Supplementary Appendix to "NAPA: Neighborhood-Assisted and Posterior-Adjusted Two-sample Inference"

Li Ma[†], Yin Xia[†], and Lexin Li[‡]

[†]*Fudan University and* [‡]*University of California at Berkeley*

This supplement contains the proofs of all theoretical results in Section S1, the technical lemmas and their proofs in Section S2, and additional numerical results and discussions in Section S3.

S1 Proofs

S1.1 Proof of Theorem 1

Proof. We show in this theorem that NAPA weight produces a better ranking of p-values than LAWS weight. As such, NAPA yields a higher power than LAWS with an mFDR no larger than LAWS. For any threshold t, we have that,

$$\begin{aligned}
\mathrm{mFDR}\left(\boldsymbol{\delta}_t^{\mathrm{LAWS}}\right) &= \frac{\mathsf{E}\left[\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{\theta(\boldsymbol{s})=0, p(\boldsymbol{s}) \leq \frac{\pi(\boldsymbol{s})}{1-\pi(\boldsymbol{s})}t\right\}\right]}{\mathsf{E}\left[\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p(\boldsymbol{s}) \leq \frac{\pi(\boldsymbol{s})}{1-\pi(\boldsymbol{s})}t\right\}\right]} \\
&= \frac{\sum_{\boldsymbol{s}\in\mathcal{S}}\left[\{1-\pi(\boldsymbol{s})\}\frac{\pi(\boldsymbol{s})}{1-\pi(\boldsymbol{s})}t\right]}{\sum_{\boldsymbol{s}\in\mathcal{S}}\left[\{1-\pi(\boldsymbol{s})\}\frac{\pi(\boldsymbol{s})}{1-\pi(\boldsymbol{s})}t\right] + \sum_{\boldsymbol{s}\in\mathcal{S}}\left[\pi(\boldsymbol{s})F_1\left\{\left.\frac{\pi(\boldsymbol{s})}{1-\pi(\boldsymbol{s})}t\right|\boldsymbol{s}\right\}\right]} \\
&= \frac{\sum_{\boldsymbol{s}\in\mathcal{S}}\pi(\boldsymbol{s})t}{\sum_{\boldsymbol{s}\in\mathcal{S}}\pi(\boldsymbol{s})t + \sum_{\boldsymbol{s}\in\mathcal{S}}\left[\pi(\boldsymbol{s})F_1\left\{\left.\frac{\pi(\boldsymbol{s})}{1-\pi(\boldsymbol{s})}t\right|\boldsymbol{s}\right\}\right]}.
\end{aligned}$$

In addition, we have,

$$
\begin{aligned}
\text{mFDR}\left(\boldsymbol{\delta}_t^{\text{NAPA}}\right) &= \frac{\mathsf{E}\left[\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{\theta(\boldsymbol{s})=0, p(\boldsymbol{s}) \le \frac{\pi(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi(\boldsymbol{s},U(\boldsymbol{s}))}t\right\}\right]}{\mathsf{E}\left[\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p(\boldsymbol{s}) \le \frac{\pi(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi(\boldsymbol{s},U(\boldsymbol{s}))}t\right\}\right]} \\
&= \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}\left(\mathsf{E}\left[\mathsf{I}\left\{\theta(\boldsymbol{s})=0, p(\boldsymbol{s}) \le \frac{\pi(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi(\boldsymbol{s},U(\boldsymbol{s}))}t\right\}|U(\boldsymbol{s})\right]\right)}{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}\left(\mathsf{E}\left[\mathsf{I}\left\{p(\boldsymbol{s}) \le \frac{\pi(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi(\boldsymbol{s},U(\boldsymbol{s}))}t\right\}|U(\boldsymbol{s})\right]\right)} \\
&= \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}\left[\{1-\pi(\boldsymbol{s},U(\boldsymbol{s}))\}\, w(\boldsymbol{s},U(\boldsymbol{s}))\right] t}{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}\left[\{1-\pi(\boldsymbol{s},U(\boldsymbol{s}))\}\, w(\boldsymbol{s},U(\boldsymbol{s}))\right] t + \sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}\left[\pi(\boldsymbol{s},U(\boldsymbol{s}))F_1\left\{w(\boldsymbol{s},U(\boldsymbol{s}))t|\boldsymbol{s}\right\}\right]} \\
&= \frac{\mathsf{E}\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \pi(\boldsymbol{s},U(\boldsymbol{s}))t\right\}}{\mathsf{E}\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \pi(\boldsymbol{s},U(\boldsymbol{s}))t\right\} + \mathsf{E}\left(\sum_{\boldsymbol{s}\in\mathcal{S}} \left[\pi(\boldsymbol{s},U(\boldsymbol{s}))F_1\left\{w(\boldsymbol{s},U(\boldsymbol{s}))t|\boldsymbol{s}\right\}\right]\right)} \\
&= \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \pi(\boldsymbol{s})t}{\sum_{\boldsymbol{s}\in\mathcal{S}} \pi(\boldsymbol{s})t + \mathsf{E}\left(\sum_{\boldsymbol{s}\in\mathcal{S}} \left[\pi(\boldsymbol{s},U(\boldsymbol{s}))F_1\left\{w(\boldsymbol{s},U(\boldsymbol{s}))t|\boldsymbol{s}\right\}\right]\right)}.
\end{aligned}
$$

Let $t = t_{\text{LAWS}}$ and note that $g_{\boldsymbol{s}}(x) = xF_1\left\{xt_{\text{LAWS}}/(1-x)|\boldsymbol{s}\right\}$ is convex for $x \le 1/(1+t_{\text{LAWS}})$. Let event $A_{\boldsymbol{s}} = \{U(\boldsymbol{s}) : \pi(\boldsymbol{s},U(\boldsymbol{s})) \le 1/(1+t_{\text{LAWS}})\}$ and denote by $A_{\boldsymbol{s}}^c$ the complement of $A_{\boldsymbol{s}}$. Then we have that,

$$
\begin{aligned}
&\mathsf{E}\left[\pi(\boldsymbol{s},U(\boldsymbol{s}))F_1\left\{w(\boldsymbol{s},U(\boldsymbol{s}))t_{\text{LAWS}}|\boldsymbol{s}\right\}\right] \\
&= \mathsf{E}\left[\pi(\boldsymbol{s},U(\boldsymbol{s}))F_1\left\{w(\boldsymbol{s},U(\boldsymbol{s}))t_{\text{LAWS}}|\boldsymbol{s}\right\}\mathsf{I}(A_{\boldsymbol{s}})\right] + \mathsf{E}\left[\pi(\boldsymbol{s},U(\boldsymbol{s}))F_1\left\{w(\boldsymbol{s},U(\boldsymbol{s}))t_{\text{LAWS}}|\boldsymbol{s}\right\}\mathsf{I}(A_{\boldsymbol{s}}^c)\right] \\
&\ge \mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\} F_1\left[\left.\frac{\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}t_{\text{LAWS}}}{1-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}}\right|\boldsymbol{s}\right]\Pr(A_{\boldsymbol{s}}) + \mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}^c\right\}\Pr(A_{\boldsymbol{s}}^c).
\end{aligned}
$$

Then, by the concavity and non-decreasing of F_1, we have uniformly for $\boldsymbol{s} \in \mathcal{S}$ that,

$$
\begin{aligned}
&F_1\left\{\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right\} - F_1\left[\left.\frac{\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}t_{\text{LAWS}}}{1-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}}\right|\boldsymbol{s}\right] \\
&\le F_1'\left[\left.\frac{\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}t_{\text{LAWS}}}{1-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}}\right|\boldsymbol{s}\right]\left[\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})} - \frac{\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}t_{\text{LAWS}}}{1-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}}\right].
\end{aligned}
$$

Note that

$$
\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})} - \frac{\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}t_{\text{LAWS}}}{1-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}} = \frac{\pi(\boldsymbol{s}) - \mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}}{\{1-\pi(\boldsymbol{s})\}\left[1-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}\right]}t_{\text{LAWS}}.
$$

Then, we have that,

$$
\begin{aligned}
&\frac{\pi(\boldsymbol{s})F_1\left\{\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right\}-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}F_1\left[\left.\frac{\mathsf{E}\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\}t_{\text{LAWS}}}{1-\mathsf{E}\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\}}\right|\boldsymbol{s}\right]\Pr(A)}{\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}^c\right\}\Pr(A_{\boldsymbol{s}}^c)}-1\\
&=\frac{\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}\Pr(A_{\boldsymbol{s}})\left(F_1\left\{\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right\}-F_1\left[\left.\frac{\mathsf{E}\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\}t_{\text{LAWS}}}{1-\mathsf{E}\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\}}\right|\boldsymbol{s}\right]\right)}{\pi(\boldsymbol{s})-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}\Pr(A)}-1\\
&\quad+F_1\left\{\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right\}\\
&\le\frac{1}{1-\pi(\boldsymbol{s})}\frac{\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}t_{\text{LAWS}}}{1-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}}F_1'\left[\left.\frac{\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}t_{\text{LAWS}}}{1-\mathsf{E}\left\{\pi(\boldsymbol{s},U(\boldsymbol{s}))|A_{\boldsymbol{s}}\right\}}\right|\boldsymbol{s}\right]-1\\
&\quad+F_1\left\{\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right\}\\
&\le\frac{1}{1-\pi(\boldsymbol{s})}\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}F_1'\left[\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right]-1+F_1\left\{\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right\} \qquad \text{(S1.1)}\\
&=\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{\{1-\pi(\boldsymbol{s})\}^2}F_1'\left[\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right]+F_1\left\{\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right\}-1, \qquad \text{(S1.2)}
\end{aligned}
$$

where (S1.1) is derived by the condition that $yF_1'(y|\boldsymbol{s})$ is non-decreasing in $[0,1]$. Recall that $g_{\boldsymbol{s}}(x)=xF_1\{xt_{\text{LAWS}}/(1-x)|\boldsymbol{s}\}$ and we take derivative with respect to x for the fixed t_{LAWS},

$$
g_{\boldsymbol{s}}'(x)=\frac{xt_{\text{LAWS}}}{(1-x)^2}F_1'\left\{\left.\frac{xt_{\text{LAWS}}}{1-x}\right|\boldsymbol{s}\right\}+F_1\left\{\left.\frac{xt_{\text{LAWS}}}{1-x}\right|\boldsymbol{s}\right\}.
$$

Then we note that we can express (S1.2) $=g_{\boldsymbol{s}}'(\pi(\boldsymbol{s}))-1$. By F_1 and $yF_1'(y|\boldsymbol{s})$ are both non-decreasing, we have $g_{\boldsymbol{s}}'(x)$ is also non-decreasing on $[0,1]$. Hence, we have

$$
g_{\boldsymbol{s}}'(\pi(\boldsymbol{s}))\le\begin{cases}g_{\boldsymbol{s}}'(1-\zeta), \text{if } \pi(\boldsymbol{s})>0.5,\\ g_{\boldsymbol{s}}'(0.5), \text{ otherwise,}\end{cases}
$$

by condition $\pi(\boldsymbol{s})\in[\zeta,1-\zeta]$ for some small constants $\zeta>0$. Recall that $\pi_1=\operatorname{Card}(\{\boldsymbol{s}\in\mathcal{S}:\pi(\boldsymbol{s})>0.5\})/\operatorname{Card}(\mathcal{S})$. Then, by $\pi_1\le\varrho$ and $\{g_{\boldsymbol{s}}'(1-\zeta)-g_{\boldsymbol{s}}'(0.5)\}/\{1-g_{\boldsymbol{s}}'(0.5)\}\le1/\varrho$ for all $\boldsymbol{s}\in\mathcal{S}$, we have $\sum_{\boldsymbol{s}\in\mathcal{S}}\{g_{\boldsymbol{s}}'(\pi(\boldsymbol{s}))-1\}\le0$, which yields

$$
\mathsf{E}\left(\sum_{\boldsymbol{s}\in\mathcal{S}}\left[\pi(\boldsymbol{s},U(\boldsymbol{s}))F_1\left\{w(\boldsymbol{s},U(\boldsymbol{s}))t_{\text{LAWS}}|\boldsymbol{s}\right\}\right]\right)\ge\sum_{\boldsymbol{s}\in\mathcal{S}}\left[\pi(\boldsymbol{s})F_1\left\{\left.\frac{\pi(\boldsymbol{s})t_{\text{LAWS}}}{1-\pi(\boldsymbol{s})}\right|\boldsymbol{s}\right\}\right].
$$

Therefore, we have that,

$$
\operatorname{mFDR}\left(\boldsymbol{\delta}_{t_{\text{LAWS}}}^{\text{NAPA}}\right)\le\operatorname{mFDR}\left(\boldsymbol{\delta}_{t_{\text{LAWS}}}^{\text{LAWS}}\right)\le\alpha.
$$

It implies that, with the oracle threshold of LAWS, NAPA yields a smaller FDR level. Therefore, we have $t_{\text{NAPA}} \geq t_{\text{LAWS}}$. Recall that for a given group of weights v, the power of $\boldsymbol{\delta}_t^v$ is defined as

$$\begin{aligned}\Psi\left(\boldsymbol{\delta}_t^v\right) &= \mathsf{E}\left[\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{\theta(\boldsymbol{s})=1, p_v(\boldsymbol{s})\leq t\right\}\right] \\ &= \sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}\left(\mathsf{E}\left[\mathsf{I}\left\{\theta(\boldsymbol{s})=1, p_v(\boldsymbol{s})\leq t\right\}|U(\boldsymbol{s})\right]\right) = \mathsf{E}\left[\sum_{\boldsymbol{s}\in\mathcal{S}} \pi(\boldsymbol{s},U(\boldsymbol{s}))F_1\{vt|\boldsymbol{s}\}\right].\end{aligned}$$

Therefore, we have,

$$\Psi\left(\boldsymbol{\delta}_{t_{\text{NAPA}}}^{\text{NAPA}}\right) \geq \Psi\left(\boldsymbol{\delta}_{t_{\text{LAWS}}}^{\text{NAPA}}\right) \geq \Psi\left(\boldsymbol{\delta}_{t_{\text{LAWS}}}^{\text{LAWS}}\right),$$

which completes the proof of Theorem 1. □

S1.2 Proof of Theorem 2

Proof. Define the event

$$\begin{aligned}\Lambda_{1,\boldsymbol{s}} = \Bigg\{ & U(\boldsymbol{s}) : \mathsf{Var}_{\{p(\boldsymbol{s}'),U(\boldsymbol{s}')\},\boldsymbol{s}'\in\mathcal{S}}\Big(\sum_{\boldsymbol{s}'\in\mathcal{S}}\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\mathsf{I}\{p(\boldsymbol{s}')>\tau\}\right]|U(\boldsymbol{s})\Big) \\ & = O\Big(\sum_{\boldsymbol{s}'\in\mathcal{S}} \mathsf{Var}_{p(\boldsymbol{s}'),U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\mathsf{I}\{p(\boldsymbol{s}')>\tau\}|U(\boldsymbol{s})\right]\Big)\Bigg\}, \\ \Lambda_{2,\boldsymbol{s}} = \Bigg\{ & U(\boldsymbol{s}) : \mathsf{Var}_{U(\boldsymbol{s}'),\boldsymbol{s}'\in\mathcal{S}}\Big(\sum_{\boldsymbol{s}'\in\mathcal{S}} K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))|U(\boldsymbol{s})\Big) \\ & = O\Big(\sum_{\boldsymbol{s}'\in\mathcal{S}} \mathsf{Var}_{U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))|U(\boldsymbol{s})\right]\Big)\Bigg\},\end{aligned}$$

Under the conditions of Theorem 2, we have $\Pr(\Lambda_{1,\boldsymbol{s}}\cap\Lambda_{2,\boldsymbol{s}}\cap\Lambda_{\boldsymbol{s}}) \to 1$ uniformly for all $\boldsymbol{s}\in\mathcal{S}$ as $\mathcal{S}\to\mathbb{S}$. Hence, we will focus on the event $\tilde{\Lambda}_{\boldsymbol{s}} = \Lambda_{1,\boldsymbol{s}}\cap\Lambda_{2,\boldsymbol{s}}\cap\Lambda_{\boldsymbol{s}}$ in the following.

