}^{1/2}\Sigma_2 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_{i_2} \\
&\quad +\sum_{j=m+2}^{m+n}\Big[\frac{4}{(mn)^2}\sum_{i_1,i_2=1}^{m}\mathbf{E}_{i_i}^{\top}W_{\boldsymbol{\theta}}^{1/2}\Sigma_2 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_{i_2} \\
&\quad +\frac{1}{(C_n^2)^2}\sum_{i_1,i_2=m+1}^{j-1}\mathbf{E}_{i_i}^{\top}W_{\boldsymbol{\theta}}^{1/2}\Sigma_2 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_{i_2}\Big]\Big\}^2 \\
&= \sigma_{m,n}^4(1+o(1)).
\end{aligned}
$$

Therefore, we have

$$
Var(\varsigma_{m,n}) = E(\varsigma_{m,n}^2) - (E(\varsigma_{m,n}))^2 = o(\sigma_{m,n}^4). \tag{A.7}
$$

By (A.6) and (A.7), we have

$$
\sigma_{m,n}^{-2}E\left\{\sum_{j=2}^{m+n}E(R_j^2|\mathscr{F}_{j-1})\right\} = \sigma_{m,n}^{-2}E(\varsigma_{m,n}) = 1,
$$

and

$$
\sigma_{m,n}^{-4}Var\left\{\sum_{j=2}^{m+n}E(R_j^2|\mathscr{F}_{j-1})\right\} = \sigma_{m,n}^{-4}Var(\varsigma_{m,n}) = o(1).
$$

Therefore, we have

$$
\frac{\sum_{j=2}^{m+n}E(R_j^2|\mathscr{F}_{j-1})}{\sigma_{m,n}^2} \xrightarrow{\mathscr{P}} 1. \tag{A.8}
$$

Note that

$$
\sum_{j=2}^{m+n}\sigma_{m,n}^{-2}E\left\{R_j^2 I(|R_j| > \epsilon\sigma_{m,n})|\mathscr{F}_{j-1})\right\} \le \sigma_{m,n}^{-4}\epsilon^{-2}\sum_{j=2}^{m+n}E\left(R_j^4|\mathscr{F}_{j-1}\right),
$$

and

$$
E(R_j^4|\mathscr{F}_{j-1}) = E\left(\left(\sum_{i=1}^{j-1}\rho_{ij}\right)^4\middle|\mathscr{F}_{j-1}\right) = E\left(\sum_{i_1,i_2,i_3,i_4=1}^{j-1}\rho_{i_1 j}\rho_{i_2 j}\rho_{i_3 j}\rho_{i_4 j}\middle|\mathscr{F}_{j-1}\right).
$$

Therefore, we have

$$E\left\{\sum_{j=2}^{m+n} E\left(R_j^4|\mathscr{F}_{j-1}\right)\right\} = \sum_{j=2}^{m+n}\sum_{i_1,i_2,i_3,i_4=1}^{j-1} E(\rho_{i_1j}\rho_{i_2j}\rho_{i_3j}\rho_{i_4j})$$
$$= O((m+n)^{-8})\sum_{j=2}^{m+n}\sum_{i_1,i_2,i_3,i_4=1}^{j-1} E(\mathbf{E}_j^\top\mathbf{E}_{i_1}\mathbf{E}_{i_2}^\top\mathbf{E}_j\mathbf{E}_j^\top\mathbf{E}_{i_3}\mathbf{E}_{i_4}^\top\mathbf{E}_j)$$
$$= O((m+n)^{-8})(3D+Q), \tag{A.9}$$

where

$$D = \sum_{j=2}^{m+n}\sum_{i_1\neq i_2}^{j-1} E(\mathbf{E}_j^\top\mathbf{E}_{i_1}\mathbf{E}_{i_1}^\top\mathbf{E}_j\mathbf{E}_j^\top\mathbf{E}_{i_2}\mathbf{E}_{i_2}^\top\mathbf{E}_j)$$

and

$$Q = \sum_{j=2}^{m+n}\sum_{i_1=1}^{j-1} E(\mathbf{E}_j^\top\mathbf{E}_{i_1})^4.$$

By applying Proposition A.1. in Chen et al. (2010), we have

$$\begin{aligned} D &= \sum_{j=2}^{m}\sum_{i_1\neq i_2}^{j-1} E(\mathbf{E}_j^\top W_{\boldsymbol{\theta}}^{1/2}\Sigma_1 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_j\mathbf{E}_j^\top W_{\boldsymbol{\theta}}^{1/2}\Sigma_1 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_j) \\ &\quad + \sum_{i_1\neq i_2}^{m} E(\mathbf{E}_{m+1}^\top W_{\boldsymbol{\theta}}^{1/2}\Sigma_1 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_{m+1}\mathbf{E}_{m+1}^\top W_{\boldsymbol{\theta}}^{1/2}\Sigma_1 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_{m+1}) \\ &\quad + \sum_{j=m+2}^{m+n}\sum_{i_1\neq i_2}^{j-1} E(\mathbf{E}_j^\top W_{\boldsymbol{\theta}}^{1/2}\Sigma_2 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_j\mathbf{E}_j^\top W_{\boldsymbol{\theta}}^{1/2}\Sigma_2 W_{\boldsymbol{\theta}}^{1/2}\mathbf{E}_j) \\ &= o\{(m+n)^8\sigma_{m,n}^4\}, \end{aligned}$$

and

$$\begin{aligned} Q &= \sum_{j=2}^{m+n}\sum_{i_1=1}^{j-1} E(\mathbf{E}_j^\top\mathbf{E}_{i_1})^4 \\ &= \sum_{j=2}^{m}\sum_{i_1=1}^{j-1} E(\mathbf{E}_j^\top\mathbf{E}_{i_1})^4 + \sum_{j=m+1}^{m+n}\sum_{i_1=1}^{m} E(\mathbf{E}_j^\top\mathbf{E}_{i_1})^4 + \sum_{j=m+1}^{m+n}\sum_{i_1=m+1}^{j-1} E(\mathbf{E}_j^\top\mathbf{E}_{i_1})^4 \\ &= Q_1+Q_2+Q_3, \end{aligned}$$

$$
\begin{aligned}
Q_1 &= \sum_{j=2}^{m}\sum_{i_1=1}^{j-1} E(\mathbf{E}_j^\top \mathbf{E}_{i_1})^4, \\
Q_2 &= \sum_{j=m+1}^{m+n}\sum_{i_1=1}^{m} E(\mathbf{E}_j^\top \mathbf{E}_{i_1})^4, \\
Q_3 &= \sum_{j=m+1}^{m+n}\sum_{i_1=m+1}^{j-1} E(\mathbf{E}_j^\top \mathbf{E}_{i_1})^4.
\end{aligned}
$$

Denote $M = \Gamma_2^\top W_{\boldsymbol{\theta}} \Gamma_1 = (M_{ij})_{k_2\times k_1}$. Then, by Condition E1 and E4, we have

$$
\begin{aligned}
E(\mathbf{E}_j^\top \mathbf{E}_{i_1})^4 &= E(\mathbf{Z}_{2j}^\top M \mathbf{Z}_{1i_1})^4, \\
&= \sum_{i=1}^{k_2}\sum_{j=1}^{k_1}(3+\Delta_1)(3+\Delta_2)M_{ij}^4 + \sum_{i=1}^{k_2}\sum_{j_1\neq j_2}^{k_1}(3+\Delta_1)M_{ij_1}^2 M_{ij_2}^2 \\
&\quad + \sum_{i_1\neq i_2}^{k_2}\sum_{j=1}^{k_1}(3+\Delta_2)M_{i_1j}^2 M_{i_2j}^2 + 9\sum_{i_1\neq i_2}^{k_2}\sum_{j_1\neq j_2}^{k_1} M_{i_1j_1}^2 M_{i_1j_2}^2 M_{i_2j_1}^2 M_{i_2j_2}^2 \\
&\leq (3+\Delta_1)(3+\Delta_2)tr^2(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2) + (3+\Delta_1)tr^2(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2) \\
&\quad + (3+\Delta_2)tr^2(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2) + 9tr\{(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2)^2\}.
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
O((m+n)^{-8})Q_2 &= O((m+n)^{-6})[O\{tr^2(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2)\} + O\{tr(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2)^2\}] \\
&= o(\sigma_{m,n}^4).
\end{aligned}
$$

Similarly, we have

$$
O((m+n)^{-8})Q_1 = o(\sigma_{m,n}^4), \quad \text{and} \quad O((m+n)^{-8})Q_3 = o(\sigma_{m,n}^4).
$$

Therefore, we have

$$
\sum_{j=2}^{m+n} \sigma_{m,n}^{-2} E\left\{R_j^2 I(|R_j| > \epsilon\sigma_{m,n})|\mathscr{F}_{j-1})\right\} = o_p(1). \tag{A.10}
$$

Based on (A.9), (A.10) and Corollary 3.1 in Hall and Heyde (1980), we have

$$
\frac{\widehat{\mathrm{RID}}_{\boldsymbol{\theta},m,n}}{\sigma_{m,n}} \xrightarrow{\mathscr{D}} \mathcal{N}(0,1).
$$

This completes the proof of Theorem 2. $\square$

Proof of Theorem 3.

$$
\begin{aligned}
& tr\{\widehat{(W_{\boldsymbol{\theta}}\Sigma_1)^2}\} \\
= \ & \frac{1}{2C_m^2} tr\left\{\sum_{i\neq j} W_{\boldsymbol{\theta}}^{1/2}(\mathbf{X}_i - \bar{\mathbf{X}}_{(i,j)})\mathbf{X}_i^\top W_{\boldsymbol{\theta}}(\mathbf{X}_j - \bar{\mathbf{X}}_{(i,j)})\mathbf{X}_j^\top W_{\boldsymbol{\theta}}^{1/2}\right\}, \\
= \ & \frac{1}{2C_m^2} tr\Big\{\sum_{i\neq j} W_{\boldsymbol{\theta}}^{1/2}(\mathbf{X}_i - \boldsymbol{\mu}_1 + \boldsymbol{\mu}_1 - \bar{\mathbf{X}}_{(i,j)})(\mathbf{X}_i - \boldsymbol{\mu}_1 + \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}} \\
& (\mathbf{X}_j - \boldsymbol{\mu}_1 + \boldsymbol{\mu}_1 - \bar{\mathbf{X}}_{(i,j)})(\mathbf{X}_j - \boldsymbol{\mu}_1 + \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
= \ & C_1 + C_2 + C_3 + C_4 + C_5 + C_6 + C_7 + C_8 + C_9 + C_{10},
\end{aligned}
$$

where

$$
\begin{aligned}
C_1 &= \frac{1}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\mathbf{X}_i - \boldsymbol{\mu}_1)(\mathbf{X}_i - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}(\mathbf{X}_j - \boldsymbol{\mu}_1)(\mathbf{X}_j - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
C_2 &= -\frac{2}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)(\mathbf{X}_i - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}(\mathbf{X}_j - \boldsymbol{\mu}_1)(\mathbf{X}_j - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
C_3 &= \frac{2}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\mathbf{X}_i - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}(\mathbf{X}_j - \boldsymbol{\mu}_1)(\mathbf{X}_j - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
C_4 &= -\frac{2}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}(\mathbf{X}_j - \boldsymbol{\mu}_1)(\mathbf{X}_j - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
C_5 &= \frac{1}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)(\mathbf{X}_i - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)(\mathbf{X}_j - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
C_6 &= -\frac{2}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\mathbf{X}_i - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)(\mathbf{X}_j - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
C_7 &= -\frac{2}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)(\mathbf{X}_j - \boldsymbol{\mu}_1)^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
C_8 &= \frac{1}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\mathbf{X}_i - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}(\mathbf{X}_j - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
C_9 &= -\frac{2}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}(\mathbf{X}_j - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}, \\
C_{10} &= \frac{1}{2C_m^2}\sum_{i\neq j} tr\Big\{W_{\boldsymbol{\theta}}^{1/2}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}(\bar{\mathbf{X}}_{(i,j)} - \boldsymbol{\mu}_1)\boldsymbol{\mu}_1^\top W_{\boldsymbol{\theta}}^{1/2}\Big\}.
\end{aligned}
$$

