Appendix A: Conditions and Proofs of Main Results

Conditions. The following technical conditions are not the weakest possible, but facilitate the derivations.

A1. For each i, $\varepsilon_{i,1}, \ldots, \varepsilon_{i,n_1}$ are i.i.d. with $\mathrm{E}(\varepsilon_{i,1}) = 0$ and $\mathrm{var}(\varepsilon_{i,1}) = \sigma^2_{\varepsilon;i} \in (0, \infty)$.

A2. For each i, $e_{i,1}, \ldots, e_{i,n_2}$ are i.i.d. with $\mathrm{E}(e_{i,1}) = 0$ and $\mathrm{var}(e_{i,1}) = \sigma^2_{e;i} \in (0, \infty)$.

A3. For each i, $(\varepsilon_{i,1}, \ldots, \varepsilon_{i,n_1})$ is independent of $(e_{i,1}, \ldots, e_{i,n_2})$.

A4. For each i, $\mathrm{E}(|\varepsilon_{i,1}|^3) < \infty$ and $\mathrm{E}(|e_{i,1}|^3) < \infty$.

A5. Two-sample t-statistics corresponding to true nulls are identically distributed.

A5′. Two-sample t-statistics corresponding to true non-nulls are identically distributed.

A6. There are constants c_1 and c_2 satisfying $0 < c_1 \le c_2 < \infty$, such that $c_1 \le n_1/n_2 \le c_2$.

A7. Two-sample t-statistics corresponding to true nulls are independent.

A7′. Two-sample t-statistics are independent.

A8. Let $F_{0;T}(\cdot; n)$ and $F_{1;T}(\cdot; n)$ denote the C.D.F. of two-sample t-statistics under the true null and non-null, respectively.

A9. The marginal C.D.F. and p.d.f. of two-sample t-statistics are $F_T(\cdot; n) = \pi_0 F_{0;T}(\cdot; n) + (1 - \pi_0) F_{1;T}(\cdot; n)$ and $f_T(\cdot; n) = F'_T(\cdot; n)$, where $f_T(t; n)$ is Lipschitz continuous in t uniformly in n.

A10. The marginal C.D.F. of true p-values $\{P_i\}$ is $F_P(\cdot; n) = \pi_0 F_{0;P}(\cdot; n) + (1 - \pi_0) F_{1;P}(\cdot; n)$, where $F_{0;P}(\cdot; n)$ is the C.D.F. of the standard uniform distribution, and $F_{1;P}(\cdot; n)$ is the C.D.F. of $\{P_i\}$ under the true non-null. Assume $F_{1;P}(t; n)$ is continuous in t.

Note that condition A5 is valid when $\{\varepsilon_{i,1} : i \in \mathcal{I}_0\}$ are identically distributed and $\{e_{i,1} : i \in \mathcal{I}_0\}$ are identically distributed. Condition A7 holds if $\{(\varepsilon_{i,1}, \ldots, \varepsilon_{i,n_1}; e_{i,1}, \ldots, e_{i,n_2}) : i \in \mathcal{I}_0\}$ are independent.

We first present Lemma 1, which will be used in proving Propositions 1, 2, 4 and 5.

Notation. For sequences $\{a_n\}_{n\ge1}$ and $\{b_n\}_{n\ge1}$, $a_n \asymp b_n$ denotes $\lim_{n\to\infty} a_n/b_n = 1$.

Lemma 1 *Assume model (2.1) and conditions* A1–A6*, and the general two-sample t-statistics* $\{T^{\text{general}}_{i;n_1,n_2}\}_{i=1}^m$ *are used. Assume* $\alpha \in (0,1)$, $m_0/m \to \pi_0 \in (0,1]$, $m \to \infty$, $n \to \infty$, *and* (m,n) *satisfies (3.1).*

(i) *If* $\mathrm{t}^{\mathrm{a}}_{\alpha;m}$ *is given in (3.2) and* (m,n) *satisfies (3.1), then*

$$\max_{i\in\mathcal{I}_0} \alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) \le \frac{\beta_{1;\alpha}}{m}\{1+o(1)\}, \quad \textit{and} \quad \sum_{i\in\mathcal{I}_0} \alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) \to \pi_0\beta_{1;\alpha}, \tag{A.1}$$

where $\beta_{1;\alpha} = -\log(1-\alpha)$.

(ii) *If* $\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}$ *is given in (3.4) and* (m,n) *satisfies (3.1), then for* $k \ge 2$,

$$\max_{i\in\mathcal{I}_0} \alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) \le \frac{\beta_{k;\alpha}}{m}\{1+o(1)\}, \quad \textit{and} \quad \sum_{i\in\mathcal{I}_0} \alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) \to \pi_0\beta_{k;\alpha}, \tag{A.2}$$

where $\beta_{k;\alpha}$ *solves the equation (3.5).*

Proof: We first show part (i). For $\mathrm{t}^{\mathrm{a}}_{\alpha;m}$ given in (3.2) and $\mathbb{N}(0,1)$ random variables $\{T^{\mathrm{a}}_i\}_{i=1}^m$, we obtain

$$\alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) = \mathrm{P}(|T^{\mathrm{a}}_i| > \mathrm{t}^{\mathrm{a}}_{\alpha;m}) = 2\{1-\Phi(\mathrm{t}^{\mathrm{a}}_{\alpha;m})\} = 1-(1-\alpha)^{1/m}, \qquad i=1,\ldots,m, \tag{A.3}$$

and as $m \to \infty$,

$$\max_{i\in\mathcal{I}_0} \alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) = 1-(1-\alpha)^{1/m} = \frac{\beta_{1;\alpha}}{m}\{1+o(1)\} = o(1), \tag{A.4}$$

$$\sum_{i\in\mathcal{I}_0} \alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) = m_0\{1-(1-\alpha)^{1/m}\} = \pi_0\beta_{1;\alpha}\{1+o(1)\}, \tag{A.5}$$

where

$$\beta_{1;\alpha} \equiv -\log(1-\alpha) \in (0,\infty). \tag{A.6}$$

For $\alpha_{i;n_1,n_2}(\mathrm{t})$, it can be rewritten as $\alpha_{i;n_1,n_2}(\mathrm{t}) = \alpha_i^{\mathrm{a}}(\mathrm{t}) + d_i(\mathrm{t})$, where

$$\begin{aligned}
d_i(\mathrm{t}) &= \alpha_i^{\mathrm{a}}(\mathrm{t})\Big\{\frac{\alpha_{i;n_1,n_2}(\mathrm{t})}{\alpha_i^{\mathrm{a}}(\mathrm{t})} - 1\Big\}, \\
|d_i(\mathrm{t})| &= \alpha_i^{\mathrm{a}}(\mathrm{t})\Big|\frac{\alpha_{i;n_1,n_2}(\mathrm{t})}{\alpha_i^{\mathrm{a}}(\mathrm{t})} - 1\Big|, \\
\max_{i\in\mathcal{I}_0} |d_i(\mathrm{t})| &\le \Big\{\max_{i\in\mathcal{I}_0} \alpha_i^{\mathrm{a}}(\mathrm{t})\Big\}\Big\{\max_{i\in\mathcal{I}_0}\Big|\frac{\alpha_{i;n_1,n_2}(\mathrm{t})}{\alpha_i^{\mathrm{a}}(\mathrm{t})} - 1\Big|\Big\}, && \text{(A.7)} \\
\Big|\sum_{i\in\mathcal{I}_0} d_i(\mathrm{t})\Big| &\le \sum_{i\in\mathcal{I}_0} |d_i(\mathrm{t})| \le \Big(\max_{i\in\mathcal{I}_0} |d_i(\mathrm{t})|\Big) m_0. && \text{(A.8)}
\end{aligned}$$

This leads to

$$\begin{aligned}
\alpha_{i;n_1,n_2}(\mathrm{t}) &\le \alpha_i^{\mathrm{a}}(\mathrm{t}) + |d_i(\mathrm{t})|, \\
\max_{i\in\mathcal{I}_0} \alpha_{i;n_1,n_2}(\mathrm{t}) &\le \max_{i\in\mathcal{I}_0} \alpha_i^{\mathrm{a}}(\mathrm{t}) + \max_{i\in\mathcal{I}_0} |d_i(\mathrm{t})|, && \text{(A.9)} \\
\sum_{i\in\mathcal{I}_0} \alpha_{i;n_1,n_2}(\mathrm{t}) &= \sum_{i\in\mathcal{I}_0} \alpha_i^{\mathrm{a}}(\mathrm{t}) + \sum_{i\in\mathcal{I}_0} d_i(\mathrm{t}). && \text{(A.10)}
\end{aligned}$$

Thus, if the condition

$$\max_{i\in\mathcal{I}_0} \Big|\frac{\alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m})}{\alpha_i^{\mathrm{a}}(\mathrm{t}^{\mathrm{a}}_{\alpha;m})} - 1\Big| = o(1) \tag{A.11}$$

holds, then (A.7), (A.4), (A.11) and (A.8) imply that

$$\max_{i\in\mathcal{I}_0} |d_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m})| \le \frac{\beta_{1;\alpha}}{m}\{1+o(1)\}o(1) = \frac{\beta_{1;\alpha}}{m}o(1), \qquad \Big|\sum_{i\in\mathcal{I}_0} d_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m})\Big| \le \pi_0\beta_{1;\alpha}o(1), \tag{A.12}$$

which combined with (A.9), (A.4), (A.10), and (A.5) gives

$$\begin{aligned}
\max_{i\in\mathcal{I}_0} \alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) &\le \frac{\beta_{1;\alpha}}{m}\{1+o(1)\} + \frac{\beta_{1;\alpha}}{m}o(1) = \frac{\beta_{1;\alpha}}{m}\{1+o(1)\}, \\
\sum_{i\in\mathcal{I}_0} \alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) &= \pi_0\beta_{1;\alpha}\{1+o(1)\} + o(1) = \pi_0\beta_{1;\alpha}\{1+o(1)\}. && \text{(A.13)}
\end{aligned}$$

Hence condition (A.11) indeed implies (A.1).