Recall that for each $\boldsymbol{s}$ and conresponding $\mathcal{I}(\tau) = \{\boldsymbol{s}'\in\mathcal{S} : p(\boldsymbol{s}')>\tau\}$,

$$\begin{aligned}1-\hat{\pi}_\tau(\boldsymbol{s},U(\boldsymbol{s})) &= \frac{1/m\sum_{\boldsymbol{s}'\in\mathcal{I}(\tau)} v_{\boldsymbol{H}}\{(\boldsymbol{s},U(\boldsymbol{s})),(\boldsymbol{s}',U(\boldsymbol{s}'))\}}{(1-\tau)/m\sum_{\boldsymbol{s}'\in\mathcal{S}} v_{\boldsymbol{H}}\{(\boldsymbol{s},U(\boldsymbol{s})),(\boldsymbol{s}',U(\boldsymbol{s}'))\}} \\ &= \frac{1/m\sum_{\boldsymbol{s}'\in\mathcal{I}(\tau)} K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))}{(1-\tau)/m\sum_{\boldsymbol{s}'\in\mathcal{S}} K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))}.\end{aligned} \quad \text{(S1.3)}$$

To show $\hat{\pi}_\tau(\boldsymbol{s}, U(\boldsymbol{s}))$ converges to $\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))$ in probability, we calculate the conditional mean square errors of the numerator and denominator of (S1.3) under event $\tilde{\Lambda}_{\boldsymbol{s}}$ respectively.

Step 1. First, we deal with the numerator of (S1.3).

Step 1.1 For the bias term, note that

$$
\begin{aligned}
&\mathsf{E}_{\{p(\boldsymbol{s}'),U(\boldsymbol{s}')\},\boldsymbol{s}'\in\mathcal{S}}\left\{\sum_{\boldsymbol{s}'\in\mathcal{I}(\tau)} K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))|U(\boldsymbol{s})\right\}\\
=&\sum_{\boldsymbol{s}'\in\mathcal{S}}\left\{\mathsf{E}_{p(\boldsymbol{s}'),U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\mathsf{I}\left\{p(\boldsymbol{s}')>\tau\right\}\right]|U(\boldsymbol{s})\right\}\\
=&\sum_{\boldsymbol{s}'\in\mathcal{S}}\mathsf{E}_{U(\boldsymbol{s}')}\left\{\mathsf{E}_{p(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\mathsf{I}\left\{p(\boldsymbol{s}')>\tau\right\}|U(\boldsymbol{s}'),U(\boldsymbol{s})\right]|U(\boldsymbol{s})\right\}\\
=&\sum_{\boldsymbol{s}'\in\mathcal{S}}\mathsf{E}_{U(\boldsymbol{s}')}\left\{\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\Pr\left\{p(\boldsymbol{s}')>\tau|U(\boldsymbol{s}'),U(\boldsymbol{s})\right\}\right]|U(\boldsymbol{s})\right\}.
\end{aligned}
$$

By Condition (C2), with probability $1-O(m^{-1})$, the Hessian $\boldsymbol{A}$ has bounded eigenvalues at $(\boldsymbol{s}'^{\intercal}, U(\boldsymbol{s}'), U(\boldsymbol{s}))^{\intercal}$ uniformly for all $\boldsymbol{s}'\in\mathcal{S}$. This, together with $\Pr\left\{p(\boldsymbol{s}')>\tau|U(\boldsymbol{s}'),U(\boldsymbol{s})\right\}\in[0,1]$, yields that, the first partial derivatives of $\Pr\left\{p(\boldsymbol{s}')>\tau|U(\boldsymbol{s}'),U(\boldsymbol{s})\right\}$ are bounded at $(\boldsymbol{s}'^{\intercal}, U(\boldsymbol{s}'), U(\boldsymbol{s}))^{\intercal}$ uniformly for all $\boldsymbol{s}'\in\mathcal{S}$. Because $\Pr\left\{p(\boldsymbol{s}')>\tau|U(\boldsymbol{s}'),U(\boldsymbol{s})\right\}$ has continuous first and second partial derivatives at $(\boldsymbol{s}'^{\intercal}, U(\boldsymbol{s}'), U(\boldsymbol{s}))^{\intercal}$, by its Taylor expansion at $(\boldsymbol{s}^{\intercal}, U(\boldsymbol{s}), U(\boldsymbol{s}))^{\intercal}$, we have with probability $1-O(m^{-1})$,

$$
\begin{aligned}
&\sum_{\boldsymbol{s}'\in\mathcal{S}}\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\Pr\left\{p(\boldsymbol{s}')>\tau|U(\boldsymbol{s}'),U(\boldsymbol{s})\right\}\right]\\
=&\Pr\left\{p(\boldsymbol{s})>\tau|U(\boldsymbol{s})\right\}\sum_{\boldsymbol{s}'\in\mathcal{S}}K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\\
+&\sum_{\boldsymbol{s}'\in\mathcal{S}}[(\boldsymbol{s}'-\boldsymbol{s})^{\intercal}, U(\boldsymbol{s}')-U(\boldsymbol{s}), 0]\boldsymbol{v}\cdot K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\\
+&O\left(\sum_{\boldsymbol{s}'\in\mathcal{S}}\|[(\boldsymbol{s}'-\boldsymbol{s})^{\intercal}, U(\boldsymbol{s}')-U(\boldsymbol{s}), 0]^{\intercal}\|_2^2 K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\right),
\end{aligned}
$$

where $\boldsymbol{v}=(v_1,\ldots,v_{b+1},v_{b+2})^{\intercal}$ satisfies $v_j=O(1)$ for $j=1,\ldots,b+2$. Recall the partition of $\boldsymbol{H}$ into $\begin{bmatrix}\boldsymbol{H}_{\boldsymbol{S}} & \boldsymbol{a}\\ \boldsymbol{a}^{\intercal} & h_U^2\end{bmatrix}$ and that $\tilde{h}=h_U^2-\boldsymbol{a}^{\intercal}\boldsymbol{H}_{\boldsymbol{S}}^{-1}\boldsymbol{a}$. Therefore, under the condition that $\boldsymbol{H}_{\boldsymbol{S}}$ is nonsingular and $\tilde{h}\neq 0$, we have that,

$$\boldsymbol{H}^{-1} = \begin{pmatrix} \boldsymbol{H}_{\boldsymbol{S}}^{-1} + \boldsymbol{H}_{\boldsymbol{S}}^{-1}\boldsymbol{a}\tilde{h}^{-1}\boldsymbol{a}^{\intercal}\boldsymbol{H}_{\boldsymbol{S}}^{-1} & -\boldsymbol{H}_{\boldsymbol{S}}^{-1}\boldsymbol{a}\tilde{h}^{-1} \\ -\tilde{h}^{-1}\boldsymbol{a}^{\intercal}\boldsymbol{H}_{\boldsymbol{S}}^{-1} & \tilde{h}^{-1} \end{pmatrix}, \boldsymbol{L} = \begin{pmatrix} \boldsymbol{H}_{\boldsymbol{S}}^{-1/2} & \boldsymbol{0} \\ \tilde{h}^{-1/2}\boldsymbol{a}^{\intercal}\boldsymbol{H}_{\boldsymbol{S}}^{-1} & -\tilde{h}^{-1/2} \end{pmatrix},$$

where $\boldsymbol{L}^{\intercal}\boldsymbol{L} = \boldsymbol{H}^{-1}$. Let $\boldsymbol{H}^{1/2}$ be the unique positive definite square root of $\boldsymbol{H}$ and $\boldsymbol{Q} = \boldsymbol{H}^{-1/2}\boldsymbol{L}^{-1}$. Then we have that $\boldsymbol{Q}^{\intercal}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^{\intercal} = \boldsymbol{I}$ and $\boldsymbol{Q}\begin{bmatrix} \boldsymbol{x} \\ y_{\boldsymbol{x}} \end{bmatrix} = \boldsymbol{H}^{-1/2}\begin{bmatrix} \boldsymbol{s}' - \boldsymbol{s} \\ U(\boldsymbol{s}') - U(\boldsymbol{s}) \end{bmatrix}$, where $\begin{bmatrix} \boldsymbol{x} \\ y_{\boldsymbol{x}} \end{bmatrix} = \boldsymbol{L}\begin{bmatrix} \boldsymbol{s}' - \boldsymbol{s} \\ U(\boldsymbol{s}') - U(\boldsymbol{s}) \end{bmatrix}$. By the symmetry of K, we have that,

$$\begin{aligned} &\int \|[(\boldsymbol{s}' - \boldsymbol{s})^{\intercal}, U(\boldsymbol{s}') - U(\boldsymbol{s}), 0]^{\intercal}\|_2^2 K_{\boldsymbol{H}}(\boldsymbol{s} - \boldsymbol{s}', U(\boldsymbol{s}) - U(\boldsymbol{s}'))d\boldsymbol{s}' \\ &= (|\boldsymbol{H}_{\boldsymbol{S}}|/|\boldsymbol{H}|)^{1/2}\int (\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})\boldsymbol{Q}^{\intercal}\boldsymbol{H}\boldsymbol{Q}(\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})^{\intercal}K\{\boldsymbol{Q}(\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})^{\intercal}\}d\boldsymbol{x}. \end{aligned}$$

By $|\boldsymbol{H}| = \tilde{h}|\boldsymbol{H}_{\boldsymbol{S}}|$, $\|\boldsymbol{H}\|_2 = O\{\mathsf{tr}(\boldsymbol{H})\}$ and Condition (C2), there exists some constant $c > 0$,

$$\begin{aligned} &\lim_{\mathcal{S}\to\mathbb{S}}\Big\{\mathsf{E}_{\{p(\boldsymbol{s}'),U(\boldsymbol{s}')\},\boldsymbol{s}'\in\mathcal{S}}\Big[m^{-1}\sum_{\boldsymbol{s}'\in\mathcal{I}(\tau)} K_{\boldsymbol{H}}(\boldsymbol{s} - \boldsymbol{s}', U(\boldsymbol{s}) - U(\boldsymbol{s}'))|U(\boldsymbol{s})\Big] \\ &\quad - \Pr\{p(\boldsymbol{s}) > \tau|U(\boldsymbol{s})\}\, m^{-1}\sum_{\boldsymbol{s}'\in\mathcal{S}} E_{U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}(\boldsymbol{s} - \boldsymbol{s}', U(\boldsymbol{s}) - U(\boldsymbol{s}'))|U(\boldsymbol{s})\right]\Big\}^2 \\ &\le c\left[\lim_{\mathcal{S}\to\mathbb{S}}\frac{1}{m}\sum_{\boldsymbol{s}'\in\mathcal{S}}\mathsf{E}_{U(\boldsymbol{s}')}\left(\{[(\boldsymbol{s}' - \boldsymbol{s})^{\intercal}, U(\boldsymbol{s}') - U(\boldsymbol{s})]^{\intercal}K_{\boldsymbol{H}}(\boldsymbol{s} - \boldsymbol{s}', U(\boldsymbol{s}) - U(\boldsymbol{s}'))\}\,|U(\boldsymbol{s})\right)\right]^2 \\ &\quad + c\left[\int (\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})(\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})^{\intercal}K\{\boldsymbol{Q}(\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})^{\intercal}\}d\boldsymbol{x}\lim_{\mathcal{S}\to\mathbb{S}}\tilde{h}^{-1/2}\mathsf{tr}(\boldsymbol{H})\right]^2 + c\left[\lim_{\mathcal{S}\to\mathbb{S}} m^{-1}|\boldsymbol{H}|^{-1/2}\right]^2, \end{aligned} \tag{S1.4}$$

uniformly in $\boldsymbol{s} \in \mathcal{S}$. Letting $(\tilde{\boldsymbol{x}}^{\intercal}, \tilde{y})^{\intercal} = \boldsymbol{Q}(\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})^{\intercal} = (\tilde{x}_1, \cdots, \tilde{x}_b, \tilde{y})^{\intercal}$, we have that,

$$\int (\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})(\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})^{\intercal}K\{\boldsymbol{Q}(\boldsymbol{x}^{\intercal}, y_{\boldsymbol{x}})^{\intercal}\}d\boldsymbol{x} = \int (\tilde{\boldsymbol{x}}^{\intercal}, \tilde{y})(\tilde{\boldsymbol{x}}^{\intercal}, \tilde{y})^{\intercal}K(\tilde{\boldsymbol{x}}, \tilde{y})d\boldsymbol{x}.$$