By Condition E3, it is straightforward to show that

$$\begin{aligned} E(C_1) &= tr\{(W_{\boldsymbol\theta}\Sigma_1)^2\}, \quad E(C_i)=0, \text{for} \quad i=2,\cdots,9, \\ E(C_{10}) &= \boldsymbol\mu_1^\top W_{\boldsymbol\theta}\Sigma_1 W_{\boldsymbol\theta}\boldsymbol\mu_1/(m-2) = o(tr\{(W_{\boldsymbol\theta}\Sigma_1)^2\}). \end{aligned}$$

Therefore, we have

$$E\left(tr\{\widehat{(W_{\boldsymbol\theta}\Sigma_1})^2\}\right) = tr\{(W_{\boldsymbol\theta}\Sigma_1)^2\}(1+o(1)). \tag{A.11}$$

Note that

$$Var\left\{tr\{\widehat{(W_{\boldsymbol\theta}\Sigma_1})^2\}\right\} \le 10\sum_{i=1}^{10} Var(C_i).$$

Since

$$\begin{aligned} EC_1^2 &= \frac{1}{(2C_m^2)^2}E\Big\{\sum_{i_1\neq j_1} tr\Big(W_{\boldsymbol\theta}^{1/2}(\mathbf{X}_{i_1}-\boldsymbol\mu_1)(\mathbf{X}_{i_1}-\boldsymbol\mu_1)^\top W_{\boldsymbol\theta}(\mathbf{X}_{j_1}-\boldsymbol\mu_1)(\mathbf{X}_{j_1}-\boldsymbol\mu_1)^\top W_{\boldsymbol\theta}^{1/2}\Big) \\ &\quad \times \sum_{i_2\neq j_2} tr\Big(W_{\boldsymbol\theta}^{1/2}(\mathbf{X}_{i_2}-\boldsymbol\mu_1)(\mathbf{X}_{i_2}-\boldsymbol\mu_1)^\top W_{\boldsymbol\theta}(\mathbf{X}_{j_2}-\boldsymbol\mu_1)(\mathbf{X}_{j_2}-\boldsymbol\mu_1)^\top W_{\boldsymbol\theta}^{1/2}\Big)\Big\}, \end{aligned}$$

then, we have

$$\begin{aligned} Var(C_1) &= \frac{2E\{[(\mathbf{X}_1-\boldsymbol\mu_1)^\top W_{\boldsymbol\theta}(\mathbf{X}_2-\boldsymbol\mu_1)]^4\}}{m(m-1)} \\ &\quad + \frac{4(m-2)E\{[(\mathbf{X}_1-\boldsymbol\mu_1)^\top W_{\boldsymbol\theta}\Sigma_1 W_{\boldsymbol\theta}(\mathbf{X}_1-\boldsymbol\mu_1)]^2\}}{m(m-1)} + o\left(tr^2\{(W_{\boldsymbol\theta}\Sigma_1)^2\}\right). \end{aligned}$$

Denote $L = \Gamma_1^\top W_{\boldsymbol\theta}\Gamma_1 = (L_{ij})$ and $U = \Gamma_1^\top W_{\boldsymbol\theta}\Sigma_1 W_{\boldsymbol\theta}\Gamma_1$. Then, by using Proposition A.1. in Chen et al. (2010) and Conditions E1, E3, and E4, we have

$$\begin{aligned} E\{[(\mathbf{X}_1-\boldsymbol\mu_1)^\top W_{\boldsymbol\theta}(\mathbf{X}_1-\boldsymbol\mu_1)]^4\} &= E(\mathbf{Z}_{11}^\top L\mathbf{Z}_{12})^4 \\ &= 3tr^2(L^2)+6tr(L^4)+6\Delta_1 tr(L^2\circ L^2)+\Delta_1^2\sum_{i,j=1}^{k_1}L_{ij}^4 \\ &= O\left(tr^2\{(W_{\boldsymbol\theta}\Sigma_1)^2\}\right), \end{aligned} \tag{A.12}$$

and

$$\begin{aligned} E\{[(\mathbf{X}_1-\boldsymbol\mu_1)^\top W_{\boldsymbol\theta}\Sigma_1 W_{\boldsymbol\theta}(\mathbf{X}_1-\boldsymbol\mu_1)]^2\} &= E(\mathbf{Z}_{11}^\top U\mathbf{Z}_{11}\mathbf{Z}_{11}^\top U\mathbf{Z}_{11}) \\ &= tr^2(U)+2tr(U^2)+\Delta_1 tr(U\circ U) \\ &= O\left(tr^2\{(W_{\boldsymbol\theta}\Sigma_1)^2\}\right), \end{aligned} \tag{A.13}$$

where $tr(S_1 \circ S_2) = (S_{1ij} S_{2ij})$ for the matrices S_1 and S_2. By (A.12) and (A.13), we have

$$Var(C_1) \;=\; o\left(tr^2\{(W_{\boldsymbol{\theta}}\Sigma_1)^2\}\right).$$

By carrying out similar procedures, we can obtain

$$Var(C_i) \;=\; o\left(tr^2\{(W_{\boldsymbol{\theta}}\Sigma_1)^2\}\right), \quad \text{for} \quad i = 2, \cdots, 10.$$

Therefore, we have

$$Var\left\{tr\{(\widehat{W_{\boldsymbol{\theta}}\Sigma_1})^2\}\right\} = o\left(tr^2\{(W_{\boldsymbol{\theta}}\Sigma_1)^2\}\right). \tag{A.14}$$

By (A.11) and (A.14), we have

$$tr\{(\widehat{W_{\boldsymbol{\theta}}\Sigma_1})^2\} \xrightarrow{\mathscr{P}} tr\{(W_{\boldsymbol{\theta}}\Sigma_1)^2\}. \tag{A.15}$$

Similarly, we have

$$tr\{(\widehat{W_{\boldsymbol{\theta}}\Sigma_2})^2\} \xrightarrow{\mathscr{P}} tr\{(W_{\boldsymbol{\theta}}\Sigma_2)^2\}, \tag{A.16}$$

and

$$tr(\widehat{W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}}\Sigma_2) \xrightarrow{\mathscr{P}} tr(W_{\boldsymbol{\theta}}\Sigma_1 W_{\boldsymbol{\theta}}\Sigma_2). \tag{A.17}$$

By (A.15), (A.16), and (A.17), we have

$$\frac{\sigma_{m,n}}{\hat{\sigma}_{m,n}} \xrightarrow{\mathscr{P}} 1. \tag{A.18}$$

By Theorem 2 and Slutsky's theorem, we complete the proof of Theorem 3. □

Proof of Theorem 4. By Theorem 2, Theorem 3, (A.18) and Conditions E1-E4, under $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, we have

$$\begin{aligned}
&\lim_{m,n,p\to\infty} P\left(\text{RID}_{\boldsymbol{\theta},m,n} \geq \hat{\sigma}_{m,n} z_\alpha\right) \\
=\;& \lim_{m,n,p\to\infty} P\left(\frac{\text{RID}_{\boldsymbol{\theta},m,n} - \text{RID}_{\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{\sigma_{m,n}} \geq \frac{\hat{\sigma}_{m,n} z_\vartheta - \text{RID}_{\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{\sigma_{m,n}}\right) \\
=\;& \lim_{m,n,p\to\infty} \Phi\left\{-z_\vartheta + \frac{(m+n)\tau(1-\tau)\text{RID}_{\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{\sqrt{2tr(\Sigma_{\boldsymbol{\theta}}^2(\tau))}}\right\} \\
&= \lim_{m,n,p\to\infty} \Phi\left\{-z_\vartheta + (m+n)\tau(1-\tau)\mathscr{P}_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2, \Sigma_1, \Sigma_2)\right\},
\end{aligned}$$

where $\Sigma_{\boldsymbol{\theta}}(\tau) = W_{\boldsymbol{\theta}}\{(1-\tau)\Sigma_1 + \tau\Sigma_2\}$, and

$$\mathscr{P}_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2, \Sigma_1, \Sigma_2) = \frac{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\top} W_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}{\sqrt{2tr(\Sigma_{\boldsymbol{\theta}}^2(\tau))}}.$$

This completes the proof of Theorem 4. □

Proof of Theorem 5. Let $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_p$ and $\lambda_1^* \leq \lambda_2^* \cdots \leq \lambda_p^*$ be eigenvalues of $W_{\boldsymbol{\theta}}$ and $\widetilde{\Sigma}(\tau)$, respectively. According to the definition of $W_{\boldsymbol{\theta}}$, we have $\lambda_1 = \cdots = \lambda_{p-1} = \beta^2$, and $\lambda_p = \beta^2 + p\alpha^2$. By some simple calculations, we have

$$(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\top} W_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \alpha^2 \left(\sum_{i=1}^{p} (\mu_{i1} - \mu_{i2}) \right)^2 + \beta^2 ||\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2||^2.$$

According to Bushell and Trustrum (1990), we have

$$tr(\Sigma_{\boldsymbol{\theta}}^2(\tau)) = tr\left\{ (W_{\boldsymbol{\theta}}\widetilde{\Sigma}(\tau))^2 \right\} \leq \sum_{i=1}^{p} (\lambda_i \lambda_i^*)^2 = \lambda_1^2 tr\{\widetilde{\Sigma}(\tau)^2\} + (\lambda_p^2 - \lambda_1^2)\lambda_p^{*2}.$$

Therefore,

$$\begin{aligned}
ARE(\beta_{RID}, \beta_{CQ}) &= \frac{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\top} W_{\boldsymbol{\theta}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\sqrt{tr\{\widetilde{\Sigma}(\tau)^2\}}}{||\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2||^2 \sqrt{tr(\Sigma_{\boldsymbol{\theta}}^2(\tau))}} \\
&\geq \frac{\left[\alpha^2 \left(\sum_{i=1}^{p} (\mu_{i1} - \mu_{i2})\right)^2 + \beta^2 ||\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2||^2\right] \sqrt{tr\{\widetilde{\Sigma}(\tau)^2\}}}{||\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2||^2 \sqrt{\lambda_1^2 tr\{\widetilde{\Sigma}(\tau)^2\} + (\lambda_p^2 - \lambda_1^2)\lambda_p^{*2}}} \\
&= \frac{1 + \frac{r^2 \left(\sum_{i=1}^{p} (\mu_{i1} - \mu_{i2})\right)^2}{||\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2||^2}}{\sqrt{1 + \frac{2pr^2(\lambda_p^*)^2}{tr\{\widetilde{\Sigma}(\tau)^2\}} + \frac{p^2 r^4 (\lambda_p^*)^2}{tr\{\widetilde{\Sigma}(\tau)^2\}}}}.
\end{aligned}$$

Since $\max\left\{pr^2(\lambda_p^*)^2, p^2 r^4 (\lambda_p^*)^2\right\} = o(tr\{\widetilde{\Sigma}(\tau)^2\})$, then we have

$$\lim_{m,n,p \to \infty} ARE(\beta_{RID}, \beta_{CQ}) \geq 1.$$

This completes the proof of Theorem 5. □

2. SIMULATION RESULTS

Example 1. The empirical powers are reported when Z_{ik} follows the following three distributions:

1. The standardized t-distribution with degrees of freedom 5, i.e., $(5/3)^{-1/2}t(5)$;

2. The standardized chi-squared distribution with degrees of freedom 4, i.e., $8^{-1/2}\{\chi^2(4)-4\}$;

3. The standardized Gamma distribution with $a=4, b=0.5$, i.e., $\Gamma(4,0.5)-2$.

The corresponding results are shown in Figures 1-12.