Now, we justify that (A.11) holds. Recall

$$\alpha_i^{\mathrm{a}}(\mathrm{t}) = \mathrm{P}(|T_i^{\mathrm{a}}| > \mathrm{t})$$

$$
\begin{aligned}
&= \mathrm{P}(T_i^{\mathrm{a}} > \mathrm{t}) + \mathrm{P}(T_i^{\mathrm{a}} < -\mathrm{t}) \\
&= \{1-\Phi(\mathrm{t})\} + \Phi(-\mathrm{t}) = 2\{1-\Phi(\mathrm{t})\}, \\
\alpha_{i;n_1,n_2}(\mathrm{t}) &= \mathrm{P}_{H_{0,i}}(|T_{i;n_1,n_2}^{\mathrm{general}}| > \mathrm{t}) \\
&= \mathrm{P}_{H_{0,i}}(T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t}) + \mathrm{P}_{H_{0,i}}(T_{i;n_1,n_2}^{\mathrm{general}} < -\mathrm{t}).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\frac{\alpha_{i;n_1,n_2}(\mathrm{t})}{\alpha_i^{\mathrm{a}}(\mathrm{t})} - 1 &= \frac{\mathrm{P}_{H_{0,i}}(T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t}) + \mathrm{P}_{H_{0,i}}(T_{i;n_1,n_2}^{\mathrm{general}} < -\mathrm{t})}{2\{1-\Phi(\mathrm{t})\}} - 1 \\
&= \frac{1}{2}\Big\{\frac{\mathrm{P}_{H_{0,i}}(T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t}) + \mathrm{P}_{H_{0,i}}(-T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t})}{1-\Phi(\mathrm{t})} - 2\Big\} \\
&= \frac{1}{2}\Big[\Big\{\frac{\mathrm{P}_{H_{0,i}}(T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t})}{1-\Phi(\mathrm{t})} - 1\Big\} + \Big\{\frac{\mathrm{P}_{H_{0,i}}(-T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t})}{1-\Phi(\mathrm{t})} - 1\Big\}\Big],
\end{aligned}
$$

and thus

$$
\begin{aligned}
&\max_{i\in\mathcal{I}_0}\Big|\frac{\alpha_{i;n_1,n_2}(\mathrm{t})}{\alpha_i^{\mathrm{a}}(\mathrm{t})} - 1\Big| \\
= \;& \frac{1}{2}\max_{i\in\mathcal{I}_0}\Big|\Big\{\frac{\mathrm{P}_{H_{0,i}}(T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t})}{1-\Phi(\mathrm{t})} - 1\Big\} + \Big\{\frac{\mathrm{P}_{H_{0,i}}(-T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t})}{1-\Phi(\mathrm{t})} - 1\Big\}\Big| \\
\le \;& \max_{i\in\mathcal{I}_0}\Big|\frac{\mathrm{P}_{H_{0,i}}(T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t})}{1-\Phi(\mathrm{t})} - 1\Big| + \max_{i\in\mathcal{I}_0}\Big|\frac{\mathrm{P}_{H_{0,i}}(-T_{i;n_1,n_2}^{\mathrm{general}} > \mathrm{t})}{1-\Phi(\mathrm{t})} - 1\Big|. \qquad \text{(A.14)}
\end{aligned}
$$

From (A.3), we observe that

$$
\Phi(\mathrm{t}_{\alpha;m}^{\mathrm{a}}) = 1 - \frac{1-(1-\alpha)^{1/m}}{2} = 1 - \frac{\beta_{1;\alpha}/2}{m}\{1+o(1)\} \to 1, \qquad \text{(A.15)}
$$

as $m \to \infty$, where $\beta_{1;\alpha}$ is as defined in (A.6), and thus we conclude $\mathrm{t}_{\alpha;m}^{\mathrm{a}} \to \infty$. To find the explicit convergence rate of $\mathrm{t}_{\alpha;m}^{\mathrm{a}}$ defined in (3.2), we use the tail probability (DasGupta, 2008, p. 655) of a $\mathbb{N}(0,1)$ distribution,

$$
1 - \Phi(\mathrm{t}_{\alpha;m}^{\mathrm{a}}) \asymp \frac{1}{\mathrm{t}_{\alpha;m}^{\mathrm{a}}}\frac{1}{\sqrt{2\pi}} e^{-(\mathrm{t}_{\alpha;m}^{\mathrm{a}})^2/2}. \qquad \text{(A.16)}
$$

Combining (A.15) and (A.16) gives $\frac{1}{\sqrt{2\pi}\,\mathrm{t}_{\alpha;m}^{\mathrm{a}} e^{(\mathrm{t}_{\alpha;m}^{\mathrm{a}})^2/2}} \asymp \frac{\beta_{1;\alpha}/2}{m}$, which is equivalent to $\sqrt{2\pi}\,\mathrm{t}_{\alpha;m}^{\mathrm{a}} e^{(\mathrm{t}_{\alpha;m}^{\mathrm{a}})^2/2} \asymp \frac{m}{\beta_{1;\alpha}/2}$. This gives $(\mathrm{t}_{\alpha;m}^{\mathrm{a}})^2 = O(\log(m))$, i.e., $\mathrm{t}_{\alpha;m}^{\mathrm{a}} = O(\{\log(m)\}^{1/2})$, which together with (3.1) gives $\mathrm{t}_{\alpha;m}^{\mathrm{a}} = o(n^{1/6})$. An application of Theorem 1.2 of Cao (2007) to (A.14), together with condition A5, give $\max_{i\in\mathcal{I}_0}|\frac{\alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m})}{\alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m})}-1| = o(1)$ as $m\to\infty$ and $n\to\infty$. Hence (A.11) is verified.

Next, we show part (ii). The critical values $\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}$ given in (3.4) satisfy

$$\alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) = \mathrm{P}(|T^{\mathrm{a}}_i| > \mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) = 2\{1-\Phi(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k})\} = \frac{\beta_{k;\alpha}}{m}, \qquad i=1,\ldots,m, \tag{A.17}$$

where (3.5) implies that $\beta_{k;\alpha}\in(0,\infty)$. Thus, we obtain

$$\max_{i\in\mathcal{I}_0}\alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) = \frac{\beta_{k;\alpha}}{m}, \qquad \sum_{i\in\mathcal{I}_0}\alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) = \frac{m_0}{m}\beta_{k;\alpha} = \pi_0\beta_{k;\alpha}+o(1).$$

Also, $\Phi(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) = 1-\frac{\beta_{k;\alpha}/2}{m}\to 1$. The rest of the proof is similar to that used in part (i). ■

Proof of Proposition 1. From (2.4) and condition A7, it suffices to consider $\mathbb{N}(0,1)$ random variables $\{T^{\mathrm{a}}_i\}_{i=1}^m$, with $\{T^{\mathrm{a}}_i : i\in\mathcal{I}_0\}$ being independent. Direct calculations give

$$\begin{aligned}
\mathrm{FWER}^{\mathrm{a}}_1(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) &= \mathrm{P}\Big(\sum_{i\in\mathcal{I}_0}\mathrm{I}(|T^{\mathrm{a}}_i|>\mathrm{t}^{\mathrm{a}}_{\alpha;m})\ge 1\Big)\\
&= \mathrm{P}(\cup_{i\in\mathcal{I}_0}\{|T^{\mathrm{a}}_i|>\mathrm{t}^{\mathrm{a}}_{\alpha;m}\})\\
&= 1-\mathrm{P}(\cap_{i\in\mathcal{I}_0}\overline{\{|T^{\mathrm{a}}_i|>\mathrm{t}^{\mathrm{a}}_{\alpha;m}\}})\\
&= 1-\prod_{i\in\mathcal{I}_0}\mathrm{P}(\overline{\{|T^{\mathrm{a}}_i|>\mathrm{t}^{\mathrm{a}}_{\alpha;m}\}})\\
&= 1-\prod_{i\in\mathcal{I}_0}\{1-\alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m})\} = 1-(1-\alpha)^{m_0/m},
\end{aligned} \tag{A.18}$$

where $1-(1-\alpha)^{m_0/m}\le\alpha$. This shows the second part of (3.3).

To show the first part of (3.3), note that derivations similar to (A.18) together with condition A7 give $\mathrm{FWER}_1(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) = 1-\prod_{i\in\mathcal{I}_0}\{1-\alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m})\}$. It thus suffices to show $\prod_{i\in\mathcal{I}_0}\{1-\alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m})\}-\prod_{i\in\mathcal{I}_0}\{1-\alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m})\} = o(1)$. From (A.18), $\prod_{i\in\mathcal{I}_0}\{1-\alpha^{\mathrm{a}}_i(\mathrm{t}^{\mathrm{a}}_{\alpha;m})\} = (1-\alpha)^{\pi_0+o(1)} = e^{-\pi_0\{-\log(1-\alpha)\}}+o(1) = e^{-\pi_0\beta_{1;\alpha}}+o(1)$, we thus will show that

$$\prod_{i\in\mathcal{I}_0}\{1-\alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m})\} = e^{-\pi_0\beta_{1;\alpha}}+o(1) \tag{A.19}$$

as $m\to\infty$ and $n\to\infty$. According to Leadbetter *et al.* (1983) (Lemma 6.1.1, p. 125), (A.19) will be deduced from (A.1). The proof is completed. ■

Proof of Proposition 2. Similar to the proof of Proposition 1, it suffices to consider $\mathbb{N}(0,1)$ random variables $\{T_i^{\mathrm{a}}\}_{i=1}^m$, with $\{T_i^{\mathrm{a}} : i \in \mathcal{I}_0\}$ being independent.

To show the first part of (3.6), note that

$$\begin{aligned}
\mathrm{FWER}_k^{\mathrm{a}}(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}}) &= \mathrm{P}(V_m^{\mathrm{a}}(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}}) \ge k) = \mathrm{P}\Big(\sum_{i\in\mathcal{I}_0} \mathrm{I}(|T_i^{\mathrm{a}}| > \mathrm{t}_{\alpha;m;k}^{\mathrm{a}}) \ge k\Big) \\
&= 1 - \mathrm{P}\Big(\sum_{i\in\mathcal{I}_0} \mathrm{I}(|T_i^{\mathrm{a}}| > \mathrm{t}_{\alpha;m;k}^{\mathrm{a}}) \le k-1\Big), \\
\mathrm{FWER}_k(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}}) &= \mathrm{P}(V_m(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}}) \ge k) = \mathrm{P}\Big(\sum_{i\in\mathcal{I}_0} \mathrm{I}(|T_{i;n_1,n_2}^{\mathrm{general}}| > \mathrm{t}_{\alpha;m;k}^{\mathrm{a}}) \ge k\Big) \\
&= 1 - \mathrm{P}\Big(\sum_{i\in\mathcal{I}_0} \mathrm{I}(|T_{i;n_1,n_2}^{\mathrm{general}}| > \mathrm{t}_{\alpha;m;k}^{\mathrm{a}}) \le k-1\Big).
\end{aligned}$$

Define by $\varphi_{V_m(\mathrm{t})}(u)$ and $\varphi_{V_m^{\mathrm{a}}(\mathrm{t})}(u)$ the characteristic functions of $V_m(\mathrm{t})$ and $V_m^{\mathrm{a}}(\mathrm{t})$ respectively, where $u \in \mathbb{R}$. It suffices to show that as $m \to \infty$ and $n \to \infty$,

$$\varphi_{V_m^{\mathrm{a}}(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}})}(u) - \varphi_{V_m(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}})}(u) = o(1). \tag{A.20}$$