Letting $\boldsymbol{Q} = \begin{bmatrix} \tilde{\boldsymbol{Q}} & \boldsymbol{e} \\ \boldsymbol{v}^{\intercal} & z \end{bmatrix}$, then $\begin{bmatrix} \tilde{\boldsymbol{x}} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{Q}} & \boldsymbol{e} \\ \boldsymbol{v}^{\intercal} & z \end{bmatrix}\begin{bmatrix} \boldsymbol{x} \\ y_{\boldsymbol{x}} \end{bmatrix}$. Henceforth, if $\tilde{y} = 0$, we have $\tilde{\boldsymbol{x}} = (\tilde{\boldsymbol{Q}} - \boldsymbol{e}\boldsymbol{v}^{\intercal}/z)\boldsymbol{x}$. By $|\boldsymbol{Q}| = |\tilde{\boldsymbol{Q}}|(z - \boldsymbol{v}^{\intercal}\tilde{\boldsymbol{Q}}^{-1}\boldsymbol{e}) = \pm 1$, we have $|\tilde{\boldsymbol{Q}} - \boldsymbol{e}\boldsymbol{v}^{\intercal}/z| = |\tilde{\boldsymbol{Q}}|(1 - \boldsymbol{v}^{\intercal}\tilde{\boldsymbol{Q}}^{-1}\boldsymbol{e}/z) = \pm 1/z$. By $\boldsymbol{Q}^{\intercal}\boldsymbol{Q} = \boldsymbol{I}$, we have $\boldsymbol{e}^{\intercal}\boldsymbol{e} + z^2 = 1$. This yields $|z| \le 1$. By Condition (C1) and the unimodality of K, we have uniformly in $i = 1, \cdots, b$ that,

$$\begin{aligned} \int \tilde{x}_i^2 K(\tilde{\boldsymbol{x}}, \tilde{y})d\boldsymbol{x} &\le |z|\int \tilde{x}_i^2 K(\tilde{\boldsymbol{x}}, 0)d\tilde{\boldsymbol{x}} \\ &\le \int_{\|\tilde{\boldsymbol{x}}\|_2 \le 1} K(\tilde{\boldsymbol{x}}, 0)d\tilde{\boldsymbol{x}} + \int_{\|\tilde{\boldsymbol{x}}\|_2 > 1} \tilde{x}_i^2 K(\tilde{\boldsymbol{x}}, 0)d\tilde{\boldsymbol{x}} < \infty. \end{aligned}$$

Meanwhile, by $\boldsymbol{Q}^\intercal\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^\intercal = \boldsymbol{I}$ again, we have $(\tilde{\boldsymbol{Q}} - \boldsymbol{e}\boldsymbol{v}^\intercal/z)^{-1} = \tilde{\boldsymbol{Q}}^\intercal$, and $\boldsymbol{v}^\intercal\tilde{\boldsymbol{Q}}^\intercal + z\boldsymbol{e}^\intercal = (1, 0, \cdots, 0)$. These together with $\boldsymbol{e}^\intercal\boldsymbol{e} + z^2 = 1$ give

$$\left\|\boldsymbol{v}^\intercal(\tilde{\boldsymbol{Q}} - \boldsymbol{e}\boldsymbol{v}^\intercal/z)^{-1}z\right\|_{\max} \le z^2\left\|\boldsymbol{v}^\intercal\tilde{\boldsymbol{Q}}^\intercal\right\|_2 \le 1 + \|z\boldsymbol{e}^\intercal\|_2 \le 2.$$

By Condition (C1), we have that,

$$\begin{aligned}
\int\tilde{y}^2 K(\tilde{\boldsymbol{x}}, \tilde{y})d\boldsymbol{x} = &\int\left|\boldsymbol{v}^\intercal\boldsymbol{x} + zy_{\boldsymbol{x}}\right|^2 K\{\boldsymbol{Q}(\boldsymbol{x}^\intercal, y_{\boldsymbol{x}})^\intercal\}d\boldsymbol{x} \\
&\le |z|\int\left|\boldsymbol{v}^\intercal(\tilde{\boldsymbol{Q}} - \boldsymbol{e}\boldsymbol{v}^\intercal/z)^{-1}\tilde{\boldsymbol{x}}\right|^2 K(\tilde{\boldsymbol{x}}, 0)d\tilde{\boldsymbol{x}} + \int|zy_{\boldsymbol{x}}|^2 K\{\boldsymbol{Q}(\boldsymbol{x}^\intercal, y_{\boldsymbol{x}})^\intercal\}d\boldsymbol{x} \\
&\le 4\int|\boldsymbol{1}_b^\intercal\tilde{\boldsymbol{x}}|^2 K(\tilde{\boldsymbol{x}}, 0)d\tilde{\boldsymbol{x}} + \int y_{\boldsymbol{x}}^2 K\{\boldsymbol{Q}(\boldsymbol{x}^\intercal, y_{\boldsymbol{x}})^\intercal\}d\boldsymbol{x} < \infty.
\end{aligned}$$

Hence, we obtain uniformly in $i = 1, \cdots, b$ that,

$$\left|\int\tilde{x}_i^2 K(\tilde{\boldsymbol{x}}, \tilde{y})d\boldsymbol{x}\right| < \infty, \left|\int\tilde{y}^2 K(\tilde{\boldsymbol{x}}, \tilde{y})d\boldsymbol{x}\right| < \infty.$$

These, together with Condition (C4), yield that the right-hand-side of (S1.4) goes to 0.

Step 1.2 We show that the variance term converges to zero. Under Condition (C3), we have

$$\begin{aligned}
&\mathsf{Var}_{\{p(\boldsymbol{s}'), U(\boldsymbol{s}')\}, \boldsymbol{s}' \in \mathcal{S}}\left\{\sum_{\boldsymbol{s}' \in \mathcal{T}(\tau)} K_{\boldsymbol{H}}(\boldsymbol{s} - \boldsymbol{s}', U(\boldsymbol{s}) - U(\boldsymbol{s}'))|U(\boldsymbol{s})\right\} \\
&= \mathsf{Var}_{\{p(\boldsymbol{s}'), U(\boldsymbol{s}')\}, \boldsymbol{s}' \in \mathcal{S}}\left(\sum_{\boldsymbol{s}' \in \mathcal{S}}\left[K_{\boldsymbol{H}}(\boldsymbol{s} - \boldsymbol{s}', U(\boldsymbol{s}) - U(\boldsymbol{s}'))\mathsf{I}\left\{p\left(\boldsymbol{s}'\right) > \tau\right\}\right]|U(\boldsymbol{s})\right) \\
&= O\left(\sum_{\boldsymbol{s}' \in \mathcal{S}}\mathsf{Var}_{p(\boldsymbol{s}'), U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}(\boldsymbol{s} - \boldsymbol{s}', U(\boldsymbol{s}) - U(\boldsymbol{s}'))\mathsf{I}\left\{p(\boldsymbol{s}') > \tau\right\}|U(\boldsymbol{s})\right]\right) \\
&= O\left(\sum_{\boldsymbol{s}' \in \mathcal{S}}\mathsf{E}_{p(\boldsymbol{s}'), U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}^2(\boldsymbol{s} - \boldsymbol{s}', U(\boldsymbol{s}) - U(\boldsymbol{s}'))\mathsf{I}\left\{p(\boldsymbol{s}') > \tau\right\}|U(\boldsymbol{s})\right]\right) \\
&= O\left\{\sum_{\boldsymbol{s}' \in \mathcal{S}}\mathsf{E}_{U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}^2\left(\boldsymbol{s} - \boldsymbol{s}', U(\boldsymbol{s}) - U(\boldsymbol{s}')\right)|U(\boldsymbol{s})\right]\right\}.
\end{aligned}$$

Note that

$$\begin{aligned}
&\int K_{\boldsymbol{H}}^2(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))d\boldsymbol{s}' = |\boldsymbol{H}_{\boldsymbol{S}}|^{1/2}|\boldsymbol{H}|^{-1}\int K^2\{\boldsymbol{Q}(\boldsymbol{x}^{\mathsf{T}}, y_{\boldsymbol{x}})^{\mathsf{T}}\}d\boldsymbol{x},\\
&\left|\int K^2\{\boldsymbol{Q}(\boldsymbol{x}^{\mathsf{T}}, y_{\boldsymbol{x}})^{\mathsf{T}}\}d\boldsymbol{x}\right| \le M|z|\int K(\tilde{\boldsymbol{x}}, 0)d\tilde{\boldsymbol{x}}\\
\le M&\int_{\|\tilde{\boldsymbol{x}}\|_2\le 1} K(\tilde{\boldsymbol{x}}, 0)d\tilde{\boldsymbol{x}} + M\int_{\|\tilde{\boldsymbol{x}}\|_2> 1} \tilde{\boldsymbol{x}}^{\mathsf{T}}\tilde{\boldsymbol{x}}K(\tilde{\boldsymbol{x}}, 0)d\tilde{\boldsymbol{x}} < \infty,
\end{aligned}$$

where $|K(\tilde{\boldsymbol{x}}, \tilde{y})| \le M < \infty$. Therefore, there exist some constants $c_1, c_2 > 0$, such that

$$\begin{aligned}
&\lim_{\mathcal{S}\to\mathbb{S}} \mathsf{Var}_{\{p(\boldsymbol{s}'),U(\boldsymbol{s}')\},\boldsymbol{s}'\in\mathcal{S}}\left\{\frac{1}{m}\sum_{\boldsymbol{s}'\in\mathcal{T}(\tau)} K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))|U(\boldsymbol{s})\right\}\\
\le\ & c_1\lim_{\mathcal{S}\to\mathbb{S}}\frac{1}{m^2}\sum_{\boldsymbol{s}'\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}^2\left(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}')\right)|U(\boldsymbol{s})\right]\\
\le\ & c_2\int K^2\{\boldsymbol{Q}(\boldsymbol{x}^{\mathsf{T}}, y_{\boldsymbol{x}})^{\mathsf{T}}\}d\boldsymbol{x}\lim_{\mathcal{S}\to\mathbb{S}} m^{-1}\tilde{h}^{-1/2}|\boldsymbol{H}|^{-1/2}. \qquad \text{(S1.5)}
\end{aligned}$$

By $m^{-1}|\boldsymbol{H}|^{-1/2} = o(\tilde{h}^{1/2})$, we have the right-hand-side of (S1.5) also goes to 0.

Combine (S1.4) and (S1.5), and then we have that

$$\begin{aligned}
&\mathsf{E}_{\{p(\boldsymbol{s}'),U(\boldsymbol{s}')\},\boldsymbol{s}'\in\mathcal{S}}\left\{\frac{1}{m}\sum_{\boldsymbol{s}'\in\mathcal{I}(\tau)} K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}')) -\right.\\
&\left.\Pr\{p(\boldsymbol{s})>\tau|U(\boldsymbol{s})\}\frac{1}{m}\sum_{\boldsymbol{s}'\in\mathcal{S}} E_{U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))\right]\middle| U(\boldsymbol{s})\right\}^2 \to 0, \qquad \text{(S1.6)}
\end{aligned}$$

in probability as $\mathcal{S}\to\mathbb{S}$.

Step 2. Next, we deal with the denominator of (S1.3).

Note that

$$\begin{aligned}
&\mathsf{E}_{\{U(\boldsymbol{s}'),\boldsymbol{s}'\in\mathcal{S}\}}\left\{\sum_{\boldsymbol{s}'\in\mathcal{S}} K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))|U(\boldsymbol{s})\right\}\\
=&\sum_{\boldsymbol{s}'\in\mathcal{S}}\mathsf{E}_{U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}', U(\boldsymbol{s})-U(\boldsymbol{s}'))|U(\boldsymbol{s})\right],
\end{aligned}$$

and that

$$\mathsf{Var}_{\{U(\boldsymbol{s}'),\boldsymbol{s}'\in\mathcal{S}\}}\left\{\sum_{\boldsymbol{s}'\in\mathcal{S}} K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}',U(\boldsymbol{s})-U(\boldsymbol{s}'))|U(\boldsymbol{s})\right\}$$
$$=O\left\{\sum_{\boldsymbol{s}'\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}^2\left(\boldsymbol{s}-\boldsymbol{s}',U(\boldsymbol{s})-U(\boldsymbol{s}')\right)|U(\boldsymbol{s})\right]\right\}.$$

Then, we can similarly obtain

$$\mathsf{E}_{\{U(\boldsymbol{s}'),\boldsymbol{s}'\in\mathcal{S}\}}\left\{\frac{1}{m}\sum_{\boldsymbol{s}'\in\mathcal{S}} K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}',U(\boldsymbol{s})-U(\boldsymbol{s}'))-\right.$$
$$\left.\frac{1}{m}\sum_{\boldsymbol{s}'\in\mathcal{S}} E_{U(\boldsymbol{s}')}\left[K_{\boldsymbol{H}}(\boldsymbol{s}-\boldsymbol{s}',U(\boldsymbol{s})-U(\boldsymbol{s}'))\right]\middle| U(\boldsymbol{s})\right\}^2\to 0, \tag{S1.7}$$

in probability as $\mathcal{S}\to\mathbb{S}$.

Step 3. By the continuous mapping theorem, combining the results of (S1.6) and (S1.7), we complete the proof of Theorem 2. □

S1.3 Proof of Theorem 3

Proof. Recall that

$$t_w=\sup_t\left\{t:\frac{\sum_{\boldsymbol{s}\in\mathcal{S}}\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))t}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}}\mathsf{I}\left\{p_w(\boldsymbol{s})\le t\right\},1\right\}}\le\alpha\right\},$$

and we reject the null hypothesis $H_0(\boldsymbol{s})$ if $p_w(\boldsymbol{s})\le t_w$ where $p_w(\boldsymbol{s})=p(\boldsymbol{s})/\frac{\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}$. Hence, to prove Theorem 3, it is sufficient to show that, uniformly for all $t\ge t_w$,

$$\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}}\mathsf{I}\left\{p_w(\boldsymbol{s})\le t,\theta(\boldsymbol{s})=0\right\}}{c\sum_{\boldsymbol{s}\in\mathcal{S}}\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))t}-1\right|\to 0, \tag{S1.8}$$

in probability for some $0<c\le 1$, where the indicator $\mathsf{I}\left\{p_w(\boldsymbol{s})\le t,\theta(\boldsymbol{s})=0\right\}$ incorporates $U(\boldsymbol{s})$ via the p-value weight.

By Genovese et al. (2006), the BH procedure controls FDR when it is applied to the weighted p-values, when the weights sum to m and they are independent of the p-values.

Algorithm S1.1 The scaled weighting procedure (SWP).

Step 1. Calculate the weights as $\tilde{w}(\boldsymbol{s}, U(\boldsymbol{s})) = \left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \frac{\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}\right\}^{-1} \frac{m\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}$, and then adjust p-values by $p_{\tilde{w}}(\boldsymbol{s}) = p(\boldsymbol{s})/\tilde{w}(\boldsymbol{s}, U(\boldsymbol{s}))$ for $\boldsymbol{s} \in \mathcal{S}$.