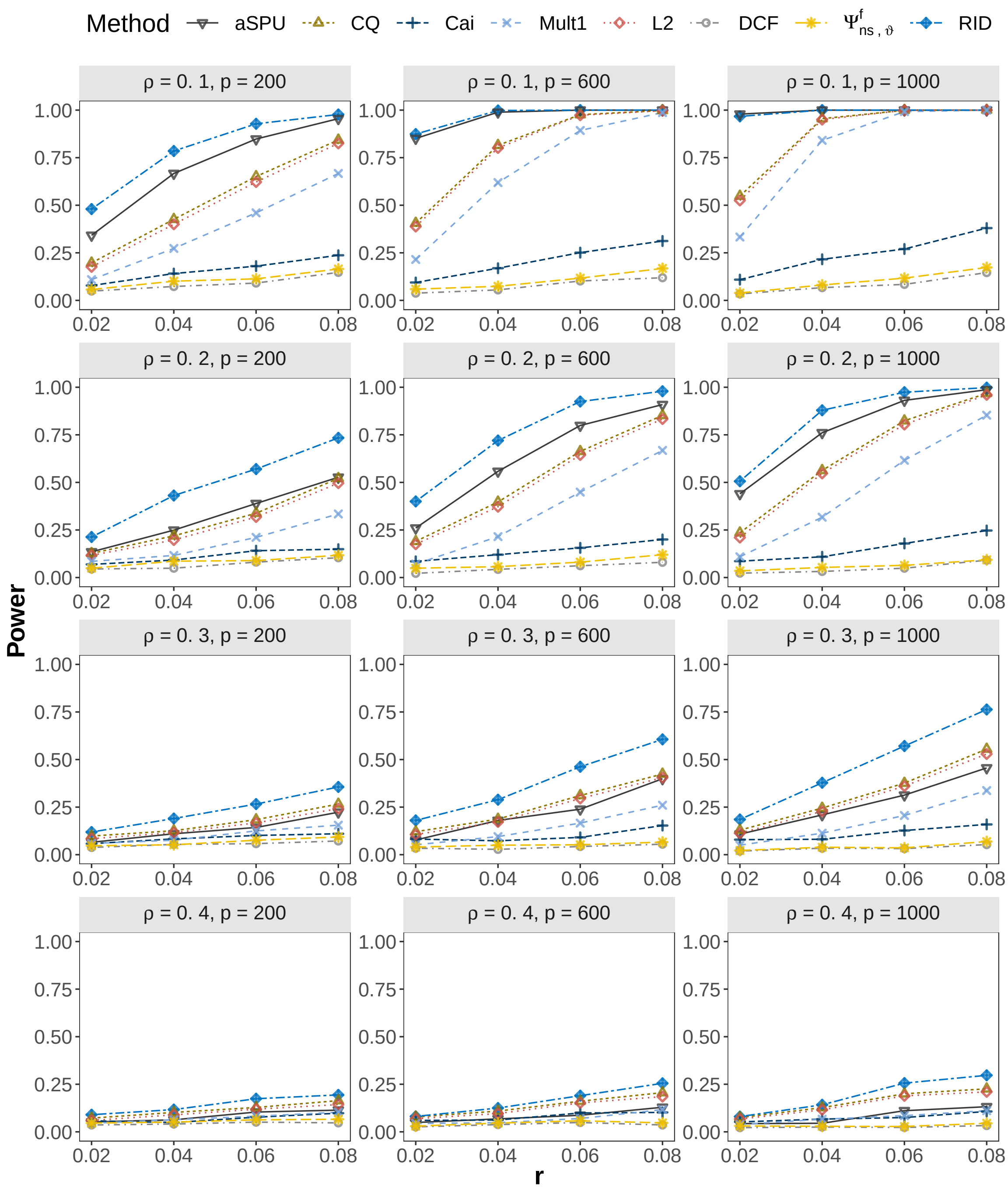

Figure 1: Empirical powers with Z_{ij} follows $(5/3)^{-1/2}t(5)$ and $m = 60, n = 80$ for **Scenario 1** under different signal levels of r and sparsity levels of ρ in **Example 1**.

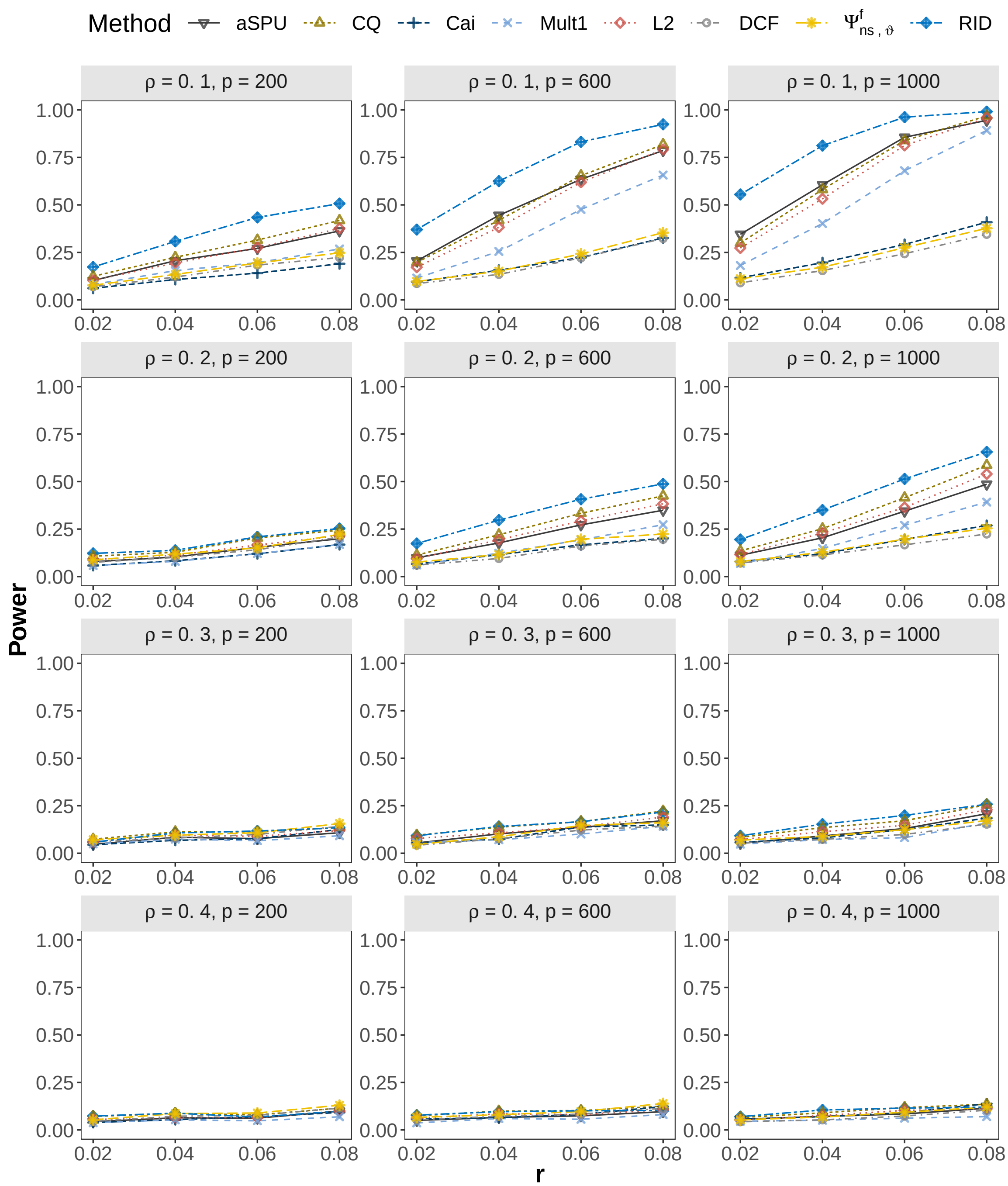

Figure 2: Empirical powers with Z_{ij} follows $(5/3)^{-1/2}t(5)$ and $m = 60, n = 80$ for **Scenario 2** under different signal levels of r and sparsity levels of ρ in **Example 1**.

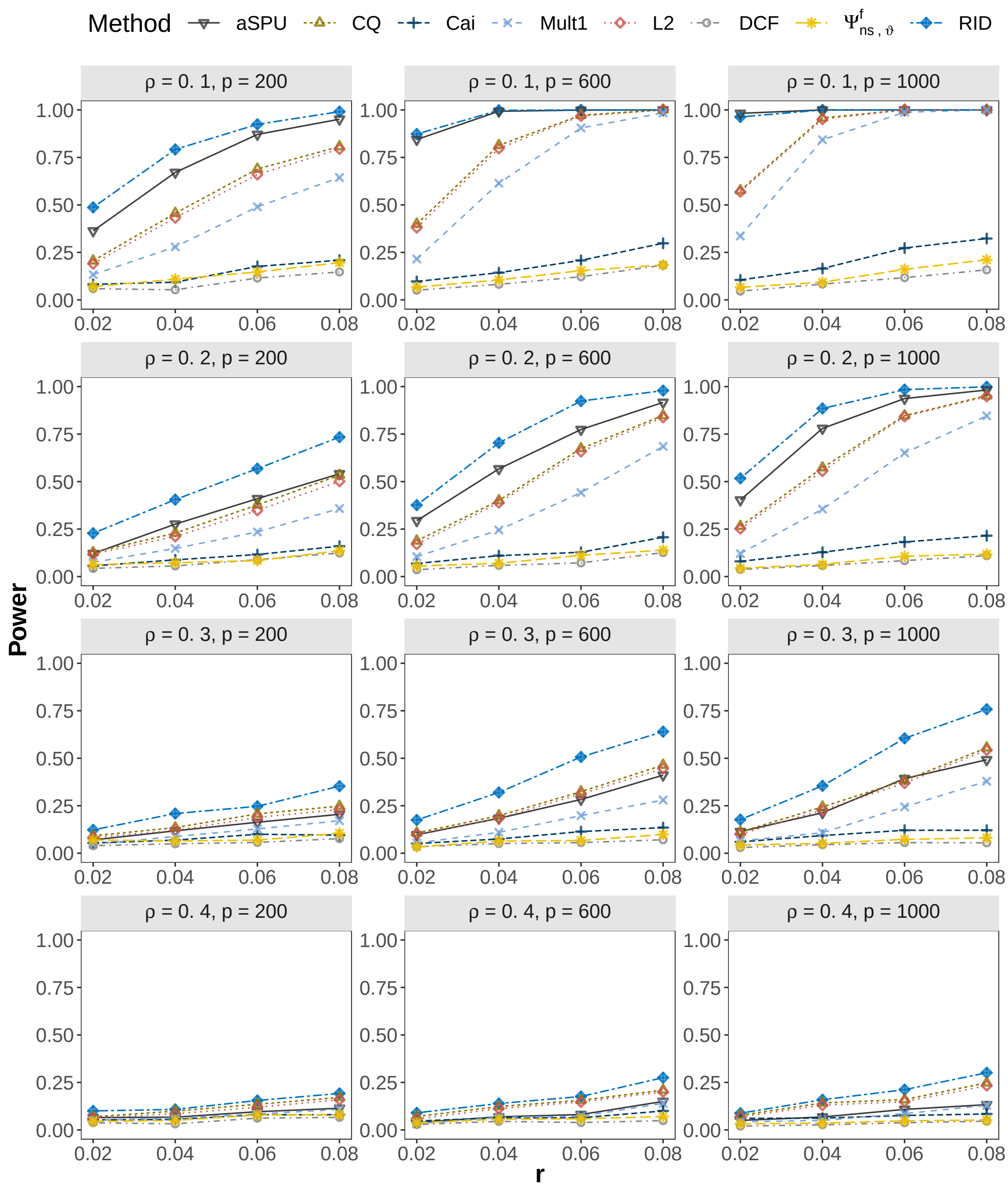

Figure 3: Empirical powers with Z_{ij} follows $(5/3)^{-1/2}t(5)$ and $m = 90, n = 120$ for **Scenario 1** under different signal levels of r and sparsity levels of ρ in **Example 1**.