Direct calculations give

$$\begin{aligned}
\varphi_{V_m^{\mathrm{a}}(\mathrm{t})}(u) &= \mathrm{E}\{e^{\mathrm{i}\,u V_m^{\mathrm{a}}(\mathrm{t})}\} = \prod_{\ell\in\mathcal{I}_0} \mathrm{E}\{e^{\mathrm{i}\,u \mathrm{I}(|T_\ell^{\mathrm{a}}|>\mathrm{t})}\} \\
&= \prod_{\ell\in\mathcal{I}_0} [\alpha_\ell^{\mathrm{a}}(\mathrm{t}) e^{\mathrm{i}\,u} + \{1 - \alpha_\ell^{\mathrm{a}}(\mathrm{t})\}] = \prod_{\ell\in\mathcal{I}_0} \{1 + \alpha_\ell^{\mathrm{a}}(\mathrm{t})(e^{\mathrm{i}\,u} - 1)\},
\end{aligned}$$

where $\mathrm{i} = \sqrt{-1}$ denotes the imaginary number. By (A.17),

$$\begin{aligned}
\max_{\ell\in\mathcal{I}_0} |\alpha_\ell^{\mathrm{a}}(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}})(e^{\mathrm{i}\,u} - 1)| &= \Big|\frac{\beta_{k;\alpha}}{m}(e^{\mathrm{i}\,u} - 1)\Big| \le 2 \times \frac{\beta_{k;\alpha}}{m} = o(1), \\
\sum_{\ell\in\mathcal{I}_0} |\alpha_\ell^{\mathrm{a}}(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}})(e^{\mathrm{i}\,u} - 1)| &= \sum_{\ell\in\mathcal{I}_0} \Big|\frac{\beta_{k;\alpha}}{m}(e^{\mathrm{i}\,u} - 1)\Big| \le 2 \times \frac{\beta_{k;\alpha}}{m} m_0 \le 2\beta_{k;\alpha} < \infty, \\
\sum_{\ell\in\mathcal{I}_0} \alpha_\ell^{\mathrm{a}}(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}})(e^{\mathrm{i}\,u} - 1) &= \Big\{\sum_{\ell\in\mathcal{I}_0} \alpha_\ell^{\mathrm{a}}(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}})\Big\}(e^{\mathrm{i}\,u} - 1) = \pi_0 \beta_{k;\alpha}(e^{\mathrm{i}\,u} - 1) + o(1).
\end{aligned}$$

According to Chung (2001) (a lemma on p. 208),

$$\varphi_{V_m^{\mathrm{a}}(\mathrm{t}_{\alpha;m;k}^{\mathrm{a}})}(u) \to \exp\{\pi_0 \beta_{k;\alpha}(e^{\mathrm{i}\,u} - 1)\} \tag{A.21}$$

as $m \to \infty$. Similarly,

$$\varphi_{V_m(\mathrm{t})}(u) = \mathrm{E}\{e^{\mathrm{i}\,u V_m(\mathrm{t})}\} = \prod_{\ell\in\mathcal{I}_0} \mathrm{E}\{e^{\mathrm{i}\,u \mathrm{I}(|T_{\ell;n_1,n_2}|>\mathrm{t})}\}$$

$$= \prod_{\ell \in \mathcal{I}_0} [\alpha_{\ell;n_1,n_2}(\mathrm{t}) e^{\mathrm{i}\,u} + \{1 - \alpha_{\ell;n_1,n_2}(\mathrm{t})\}] = \prod_{\ell \in \mathcal{I}_0} \{1 + \alpha_{\ell;n_1,n_2}(\mathrm{t})(e^{\mathrm{i}\,u} - 1)\},$$

Note that as $m \to \infty$ and $n \to \infty$, an application of (A.2) gives

$$\begin{aligned}
\max_{\ell \in \mathcal{I}_0} |\alpha_{\ell;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k})(e^{\mathrm{i}\,u} - 1)| &= o(1), \\
\sum_{\ell \in \mathcal{I}_0} |\alpha_{\ell;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k})(e^{\mathrm{i}\,u} - 1)| &\leq M < \infty, \\
\sum_{\ell \in \mathcal{I}_0} \alpha_{\ell;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k})(e^{\mathrm{i}\,u} - 1) &\to \pi_0 \beta_{k;\alpha}(e^{\mathrm{i}\,u} - 1),
\end{aligned}$$

Applying Chung (2001) (a lemma on p. 208) again implies

$$\varphi_{V_m(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k})}(u) \to \exp\{\pi_0 \beta_{k;\alpha}(e^{\mathrm{i}\,u} - 1)\}. \tag{A.22}$$

Thus (A.21) and (A.22) imply (A.20).

To show the second part of (3.6), note that (A.21) yields $V^{\mathrm{a}}_m(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) \xrightarrow{\mathcal{D}} \text{Poisson}(\pi_0 \beta_{k;\alpha})$, where Poisson($\beta$) denotes the Poisson random variable with the parameter β. Thus as $m \to \infty$,

$$\begin{aligned}
\text{FWER}^{\mathrm{a}}_k(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) &= \mathrm{P}(\text{Poisson}(\pi_0 \beta_{k;\alpha}) \geq k) + o(1) \\
&= G_k(\pi_0 \beta_{k;\alpha}) + o(1).
\end{aligned}$$

Since $G_k(\beta)$ is monotone increasing in $\beta \in (0, \infty)$, we obtain $G_k(\pi_0 \beta_{k;\alpha}) \leq G_k(\beta_{k;\alpha})$. This combined with (3.5) completes the proof. ■

Proof of Proposition 3. Consider $H_{1,i} : \mu_{X;i} > \mu_{Y;i}$; the two-sided alternative can be treated similarly. It suffices to show

$$\varsigma^{\mathrm{a}}_{\alpha;n} - \varsigma_{\alpha;n} = o_n(1), \tag{A.23}$$

$$\widehat{\text{FDR}}(\tau^{\mathrm{a}}_{\alpha;m;n}) = \alpha + o_n(1) + O_{\mathrm{P}}(m^{-1/2}), \tag{A.24}$$

where $o_n(1)$ denotes a term converging to zero as $n \to \infty$.

To show (A.23), let $c_n = F^{-1}_{0;T}(1 - \varsigma^{\mathrm{a}}_{\alpha;n}; n)$ and $d_n = \Phi^{-1}(1 - \varsigma^{\mathrm{a}}_{\alpha;n})$. Then

$$1 - F_{0;T}(c_n; n) = \varsigma^{\mathrm{a}}_{\alpha;n} = 1 - \Phi(d_n). \tag{A.25}$$

By condition (3.7) and Cao (2007) (Theorem 1.2), we have

$$\frac{1 - F_{0;T}(d_n; n)}{1 - \Phi(d_n)} \to 1. \tag{A.26}$$

Since $T_{i;n_1,n_2}^{\text{general}} \xrightarrow{\mathcal{D}} \mathbb{N}(0, 1)$ under $H_{0,i}$, we have $1 - F_{0;T}(x; n) \to 1 - \Phi(x)$ for any x. By (A.25),

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1 - \Phi(c_n)}{1 - \Phi(d_n)} = \lim_{n\to\infty} \frac{1 - F_{0;T}(c_n; n)}{1 - F_{0;T}(d_n; n)} \\
&= \lim_{n\to\infty} \frac{1 - F_{0;T}(c_n; n)}{1 - \Phi(d_n)} \frac{1 - \Phi(d_n)}{1 - F_{0;T}(d_n; n)} = \lim_{n\to\infty} \frac{1 - \Phi(d_n)}{1 - F_{0;T}(d_n; n)} = 1,
\end{aligned}$$

which implies

$$c_n - d_n = F_{0;T}^{-1}(1 - \varsigma_{\alpha;n}^{\text{a}}; n) - \Phi^{-1}(1 - \varsigma_{\alpha;n}^{\text{a}}) = o_n(1). \tag{A.27}$$

Then Jing *et al.* (2014) (result (A.6)) together with (A.27) imply $H(\varsigma_{\alpha;n}^{\text{a}}; n) - H(\varsigma_{\alpha;n}; n) = o_n(1)$. Since $H'(t; n)$ is bounded below for t in an open interval with endpoints $\varsigma_{\alpha;n}$ and $\varsigma_{\alpha;n}^{\text{a}}$, $\varsigma_{\alpha;n}^{\text{a}} - \varsigma_{\alpha;n} = o_n(1)$ holds.

We now show (A.24). By the definition of $\tau_{\alpha;m;n}$, $\widehat{\text{FDR}}(\tau_{\alpha;m;n}) = \alpha$, which yields

$$\begin{aligned}
\widehat{\text{FDR}}(\tau_{\alpha;m;n}^{\text{a}}) - \alpha &= \widehat{\text{FDR}}(\tau_{\alpha;m;n}^{\text{a}}) - \widehat{\text{FDR}}(\varsigma_{\alpha;n}^{\text{a}}) \\
&\quad + \widehat{\text{FDR}}(\varsigma_{\alpha;n}^{\text{a}}) - \widehat{\text{FDR}}(\varsigma_{\alpha;n}) + \widehat{\text{FDR}}(\varsigma_{\alpha;n}) - \widehat{\text{FDR}}(\tau_{\alpha;m;n}).
\end{aligned} \tag{A.28}$$

Utilizing Jing *et al.* (2014) (results (A.10), (A.11) and (A.9)) yields

$$\begin{aligned}
\widehat{\text{FDR}}(\tau_{\alpha;m;n}^{\text{a}}) - \widehat{\text{FDR}}(\varsigma_{\alpha;n}^{\text{a}}) &= O_{\text{P}}(m^{-1/2}), \\
\widehat{\text{FDR}}(\varsigma_{\alpha;n}^{\text{a}}) - \widehat{\text{FDR}}(\varsigma_{\alpha;n}) &= o_n(1) + O_{\text{P}}(m^{-1/2}), \\
\widehat{\text{FDR}}(\tau_{\alpha;m;n}) - \widehat{\text{FDR}}(\varsigma_{\alpha;n}) &= O_{\text{P}}(m^{-1/2}),
\end{aligned} \tag{A.29}$$

respectively, where the second equality also utilizes (A.23). Substituting (A.29) into (A.28), we get (A.24).