Step 2. Obtain the data-driven threshold

$$t_{\tilde{w}} = \sup_t \left\{ 0 \le t \le 1 : \frac{mt}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p_{\tilde{w}}(\boldsymbol{s}) \le t\right\}, 1\right\}} \le \alpha \right\}.$$

Step 3. Reject $H_0(\boldsymbol{s})$ if $p_{\tilde{w}}(\boldsymbol{s}) \le t_{\tilde{w}}$ for all $\boldsymbol{s} \in \mathcal{S}$.

Therefore, we first propose a scaled weighting procedure (SWP) in Algorithm S1.1, and show that it controls FDR and FDP asymptotically in Lemma 2. We then further prove (S1.8).

Recall that we aim to show the false rejections $\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p_w(\boldsymbol{s}) \le t, \theta(\boldsymbol{s}) = 0\right\}$ is close to $c\sum_{\boldsymbol{s}\in\mathcal{S}} \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))t$ for some $0 < c \le 1$. However, it is difficult to achieve this goal directly. Hence, we introduce two intermediate quantities as a bridge. Specifically, let $B_{\boldsymbol{s}} = \{U(\boldsymbol{s}) : \pi(\boldsymbol{s}, U(\boldsymbol{s})) \in [\xi, 1-\xi]$ and $|U(\boldsymbol{s}) - \mu(\boldsymbol{s})| \le (2\log m)^{1/2}\}$, and then we divide the proof into three steps. First, by Lemma 2 and comparing the p-value thresholds of NAPA and SWP, we show that, given $\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}$, the false rejection $\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p_w(\boldsymbol{s}) \le t, \theta(\boldsymbol{s}) = 0\right\}$ is close to its expectation over $\{(p(\boldsymbol{s}), U(\boldsymbol{s})) : \boldsymbol{s} \in \mathcal{S}\}$, $\sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})}\big[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\big]$. Next, we further show that $\sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})}\big[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\big]$ is close to its expectation over $\{(\theta(\boldsymbol{s})) : \boldsymbol{s} \in \mathcal{S}\}$, $\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})}[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \Pr\{\theta(\boldsymbol{s}) = 0|U(\boldsymbol{s})\}]$, by dealing with the randomness of $\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}$. Finally, we show that $\sum_{\boldsymbol{s}\in\mathcal{S}} \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))t$ can conservatively estimate $\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})}[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \Pr\left\{\theta(\boldsymbol{s}) = 0|U(\boldsymbol{s})\right\}]$. By combining these steps, we establish the asymptotic FDP and FDR control of the NAPA procedure.

Step 1. Firstly, we show that given $\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}$, $\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p_w(\boldsymbol{s}) \le t, \theta(\boldsymbol{s}) = 0\right\}$ is close to $\sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]$ for all $t \ge t_w$.

Note that, by the proof of Lemma 2, we have uniformly for all $t \geq t_{\tilde{w}}$ that,

$$\left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\{p_{\tilde{w}}(\boldsymbol{s}) \leq t, \theta(\boldsymbol{s}) = 0\}}{\sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_{\tilde{w}}(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\} \right]} - 1 \right| \to 0, \tag{S1.9}$$

in probability as $\mathcal{S} \to \mathbb{S}$, where $p_{\tilde{w}}(\boldsymbol{s}) = p_w(\boldsymbol{s}) m^{-1} \sum_{\boldsymbol{s}\in\mathcal{S}} \frac{\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}$.

Letting $\tilde{t} = \left\{ \sum_{\boldsymbol{s}\in\mathcal{S}} \frac{\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))} \right\}^{-1} mt$, then we have that,

$$\frac{mt}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p_{\tilde{w}}(\boldsymbol{s}) \leq t\right\}, 1\right\}} = \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \frac{\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))} \tilde{t}}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p_w(\boldsymbol{s}) \leq \tilde{t}\right\}, 1\right\}}.$$

Thus, the corresponding threshold for $p_w(\boldsymbol{s})$ in SWP is

$$\tilde{t}_{\tilde{w}} = \sup_t \left\{ 0 \leq t \leq 1 : \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \frac{\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))} t}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p_w(\boldsymbol{s}) \leq t\right\}, 1\right\}} \leq \alpha \right\}.$$

We have that $t_{\tilde{w}} = \tilde{t}_{\tilde{w}} \cdot m^{-1} \sum_{\boldsymbol{s}\in\mathcal{S}} \frac{\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}{1-\pi_\tau(\boldsymbol{s},U(\boldsymbol{s}))}$, which implies that (S1.9) holds for $t > \tilde{t}_{\tilde{w}}$ with p_w. Recall that

$$t_w = \sup_t \left\{ 0 \leq t \leq 1 : \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s})) t}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p_w(\boldsymbol{s}) \leq t\right\}, 1\right\}} \leq \alpha \right\}.$$

Comparing the definition of t_w and $\tilde{t}_{\tilde{w}}$, we see that $\tilde{t}_{\tilde{w}} \leq t_w$. Then by (S1.9), we have uniformly for all $t \geq t_w$,

$$\left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{p_w(\boldsymbol{s}) \leq t, \theta(\boldsymbol{s}) = 0\right\}}{\sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\} \right]} - 1 \right| \to 0, \tag{S1.10}$$

in probability as $\mathcal{S} \to \mathbb{S}$.

Step 2. Next, we show that $\sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\} \right]$ is close to $\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \Pr\left\{\theta(\boldsymbol{s}) = 0 | U(\boldsymbol{s})\right\} \right]$.

More specifically, by Lemma 1, we have

$$\begin{aligned}
&\mathsf{E}_{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}} \sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\} \right] \\
&= \sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{\theta(\boldsymbol{s}),U(\boldsymbol{s})} \left[\mathsf{I}\{\theta(\boldsymbol{s}) = 0\} \Pr\{p_w(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \right] \\
&= \{1 + o(1)\} \sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \Pr\left\{\theta(\boldsymbol{s}) = 0 | U(\boldsymbol{s})\right\} \right],
\end{aligned}$$

where the $o(1)$ is in the limit of $\mathcal{S} \to \mathbb{S}$.

Also note that, by Condition (C7) and the proof of Lemma 2,

$$\frac{\mathsf{Var}_{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}} \sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]}{\left(\mathsf{E}_{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}} \sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]\right)^2}$$
$$= \frac{\mathsf{Var}_{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}} \sum_{\boldsymbol{s}\in\mathcal{S}} \left(\mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] \mathsf{I}\{\theta(\boldsymbol{s}) = 0\}\right)}{\left\{\mathsf{E}_{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}} \sum_{\boldsymbol{s}\in\mathcal{S}} \left(\mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] \mathsf{I}\{\theta(\boldsymbol{s}) = 0\}\right)\right\}^2}$$
$$= O\left(\frac{\mathsf{Var}_{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}} \sum_{\boldsymbol{s}\in\mathcal{S}} \left(\mathsf{E}_{U(\boldsymbol{s})} \left[w(\boldsymbol{s}, U(\boldsymbol{s}))\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] \mathsf{I}\left\{\theta(\boldsymbol{s}) = 0\right\}\right)}{m^2}\right) = o(1).$$

Therefore, we have

$$\left|\frac{\sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]}{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \Pr\left\{\theta(\boldsymbol{s}) = 0|U(\boldsymbol{s})\right\}\right]} - 1\right| \to 0, \tag{S1.11}$$

in probability.

Step 3. Finally, we show that $\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \Pr\left\{\theta(\boldsymbol{s}) = 0|U(\boldsymbol{s})\right\}\right]$ can be further conservatively estimated by $\sum_{\boldsymbol{s}\in\mathcal{S}} \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))t$.

By the proof of Lemma 2, under event $B_{\boldsymbol{s}}$, the thresholds for $\{p_w(\boldsymbol{s}), U(\boldsymbol{s})\}$ is in the corresponding range of $\{T(\boldsymbol{s}), U(\boldsymbol{s})\}$ where Lemma 1 holds. Then we have uniformly for all $t \ge t_w$,

$$\begin{aligned}
&\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \le t|\theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \Pr\left\{\theta(\boldsymbol{s}) = 0|U(\boldsymbol{s})\right\}\right] \\
=&\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left\{\left[\{1 + o(1)\}\frac{\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))}{1 - \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))} t + O(m^{-2})\right] \{1 - \pi(\boldsymbol{s}, U(\boldsymbol{s}))\}\right\} \qquad \text{(S1.12)} \\
\le&\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\{1 + o(1)\}\frac{\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))}{1 - \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))} t\{1 - \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))\} + O(m^{-2})\right] \\
=&\{1 + o(1)\}\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})}\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))t + O(m^{-1}) \\
=&\{1 + o(1)\}\mathsf{E}_{U(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}} \sum_{\boldsymbol{s}\in\mathcal{S}} \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))t,
\end{aligned}$$

where $o(1)$ is in the limit of $\mathcal{S} \to \mathbb{S}$, and (S1.12) is obtained by Lemma 1.

Also, by Condition (C7), $\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))$ has bounded first derivatives and is hence continuous in $U(\boldsymbol{s})$. By the proof of Proposition 1, we have $\{U(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\} \to N(\{\mu(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}, \boldsymbol{\Sigma})$ in distribution, where $\boldsymbol{\Sigma} = (\sigma_{\boldsymbol{s},\boldsymbol{l}}), \sigma_{\boldsymbol{s},\boldsymbol{l}} = (r_{\boldsymbol{s},\boldsymbol{l};1} + \{\vartheta(\boldsymbol{s})\vartheta(\boldsymbol{l})\}^{1/2} r_{\boldsymbol{s},\boldsymbol{l};2})/(\{1+\vartheta(\boldsymbol{s})\}\{1+\vartheta(\boldsymbol{l})\})^{1/2}$. Let $g(\{U(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}) = \sum_{\boldsymbol{s}\in\mathcal{S}} \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))$. By delta method, we have $g(\{U(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}) \to N(g(\{\mu(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}), \sigma_g^2)$ in distribution with

$$\sigma_g^2 = \nabla g(\{\mu(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\})^\intercal \boldsymbol{\Sigma} \nabla g(\{\mu(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}) = O\left(\max\{\lambda_1(\boldsymbol{R}_1), \lambda_1(\boldsymbol{R}_2)\} m\right).$$

Then, under Condition (C6), we have

$$\mathsf{Var}_{U(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}} \sum_{\boldsymbol{s}\in\mathcal{S}} \{\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))t\} = o\left(\mathsf{E}^2_{U(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}} \sum_{\boldsymbol{s}\in\mathcal{S}} \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))t\right).$$

Hence, we have

$$\left| \frac{c\sum_{\boldsymbol{s}\in\mathcal{S}} \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))t}{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\{p_w(\boldsymbol{s}) \le t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \Pr\{\theta(\boldsymbol{s}) = 0 | U(\boldsymbol{s})\}\right]} - 1 \right| \to 0, \tag{S1.13}$$

in probability for some $0 < c \le 1$.

Finally, by combining (S1.10), (S1.11) and (S1.13), (S1.8) is proved, which completes the proof of Theorem 3. □

S1.4 Proof of Theorem 4

Proof. Note that, to show Theorem 4, it suffices to show,

$$\left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\{p_{\hat{w}}(\boldsymbol{s}) \le t, \theta(\boldsymbol{s}) = 0\}}{c\sum_{\boldsymbol{s}\in\mathcal{S}} \hat{\pi}_\tau(\boldsymbol{s}, U(\boldsymbol{s}))t} - 1 \right| \to 0, \tag{S1.14}$$

in probability for some $0 < c \le 1$.

First, define the events $\hat{B}_{\boldsymbol{s}} = B_{\boldsymbol{s}} \cap \{U(\boldsymbol{s}) : \hat{\pi}_\tau(\boldsymbol{s}, U(\boldsymbol{s})) = \{1 + o(1)\}\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))\}$ for all $\boldsymbol{s} \in \mathcal{S}$. Then, under Theorem 3 and Condition (C7), we have $\Pr(\hat{B}_{\boldsymbol{s}}) \to 1$ for all $\boldsymbol{s} \in \mathcal{S}$. Hence, it is easy to see that Lemma 2 still holds with estimated weights $\{\hat{\pi}_\tau(\boldsymbol{s}, U(\boldsymbol{s})) : \boldsymbol{s} \in \mathcal{S}\}$. Therefore, similar to the proof of Theorem 3, we have, uniformly for all $t \ge t_{\hat{w}}$,

$$\left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\{p_{\hat{w}}(\boldsymbol{s}) \le t, \theta(\boldsymbol{s}) = 0\}}{\sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\{p_{\hat{w}}(\boldsymbol{s}) \le t | \theta(\boldsymbol{s}) = 0, \hat{B}_{\boldsymbol{s}}\} \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]} - 1 \right| \to 0, \tag{S1.15}$$

in probability as $\mathcal{S} \to \mathbb{S}$.

Then, by Condition (C7), Lemma 2 and Theorem 2, we can similarly obtain that,

$$\left| \frac{\sum_{\theta(\boldsymbol{s})=0} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_{\hat{w}}(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, \hat{B}_{\boldsymbol{s}}\} \big| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\} \right]}{\sum_{\boldsymbol{s} \in \mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_{\hat{w}}(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, \hat{B}_{\boldsymbol{s}}\} \Pr\{\theta(\boldsymbol{s}) = 0 | U(\boldsymbol{s})\} \right]} - 1 \right| \to 0, \qquad \text{(S1.16)}$$

in probability.