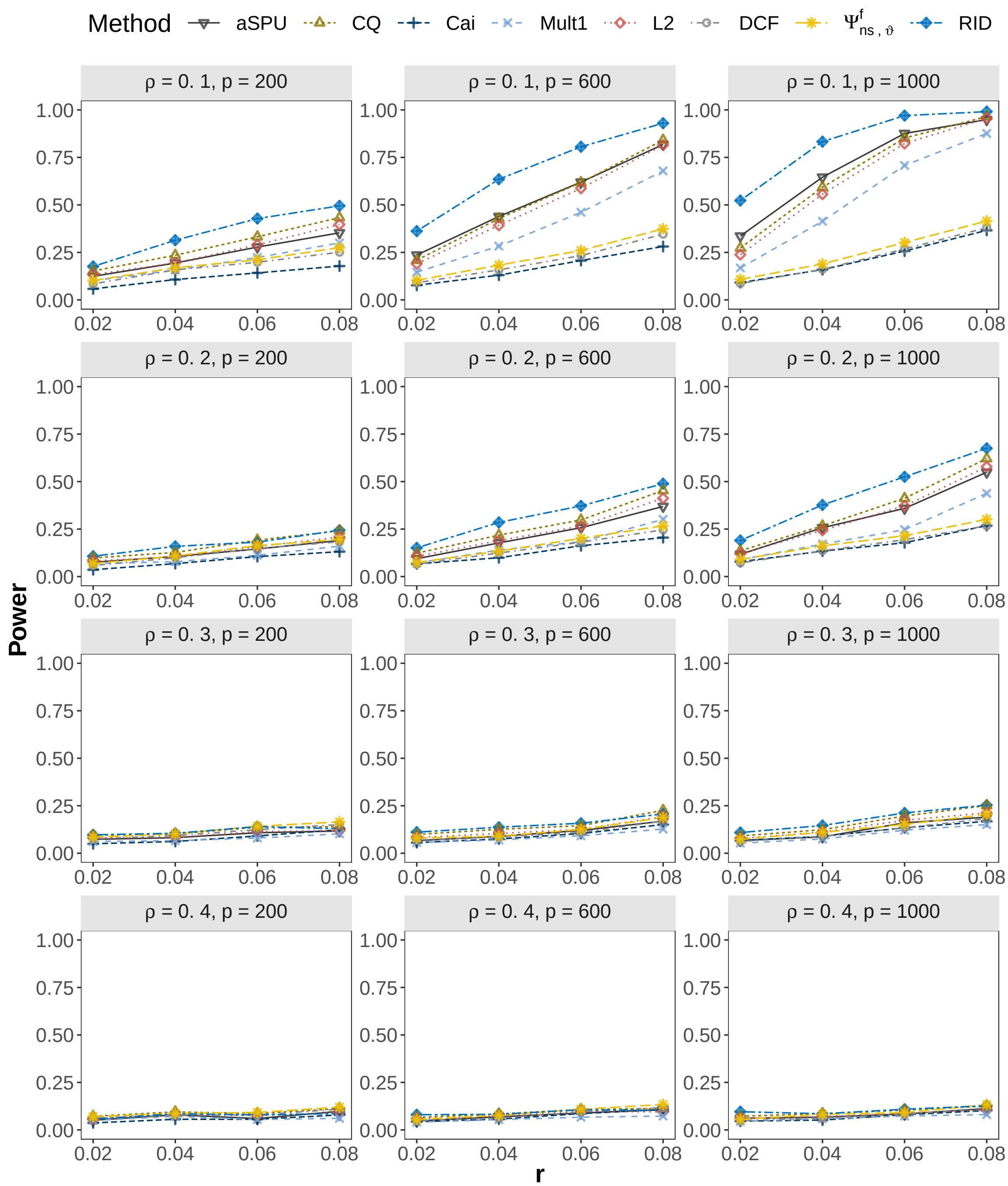

Figure 4: Empirical powers with Z_{ij} follows $(5/3)^{-1/2}t(5)$ and $m = 90, n = 120$ for **Scenario 2** under different signal levels of r and sparsity levels of ρ in **Example 1**.

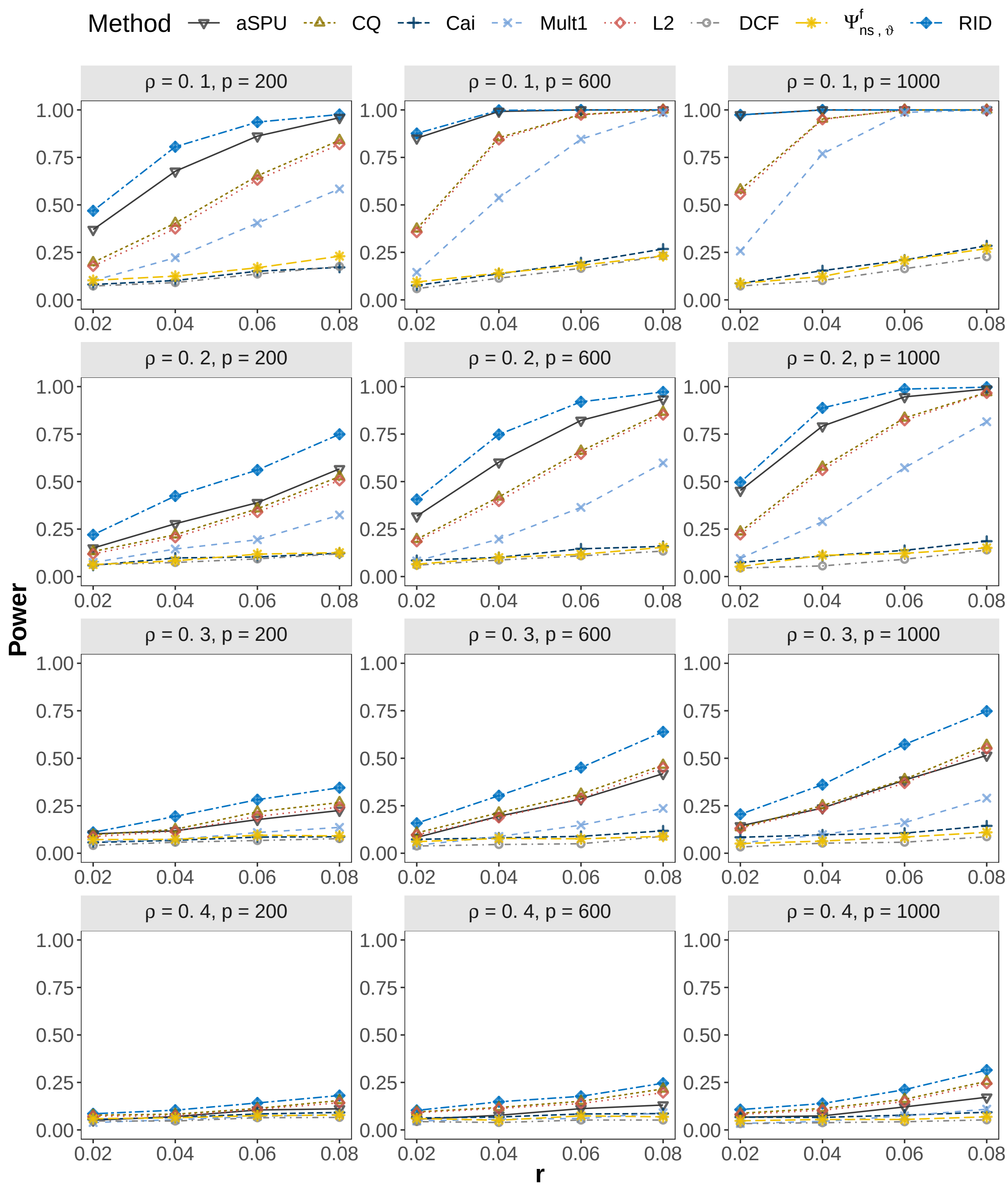

Figure 5: Empirical powers with Z_{ij} follows $8^{-1/2}\{\chi^2(4) - 4\}$ and $m = 60, n = 80$ for **Scenario 1** under different signal levels of r and sparsity levels of ρ in **Example 1**.

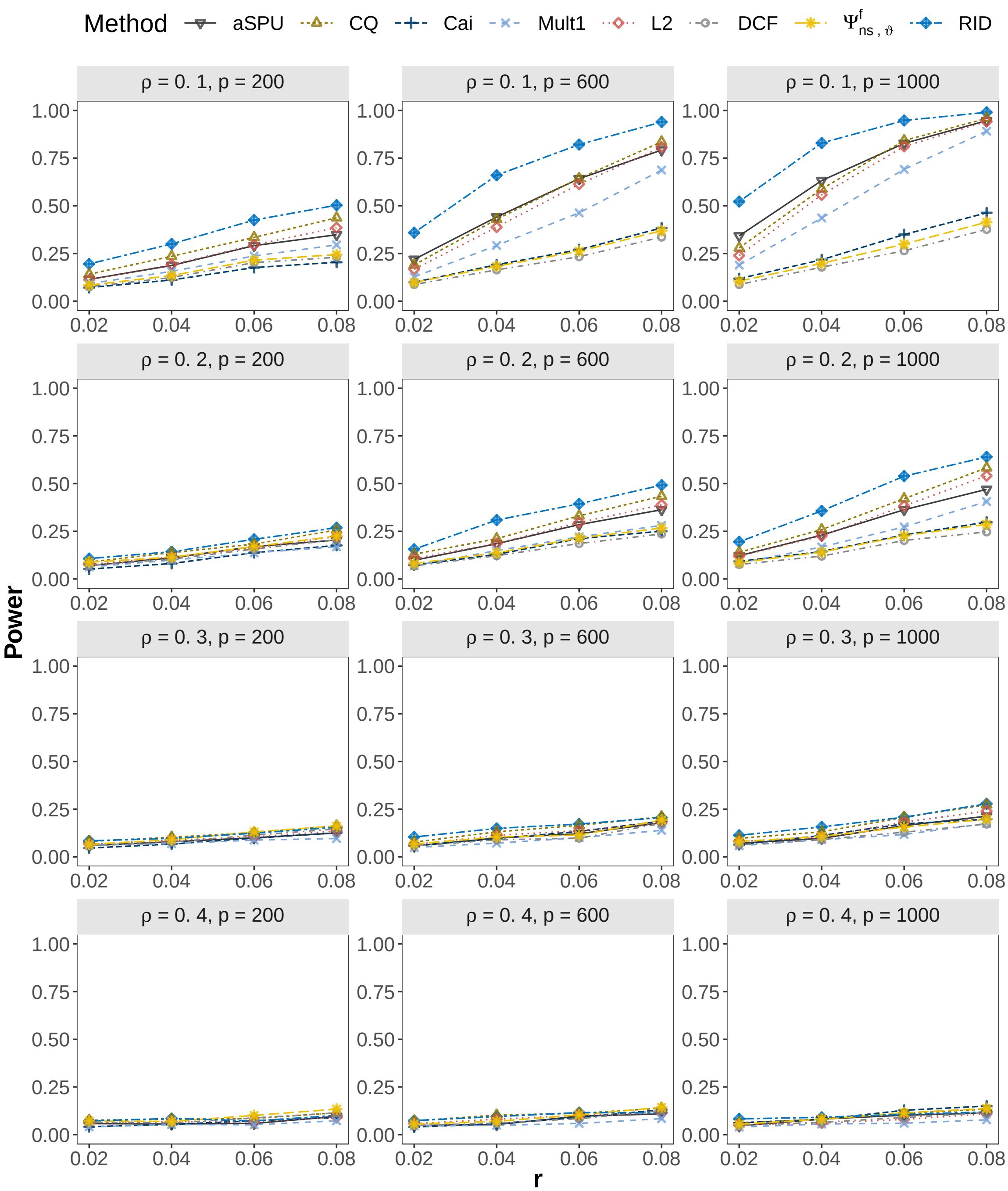

Figure 6: Empirical powers with Z_{ij} follows $8^{-1/2}\{\chi^2(4) - 4\}$ and $m = 60, n = 80$ for **Scenario 2** under different signal levels of r and sparsity levels of ρ in **Example 1**.