Finally, an application of Storey *et al.* (2004) (Theorem 6) shows that

$$\text{P}(\text{FDR}(\tau_{\alpha;m;n}^{\text{a}}) \le \widehat{\text{FDR}}(\tau_{\alpha;m;n}^{\text{a}})) \to 1. \tag{A.30}$$

By (A.30), together with (A.24), we obtain $\text{FDR}(\tau_{\alpha;m;n}^{\text{a}}) \le \alpha + o(1)$. This completes the proof. ■

Proof of Proposition 4. For the critical value $\mathrm{t}^{\mathrm{a}}_{\alpha;m}$ given in (3.2), we observe

$$\begin{aligned}\mathrm{FWER}_1(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) &= \mathrm{P}(\cup_{i\in\mathcal{I}_0}\{|T^{\mathrm{general}}_{i;n_1,n_2}| > \mathrm{t}^{\mathrm{a}}_{\alpha;m}\}) \\ &\le \sum_{i\in\mathcal{I}_0}\mathrm{P}(|T^{\mathrm{general}}_{i;n_1,n_2}| > \mathrm{t}^{\mathrm{a}}_{\alpha;m}) \\ &= \sum_{i\in\mathcal{I}_0}\alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m}) \\ &= \pi_0\beta_{1;\alpha} + o(1),\end{aligned}$$

where the last equality comes from (A.1). ■

Proof of Proposition 5. For the critical value $\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}$ given in (3.4), an application of Markov inequality gives

$$\begin{aligned}\mathrm{FWER}_k(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k}) &\le \frac{\mathrm{E}\{V_m(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k})\}}{k} \\ &= \frac{\sum_{i\in\mathcal{I}_0}\mathrm{P}(|T^{\mathrm{general}}_{i;n_1,n_2}| > \mathrm{t}^{\mathrm{a}}_{\alpha;m;k})}{k} \\ &= \frac{\sum_{i\in\mathcal{I}_0}\alpha_{i;n_1,n_2}(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k})}{k} \\ &= \pi_0\beta_{k;\alpha}/k + o(1),\end{aligned}$$

where the last equality is obtained from (A.2). ■

Derivation of (2.6). Under $H_{0,i}$ in (2.2), (2.5) becomes

$$\begin{aligned}T^{\mathrm{pool}}_{i;n_1,n_2} &= \frac{\bar{\varepsilon}_i - \bar{e}_i}{s_{\mathrm{pool}_{X;Y};i}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \\ &= \frac{\bar{\varepsilon}_i - \bar{e}_i}{\sqrt{\frac{\sigma^2_{\varepsilon;i}}{n_1}+\frac{\sigma^2_{e;i}}{n_2}}}\frac{\sqrt{(\frac{\sigma^2_{\varepsilon;i}}{n_1}+\frac{\sigma^2_{e;i}}{n_2})/(\frac{1}{n_1}+\frac{1}{n_2})}}{s_{\mathrm{pool}_{X;Y};i}} \\ &= \frac{\bar{\varepsilon}_i - \bar{e}_i}{\sqrt{\frac{\sigma^2_{\varepsilon;i}}{n_1}+\frac{\sigma^2_{e;i}}{n_2}}}\frac{\sqrt{(1-\rho)\sigma^2_{\varepsilon;i}+\rho\sigma^2_{e;i}}}{s_{\mathrm{pool}_{X;Y};i}}\{1+o(1)\} \qquad \text{(A.31)}\end{aligned}$$

as $n_1 \to \infty$ and $n_2 \to \infty$, where $\bar{\varepsilon}_i = \sum_{j=1}^{n_1}\varepsilon_{i,j}/n_1$ and $\bar{e}_i = \sum_{j=1}^{n_2} e_{i,j}/n_2$.

(i) By CLT, $\frac{\bar{\varepsilon}_i-\bar{e}_i}{\sqrt{\frac{\sigma^2_{\varepsilon;i}}{n_1}+\frac{\sigma^2_{e;i}}{n_2}}} \xrightarrow{\mathcal{D}} \mathbb{N}(0,1)$.

(ii) By law of large numbers, $s^2_{X;i} \xrightarrow{\mathrm{P}} \sigma^2_{\varepsilon;i}$ and $s^2_{Y;i} \xrightarrow{\mathrm{P}} \sigma^2_{e;i}$, and thus

$$\begin{aligned} s^2_{\mathrm{pool}_{X;Y};i} &= \frac{(n_1-1)s^2_{X;i} + (n_2-1)s^2_{Y;i}}{(n_1+n_2-2)} \\ &\xrightarrow{\mathrm{P}} \rho\sigma^2_{\varepsilon;i} + (1-\rho)\sigma^2_{e;i}. \end{aligned}$$

This combined with (A.31) and (2.7) implies that $T^{\mathrm{pool}}_{i;n_1,n_2} \xrightarrow{\mathcal{D}} \mathbb{N}(0,1) \cdot \sqrt{\frac{(1-\rho)\sigma^2_{\varepsilon;i}+\rho\sigma^2_{e;i}}{\rho\sigma^2_{\varepsilon;i}+(1-\rho)\sigma^2_{e;i}}} = \mathbb{N}(0,1) \cdot \sigma_{\rho;\theta_{(\varepsilon,e);i}}$. ■

Appendix B: Extensions of Models (4.10) and (4.12)

More generally, consider observations $\{X_{i,j}\}$ and $\{Y_{i,j}\}$ described by the following model:

$$\begin{aligned} X_{i,j} &= \mu_{X;i} + \varepsilon_{i,j} + \boldsymbol{\gamma}^T_{X;i}\boldsymbol{w}_i, \quad 1 \le i \le m, \ 1 \le j \le n_1, \\ Y_{i,j} &= \mu_{Y;i} + e_{i,j} + \boldsymbol{\gamma}^T_{Y;i}\boldsymbol{w}_i, \quad 1 \le i \le m, \ 1 \le j \le n_2, \end{aligned} \tag{B.1}$$

where $\boldsymbol{w}_i$ are unobserved d_w-dimensional random vectors, with $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\} \overset{\text{i.i.d.}}{\sim} \mathbb{N}(\mathbf{0}, \Sigma_{\boldsymbol{w}})$; for each i, errors $\{\varepsilon_{i,1}, \ldots, \varepsilon_{i,n_1}\} \overset{\text{i.i.d.}}{\sim} \mathbb{N}(0, \sigma^2_{\varepsilon;i})$, errors $\{e_{i,1}, \ldots, e_{i,n_2}\} \overset{\text{i.i.d.}}{\sim} \mathbb{N}(0, \sigma^2_{e;i})$, and $\{(\varepsilon_{i,1}, \ldots, \varepsilon_{i,n_1}), (e_{i,1}, \ldots, e_{i,n_2}), \boldsymbol{w}_i\}$ are mutually independent; $\{(\varepsilon_{i,1}, \ldots, \varepsilon_{i,n_1}; e_{i,1}, \ldots, e_{i,n_2}; \boldsymbol{w}_i) : i \in \mathcal{I}_0\}$ are independent. Clearly, the factor $\boldsymbol{w}_i$ describes both the dependence between the X-group and Y-group, the dependence within the X-group, as well as the dependence within the Y-group, where the amount of the dependence is described by non-random parameters $\boldsymbol{\gamma}_{X;i}$ and $\boldsymbol{\gamma}_{Y;i}$. As seen from (B.2) and (B.3), test statistics (using either $\{T^{\mathrm{general}}_{i;n_1,n_2}\}$ or $\{T^{\mathrm{pool}}_{i;n_1,n_2}\}$ or $T^{\mathrm{pool;A}}_{i;n_1,n_2}$) associated with true nulls continue to be independent.

Case (i). The case of $\boldsymbol{\gamma}_{X;i} = \boldsymbol{\gamma}_{Y;i}$, which includes Model (4.10), indicates that the influence of common factors $\boldsymbol{w}_i$ are identical between the X-group and Y-group. The conclusions on $T^{\mathrm{general}}_{i;n_1,n_2}$, $T^{\mathrm{pool}}_{i;n_1,n_2}$ and $T^{\mathrm{pool;A}}_{i;n_1,n_2}$ are identical to those in Section 4.2.

Case (ii). The case of $\boldsymbol{\gamma}_{X;i} \ne \boldsymbol{\gamma}_{Y;i}$, which includes Model (4.12), indicates that the common factors $\boldsymbol{w}_i$ in the X-group and Y-group are different. The conclusions on $T^{\mathrm{general}}_{i;n_1,n_2}$, $T^{\mathrm{pool}}_{i;n_1,n_2}$ and $T^{\mathrm{pool;A}}_{i;n_1,n_2}$ are identical to those in Section 4.3.

Detailed discussions on the performance of $T^{\text{general}}_{i;n_1,n_2}$, $T^{\text{pool}}_{i;n_1,n_2}$ and $T^{\text{pool;A}}_{i;n_1,n_2}$ are given below.

Case (i): $\boldsymbol{\gamma}_{X;i} = \boldsymbol{\gamma}_{Y;i}$ in model (B.1)**.** This case means that the influence of common factors $\boldsymbol{w}_i$ are identical between the X- and Y-groups. It follows that two-sample t-statistics under $H_{0,i}$ reduce to the following forms,

$$T^{\text{general}}_{i;n_1,n_2} = \frac{\bar{\varepsilon}_i - \bar{e}_i}{\sqrt{s^2_{\varepsilon;i}/n_1 + s^2_{e;i}/n_2}}, \quad T^{\text{pool}}_{i;n_1,n_2} = \frac{\bar{\varepsilon}_i - \bar{e}_i}{s_{\text{pool}_{\varepsilon;e};i}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad T^{\text{pool;A}}_{i;n_1,n_2} = \frac{T^{\text{pool}}_{i;n_1,n_2}}{\sigma_{\rho;\widehat{\theta}_{(\varepsilon,e);i}}}. \tag{B.2}$$

It is interesting to note that **Case (i)** involves dependence between different groups, as well as within a same group, but test statistics (using either $\{T^{\text{general}}_{i;n_1,n_2}\}$ or $\{T^{\text{pool}}_{i;n_1,n_2}\}$ or $T^{\text{pool;A}}_{i;n_1,n_2}$) associated with true nulls are independent.

Under this special case, we can show two distributional results below for the "***general***" two-sample t-statistic $T^{\text{general}}_{i;n_1,n_2}$ under $H_{0,i}$:

(c1′) if $\sigma^2_{\varepsilon;i} = \sigma^2_{e;i}$ and $n_1 = n_2$, then $T^{\text{general}}_{i;n_1,n_2} \sim t_{2n_1-2}$;

(c2′) if $n_1 \to \infty$ and $n_2 \to \infty$, then $T^{\text{general}}_{i;n_1,n_2} \xrightarrow{\mathcal{D}} \mathbb{N}(0, 1)$.

Hence, conclusions of Propositions 1–3 carry through to the "***general***" two-sample t-statistics $\{T^{\text{general}}_{i;n_1,n_2}\}^m_{i=1}$.

As a comparison, for the "*pooled*" two-sample t-statistic $T^{\text{pool}}_{i;n_1,n_2}$ under $H_{0,i}$, we make two conclusions below.