Finally, by Lemma 1 and Theorem 2, we have uniformly for all $t \geq t_{\hat{w}}$ that,

$$\begin{aligned}
&\sum_{\boldsymbol{s} \in \mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, \hat{B}_{\boldsymbol{s}}\} \Pr\{\theta(\boldsymbol{s}) = 0 | U(\boldsymbol{s})\} \right] \\
=& \sum_{\boldsymbol{s} \in \mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\{1 + o(1)\}^2 \frac{\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))}{1 - \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))} t \{1 - \pi(\boldsymbol{s}, U(\boldsymbol{s}))\} + O(m^{-2}) \right] \\
\leq& \sum_{\boldsymbol{s} \in \mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\{1 + o(1)\} \frac{\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))}{1 - \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))} t \{1 - \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))\} + O(m^{-2}) \right] \\
=& \{1 + o(1)\} \sum_{\boldsymbol{s} \in \mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} [\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s}))] t + O(m^{-1}) \\
=& \{1 + o(1)\} \, \mathsf{E}_{U(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}} \sum_{\boldsymbol{s} \in \mathcal{S}} \hat{\pi}_\tau(\boldsymbol{s}, U(\boldsymbol{s})) t,
\end{aligned}$$

where $o(1)$ is in the limit of $\mathcal{S} \to \mathbb{S}$. Similarly, we also obtain

$$\mathsf{Var}_{U(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}} \sum_{\boldsymbol{s} \in \mathcal{S}} \{\hat{\pi}_\tau(\boldsymbol{s}, U(\boldsymbol{s})) t\} = o\left(\mathsf{E}^2_{U(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}} \sum_{\boldsymbol{s} \in \mathcal{S}} \hat{\pi}_\tau(\boldsymbol{s}, U(\boldsymbol{s})) t \right).$$

Then we have

$$\left| \frac{c \sum_{\boldsymbol{s} \in \mathcal{S}} \hat{\pi}_\tau(\boldsymbol{s}, U(\boldsymbol{s})) t}{\sum_{\boldsymbol{s} \in \mathcal{S}} \mathsf{E}_{U(\boldsymbol{s})} \left[\Pr\{p_w(\boldsymbol{s}) \leq t | \theta(\boldsymbol{s}) = 0, \hat{B}_{\boldsymbol{s}}\} \Pr\{\theta(\boldsymbol{s}) = 0 | U(\boldsymbol{s})\} \right]} - 1 \right| \to 0, \qquad \text{(S1.17)}$$

for some $0 < c \leq 1$.

By combining (S1.15), (S1.16) and (S1.17), (S1.14) is proved and we complete the proof of Theorem 4. □

S1.5 Proof of Proposition 1

Proof. We aim to verify the asymptotic normality of $T(\boldsymbol{s})$ and $U(\boldsymbol{s})$ and show the asymptotic independence between $T(\boldsymbol{s})$ and $U(\boldsymbol{s})$, under the null. Recall the fact

$$\frac{G\left(t+o\{(\log m)^{-1/2}\}\right)}{G(t)}=1+o(1), \quad \frac{\phi\left(t+o\{(\log m)^{-1/2}\}\right)}{\phi(t)}=1+o(1), \qquad \text{(S1.18)}$$

uniformly in $0 \le t \le c_0(\log m)^{1/2}$ for any constant $c_0 > 0$, where $\phi(\cdot)$ is the probability density function of a standard normal variable. Then we can use normal variables to approximate $T(\boldsymbol{s})$ and $U(\boldsymbol{s})$ simultaneously. Without loss of generality, we assume that $\sigma^2_{\boldsymbol{s},1} \asymp \sigma^2_{\boldsymbol{s},2} \asymp 1$, for all $\boldsymbol{s} \in \mathcal{S}$ and $d = 1, 2$, throughout.

We divide the proof into three steps. First, we replace the estimated variances with the truth and obtain $\{\tilde{T}(\boldsymbol{s}), \tilde{U}(\boldsymbol{s})\}$ with zero mean and unit variance. Next, we truncate $\{\tilde{T}(\boldsymbol{s}), \tilde{U}(\boldsymbol{s})\}$ to $\{\hat{T}(\boldsymbol{s}), \hat{U}(\boldsymbol{s})\}$, and show the difference is negligible. Finally, we use normal variables to approximate $\{\hat{T}(\boldsymbol{s}), \hat{U}(\boldsymbol{s})\}$, and derive the asymptotic results.

Step 1. Following the proof of Lemma 2 of Cai and Liu (2011), for any constant $C > 0$, there exists a constant $C_0 > 0$, such that

$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}}\left|\hat{\sigma}^2_{\boldsymbol{s},d}-\sigma^2_{\boldsymbol{s},d}\right| \ge C_0\sqrt{\log m/n_d}\right\} = O(m^{-C}),$$

for $d = 1, 2$. We then replace the estimators with the true variances, and define

$$\tilde{T}(\boldsymbol{s}) = \frac{\bar{Y}_1(\boldsymbol{s})-\bar{Y}_2(\boldsymbol{s})}{\left\{\sigma^2_{\boldsymbol{s},1}/n_1+\sigma^2_{\boldsymbol{s},2}/n_2\right\}^{1/2}},$$
$$\tilde{U}(\boldsymbol{s}) = \frac{\bar{Y}_1(\boldsymbol{s})-\beta_1(\boldsymbol{s})+\kappa(\boldsymbol{s})\{\bar{Y}_2(\boldsymbol{s})-\beta_2(\boldsymbol{s})\}}{\left\{\sigma^2_{\boldsymbol{s},1}/n_1+\kappa^2(\boldsymbol{s})\sigma^2_{\boldsymbol{s},2}/n_2\right\}^{1/2}},$$

where $\kappa(\boldsymbol{s}) = (n_2\sigma^2_{\boldsymbol{s},1})/(n_1\sigma^2_{\boldsymbol{s},2})$. Recall the condition that $\log m = o(n^{1/5})$, we have, for any constant $C > 0$, there exists $b_m = o\{(\log m)^{-1/2}\}$, such that

$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\left|T(\boldsymbol{s})-\tilde{T}(\boldsymbol{s})\right| \ge b_m\right\} = O(m^{-C}), \qquad \text{(S1.19)}$$
$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\left|[U(\boldsymbol{s})-\mu(\boldsymbol{s})]-\tilde{U}(\boldsymbol{s})\right| \ge b_m\right\} = O(m^{-C}), \qquad \text{(S1.20)}$$

where $\mu(\boldsymbol{s}) = (1 + o\{(\log m)^{-1}\}) \, \mathsf{E}\{U(\boldsymbol{s})\}$. This establishes the asymptotic normality of $T(\boldsymbol{s})$ and $U(\boldsymbol{s})$. It remains to show the asymptotic independence, for which it suffices to show

$$\lim_{\epsilon \to 0} \frac{\Pr\{|T(\boldsymbol{s})| \geq t, |U(\boldsymbol{s}) - u| \leq \epsilon\}}{\Pr\{|U(\boldsymbol{s}) - u| \leq \epsilon\}} = \{1 + o(1)\}G(t) + O(m^{-C_4}), \tag{S1.21}$$

uniformly in $t = O\{(\log m)^{1/2}\}$, $|U(\boldsymbol{s}) - \mu(\boldsymbol{s})| = O\{(\log m)^{1/2}\}$, and all $\boldsymbol{s} \in \mathcal{S}_0$.

Step 2. Next, we reconstruct $\tilde{T}(\boldsymbol{s})$ and $\tilde{U}(\boldsymbol{s})$. Let $n_2/n_1 \leq K_1$ with $K_1 \geq 1$, and define $Z_k(\boldsymbol{s}) = (n_2/n_1)\{Y_{k,1}(\boldsymbol{s}) - \beta_1(\boldsymbol{s})\}$ for $1 \leq k \leq n_1$, $Z_k(\boldsymbol{s}) = -\{Y_{k-n_1,2}(\boldsymbol{s}) - \beta_2(\boldsymbol{s})\}$ for $n_1 + 1 \leq k \leq n_1 + n_2$. We can rewrite $\tilde{T}(\boldsymbol{s})$ as,

$$\tilde{T}(\boldsymbol{s}) = \frac{\sum_{k=1}^{n_1+n_2} Z_k(\boldsymbol{s})}{\left(n_2^2\sigma_{\boldsymbol{s},1}^2/n_1 + n_2\sigma_{\boldsymbol{s},2}^2\right)^{1/2}}.$$

Define the truncated variable, $\hat{Z}_k(\boldsymbol{s}) = Z_k(\boldsymbol{s})\mathsf{I}\{|Z_k(\boldsymbol{s})| \leq \tau_n|\} - \mathsf{E}\left[Z_k(\boldsymbol{s})\mathsf{I}\{|Z_k(\boldsymbol{s})| \leq \tau_n|\}\right]$, and write

$$\hat{T}(\boldsymbol{s}) = \frac{\sum_{k=1}^{n_1+n_2} \hat{Z}_k(\boldsymbol{s})}{\left(n_2^2\sigma_{\boldsymbol{s},1}^2/n_1 + n_2\sigma_{\boldsymbol{s},2}^2\right)^{1/2}},$$

where $\tau_n = (C_3 + 2)K_1/C_1 \log(m \vee n)$, with C_3 specified later. We see that,

$$\tilde{T}(\boldsymbol{s}) - \hat{T}(\boldsymbol{s}) = \frac{\sum_{k=1}^{n_1+n_2} \left[Z_k(\boldsymbol{s})\mathsf{I}\{|Z_k(\boldsymbol{s})| > \tau_n\}\right]}{\left(n_2^2\sigma_{\boldsymbol{s},1}^2/n_1 + n_2\sigma_{\boldsymbol{s},2}^2\right)^{1/2}} - \frac{\sum_{k=1}^{n_1+n_2} \mathsf{E}\left[Z_k(\boldsymbol{s})\mathsf{I}\{|Z_k(\boldsymbol{s})| > \tau_n\}\right]}{\left(n_2^2\sigma_{\boldsymbol{s},1}^2/n_1 + n_2\sigma_{\boldsymbol{s},2}^2\right)^{1/2}}.$$

For the second term, by $n_1 \asymp n_2 \asymp n$, there exists some constants c_0, c_1, c_2, such that

$$\begin{aligned}
&\max_{\boldsymbol{s} \in \mathcal{S}_0} n^{-1/2} \sum_{k=1}^{n_1+n_2} \mathsf{E}\left[|Z_k(\boldsymbol{s})|\mathsf{I}\{|Z_k(\boldsymbol{s})| > \tau_n\}\right] \\
\leq &c_0 n^{1/2} \max_{1 \leq k \leq n_1+n_2} \max_{\boldsymbol{s} \in \mathcal{S}} \mathsf{E}\left[|Z_k(\boldsymbol{s})|\mathsf{I}\{|Z_k(\boldsymbol{s})| > \tau_n\}\right] \\
\leq &c_0 n^{1/2}(m+n)^{-\frac{C_3+2}{2}} \max_{1 \leq k \leq n_1+n_2} \max_{\boldsymbol{s} \in \mathcal{S}} \mathsf{E}\left[|Z_k(\boldsymbol{s})| \exp\{(C_1/2K_1)|Z_k(\boldsymbol{s})|\}\right] && \text{(S1.22)} \\
\leq &c_1 n^{1/2}(m+n)^{-\frac{C_3+2}{2}} \max_{1 \leq k \leq n_1+n_2} \max_{\boldsymbol{s} \in \mathcal{S}} \left\{\mathsf{E}|(C_1/K_1)Z_k(\boldsymbol{s})|^2\right\}^{1/2} \left(\mathsf{E}\left[\exp\{(C_1/K_1)|Z_k(\boldsymbol{s})|\}\right]\right)^{1/2} && \text{(S1.23)} \\
\leq &c_2 n^{1/2}(m+n)^{-\frac{C_3+2}{2}}, && \text{(S1.24)}
\end{aligned}$$

where (S1.22) comes from $\mathsf{I}\{x \geq a\} \leq e^x/e^a$, (S1.23) by Cauchy-Schwartz inequality, and (S1.24) from $(C_1/K_1)|Z_k(\boldsymbol{s})| \leq C_1|Y_{k,d}(\boldsymbol{s}) - \beta_d(\boldsymbol{s})|$ by Condition (C10), and the inequality $x^2 \leq e^x$ for $x > 0$. Note that the term (S1.24) is $o\{(\log m)^{-1/2}\}$. Then there exists some $b_m = o\{(\log m)^{-1/2}\}$, such that

$$\begin{aligned}
\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\left|\tilde{T}(\boldsymbol{s}) - \hat{T}(\boldsymbol{s})\right| \geq b_m\right\} &\leq \Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\max_{1\leq k\leq n_1+n_2}|Z_k(\boldsymbol{s})| > \tau_n\right\} \\
&\leq mn\max_{\boldsymbol{s}\in\mathcal{S}_0}\Pr\left\{|Z_k(\boldsymbol{s})| > \tau_n\right\} \\
&\leq mn\exp\left\{-(C_1/K_1)\tau_n\right\}\mathsf{E}\exp\left\{(C_1/K_1)|Z_k(\boldsymbol{s})|\right\} \\
&= O(m^{-C_3}). \qquad \text{(S1.25)}
\end{aligned}$$

Similarly, define $V_k(\boldsymbol{s}) = (n_2/n_1)\{Y_{k,1}(\boldsymbol{s}) - \beta_1(\boldsymbol{s})\}$ for $1 \leq k \leq n_1$, and $V_k(\boldsymbol{s}) = \kappa(\boldsymbol{s})\{Y_{k,2}(\boldsymbol{s}) - \beta_2(\boldsymbol{s})\}$ for $n_1 + 1 \leq k \leq n_1 + n_2$, where $\kappa(\boldsymbol{s}) = (n_2\sigma^2_{\boldsymbol{s},1})/(n_1\sigma^2_{\boldsymbol{s},2})$. Then,

$$\tilde{U}(\boldsymbol{s}) = \frac{\sum_{k=1}^{n_1+n_2} V_k(\boldsymbol{s})}{\left\{n_2^2\sigma^2_{\boldsymbol{s},1}/n_1 + n_2\kappa^2(\boldsymbol{s})\sigma^2_{\boldsymbol{s},2}\right\}^{1/2}}.$$

Define

$$\hat{U}(\boldsymbol{s}) = \frac{\sum_{k=1}^{n_1+n_2} \hat{V}_k(\boldsymbol{s})}{\left\{n_2^2\sigma^2_{\boldsymbol{s},1}/n_1 + n_2\kappa^2(\boldsymbol{s})\sigma^2_{\boldsymbol{s},2}\right\}^{1/2}},$$

with the truncated variable $\hat{V}_k(\boldsymbol{s}) = V_k(\boldsymbol{s})\mathsf{I}\left\{|V_k(\boldsymbol{s})| \leq \tau_n|\right\} - \mathsf{E}\left[V_k(\boldsymbol{s})\mathsf{I}\left\{|V_k(\boldsymbol{s})| \leq \tau_n|\right\}\right]$. Then we obtain that,

$$\begin{aligned}
\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\left|\tilde{U}(\boldsymbol{s}) - \hat{U}(\boldsymbol{s})\right| \geq b_m\right\} &\leq \Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\max_{1\leq k\leq n_1+n_2}|V_k(\boldsymbol{s})| \geq \tau_n\right\} \\
&= O(m^{-C_3}), \qquad \text{(S1.26)}
\end{aligned}$$

for some $b_m = o\{(\log m)^{-1/2}\}$. Combining (S1.19), (S1.20), (S1.25), and (S1.26), we obtain that,