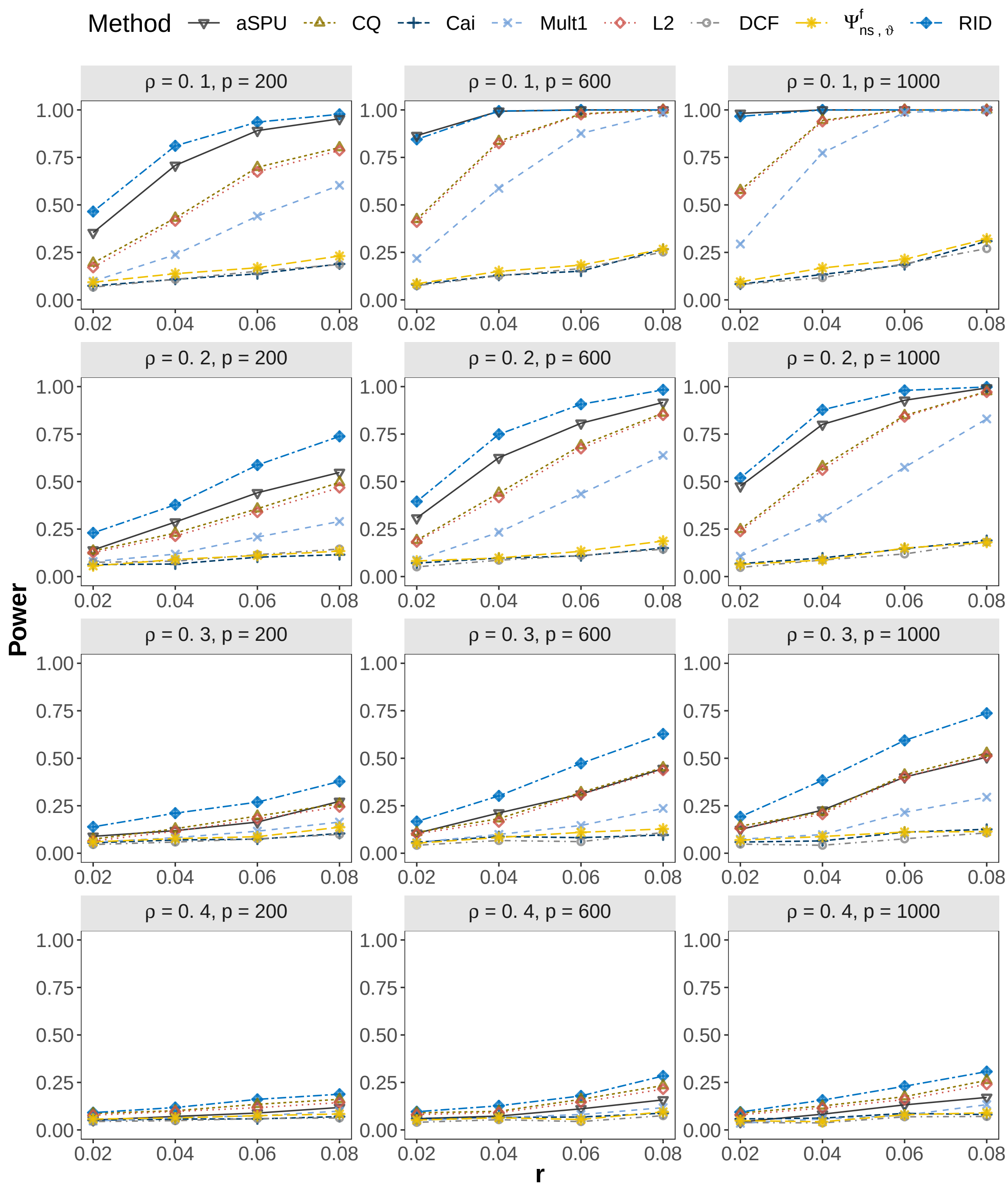

Figure 7: Empirical powers with Z_{ij} follows $8^{-1/2}\{\chi^2(4) - 4\}$ and $m = 90, n = 120$ for **Scenario 1** under different signal levels of r and sparsity levels of ρ in **Example 1**.

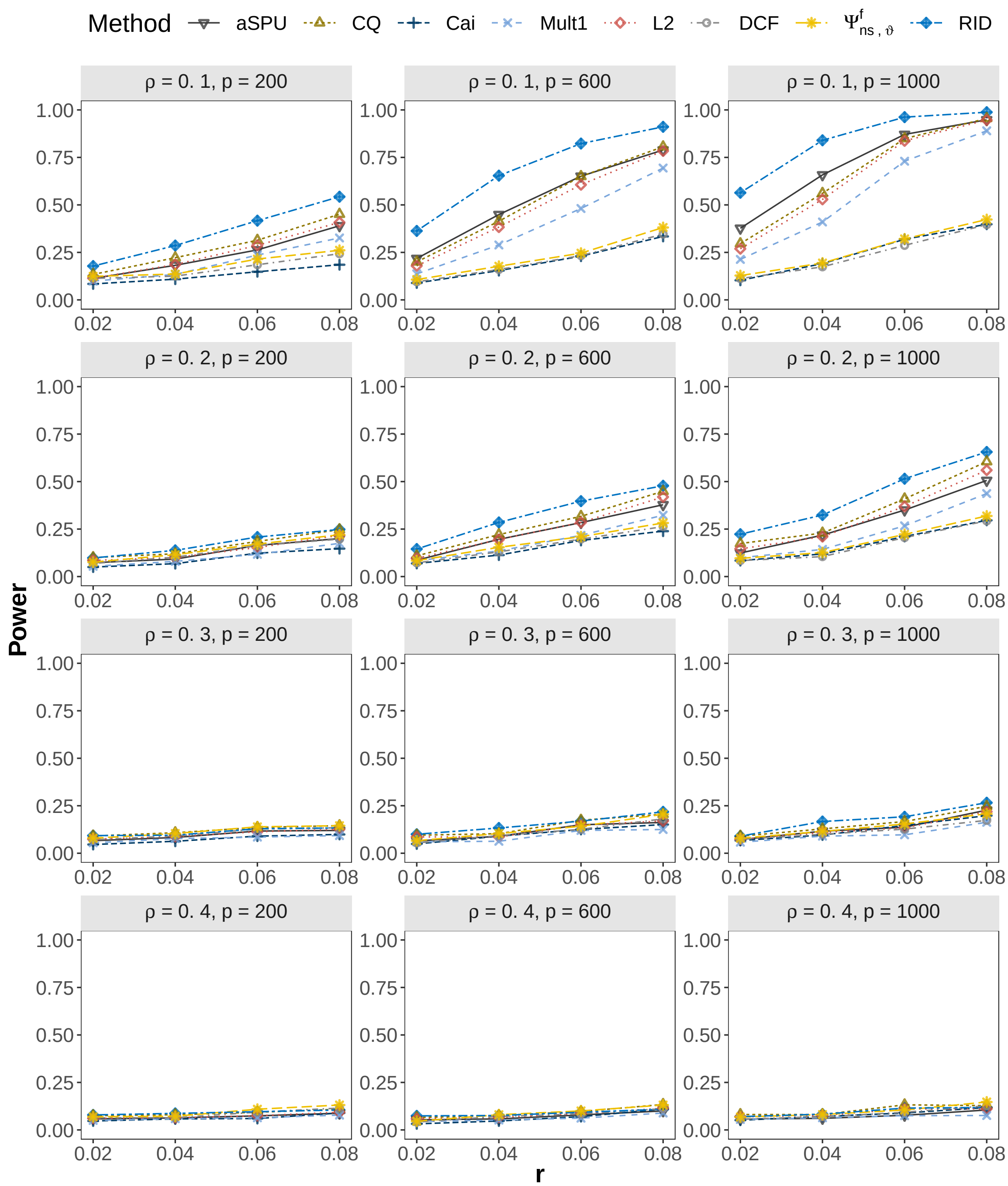

Figure 8: Empirical powers with Z_{ij} follows $8^{-1/2}\{\chi^2(4) - 4\}$ and $m = 90, n = 120$ for **Scenario 2** under different signal levels of r and sparsity levels of ρ in **Example 1**.

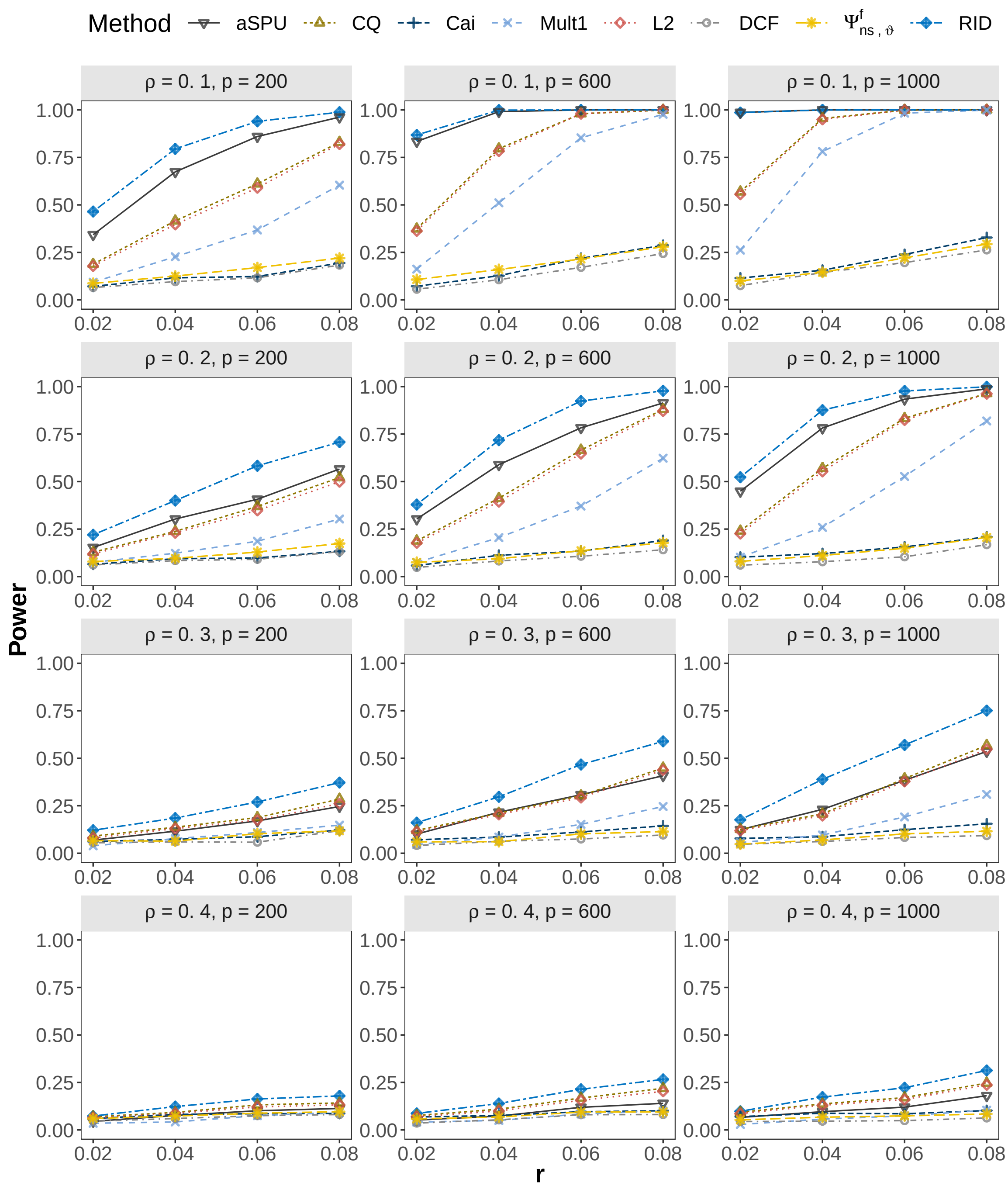

Figure 9: Empirical powers with Z_{ij} follows $\Gamma(4, 0.5) - 2$ and $m = 60, n = 80$ for **Scenario 1** under different signal levels of r and sparsity levels of ρ in **Example 1**.