(d1′) If $\sigma^2_{\varepsilon;i} = \sigma^2_{e;i}$, then $T^{\text{pool}}_{i;n_1,n_2} \sim t_{n_1+n_2-2}$. In this case, the results in Propositions 1–3 continue to apply for the "*pooled*" choice $\{T^{\text{pool}}_{i;n_1,n_2}\}^m_{i=1}$.

(d2′) If $n_1 \to \infty$ and $n_2 \to \infty$ such that $n_1/(n_1 + n_2) \to \rho \in (0, 1)$, then (2.6) gives $T^{\text{pool}}_{i;n_1,n_2} \xrightarrow{\mathcal{D}} \mathbb{N}(0, \sigma^2_{\rho;\theta_{(\varepsilon,e);i}})$. Similar to the discussion in Section 3.2, there will be no guarantee in the case of $\sigma_{\rho;\theta_{(\varepsilon,e);i}} > 1$ for achieving level bounds α in (2.11) and (2.12) using $\{T^{\text{pool}}_{i;n_1,n_2}\}^m_{i=1}$.

But according to (B.2), the "***adaptively pooled***" version satisfies $T_{i;n_1,n_2}^{\text{pool;A}} \xrightarrow{\mathcal{D}} \mathbb{N}(0,1)$, and thus the $\mathbb{N}(0,1)$ calibration remains valid for $\{T_{i;n_1,n_2}^{\text{pool;A}}\}_{i=1}^m$.

Case (ii): $\boldsymbol{\gamma}_{X;i} \neq \boldsymbol{\gamma}_{Y;i}$ in model (B.1). This case means that the common factors $\boldsymbol{w}_i$ in the X-group and Y-group are different. The explicit forms of two-sample t-statistics can be derived as follows,

$$T_{i;n_1,n_2}^{\text{general}} = \frac{\overline{\varepsilon}_i - \overline{e}_i + (\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})^T \boldsymbol{w}_i}{\sqrt{s_{\varepsilon;i}^2/n_1 + s_{e;i}^2/n_2}}, \quad T_{i;n_1,n_2}^{\text{pool}} = \frac{\overline{\varepsilon}_i - \overline{e}_i + (\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})^T \boldsymbol{w}_i}{s_{\text{pool}_{\varepsilon;e};i}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}, \quad T_{i;n_1,n_2}^{\text{pool;A}} = \frac{T_{i;n_1,n_2}^{\text{pool}}}{\sigma_{\rho;\widehat{\theta}_{(\varepsilon,e);i}}}. \tag{B.3}$$

which differ from those in (B.2). Again, dependence between different groups, as well as within a same group, exist in the dataset, where the extent of dependence is captured by the magnitude of $(\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})^T \boldsymbol{w}_i \sim \mathbb{N}(0, (\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})^T \Sigma_{\boldsymbol{w}} (\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i}))$, but two-sample t-statistics associated with true nulls remain independent.

In the context of (B.3), we can show two results for the null distribution of the "***general***" two-sample t-statistic $T_{i;n_1,n_2}^{\text{general}}$:

(e1′) if $\sigma_{\varepsilon;i}^2 = \sigma_{e;i}^2 = \sigma_i^2$ and $n_1 = n_2$, then

$$T_{i;n_1,n_2}^{\text{general}} \sim t_{2n_1-2} \times f_1', \quad \text{where } f_1' = \sqrt{1 + \frac{n_1}{2}\frac{(\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})^T \Sigma_{\boldsymbol{w}} (\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})}{\sigma_i^2}};$$

(e2′) if $n_1 \to \infty$ and $n_2 \to \infty$, then

$$T_{i;n_1,n_2}^{\text{general}} = Z \times f_2' \{1 + o_{\mathrm{P}}(1)\} \xrightarrow{\mathrm{P}} \infty, \quad \text{where } f_2' = \sqrt{1 + \frac{n_1 n_2 (\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})^T \Sigma_{\boldsymbol{w}} (\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})}{n_2 \sigma_\varepsilon^2 + n_1 \sigma_e^2}}, \tag{B.4}$$

$Z \sim \mathbb{N}(0,1)$ and $\xrightarrow{\mathrm{P}}$ denotes converges in probability.

We can also show that $T_{i;n_1,n_2}^{\text{pool;A}}$ has the same limit null distribution as $T_{i;n_1,n_2}^{\text{general}}$. For the null distribution of the "*pooled*" two-sample t-statistic $T_{i;n_1,n_2}^{\text{pool}}$, we make two conclusions below.

(f1′) If $\sigma_{\varepsilon;i}^2 = \sigma_{e;i}^2 = \sigma_i^2$, then

$$T_{i;n_1,n_2}^{\text{pool}} \sim t_{n_1+n_2-2} \times f_3', \quad \text{where } f_3' = \sqrt{1 + \frac{n_1 n_2}{n_1+n_2} \frac{(\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})^T \Sigma_{\boldsymbol{w}} (\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})}{\sigma_i^2}}.$$

(f2′) If $n_1 \to \infty$ and $n_2 \to \infty$ such that $n_1/(n_1+n_2) \to \rho \in (0,1)$, then

$$T_{i;n_1,n_2}^{\text{pool}} = Z \times f_4'\{1 + o_{\mathrm{P}}(1)\} \xrightarrow{\mathrm{P}} \infty, \quad \text{where } f_4' = \sqrt{\sigma_{\rho;\theta_{(\varepsilon,e);i}}^2 + \frac{n_1 n_2}{n_1+n_2} \frac{(\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})^T \Sigma_{\boldsymbol{w}} (\boldsymbol{\gamma}_{X;i} - \boldsymbol{\gamma}_{Y;i})}{\rho\sigma_\varepsilon^2 + (1-\rho)\sigma_e^2}}. \tag{B.5}$$

Thus, conclusions of Propositions 1–3 will fail for two-sample t-statistics $\{T_{i;n_1,n_2}^{\text{general}}\}_{i=1}^m$, since the factor f_2' in (B.4) invariably exceeds one. As a comparison, Propositions 1–3 may fail for $\{T_{i;n_1,n_2}^{\text{pool}}\}_{i=1}^m$, particularly when the factor f_4' in (B.5) substantially exceeds one. In the case of $f_2' > f_4'$, the "***adaptively pooled***" versions $\{T_{i;n_1,n_2}^{\text{pool;A}}\}_{i=1}^m$ will not ameliorate $\{T_{i;n_1,n_2}^{\text{pool}}\}_{i=1}^m$.

Appendix C: Figures and Tables in the Paper

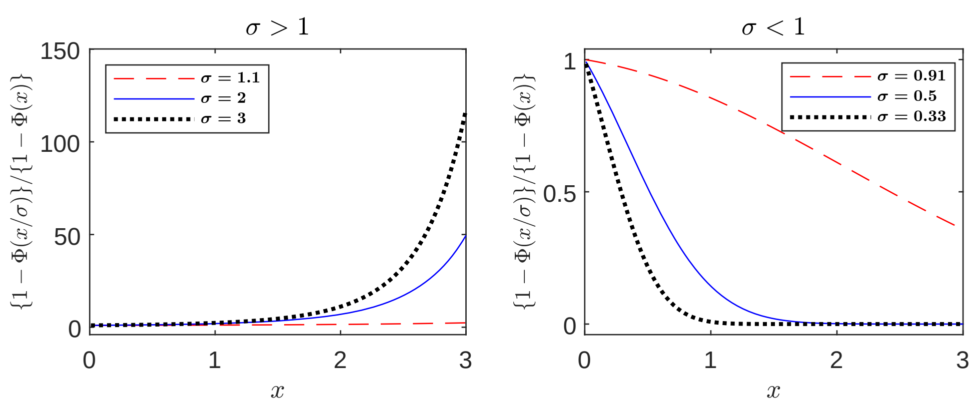


Figure 1: *Plots of* $\{1-\Phi(x/\sigma)\}/\{1-\Phi(x)\}$ *versus* x. *Left panel:* $\sigma > 1$; *right panel:* $\sigma < 1$.

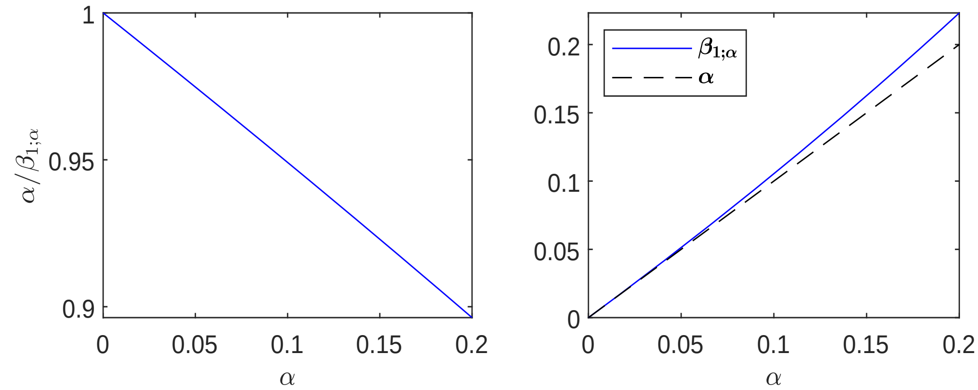
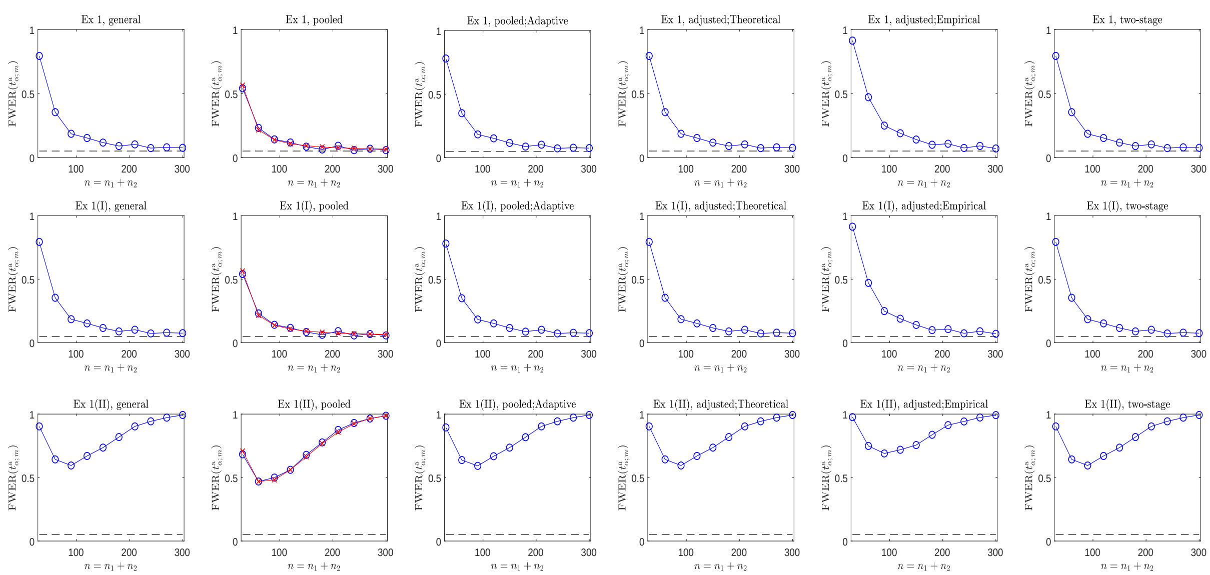