$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\left|T(\boldsymbol{s}) - \hat{T}(\boldsymbol{s})\right| \geq b_m\right\} = O(m^{-C_3}), \qquad \text{(S1.27)}$$

$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\left|[U(\boldsymbol{s}) - \mu(\boldsymbol{s})] - \hat{U}(\boldsymbol{s})\right| \geq b_m\right\} = O(m^{-C_3}). \qquad \text{(S1.28)}$$

Step 3. Now we can apply normal approximation to $\{(\hat{T}(\boldsymbol{s}), \hat{U}(\boldsymbol{s})) : \boldsymbol{s} \in \mathcal{S}_0\}$. Denote by $\boldsymbol{N} = (N_1, N_2)$ a normal random vector with $\mathsf{E}(\boldsymbol{N}) = 0$ and $\mathsf{Cov}\,(\boldsymbol{N}) = \mathsf{Cov}(\boldsymbol{W}_k)$, where

$$\boldsymbol{W}_k = (W_{k,1}, W_{k,2}) = \left(\frac{\hat{Z}_k(\boldsymbol{s})}{\left\{n_2\sigma^2_{\boldsymbol{s},1}/n_1 + \sigma^2_{\boldsymbol{s},2}\right\}^{1/2}}, \frac{\hat{V}_k(\boldsymbol{s})}{\left\{n_2\sigma^2_{\boldsymbol{s},1}/n_1 + \kappa^2(\boldsymbol{s})\sigma^2_{\boldsymbol{s},2}\right\}^{1/2}} \right).$$

Then we have that,

$$\Pr_{H_0}\left\{|\hat{T}(\boldsymbol{s})| \geq t, |\hat{U}(\boldsymbol{s}) - \tilde{u}| \leq \epsilon\right\} = \Pr\left\{\left|n_2^{-1/2}\sum_{k=1}^{n_1+n_2} W_{k,1}\right| \geq t, \left|n_2^{-1/2}\sum_{k=1}^{n_1+n_2} W_{k,2} - \tilde{u}\right| \leq \epsilon\right\},$$

where $\tilde{u} = u - \mu(\boldsymbol{s})$. By Theorem 1.1 in Zaitsev (1987), we have that,

$$\begin{aligned}
&\Pr_{H_0}\left\{|\hat{T}(\boldsymbol{s})| \geq t, |\hat{U}(\boldsymbol{s}) - \tilde{u}| \leq \epsilon\right\} \\
\leq &\Pr\left\{|N_1| \geq t - \epsilon_n(\log m)^{-1/2}, |N_2 - \tilde{u}| \leq \epsilon + \epsilon_n(\log m)^{-1/2}\right\} \\
&+ c_1 \exp\left\{-\frac{n^{1/2}\epsilon_n}{c_2\tau_n(\log m)^{1/2}}\right\},
\end{aligned}$$

and

$$\begin{aligned}
&\Pr_{H_0}\left\{|\hat{T}(\boldsymbol{s})| \geq t, |\hat{U}(\boldsymbol{s}) - \tilde{u}| \leq \epsilon\right\} \\
\geq &\Pr\left\{|N_1| \geq t + \epsilon_n(\log m)^{-1/2}, |N_2 - \tilde{u}| \leq \epsilon - \epsilon_n(\log m)^{-1/2}\right\} \\
&- c_1 \exp\left\{-\frac{n^{1/2}\epsilon_n}{c_2\tau_n(\log m)^{1/2}}\right\},
\end{aligned}$$

where $c_1 > 0$ and $c_2 > 0$ are constants, and $\epsilon_n \to 0$ is to be specified later. Similar to the proof of Lemma 3, by Condition (C10) and an appropriate choice of τ_n, we have $\|\mathsf{Cov}(\boldsymbol{W}_k) - \boldsymbol{I}\|_2 = o\left\{(\log m)^{-1}\right\}$. Combined with (S1.27) and (S1.28), we have that,

$$\begin{aligned}
&\Pr_{H_0}\left\{|T(\boldsymbol{s})| \geq t, |U(\boldsymbol{s}) - u| \leq \epsilon\right\} \\
\leq &\{1 + o(1)\}G\left\{t - \epsilon_n(\log m)^{-1/2} - b_m\right\} \cdot 2\left\{\epsilon + \epsilon_n(\log m)^{-1/2} + b_m\right\}\phi(u_1) \\
&+ c_1 \exp\left\{-\frac{n^{1/2}\epsilon_n}{c_2\tau_n(\log m)^{1/2}}\right\} + O\left(m^{-C_3}\right),
\end{aligned} \tag{S1.29}$$

and

$$\begin{aligned}
&\Pr_{H_0}\left\{|T(\boldsymbol{s})| \geq t, |U(\boldsymbol{s}) - u| \leq \epsilon\right\} \\
\geq &\{1 + o(1)\}G\left\{t + \epsilon_n(\log m)^{-1/2} + b_m\right\} \cdot 2\left\{\epsilon - \epsilon_n(\log m)^{-1/2} - b_m\right\}\phi(u_2) \\
&- c_1 \exp\left\{-\frac{n^{1/2}\epsilon_n}{c_2\tau_n(\log m)^{1/2}}\right\} - O\left(m^{-C_3}\right),
\end{aligned} \tag{S1.30}$$

where $u_1 = \arg\max \phi(\check{u})$ for $\check{u} \in [\tilde{u} - \epsilon - \epsilon_n(\log m)^{-1/2} - b_m, \tilde{u} + \epsilon + \epsilon_n(\log m)^{-1/2} + b_m]$ and $u_2 = \arg\min \phi(\check{u})$ for $\check{u} \in [\tilde{u} - \epsilon + \epsilon_n(\log m)^{-1/2} + b_m, \tilde{u} + \epsilon - \epsilon_n(\log m)^{-1/2} - b_m]$. Meanwhile, by Theorem 1.1 in Zaitsev (1987) and (S1.28) again, we have that,

$$\begin{aligned}\Pr_{H_0}\{|U(\boldsymbol{s}) - u| \leq \epsilon\} \leq & 2\{1 + o(1)\}\left\{\epsilon + \epsilon_n(\log m)^{-1/2} + b_m\right\}\phi(u_1) \\ & + c_1 \exp\left\{-\frac{n^{1/2}\epsilon_n}{c_2\tau_n(\log m)^{1/2}}\right\} + O\left(m^{-C_3}\right),\end{aligned} \tag{S1.31}$$

and

$$\begin{aligned}\Pr_{H_0}\{|U(\boldsymbol{s}) - u| \leq \epsilon\} \geq & 2\{1 + o(1)\}\left\{\epsilon - \epsilon_n(\log m)^{-1/2} - b_m\right\}\phi(u_2) \\ & - c_1 \exp\left\{-\frac{n^{1/2}\epsilon_n}{c_2\tau_n(\log m)^{1/2}}\right\} - O\left(m^{-C_3}\right).\end{aligned} \tag{S1.32}$$

Now let ϵ_n approach zero sufficiently slowly, such that both $\epsilon_n^{1/2}(\log m)^{-1/2} \gg b_m$, and $\frac{n^{1/2}\epsilon_n}{c_2\tau_n(\log m)^{1/2}} \gg \log m$ hold. The latter means $c_1 \exp\left\{-\frac{n^{1/2}\epsilon_n}{c_2\tau_n(\log m)^{1/2}}\right\} = O\left(m^{-C}\right)$ for any constant $C > 0$ under $\log m = o\left(n^{1/5}\right)$. Let $\epsilon = \epsilon_n^{1/2}(\log m)^{-1/2}$, by $|\tilde{u}| = O\{(\log m)^{1/2}\}$, we have $|u_1 - u_2| = O\left\{\epsilon_n^{1/2}(\log m)^{-1/2}\right\}$ which gives $\phi(u_1)/\phi(u_2) = 1 + o(1)$ by (S1.18). Recall (S1.21), and combine (S1.29), (S1.30), (S1.31), and (S1.32). Then for $t = O\{(\log m)^{1/2}\}$ and $|U(\boldsymbol{s}) - \mu(\boldsymbol{s})| = O\{(\log m)^{1/2}\}$, for any constant $C_4 > 0$, we choose C_3 sufficiently large, such that

$$\begin{aligned}&\Pr_{H_0}\{|T(\boldsymbol{s})| \geq t |\, U(\boldsymbol{s})\} \\ \leq & \lim_{\epsilon\to 0} \frac{\{1 + o(1)\}G\left\{t - \epsilon_n(\log m)^{-1/2} - b_m\right\} \cdot 2\left\{\epsilon + \epsilon_n(\log m)^{-1/2} + b_m\right\}\phi(u_1) + O\left(m^{-C_3}\right)}{2\{1 + o(1)\}\left\{\epsilon - \epsilon_n(\log m)^{-1/2} - b_m\right\}\phi(u_2) - O\left(m^{-C_3}\right)} \\ = & \{1 + o(1)\}G(t) + O(m^{-C_4}),\end{aligned}$$

and

$$\begin{aligned}&\Pr_{H_0}\{|T(\boldsymbol{s})| \geq t |\, U(\boldsymbol{s})\} \\ \geq & \lim_{\epsilon\to 0} \frac{\{1 + o(1)\}G\left\{t + \epsilon_n(\log m)^{-1/2} + b_m\right\} \cdot 2\left\{\epsilon - \epsilon_n(\log m)^{-1/2} - b_m\right\}\phi(u_2) - O\left(m^{-C_3}\right)}{2\{1 + o(1)\}\left\{\epsilon + \epsilon_n(\log m)^{-1/2} + b_m\right\}\phi(u_1) + O\left(m^{-C_3}\right)} \\ = & \{1 + o(1)\}G(t) - O(m^{-C_4}).\end{aligned}$$

Combining the upper and lower bounds completes the proof of Proposition 1. □

S2 Technical Lemmas

We collect the technical lemmas and provide their proofs in this section. Lemma 1 presents the conditional independence between $T(\boldsymbol{s})$ and $U(\boldsymbol{s})$ under the null. Its proof is similar to that of Proposition 1 and is thus omitted. Lemma 2 establishes the theoretical properties of the scaled weighting procedure proposed in Algorithm S1.1. Lemma 3 is needed for (S2.19) in the proof of Lemma 2.

Lemma 1. *Suppose Condition (C5) holds. Then we have*

$$\Pr\left\{|T(\boldsymbol{s})| \geq t \middle| U(\boldsymbol{s})\right\} = \{1+o(1)\}G(t) + O(m^{-2}),$$

uniformly in $0 \leq t \leq (2\log m)^{1/2}$, $|U(\boldsymbol{s}) - \mu(\boldsymbol{s})| \leq (2\log m)^{1/2}$ *and all* $\boldsymbol{s} \in \mathcal{S}_0$.

Lemma 2. *Suppose Conditions (C5)-(C8) hold. We have, for any* $\epsilon > 0$,

$$\overline{\lim_{\mathcal{S}\to\mathbb{S}}} \operatorname{FDR}\left(\boldsymbol{\delta}^{SWP}_{t_{\tilde{w}}}\right) \leq \alpha m_0/m, \textit{and} \lim_{\mathcal{S}\to\mathbb{S}} \Pr\left\{\operatorname{FDP}\left(\boldsymbol{\delta}^{SWP}_{t_{\tilde{w}}}\right) \leq \alpha m_0/m + \epsilon\right\} = 1.$$

Proof. To deal with the dependency among the tests, we base our analysis on z-values $z_{\tilde{w}}(\boldsymbol{s}) = \Phi^{-1}\{1 - p_{\tilde{w}}(\boldsymbol{s})/2\}$ instead of the weighted p-values, where $\Phi(\cdot)$ is the CDF of a standard normal random variable. Note that by Lemma 3 of Xia et al. (2020), SWP is equivalent to rejecting $H_0(\boldsymbol{s})$ if $|z_{\tilde{w}}(\boldsymbol{s})| \geq t_z$ for all $\boldsymbol{s} \in \mathcal{S}$, where

$$t_z = \inf_t \left\{ t \geq 0 : \frac{mG(t)}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathbb{I}\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t\right\}, 1\right\}} \leq \alpha \right\}, \tag{S2.1}$$

and $G(t) = 2\{1 - \Phi(t)\}$. Define the events

$$B_1 = \{U(\boldsymbol{s}) : \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s})) \in [\xi, 1-\xi] \text{ for all } \boldsymbol{s} \in \mathcal{S}\},$$
$$B_2 = \left\{ \{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\} : m_0 = \sum_{\boldsymbol{s}\in\mathcal{S}} \mathbb{I}\{\theta(\boldsymbol{s}) = 0\} \asymp m \right\},$$

and then we have $\Pr(B_1) \to 1$ by Condition (C7). Again by Condition (C7), we have,

$$1 - o(1) \le \Pr(B_2) = \Pr(B_2|B_1)\Pr(B_1) + \Pr(B_2|B_1^c)\Pr(B_1^c) \le 1 + o(1),$$

which gives $\Pr(B_2) \to 1$. It hence suffices to focus on event B_2 and show, given $\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}$, $\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t\}$ is close to $cmG(t)$ for some $0 < c \le 1$.

Defining $B_{\boldsymbol{s}} = \{U(\boldsymbol{s}) : \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s})) \in [\xi, 1-\xi]$ and $|U(\boldsymbol{s}) - \mu(\boldsymbol{s})| \le (2\log m)^{1/2}\}$ for $\boldsymbol{s} \in \mathcal{S}_0$, we divide the following proof into three steps. First, we show that t_z in (S2.1) is obtained in the range $[0, t_m]$, where $t_m = (2\log m - 2\log\log m)^{1/2}$. Therefore it suffices to show for $t \in [0, t_m]$, $\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t\}$ is close to $cmG(t)$ for some $0 < c \le 1$. Next, by Theorem 1 in Genovese et al. (2006), we show the conditional expectation $\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\middle|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]$ is close to $cmG(t)$ for some $0 < c \le 1$. Finally, we prove the false discoveries $\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t\}$ is close to $\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}$ $\left[\Pr\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\middle|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]$, by showing that only the weakly correlated tests play the dominant role.

Step 1. First, we show that t_z in (S2.1) is attained in the range $[0, t_m]$, where $t_m = (2\log m - 2\log\log m)^{1/2}$. This is essential for the FDP control in (S2.21) and (S2.22).