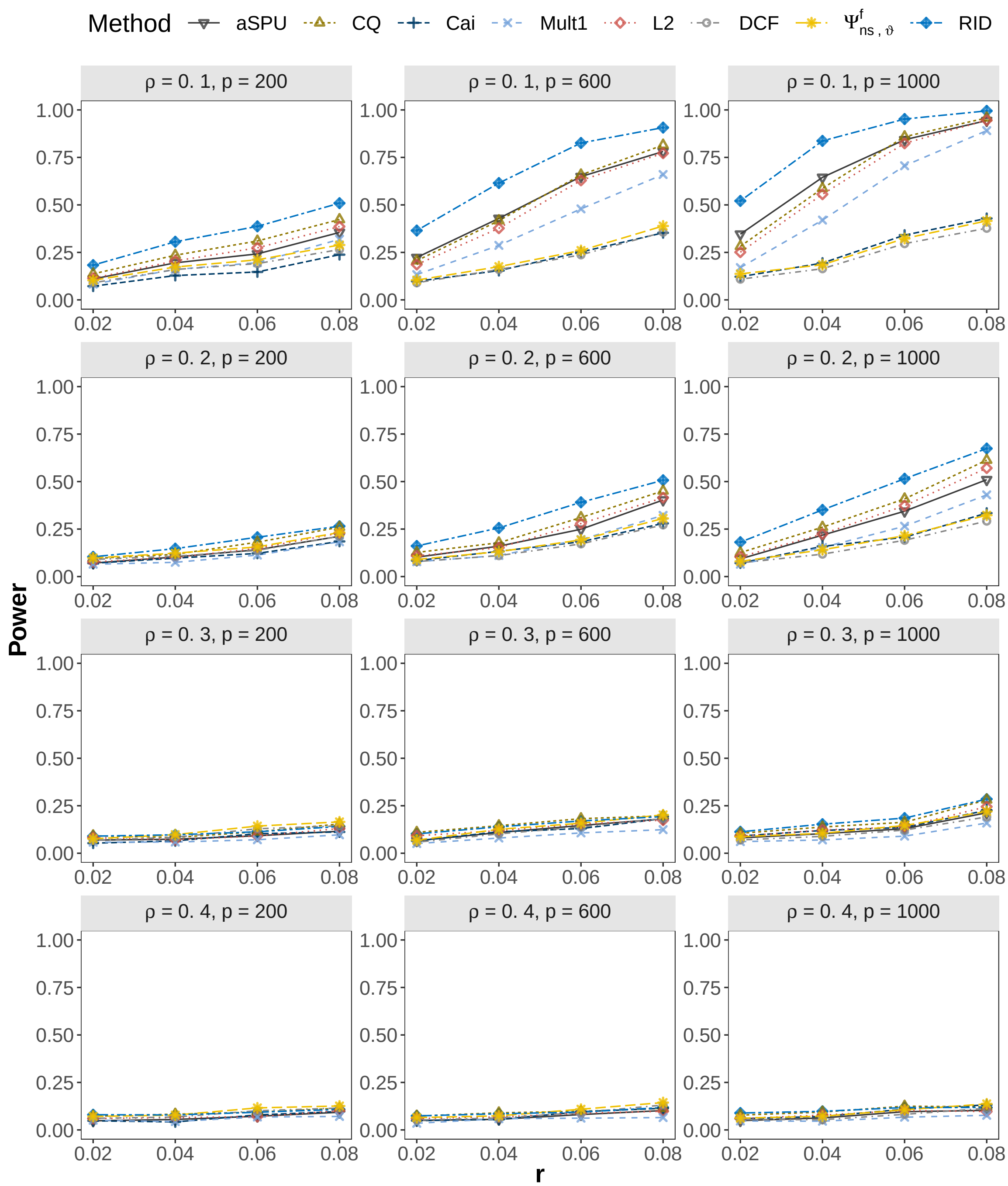

Figure 10: Empirical powers with Z_{ij} follows $\Gamma(4, 0.5) - 2$ and $m = 60, n = 80$ for **Scenario 2** under different signal levels of r and sparsity levels of ρ in **Example 1**.

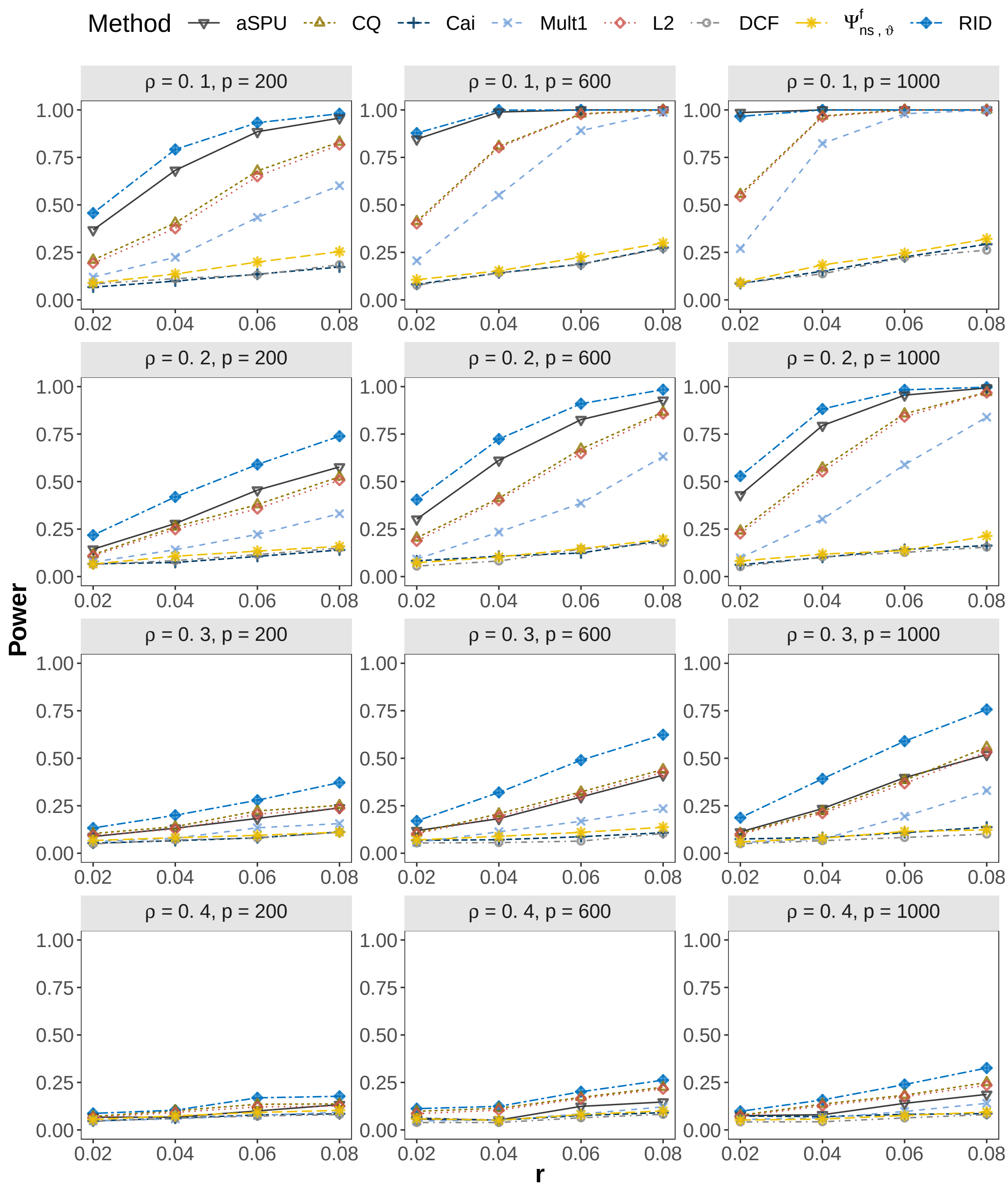

Figure 11: Empirical powers with Z_{ij} follows $\Gamma(4, 0.5) - 2$ and $m = 90, n = 120$ for **Scenario 1** under different signal levels of r and sparsity levels of ρ in **Example 1**.

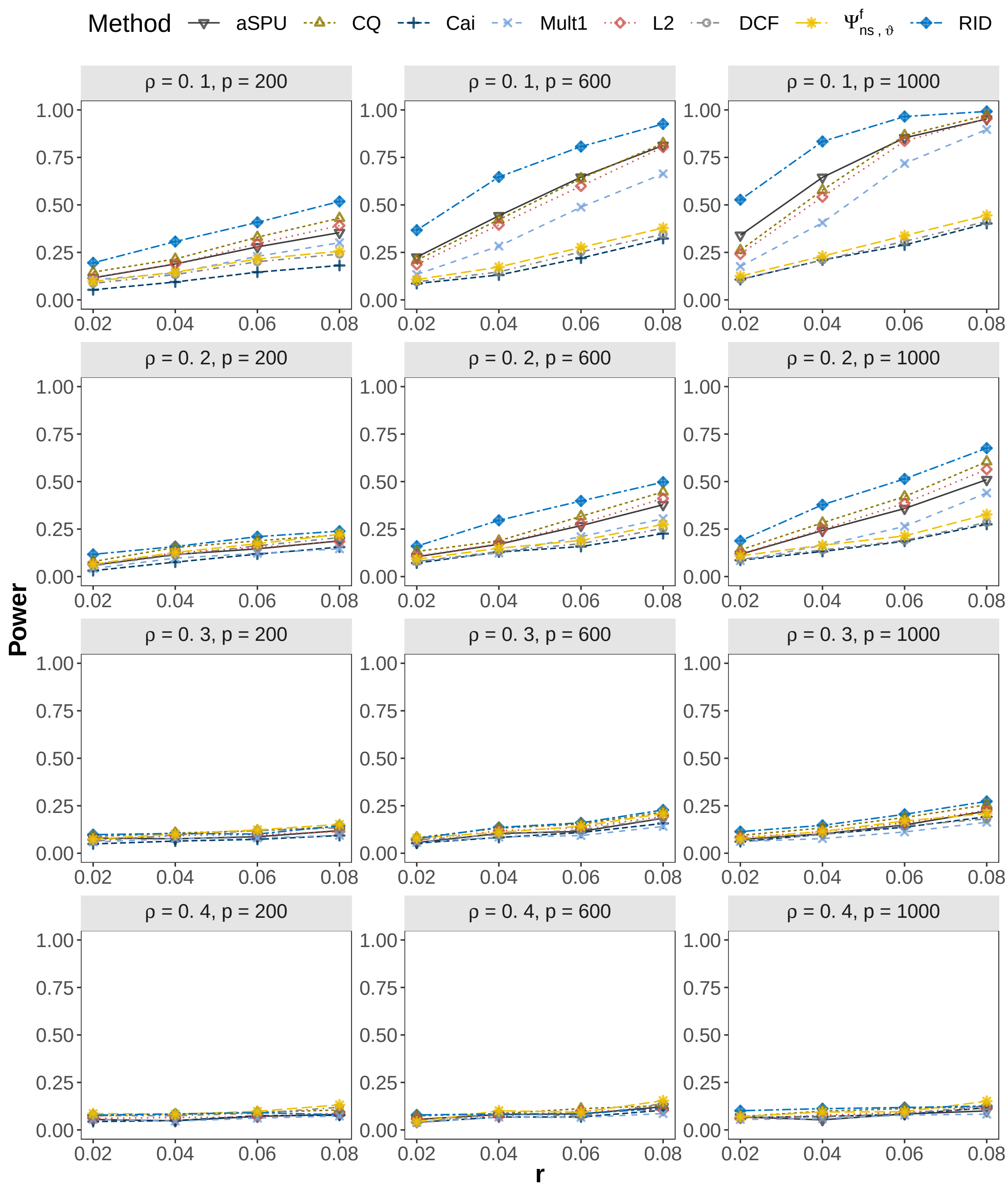

Figure 12: Empirical powers with Z_{ij} follows $\Gamma(4, 0.5) - 2$ and $m = 90, n = 120$ for **Scenario 2** under different signal levels of r and sparsity levels of ρ in **Example 1**.

Example 2. The empirical powers are reported when Z_{ik} follows the following three distributions:

1. The standardized t-distribution with degrees of freedom 5, i.e., $(5/3)^{-1/2}t(5)$;
2. The standardized chi-squared distribution with degrees of freedom 4, i.e., $8^{-1/2}\{\chi^2(4)-4\}$;
3. The standardized Gamma distribution with $a = 4, b = 0.5$, i.e., $\Gamma(4, 0.5) - 2$.

The corresponding results are shown in Figures 13-15.