Figure 2: *Left panel: plot of* $\alpha/\beta_{1;\alpha}$ *versus* α*. Right panel: compare plots of* $\beta_{1;\alpha}$ *and* α *versus* α*.*



(a) general (b) pool (c) pool; A (d) adjust; T (e) adjust; E (f) 2_stage

Figure 3: **(Empirical estimates of** $\mathrm{FWER}(\mathrm{t}^{\mathrm{a}}_{\alpha;m})$ **(using** $\circ$**).)** *The horizontal dashed line indicates* α*. Two-sample t-tests in columns* (a)–(f) *are* $T^{\mathrm{general}}_{i;n_1,n_2}$ *in (2.3),* $T^{\mathrm{pool}}_{i;n_1,n_2}$ *in (2.5),* $T^{\mathrm{pool;A}}_{i;n_1,n_2}$ *in (3.12),* $T^{\mathrm{adjust;T}}_{i;n_1,n_2}$ *in (3.13),* $T^{\mathrm{adjust;E}}_{i;n_1,n_2}$ *in (3.16),* $T^{\mathrm{2\_stage}}_{i;n_1,n_2}$ *in (3.19). Top row panels: for **Example*** 1*; middle row panels: for **Example*** 1(I)*; bottom row panels: for **Example*** 1(II).

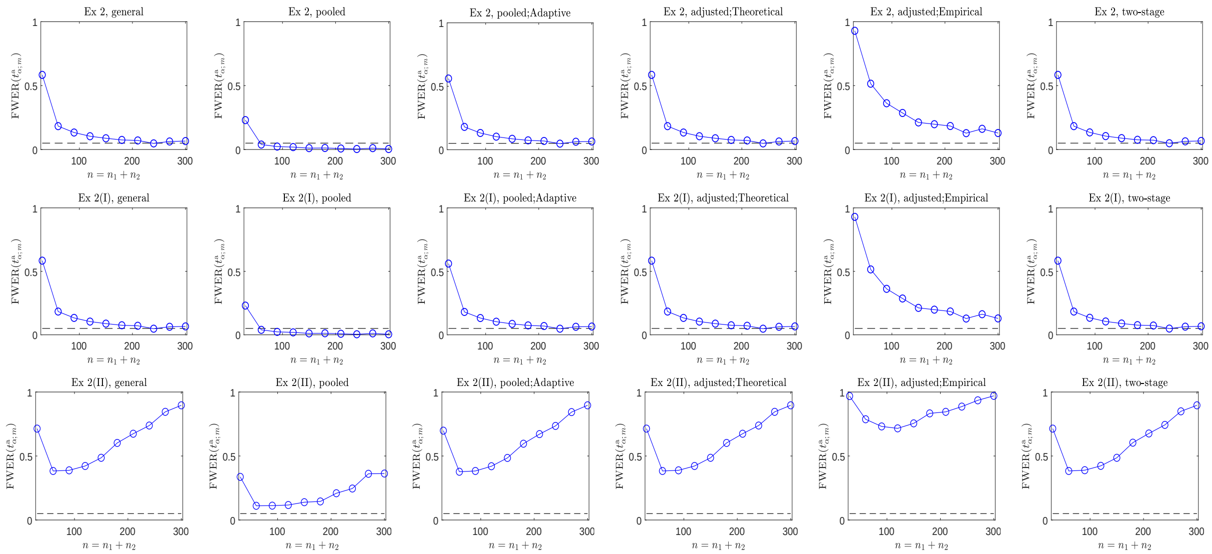


Figure 4: *The caption is similar to that of Figure* 3, *except for* ***Example*** 2, ***Example*** 2(I), ***Example*** 2(II).

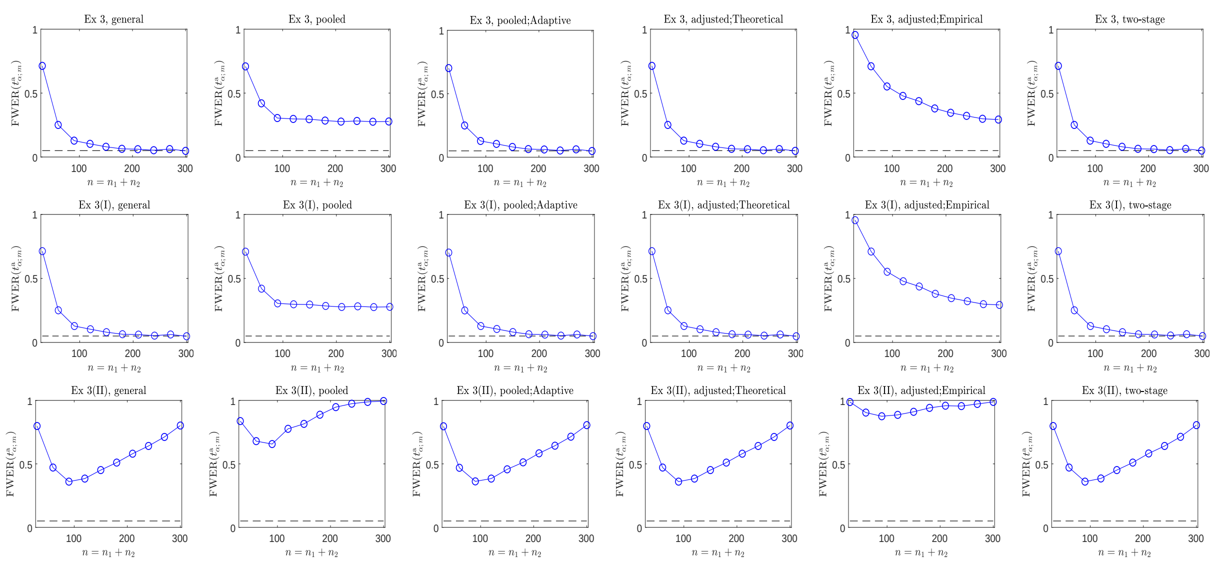


Figure 5: *The caption is similar to that of Figure* 3, *except for* ***Example*** 3, ***Example*** 3(I), ***Example*** 3(II).

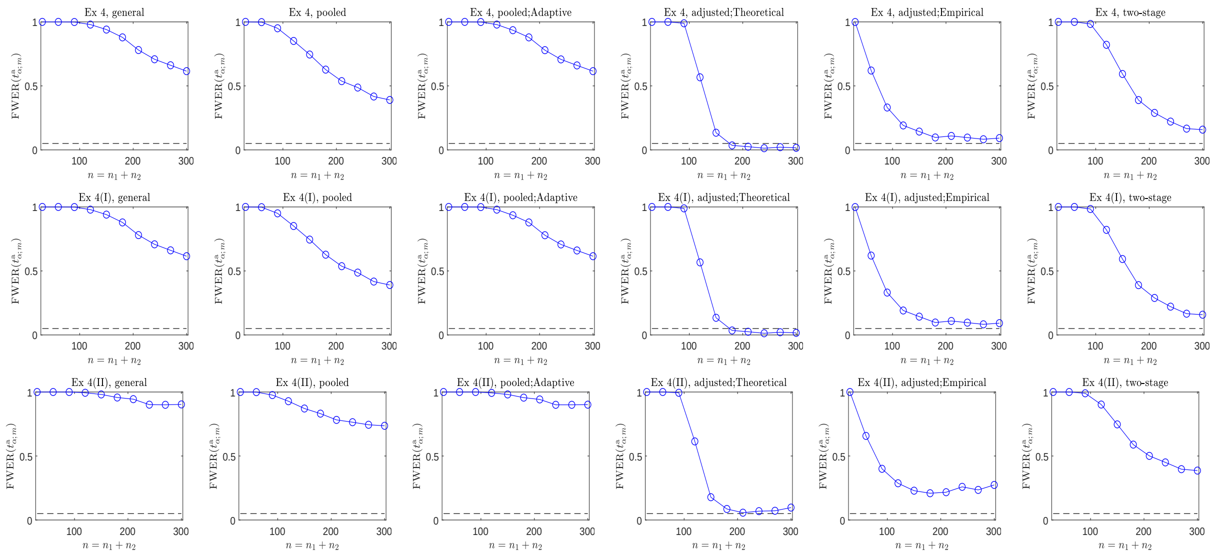


Figure 6: *The caption is similar to that of Figure* 3*, except for* ***Example*** 4*,* ***Example*** 4(I)*,* ***Example*** 4(II)*.*