Note that by Condition (C8), we have $|T(\boldsymbol{s})| \ge (\log m)^{1/2+\nu/2}$ for those $\boldsymbol{s} \in \mathcal{S}_\nu$ with probability tending to 1. Also note that under event B_1, $\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s})) \in [\xi, 1-\xi]$ for a sufficiently small constant $\xi > 0$, which gives $c \le \tilde{w}(\boldsymbol{s}, U(\boldsymbol{s})) \le C$ for some $0 < c < C < \infty$. Then,

$$\begin{aligned}
&\Pr\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{|z_{\tilde{w}}(\boldsymbol{s})| \ge (2\log m)^{1/2}\right\} \ge \left\{1/\left(c_\pi^{1/2}\alpha\right) + \varepsilon\right\}(\log m)^{1/2}\right\} \\
&\ge \Pr\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{|z_{\tilde{w}}(\boldsymbol{s})| \ge (2\log m)^{1/2}\right\} \ge \left\{1/\left(c_\pi^{1/2}\alpha\right) + \varepsilon\right\}(\log m)^{1/2}\middle| B_1\right\}\Pr(B_1) \\
&\ge \Pr\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{|T(\boldsymbol{s})| \ge (\log m)^{1/2+\nu/2}\right\} \ge \left\{1/\left(c_\pi^{1/2}\alpha\right) + \varepsilon\right\}(\log m)^{1/2}\middle| B_1\right\}\Pr(B_1) \\
&= \Pr\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{|T(\boldsymbol{s})| \ge (\log m)^{1/2+\nu/2}\right\} \ge \left\{1/\left(c_\pi^{1/2}\alpha\right) + \varepsilon\right\}(\log m)^{1/2}\right\} - o(1) \\
&= 1 - o(1).
\end{aligned}$$

Recall that $t_m = (2\log m - 2\log\log m)^{1/2}$, and $1-\Phi(t_m) \sim 1/\left\{(2c_\pi)^{1/2} t_m\right\}\exp(-t_m^2/2)$. We have $\Pr(0 \le t_z \le t_m) \to 1$. Then, we only need to prove uniformly for $0 \le t \le t_m$, given $\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}$, there exists some $0 < c \le 1$, such that

$$\left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t\} - cm_0 G(t)}{cm_0 G(t)} \right| \to 0, \tag{S2.2}$$

in probability.

Step 2. Next, we show that $\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\middle|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]$ is close to $cm_0G(t)$ for some $0 < c \le 1$, by considering the independent setting. Following Theorem 1 of Genovese et al. (2006) and Lemma 1, we obtain that $\mathsf{E}[\mathrm{FDP}(t)|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}] \le \{1+o(1)\}\alpha m_0/m + O(m^{-1})$. This, together with (S2.1), gives that,

$$\mathsf{E}_{\{z_{\tilde{w}}(\boldsymbol{s}), U(\boldsymbol{s})\}, \boldsymbol{s}\in\mathcal{S}}\left[\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t_z\}}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t_z\}, 1\right\}}\middle|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] \le \{1+o(1)\}\alpha m_0/m. \tag{S2.3}$$

Note that, by Condition (C5), similar to the proof of Proposition 1, we can obtain that

$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0} |U(\boldsymbol{s}) - \mu(\boldsymbol{s})| \ge (2\log m)^{1/2}\right\} = o(1).$$

Let $B_3 = \left\{U(\boldsymbol{s}) : |U(\boldsymbol{s}) - \mu(\boldsymbol{s})| \le (2\log m)^{1/2} \text{ for all } \boldsymbol{s} \in \mathcal{S}_0\right\}$. Then, it gives

$$\Pr(B_1 \cap B_3) = 1 - \Pr(B_1^c \cup B_3^c) \ge 1 - o(1),$$

and further $\Pr(B_{\boldsymbol{s}}) \ge \Pr(B_1 \cap B_3) \ge 1 - o(1)$. Then, we have

$$\begin{aligned}
&\mathsf{E}_{\{z_{\tilde{w}}(\boldsymbol{s}), U(\boldsymbol{s})\}, \boldsymbol{s}\in\mathcal{S}}\left[\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t\} \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] \\
=& \sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{z_{\tilde{w}}(\boldsymbol{s}), U(\boldsymbol{s})}\left[\mathsf{I}\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t\} \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] \\
=& \sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left(\mathsf{E}_{z_{\tilde{w}}(\boldsymbol{s})}\left[\mathsf{I}\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t\} \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}, U(\boldsymbol{s})\right] \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right) \\
=& \sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\}\Pr(B_{\boldsymbol{s}}) + O\left(\Pr(\{B_{\boldsymbol{s}}\}^c)\right)\middle|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] \\
=& \{1+o(1)\}\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\{|z_{\tilde{w}}(\boldsymbol{s})| \ge t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\} \middle| \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right].
\end{aligned}$$

Also, following Genovese et al. (2006) again, when $\{z_{\tilde{w}}(\boldsymbol{s}), U(\boldsymbol{s})\}$'s are independent across $\boldsymbol{s} \in \mathcal{S}$, by later (S2.10) and Lemma 1, we have uniformly in $0 \leq t \leq t_m$,

$$\begin{aligned} &\frac{\mathsf{Var}_{\{z_{\tilde{w}}(\boldsymbol{s}),U(\boldsymbol{s})\},\boldsymbol{s}\in\mathcal{S}}\left[\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t\right\} \big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]}{\left(\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\right\} \big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]\right)^2} \\ &\leq \left(\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\right\} \big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]\right)^{-1} = o(1). \end{aligned}$$

Therefore, we have

$$\left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t\right\}}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\right\} \big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]} - 1 \right| \to 0,$$

in probability. Now we can estimate the threshold by

$$\hat{t}^o = \inf\left\{ t \geq 0 : \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\right\} \big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq \hat{t}^o\right\}, 1\right\}} \leq \alpha \right\}, \tag{S2.4}$$

which gives that,

$$\mathsf{E}_{\{z_{\tilde{w}}(\boldsymbol{s}),U(\boldsymbol{s})\},\boldsymbol{s}\in\mathcal{S}}\left[\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq \hat{t}^o\right\}}{\max\left\{\sum_{\boldsymbol{s}\in\mathcal{S}} \mathsf{I}\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq \hat{t}^o\right\}, 1\right\}} \big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] \to \alpha. \tag{S2.5}$$

Comparing (S2.3) with (S2.5), we see that (S2.1) is asymptotically more conservative than (S2.4). Therefore, we have,

$$\left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\right\} \big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]}{c m_0 G(t)} - 1 \right| \to 0, \tag{S2.6}$$

in probability for some $0 < c \leq 1$ uniformly in $0 \leq t \leq t_m$.

Step 3. Finally, it remains to show that, with dependent $\{(z_{\tilde{w}}(\boldsymbol{s}), U(\boldsymbol{s})) : \boldsymbol{s} \in \mathcal{S}_0\}$, uniformly in $0 \leq t \leq t_m$,

$$\left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t\right\}}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{U(\boldsymbol{s})}\left[\Pr\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\right\} \big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]} - 1 \right| \to 0, \tag{S2.7}$$

in probability. We further divide the proof of this step into two sub-steps to handle the dependency structure among $\{(z_{\tilde{w}}(\boldsymbol{s}), U(\boldsymbol{s})) : \boldsymbol{s} \in \mathcal{S}_0\}$. First, we approximate $\{(T(\boldsymbol{s}), U(\boldsymbol{s})) :$

$\boldsymbol{s} \in \mathcal{S}_0\}$ with some standardized sum of influence functions, $\{(V(\boldsymbol{s}), W(\boldsymbol{s})) : \boldsymbol{s} \in \mathcal{S}_0\}$, for further normal approximation. Then, we divide the null set into several subsets to calculate the conditional mean squared error and show the highly correlated sets are negligible.

Step 3.1. Define

$$V(\boldsymbol{s}) = \frac{\sum_{k=1}^{n} Z_k(\boldsymbol{s})}{\mathsf{Var}\left\{\sum_{k=1}^{n} Z_k(\boldsymbol{s})\right\}^{1/2}},$$

$$W(\boldsymbol{s}) = \frac{\sum_{k=1}^{n_1} Z_k(\boldsymbol{s}) - \vartheta(\boldsymbol{s}) \sum_{k=n_1+1}^{n} Z_k(\boldsymbol{s})}{\mathsf{Var}\left\{\sum_{k=1}^{n_1} Z_k(\boldsymbol{s}) - \vartheta(\boldsymbol{s}) \sum_{k=n_1+1}^{n} Z_k(\boldsymbol{s})\right\}^{1/2}}.$$

Then, under Condition (C5), we have for some constant $C_2 > 5$,

$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0} |T(\boldsymbol{s}) - V(\boldsymbol{s})| \geq b_m\right\} = O(m^{-C_2}), \tag{S2.8}$$

$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0} |[U(\boldsymbol{s}) - \mu(\boldsymbol{s})] - W(\boldsymbol{s})| \geq b_m\right\} = O(m^{-C_2}), \tag{S2.9}$$

where $b_m = o\left\{(\log m)^{-1/2}\right\}$. Note that under event $B_{\boldsymbol{s}}$ where $\pi_\tau(\boldsymbol{s}, U(\boldsymbol{s})) \in [\xi, 1-\xi]$ for some constant $0 < \xi < 1$, we have

$$G^{-1}\left[G\left\{(a_1 \log m + a_2 \log\log m)^{1/2}\right\} \tilde{w}\left(\boldsymbol{s}, U(\boldsymbol{s})\right)\right] = (a_1 \log m + a_2 \log\log m + a_3)^{1/2}, \tag{S2.10}$$

for some constants a_1, a_2 and a_3. Recall that $z_{\tilde{w}}(\boldsymbol{s}) = G^{-1}\left\{p(\boldsymbol{s})/\tilde{w}(\boldsymbol{s}, U(\boldsymbol{s}))\right\}$ and $|T(\boldsymbol{s})| = G^{-1}\left\{p(\boldsymbol{s})\right\}$. Then $T(\boldsymbol{s})$ and $z_{\tilde{w}}(\boldsymbol{s})$ share the same order under event $B_{\boldsymbol{s}}$. By (S2.8) and the fact

$$G\left(t + o\left\{(\log m)^{-1/2}\right\}\right)/G(t) = 1 + o(1), \tag{S2.11}$$

uniformly in $0 \leq t \leq a_0(\log m)^{1/2}$ for any constant $a_0 > 0$, we have that,

$$G\left(|T(\boldsymbol{s})|\right)/\tilde{w}(\boldsymbol{s}, U(\boldsymbol{s})) = \{1 + o(1)\} G(|V(\boldsymbol{s})|)/\tilde{w}(\boldsymbol{s}, U(\boldsymbol{s})). \tag{S2.12}$$

Let $V_{\tilde{w}}(\boldsymbol{s}) = G^{-1}\{G(|V(\boldsymbol{s})|)/\tilde{w}(\boldsymbol{s}, U(\boldsymbol{s}))\}$. Recall $z_{\tilde{w}}(\boldsymbol{s}) = G^{-1}\left\{G\left(|T(\boldsymbol{s})|\right)/\tilde{w}(\boldsymbol{s}, U(\boldsymbol{s}))\right\}$. Then, by (S2.8), (S2.9), (S2.11) and (S2.12), and the proof of Proposition 1, we have that,

$$\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Pr\left\{|z_{\tilde{w}}(\boldsymbol{s})| \geq t | \theta(\boldsymbol{s}) = 0, B_{\boldsymbol{s}}\right\} = \{1 + o(1)\} \sum_{\boldsymbol{s}\in\mathcal{S}_0} \Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t | \theta(\boldsymbol{s}) = 0, \tilde{B}_{\boldsymbol{s}}\}, \tag{S2.13}$$

where $\tilde{B}_{\boldsymbol{s}} = \{U(\boldsymbol{s}) : \pi_\tau(\boldsymbol{s}, U(\boldsymbol{s})) \in [\xi, 1-\xi]\ \} \cap \{W(\boldsymbol{s}) : |W(\boldsymbol{s})| \le (2\log m)^{1/2}\}$. By the delta method, $V(\boldsymbol{s})$ and $V_{\tilde{w}}(\boldsymbol{s})$ share the same correlation structures. Then, it remains to show that, uniformly in $0 \le t \le t_m$,

$$\sup_{0\le t\le t_m} \left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t\}}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{W(\boldsymbol{s})}\left[\Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t | \theta(\boldsymbol{s}) = 0, \tilde{B}_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right]} - 1 \right| \to 0, \quad \text{(S2.14)}$$

in probability.

Step 3.2. To show (S2.14), we first discretize the range $[0, t_m]$, then divide the pairs of the null sets into several subsets: the pairs that share the same indices $\mathcal{S}_{01}$, the set of highly correlated pairs $\mathcal{S}_{02}$, and the set of weakly correlated pairs $\mathcal{S}_{03}$. We prove (S2.14) by showing the first two subsets are negligible, and $\mathcal{S}_{03}$ plays the dominant role.

More specifically, we divide the range $[0, t_m]$ into q sections with length no larger than $\upsilon_m = 1/\sqrt{\log m \log\log m}$. Let $0 = t_0 < t_1 < \ldots < t_q = t_m$, such that $t_\iota - t_{\iota-1} = \upsilon_m$ for $1 \le \iota \le q-1$, and $t_q - t_{q-1} \le \upsilon_m$, which gives $q \sim t_m/\upsilon_m$. Let $\Upsilon_{\boldsymbol{s}}(t) = \mathsf{E}_{W(\boldsymbol{s})}\left[\Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t | \theta(\boldsymbol{s}) = 0, \tilde{B}_{\boldsymbol{s}}\}\big|\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}\right]$. By Proposition 1, we have

$$\Upsilon_{\boldsymbol{s}}(t) = \{1 + o(1)\}G(t). \quad \text{(S2.15)}$$

For any t such that $t_{\iota-1} \le t \le t_\iota$, we have, for any $\boldsymbol{s} \in \mathcal{S}_0$,

$$\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t\}}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t)} \le \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t_{\iota-1}\}}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t_{\iota-1})} \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t_{\iota-1})}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t_\iota)},$$
$$\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t\}}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t)} \ge \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t_\iota\}}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t_\iota)} \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t_\iota)}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t_{\iota-1})}.$$

By (S2.15), we have $\Upsilon_{\boldsymbol{s}}(t_\iota)/\Upsilon_{\boldsymbol{s}}(t_{\iota-1}) = 1 + o(1)$. Thus, we only need to prove that,

$$\max_{0\le\iota\le q} \left| \frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0} [\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t_\iota\} - \Upsilon_{\boldsymbol{s}}(t_\iota)]}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t_\iota)} \right| \to 0, \quad \text{(S2.16)}$$

in probability.