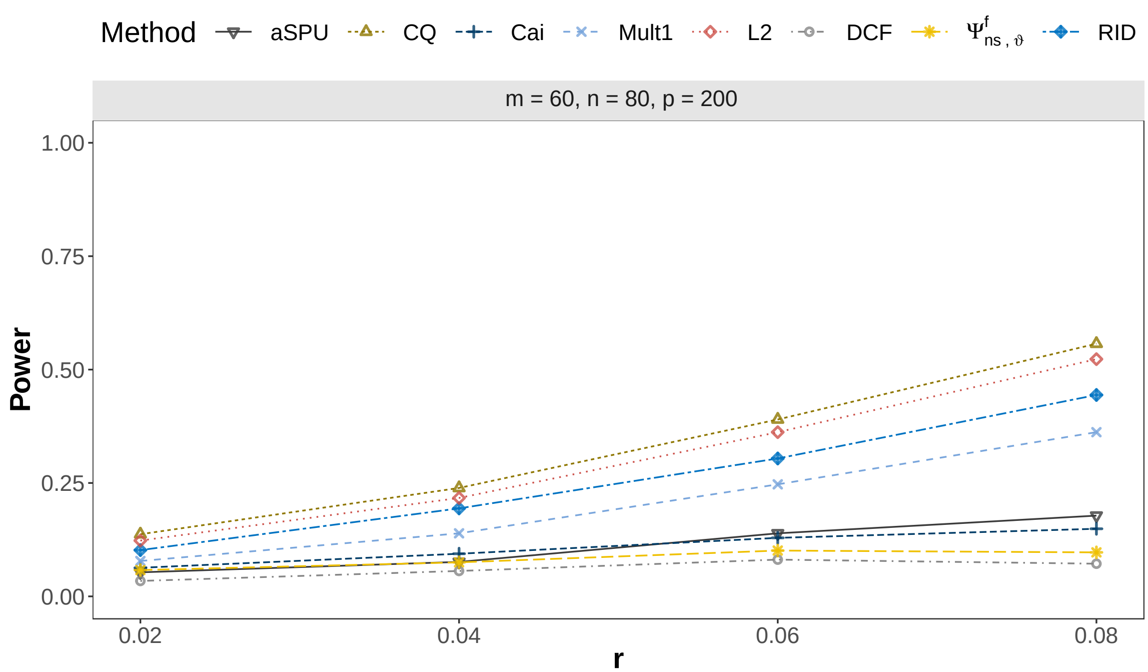

Figure 13: Empirical power when Z_{ij} follows $(5/3)^{-1/2}t(5)$, $(m,n) = (60,80)$, $p = 200$ under different signal levels of r in **Example 2**.

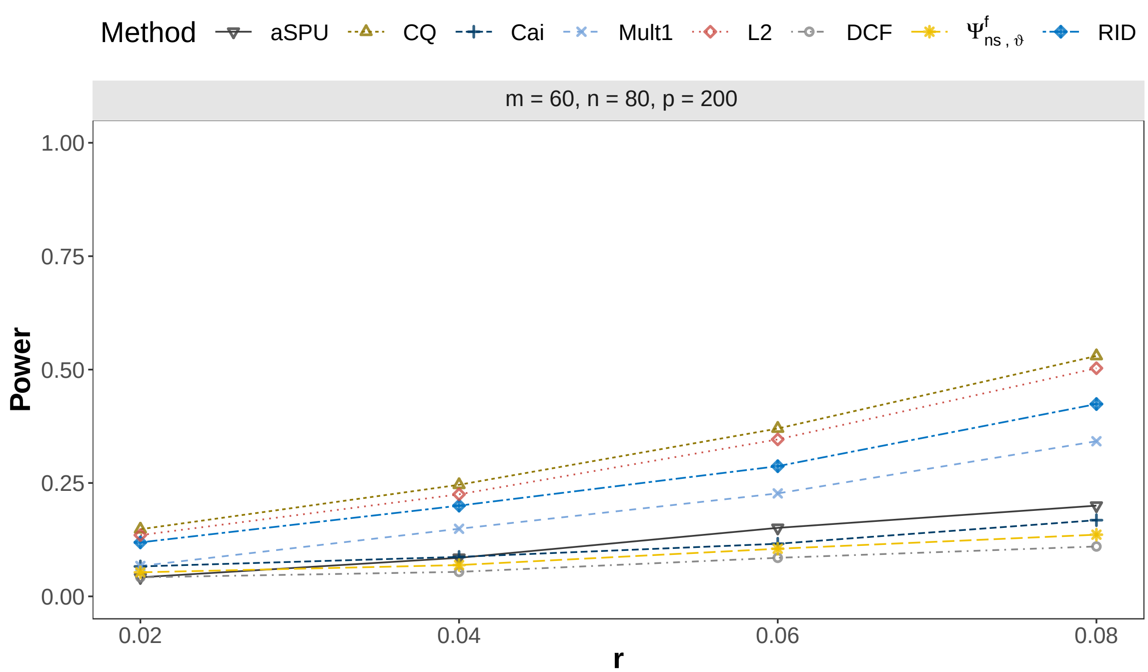

Figure 14: Empirical power when Z_{ij} follows $8^{-1/2}\{\chi^2(4) - 4\}$, $(m, n) = (60, 80)$, $p = 200$ under different signal levels of r in **Example 2**.

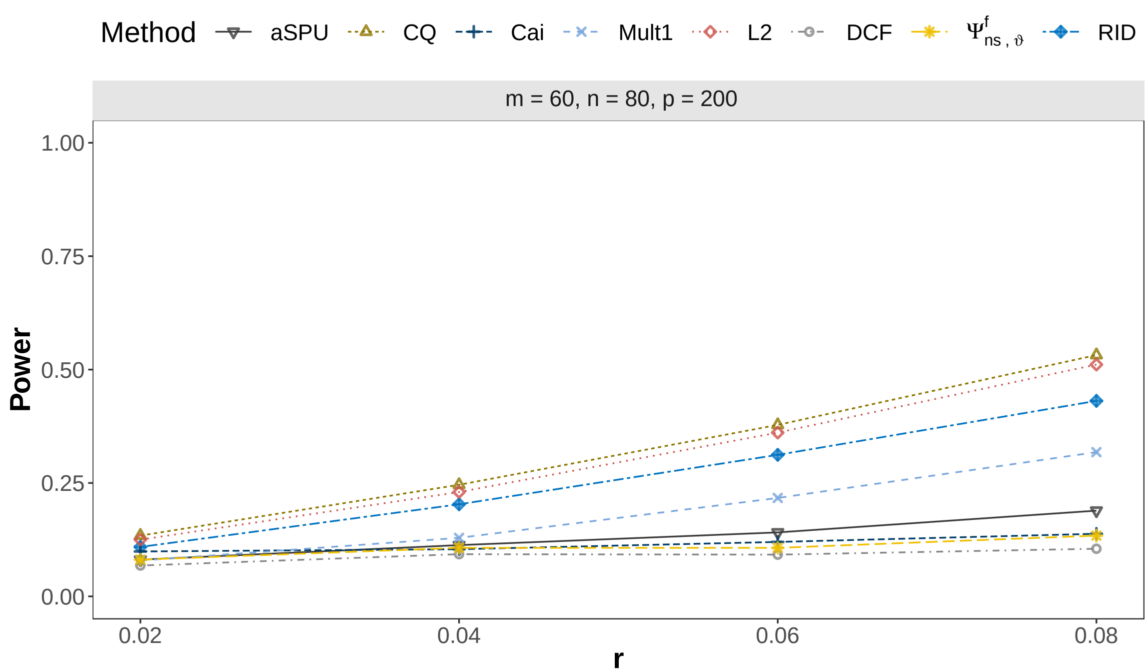

Figure 15: Empirical power when Z_{ij} follows $\Gamma(4, 0.5) - 2$, $(m, n) = (60, 80)$, $p = 200$ under different signal levels of r in **Example 2**.

Example 3. The empirical powers are reported when Z_{ik} follows the following three distributions:

1. The standardized t-distribution with degrees of freedom 5, i.e., $(5/3)^{-1/2}t(5)$;

2. The standardized chi-squared distribution with degrees of freedom 4, i.e., $8^{-1/2}\{\chi^2(4) - 4\}$;

3. The standardized Gamma distribution with $a = 4, b = 0.5$, i.e., $\Gamma(4, 0.5) - 2$.

The corresponding results are shown in Figures 16-18.

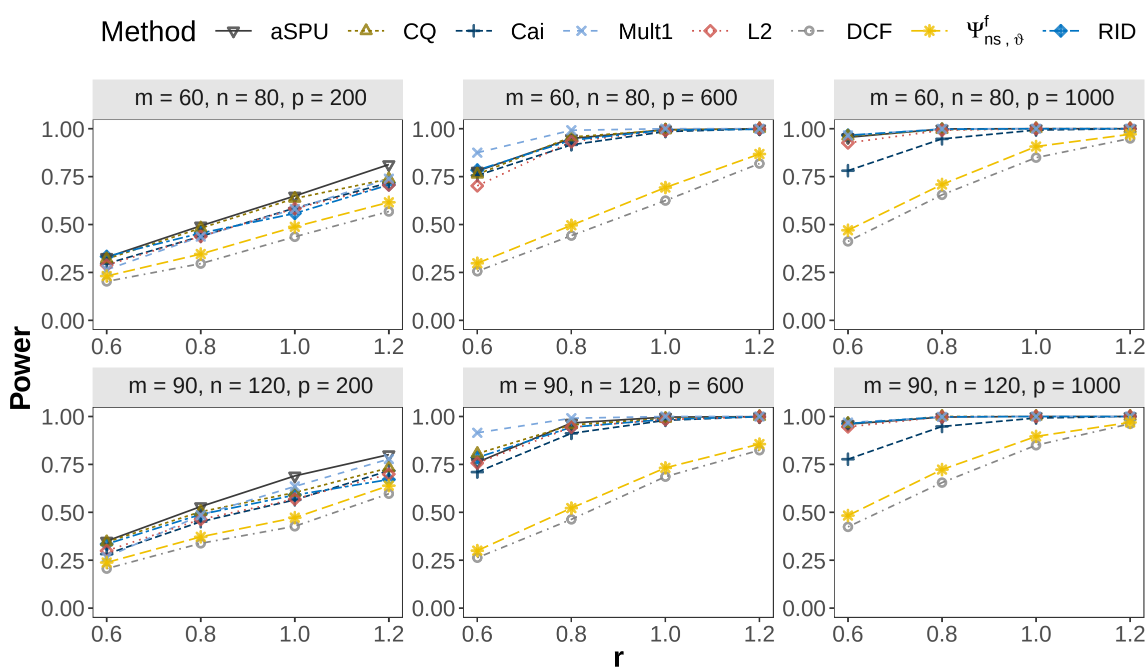

Figure 16: Empirical power when Z_{ij} follows $(5/3)^{-1/2}t(5)$, $(m, n) = (60, 80)$ and $(m, n) = (90, 120)$ under different signal levels of r in **Example 3**.

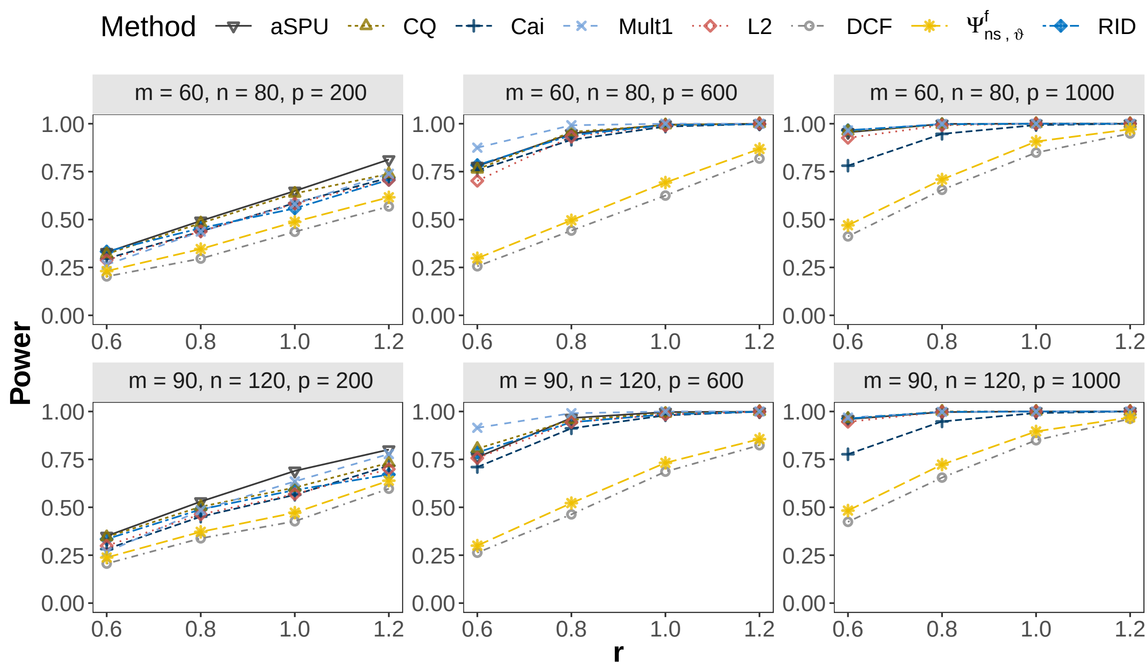

Figure 17: Empirical power when Z_{ij} follows $8^{-1/2}\{\chi^2(4)-4\}$, $(m,n)=(60,80)$ and $(m,n)=(90,120)$ under different signal levels of r in **Example 3**.