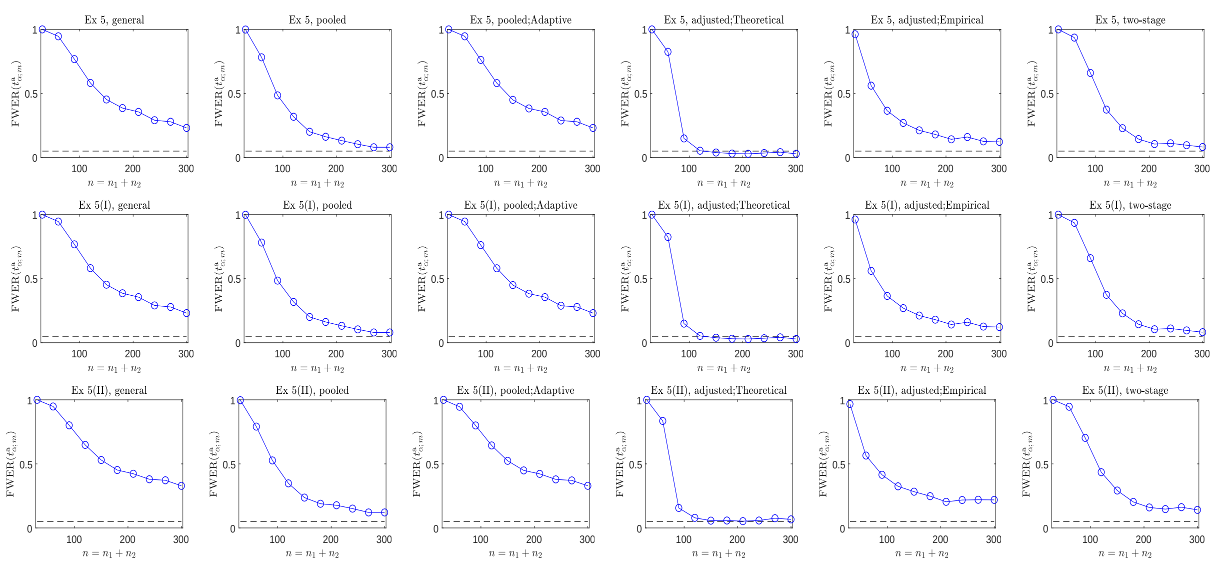
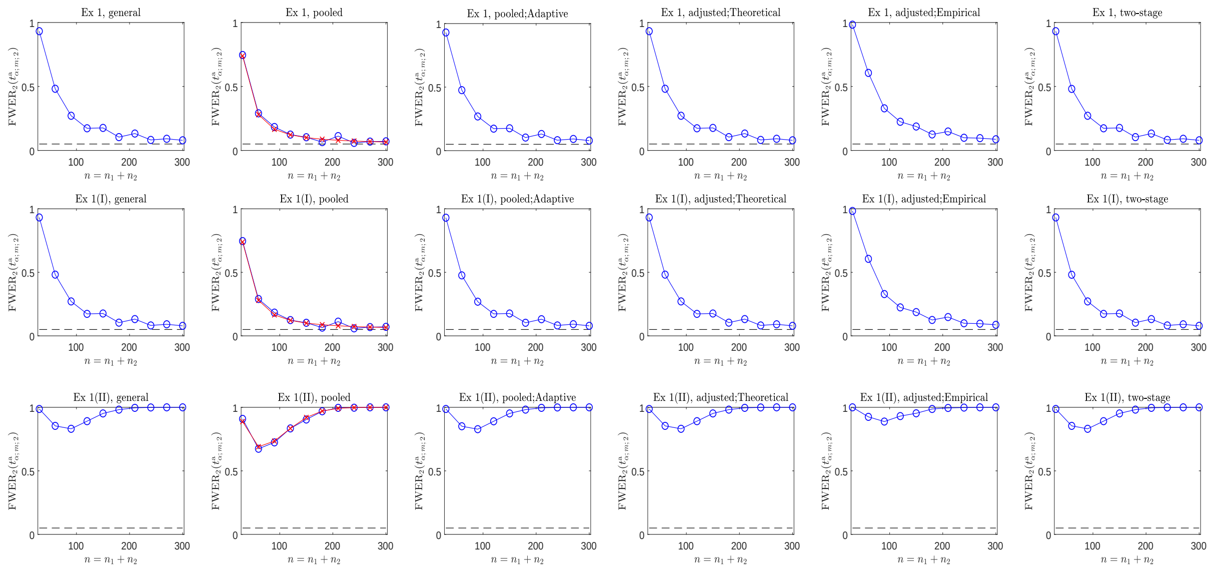


Figure 7: *The caption is similar to that of Figure* 3*, except for* ***Example*** 5*,* ***Example*** 5(I)*,* ***Example*** 5(II)*.*



(a) general (b) pool (c) pool; A (d) adjust; T (e) adjust; E (f) 2_stage

Figure 8: **(Empirical estimates of** $\mathrm{FWER}_k(\mathrm{t}^{\mathrm{a}}_{\alpha;m;k})$ **(using** $\circ$**) with** $k = 2$**.)** *The horizontal dashed line indicates* α*. Two-sample t-tests in columns* (a)–(f) *are* $T^{\text{general}}_{i;n_1,n_2}$ *in (2.3),* $T^{\text{pool}}_{i;n_1,n_2}$ *in (2.5),* $T^{\text{pool;A}}_{i;n_1,n_2}$ *in (3.12),* $T^{\text{adjust;T}}_{i;n_1,n_2}$ *in (3.13),* $T^{\text{adjust;E}}_{i;n_1,n_2}$ *in (3.16),* $T^{\text{2\_stage}}_{i;n_1,n_2}$ *in (3.19). Top row panels: for* ***Example*** 1*; middle row panels: for* ***Example*** 1(I)*; bottom row panels: for* ***Example*** 1(II).

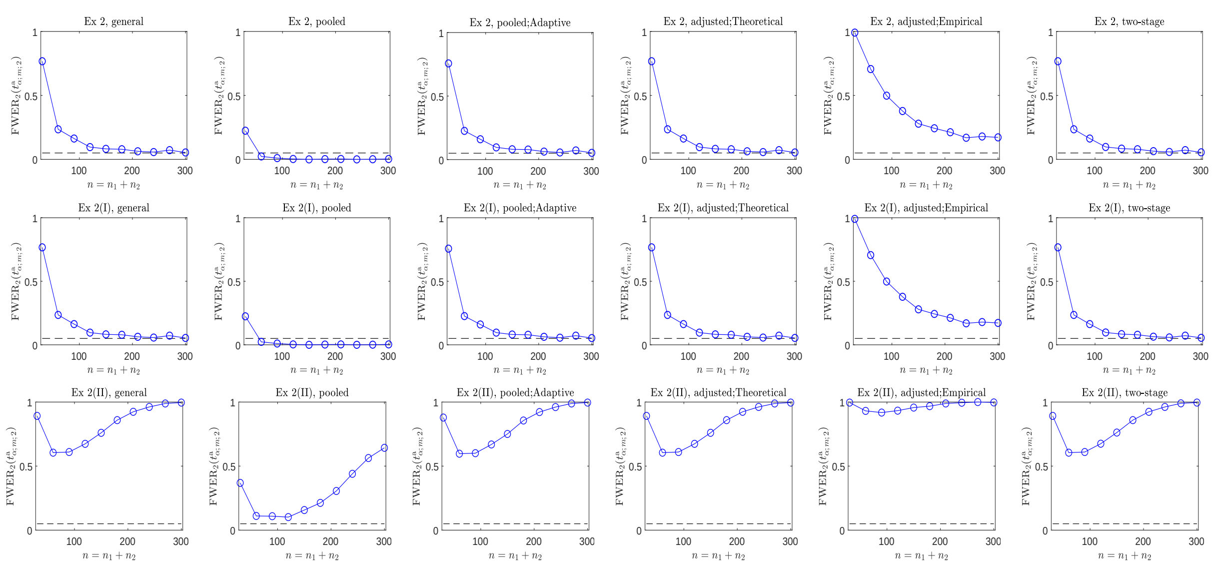


Figure 9: *The caption is similar to that of Figure* 8*, except for* ***Example*** 2*,* ***Example*** 2(I)*,* ***Example*** 2(II).

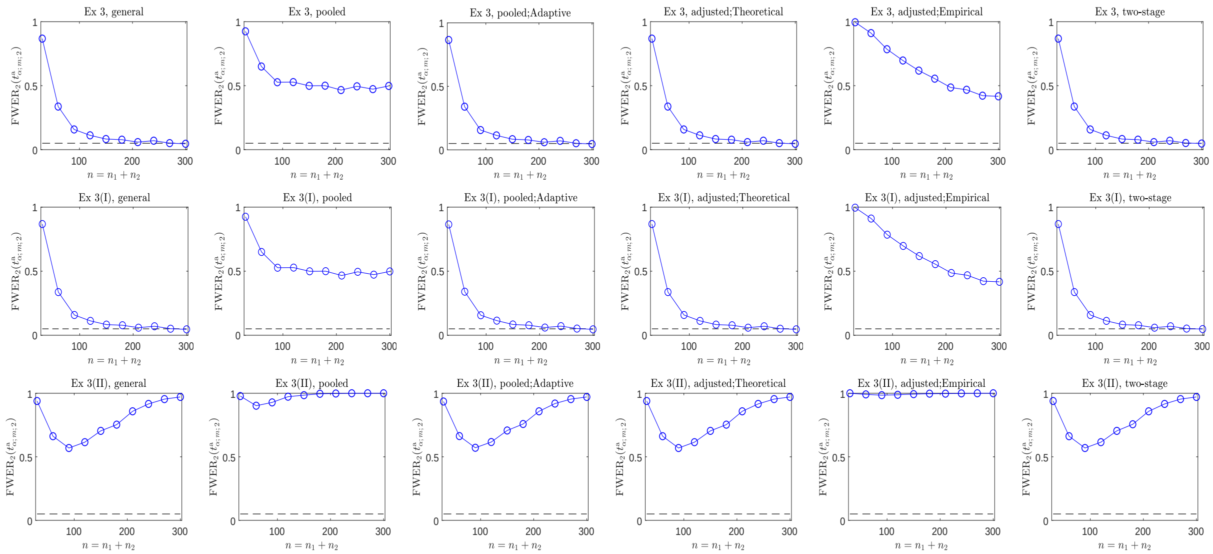


Figure 10: *The caption is similar to that of Figure* 8*, except for* ***Example*** 3, ***Example*** 3(I), ***Example*** 3(II).

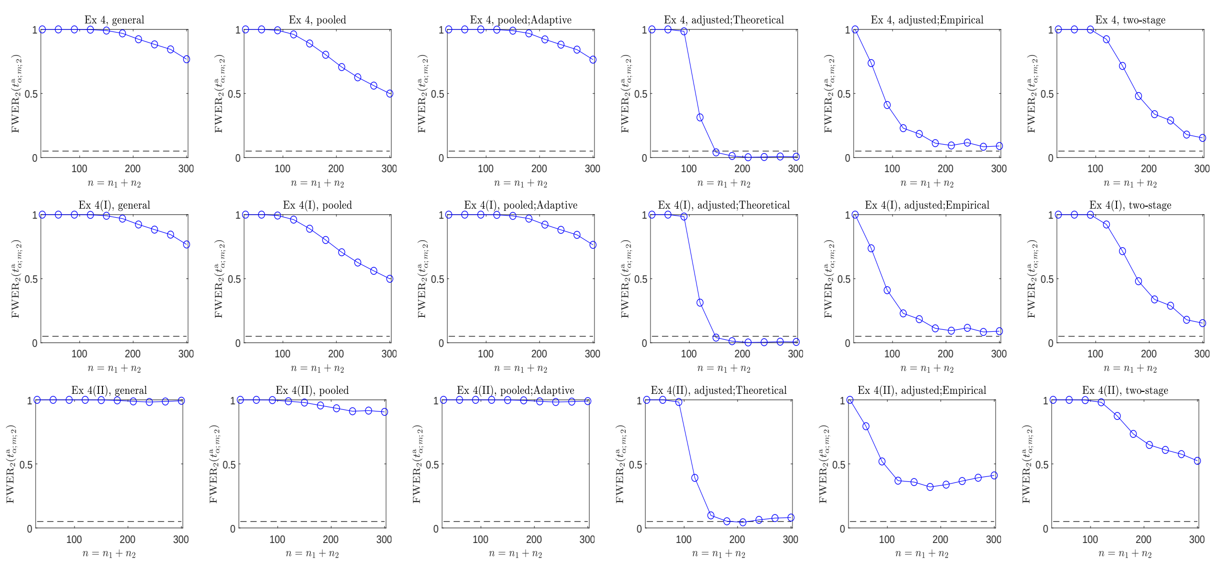


Figure 11: *The caption is similar to that of Figure* 8*, except for* ***Example*** 4, ***Example*** 4(I), ***Example*** 4(II).