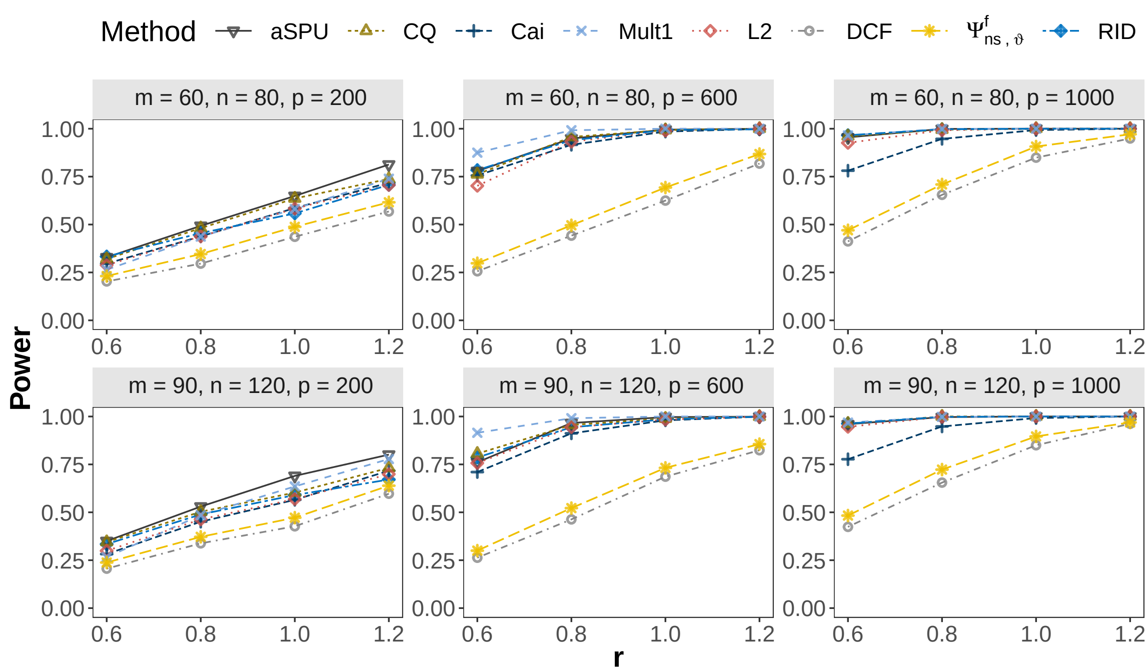

Figure 18: Empirical power when Z_{ij} follows $\Gamma(4, 0.5) - 2$, $(m, n) = (60, 80)$ and $(m, n) = (90, 120)$ under different signal levels of r in **Example 3**.

3. REAL DATA ANALYSIS

The p-values of the various methods applied to the breast cancer data and the differences in gene expression levels between the two groups in other chromosomes are shown in Table 1 and Figures 19. Because the dimensions are lower than 50 and the default candidate bandwidths of aSPU are uniformly chosen as 50, the p-values in chromosomes 19 and 21 can not be calculated.

Table 1: The p-values of the various methods applied to the breast cancer data in other chromosomes.

	Method							
Chromosome	aSPU	CQ	Cai	Mult1	L2	DCF	$\Psi^f_{ns,\vartheta}$	RID
1	1.13×10^{-1}	2.76×10^{-1}	3.93×10^{-2}	1.8×10^{-3}	2.38×10^{-1}	1.61×10^{-1}	1.40×10^{-1}	4.91×10^{-1}
2	1.36×10^{-1}	1.23×10^{-1}	4.74×10^{-2}	4.7×10^{-3}	1.46×10^{-1}	1.86×10^{-1}	1.87×10^{-1}	3.04×10^{-1}
4	1.95×10^{-2}	4.60×10^{-2}	6.53×10^{-3}	4.9×10^{-3}	7.01×10^{-2}	1.96×10^{-2}	2.08×10^{-2}	3.41×10^{-1}
5	3.03×10^{-2}	1.99×10^{-2}	1.02×10^{-2}	2.0×10^{-3}	5.11×10^{-2}	2.83×10^{-1}	2.36×10^{-1}	1.85×10^{-1}
8	0	0	1.53×10^{-7}	0	6.88×10^{-15}	9.80×10^{-3}	6.80×10^{-3}	0
9	4.17×10^{-1}	3.86×10^{-1}	1.65×10^{-1}	3.05×10^{-1}	3.41×10^{-1}	2.70×10^{-1}	2.56×10^{-1}	4.43×10^{-1}
10	6.80×10^{-1}	1.68×10^{-1}	3.69×10^{-1}	4.74×10^{-2}	1.79×10^{-1}	3.71×10^{-1}	3.57×10^{-1}	5.18×10^{-1}
11	2.26×10^{-8}	4.62×10^{-8}	2.01×10^{-3}	0	3.16×10^{-4}	7.91×10^{-3}	7.20×10^{-3}	3.65×10^{-3}
12	3.01×10^{-5}	1.10×10^{-3}	2.87×10^{-3}	2×10^{-4}	1.04×10^{-2}	7.01×10^{-3}	5.20×10^{-3}	8.90×10^{-2}
13	8.41×10^{-1}	5.47×10^{-1}	4.58×10^{-1}	3.44×10^{-1}	4.88×10^{-1}	5.45×10^{-1}	5.02×10^{-1}	6.69×10^{-1}
14	0	0	1.31×10^{-5}	0	3.56×10^{-7}	4.22×10^{-3}	2.40×10^{-3}	1.22×10^{-11}
15	1.57×10^{-2}	9.70×10^{-5}	1.93×10^{-2}	4.07×10^{-2}	6.87×10^{-3}	1.48×10^{-1}	1.28×10^{-1}	2.97×10^{-2}
16	1.03×10^{-4}	4.42×10^{-5}	2.32×10^{-2}	1×10^{-4}	2.67×10^{-3}	1.13×10^{-1}	1.00×10^{-1}	1.52×10^{-2}
17	6.05×10^{-2}	3.05×10^{-3}	1.19×10^{-1}	1.9×10^{-3}	2.15×10^{-2}	1.07×10^{-1}	8.92×10^{-2}	1.13×10^{-1}
18	3.18×10^{-1}	2.59×10^{-1}	1.20×10^{-1}	1.70×10^{-1}	2.34×10^{-1}	1.68×10^{-1}	1.27×10^{-1}	4.79×10^{-1}
19	$\backslash$	2.36×10^{-1}	6.49×10^{-1}	7.33×10^{-1}	2.21×10^{-1}	4.04×10^{-1}	3.86×10^{-1}	4.13×10^{-1}
20	2.38×10^{-3}	6.88×10^{-3}	7.95×10^{-4}	4.6×10^{-3}	2.53×10^{-2}	5.28×10^{-2}	4.80×10^{-2}	5.65×10^{-2}
21	$\backslash$	2.06×10^{-5}	4.28×10^{-3}	6×10^{-4}	2.38×10^{-3}	5.41×10^{-2}	4.36×10^{-2}	2.43×10^{-5}
22	4.05×10^{-2}	5.71×10^{-3}	1.37×10^{-2}	1.4×10^{-3}	2.08×10^{-2}	3.46×10^{-2}	3.08×10^{-2}	4.85×10^{-3}

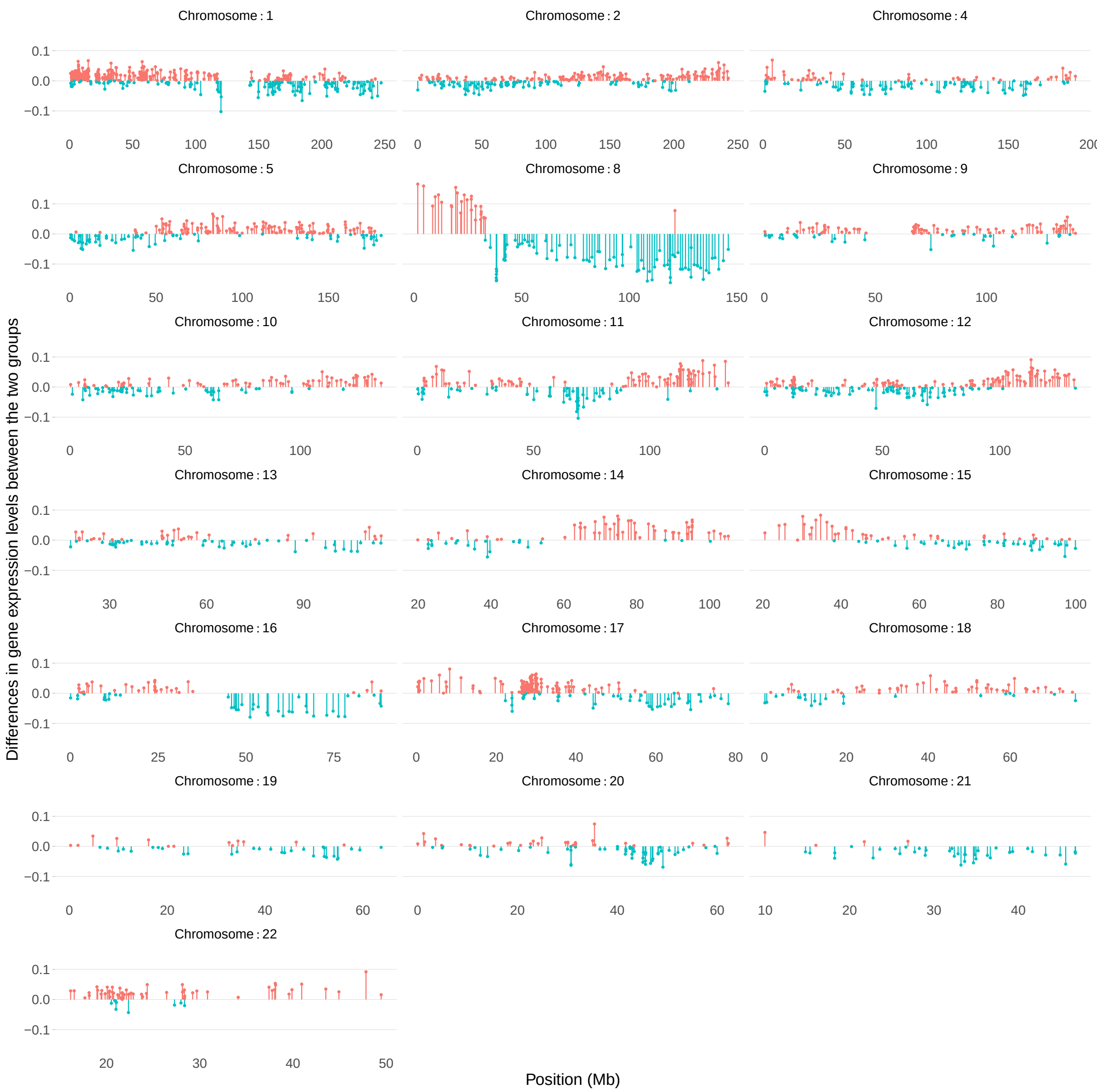

Figure 19: Differences in the ratios of Cy5/Cy3 signals between the two groups in other chromosomes.

REFERENCES

Bushell, P., and Trustrum, G. (1990), "Trace inequalities for positive definite matrix power products," *Linear Algebra and its Applications*, 132, 173–178.

Chen, S. X., Zhang, L. X., and Zhong, P. S. (2010), "Tests for high-dimensional covariance matrices," *Journal of the American Statistical Association*, 105(490), 810–819.

Hall, P., and Heyde, C. C. (1980), *Martingale Limit Theory and Its Application*, New York: Academic Press.