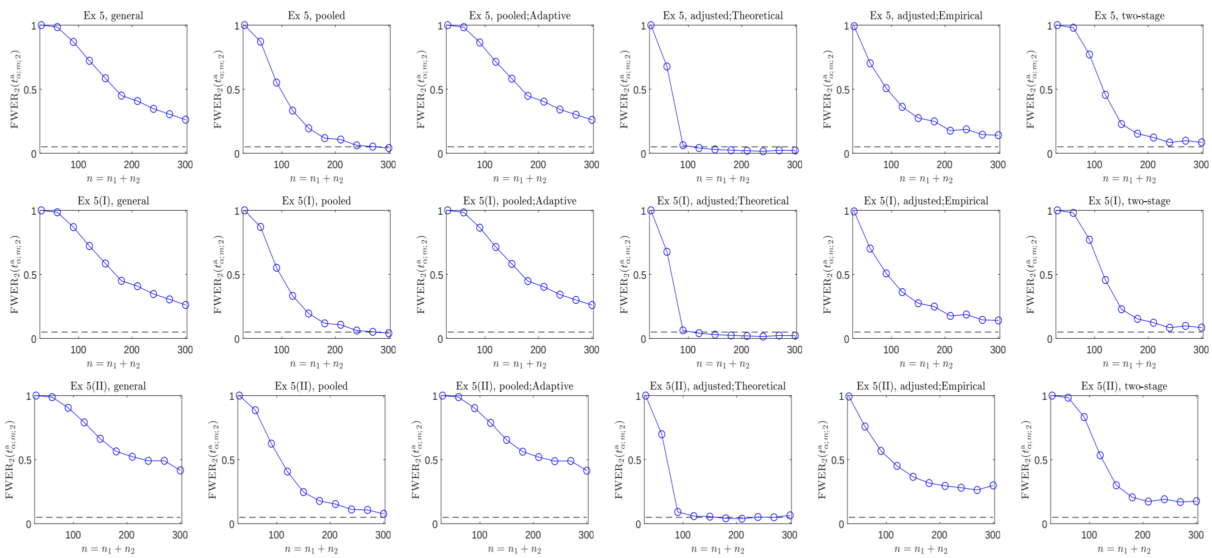


Figure 12: *The caption is similar to that of Figure* 8*, except for* ***Example*** 5, ***Example*** 5(I), ***Example*** 5(II).

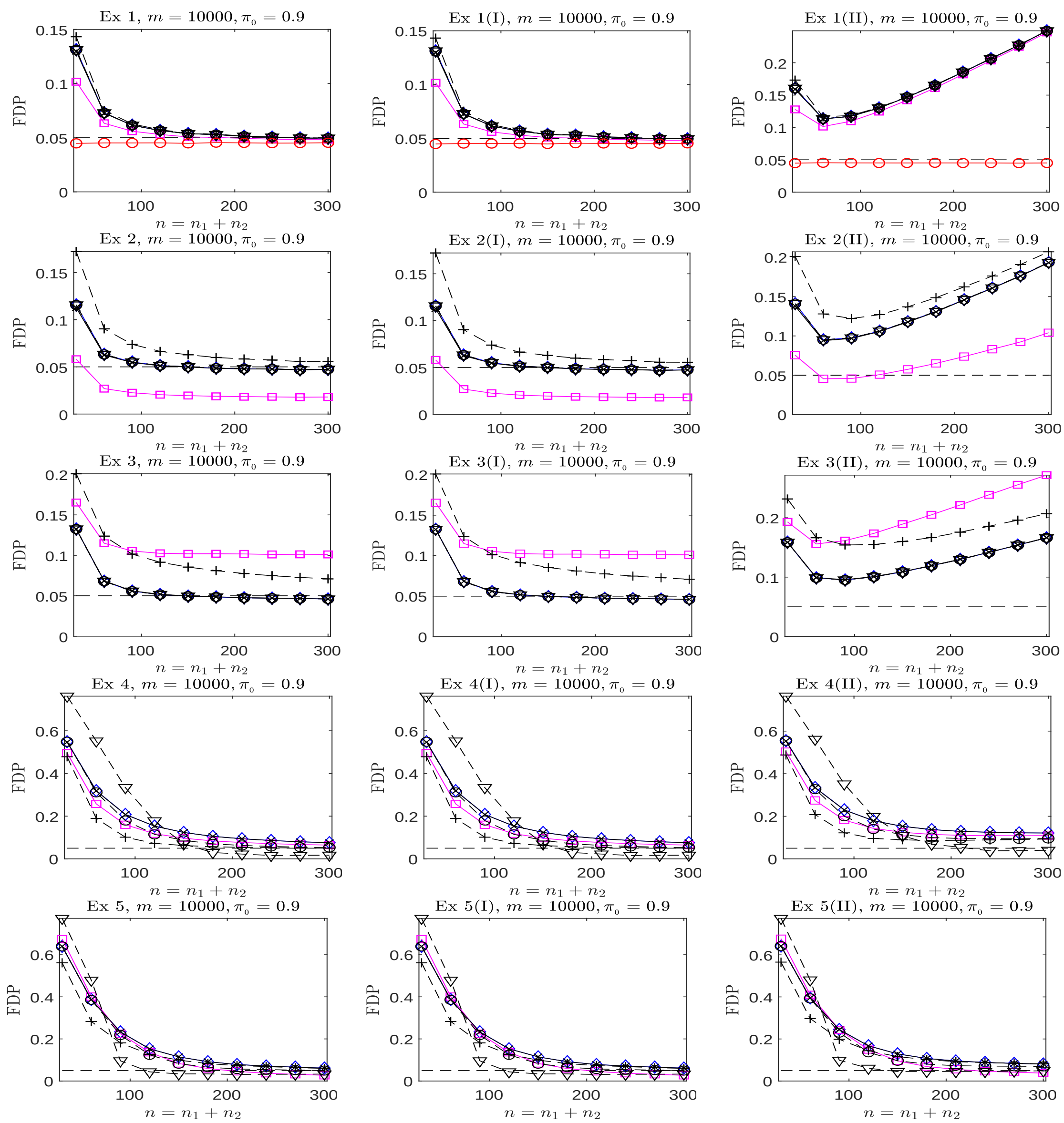


Figure 13: **(Calculated** FDP **of the** BH **procedure.)** *The p-values are calculated via the* $\mathbb{N}(0,1)$ *(using* $-\diamond$*) for* $T^{\text{general}}_{i;n_1,n_2}$*, (exact* $t_{n_1+n_2-2}$*-distribution (using red* $-\circ$*) for* $T^{\text{pool}}_{i;n_1,n_2}$ *in* ***Example*** *1),* $\mathbb{N}(0,1)$ *(using* $-\square$*) for* $T^{\text{pool}}_{i;n_1,n_2}$*,* $\mathbb{N}(0,1)$ *(using* $-\times$*) for* $T^{\text{pool;A}}_{i;n_1,n_2}$*,* $\mathbb{N}(0,1)$ *(using* $--\nabla$ *) for* $T^{\text{adjust;T}}_{i;n_1,n_2}$*,* $\mathbb{N}(0,1)$ *(using* $--+$ *) for* $T^{\text{adjust;E}}_{i;n_1,n_2}$*,* $\mathbb{N}(0,1)$ *(using* $--\circ$ *) for* $T^{\text{2\_stage}}_{i;n_1,n_2}$*. The horizontal dashed line indicates* α*.*

Table 1: *Quantities in simulations examples in (5.1).*

Example	$\sigma^2_{\rho;\theta_{(\varepsilon,e)};i}$	$\mu_{3,X;i}/n_1^2 - \mu_{3,Y;i}/n_2^2$
1	1	0
2	0.8	0
3	1.25	0
4	1	$16/n_1^2 + 16/n_2^2$
5	1.9769	$32/n_1^2 - (2b_i - 1) \times 128/n_2^2$

Table 2: *Number of genes called differentially expressed at* $\alpha = 0.05$.

data	Efron (2010)	Kim *et al.* (2007)	Bourgon *et al.* (2010)
$m; n_1; n_2$	$6033; 50; 52$	$8648; 27; 17$	$12625; 37; 42$
$T^{\text{general}}_{i;n_1,n_2}$ via $\mathbb{N}(0,1)$	51	565	214
$T^{\text{pool}}_{i;n_1,n_2}$ via $t_{n_1+n_2-2}$	21	196	169
$T^{\text{pool}}_{i;n_1,n_2}$ via $\mathbb{N}(0,1)$	51	436	210
$T^{\text{pool;A}}_{i;n_1,n_2}$ via $\mathbb{N}(0,1)$	51	563	215
$T^{\text{2\_stage}}$ via $\mathbb{N}(0,1)$	50	565	213

References

Bourgon, R., Gentleman, R. and Huber, W. (2010). Independent filtering increases detection power for high-throughput experiments. *Proceedings of the National Academy of Sciences*, 107, 9546–9551.

Cao, H. (2007). Moderate deviations for two sample t-statistics. *ESAIM Probab. Stat.*, 11, 264–271.

Chung, K.L. (2001). *A Course in Probability Theory*, Third Edition. Academic Press.

DasGupta, A. (2008). *Asymptotic Theory of Statistics and Probability*, Springer.

Efron, B. (2010). *Large-Scale Inference. Empirical Bayes methods for estimation, testing, and prediction.* Institute of Mathematical Statistics (IMS) Monographs, 1. Cambridge University Press, Cambridge.

Jing, B., Kong, X., and Zhou, W. (2014). FDR control in multiple testing under non-normality. *Statist. Sinica*, 24, 1879–1899.

Kim, J.H., Dhanasekaran, S.M., Mehra, R., Tomlins, S.A., Gu, W., Yu, J., Kumar-Sinha, C., Cao, X., Dash, A., Wang, L. et al. (2007). Integrative analysis of genomic aberrations associated with prostate cancer progression. *Cancer Research*, 67, 8229–8239.

Leadbetter, M.R., Lindgren, G. and Rootzen, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, N.Y..

Storey, J. D., Taylor, J. E., and Siegmund, D. (2004). Strong control, conservative point estimation, and simultaneous conservative consistency of false discovery rates: a unified approach. *J. Roy. Statist. Soc. Ser.* B, 66, 187–205.