Note that

$$
\begin{aligned}
&\Pr\left\{\max_{0\le\iota\le q}\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\left[\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t_\iota\}-\Upsilon_{\boldsymbol{s}}\left(t_\iota\right)\right]}{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}\left(t_\iota\right)}\right|\ge\epsilon\Big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right\}\\
\le&\sum_{\iota=0}^{q}\Pr\left\{\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\left[\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t_\iota\}-\Upsilon_{\boldsymbol{s}}\left(t_\iota\right)\right]}{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}\left(t_\iota\right)}\right|\ge\epsilon\Big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right\}\\
\le&\frac{1}{\upsilon_m}\int_0^{t_m}\Pr\left\{\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\left[\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t_\iota\}-\Upsilon_{\boldsymbol{s}}\left(t_\iota\right)\right]}{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}\left(t_\iota\right)}\right|\ge\epsilon\Big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right\}dt_\iota\\
&+\sum_{\iota=q-1}^{q}\Pr\left\{\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\left[\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t_\iota\}-\Upsilon_{\boldsymbol{s}}\left(t_\iota\right)\right]}{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}\left(t_\iota\right)}\right|\ge\epsilon\Big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right\}.
\end{aligned}
$$

Then it remains to show that, for any constant $\epsilon>0$,

$$
\int_0^{t_m}\Pr\left\{\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\left[\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t\}-\Upsilon_{\boldsymbol{s}}(t)\right]}{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}(t)}\right|\ge\epsilon\Big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right\}dt=o(\upsilon_m), \tag{S2.17}
$$

$$
\sup_{0\le t\le t_m}\Pr\left\{\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\left[\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t\}-\Upsilon_{\boldsymbol{s}}(t)\right]}{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}(t)}\right|\ge\epsilon\Big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right\}=o(1). \tag{S2.18}
$$

By conditional Markov's inequality, we have that, for any constant $\epsilon>0$,

$$
\begin{aligned}
&\Pr\left\{\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t\}-\Upsilon_{\boldsymbol{s}}(t)}{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}(t)}\right|\ge\epsilon\Big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right\}\\
\le&\,\mathsf{E}_{\{V_{\tilde{w}}(\boldsymbol{s}),W(\boldsymbol{s})\},\boldsymbol{s}\in\mathcal{S}_0}\left\{\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t\}-\Upsilon_{\boldsymbol{s}}(t)}{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}(t)}\right|^2\big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right\}/\epsilon^2,
\end{aligned}
$$

where

$$
\begin{aligned}
&\mathsf{E}_{\{V_{\tilde{w}}(\boldsymbol{s}),W(\boldsymbol{s})\},\boldsymbol{s}\in\mathcal{S}_0}\left\{\left|\frac{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t\}-\Upsilon_{\boldsymbol{s}}(t)}{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}(t)}\right|^2\big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right\}\\
=&\frac{\sum_{\boldsymbol{s},\boldsymbol{l}\in\mathcal{S}_0}\mathsf{E}_{V_{\tilde{w}}(\boldsymbol{s}),W(\boldsymbol{s}),V_{\tilde{w}}(\boldsymbol{l}),W(\boldsymbol{l})}\left[\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t\}\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{l})|\ge t\}\big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}(t)\right\}^2}\\
&-\frac{\sum_{\boldsymbol{s},\boldsymbol{l}\in\mathcal{S}_0}\prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}}\mathsf{E}_{V_{\tilde{w}}(\boldsymbol{b}),W(\boldsymbol{b})}\left[\mathsf{I}\{|V_{\tilde{w}}(\boldsymbol{b})|\ge t\}\big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}(t)\right\}^2}\\
=&\frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_0}\mathsf{E}_{W(\boldsymbol{s}),W(\boldsymbol{l})}\left[\Pr\{|V_{\tilde{w}}(\boldsymbol{s})|\ge t,|V_{\tilde{w}}(\boldsymbol{l})|\ge t|\theta(\boldsymbol{s})=0,\theta(\boldsymbol{l})=0,\tilde{B}_{\boldsymbol{s}},\tilde{B}_{\boldsymbol{l}}\}\big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}(t)\right\}^2}\\
&-\frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_0}\prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}}\mathsf{E}_{W(\boldsymbol{b})}\left[\Pr\{|V_{\tilde{w}}(\boldsymbol{b})|\ge t|\theta(\boldsymbol{b})=0,\tilde{B}_{\boldsymbol{b}}\}\big|\{\theta(\boldsymbol{s}),\boldsymbol{s}\in\mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0}\Upsilon_{\boldsymbol{s}}(t)\right\}^2}
\end{aligned}
$$

Recall $\Gamma_{\boldsymbol{s}}(\gamma) = \{\boldsymbol{l} : \boldsymbol{l} \in \mathcal{S}, |r_{\boldsymbol{s},\boldsymbol{l};d}| \geq (\log m)^{-2-\gamma}$ for $d = 1$ or $2\}$. We divide the null indices into three subsets: $\mathcal{S}_{01} = \{(\boldsymbol{s}, \boldsymbol{l}) : \boldsymbol{s}, \boldsymbol{l} \in \mathcal{S}_0, \boldsymbol{s} = \boldsymbol{l}\}$, $\mathcal{S}_{02} = \{(\boldsymbol{s}, \boldsymbol{l}) : \boldsymbol{s}, \boldsymbol{l} \in \mathcal{S}_0, \boldsymbol{s} \in \Gamma_{\boldsymbol{l}}(\gamma),$ or $\boldsymbol{l} \in \Gamma_{\boldsymbol{s}}(\gamma)\}$, and $\mathcal{S}_{03} = \{(\boldsymbol{s}, \boldsymbol{l}) : \boldsymbol{s}, \boldsymbol{l} \in \mathcal{S}_0\} \backslash (\mathcal{S}_{01} \cup \mathcal{S}_{02})$.

Note that for $\mathcal{S}_{01}$, by the proof of Proposition 1, we have that, for some constant C,

$$\frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{01}} \mathsf{E}_{W(\boldsymbol{s})} \left[\left(\Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t|\theta(\boldsymbol{s}) = 0, \tilde{B}_{\boldsymbol{s}}\} + \Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t|\theta(\boldsymbol{s}) = 0, \tilde{B}_{\boldsymbol{s}}\}\right)^2 \Big|\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t)\right\}^2}$$

$$\leq \frac{2}{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t)} \leq \frac{C}{mG(t)},$$

where the last inequality comes from (S2.15).

For $\mathcal{S}_{02}$, by the condition $\max_{\boldsymbol{s}\in\mathcal{S}_0} \text{Card}\,(\Gamma_{\boldsymbol{s}}(\gamma)) \asymp 1$ under Condition (C6), we have that,

$$\frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{02}} \mathsf{E}_{W(\boldsymbol{s}),W(\boldsymbol{l})} \left[\Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t, |V_{\tilde{w}}(\boldsymbol{l})| \geq t|\theta(\boldsymbol{s}) = 0, \theta(\boldsymbol{l}) = 0, \tilde{B}_{\boldsymbol{s}}, \tilde{B}_{\boldsymbol{l}}\}\Big|\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t)\right\}^2}$$

$$- \frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{02}} \prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}} \mathsf{E}_{W(\boldsymbol{b})} \left[\Pr\{|V_{\tilde{w}}(\boldsymbol{b})| \geq t|\theta(\boldsymbol{b}) = 0, \tilde{B}_{\boldsymbol{b}}\}\Big|\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t)\right\}^2}$$

$$\leq \frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{02}} \mathsf{E}_{W(\boldsymbol{s})} \left[\left(\Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t|\theta(\boldsymbol{s}) = 0, \tilde{B}_{\boldsymbol{s}}\} + \Pr^2\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t|\theta(\boldsymbol{s}) = 0, \tilde{B}_{\boldsymbol{s}}\}\right) \Big|\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t)\right\}^2}$$

$$\leq \frac{C}{mG(t)}.$$

For $\mathcal{S}_{03}$, by Lemma 3 that we prove after this lemma, we have that,

$$\frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{03}} \mathsf{E}_{W(\boldsymbol{s}),W(\boldsymbol{l})} \left[\Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t, |V_{\tilde{w}}(\boldsymbol{l})| \geq t|\theta(\boldsymbol{s}) = 0, \theta(\boldsymbol{l}) = 0, \tilde{B}_{\boldsymbol{s}}, \tilde{B}_{\boldsymbol{l}}\}\Big|\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t)\right\}^2}$$

$$- \frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{03}} \prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}} \mathsf{E}_{W(\boldsymbol{b})} \left[\Pr\{|V_{\tilde{w}}(\boldsymbol{b})| \geq t|\theta(\boldsymbol{b}) = 0, \tilde{B}_{\boldsymbol{b}}\}\Big|\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \Upsilon_{\boldsymbol{s}}(t)\right\}^2}$$

$$= \frac{O\left\{(\log m)^{-1-\gamma_1}\right\} \sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{03}} \prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}} \mathsf{E}_{W(\boldsymbol{b})} \left[\Pr\{|V_{\tilde{w}}(\boldsymbol{b})| \geq t|\theta(\boldsymbol{b}) = 0, \tilde{B}_{\boldsymbol{b}}\}\Big|\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}\right]}{\sum_{\boldsymbol{s},\boldsymbol{l}\in\mathcal{S}_0} \prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}} \mathsf{E}_{W(\boldsymbol{b})} \left[\Pr\{|V_{\tilde{w}}(\boldsymbol{b})| \geq t|\theta(\boldsymbol{b}) = 0, \tilde{B}_{\boldsymbol{b}}\}\Big|\{\theta(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}\right]} \tag{S2.19}$$

$$= O\left\{(\log m)^{-1-\gamma_1}\right\},$$

where $\gamma_1 = \min(\gamma, 1/2)$.

Note that

$$\int_0^{t_m} C(\log m)^{-1-\gamma_1}/\epsilon^2 dt = Ct_m(\log m)^{-1-\gamma_1}/\epsilon^2 = o(\upsilon_m). \tag{S2.20}$$

Also, by $1-\Phi(t) > t/\left\{(2c_\pi)^{1/2}(1+t^2)\right\}\exp(-t^2/2)$ for $t>0$ and monotonicity of $G(t)$, we have

$$\begin{aligned}\int_0^{t_m} \frac{C/\epsilon^2}{mG(t)}dt &\le \int_0^1 \frac{C'}{m}dt + \int_1^{t_m} \frac{C'te^{t^2/2}}{m}dt \\ &= \frac{C'}{m} + \frac{C'}{m}\, e^{t^2/2}\Big|_1^{t_m} = \frac{C'}{m} + \frac{C'}{m}\left(\frac{m}{\log m} - e^{1/2}\right) = o(\upsilon_m),\end{aligned} \tag{S2.21}$$

for any constant $\epsilon > 0$. Combining (S2.20) and (S2.21) proves (S2.17). Note that

$$C(\log m)^{-1-\gamma_1}/\epsilon^2 = o(1), \quad \frac{C}{\epsilon^2 mG(t)} = o(1), \tag{S2.22}$$

which proves (S2.18). By combining (S2.17) and (S2.18), we prove (S2.16).

This completes the proof of Lemma 2. □

Lemma 3. *Under the conditions of Lemma 2, letting* $\gamma_1 = \min(\gamma, 1/2)$*, we have that,*

$$\begin{aligned}&\max_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{03}} \mathsf{E}_{W(\boldsymbol{s}),W(\boldsymbol{l})}\left[\Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t, |V_{\tilde{w}}(\boldsymbol{l})| \ge t | \theta(\boldsymbol{s}) = 0, \theta(\boldsymbol{l}) = 0, \tilde{B}_{\boldsymbol{s}}, \tilde{B}_{\boldsymbol{l}}\}\big|\{\theta(\boldsymbol{s}), \boldsymbol{s}\in\mathcal{S}\}\right] \\ &= \left[1 + O\left\{(\log m)^{-1-\gamma_1}\right\}\right]\prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}} \mathsf{E}_{W(\boldsymbol{b})}\left[\Pr\{|V_{\tilde{w}}(\boldsymbol{b})| \ge t | \theta(\boldsymbol{b}) = 0, \tilde{B}_{\boldsymbol{b}}\}\big|\{\theta(\boldsymbol{s}), \boldsymbol{s}\in\mathcal{S}\}\right],\end{aligned}$$

uniformly for $0 \le t \le (2\log m)^{1/2}$.

Proof. We first note that it suffices to show

$$\begin{aligned}&\mathsf{E}_{W(\boldsymbol{s}),W(\boldsymbol{l})}\left[\lim_{\epsilon\to 0}\frac{\Pr_{H_0}\{|V_{\tilde{w}}(\boldsymbol{s})| \ge t, |V_{\tilde{w}}(\boldsymbol{l})| \ge t, |W(\boldsymbol{s}) - \tilde{u}_{\boldsymbol{s}}| \le \epsilon, |W(\boldsymbol{l}) - \tilde{u}_{\boldsymbol{l}}| \le \epsilon\}}{\Pr_{H_0}\{|W(\boldsymbol{s}) - \tilde{u}_{\boldsymbol{s}}| \le \epsilon, |W(\boldsymbol{l}) - \tilde{u}_{\boldsymbol{l}}| \le \epsilon\}}\big|\{\theta(\boldsymbol{s}), \boldsymbol{s}\in\mathcal{S}\}\right] \\ &= \left[1 + O\left\{(\log m)^{-1-\gamma_1}\right\}\right]\prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}} \mathsf{E}_{W(\boldsymbol{b})}\left[\Pr_{H_0}\{|V_{\tilde{w}}(\boldsymbol{b})| \ge t | W(\boldsymbol{b})\}\,\big|\{\theta(\boldsymbol{s}), \boldsymbol{s}\in\mathcal{S}\}\right],\end{aligned}$$

uniformly for $0 \le t \le (2\log m)^{1/2}$, $|W(\boldsymbol{s})| \le (2\log m)^{1/2}$, and $(\boldsymbol{s},\boldsymbol{l}) \in \mathcal{S}_{03}$.

Toward that end, we first truncate the statistics similarly as in the proof of Proposition 1. Define

$$\hat{V}(\boldsymbol{s}) = \frac{\sum_{k=1}^{n} \hat{Z}_k(\boldsymbol{s})}{\mathsf{Var}\left\{\sum_{k=1}^{n} \hat{Z}_k(\boldsymbol{s})\right\}^{1/2}},$$
$$\hat{W}(\boldsymbol{s}) = \frac{\sum_{k=1}^{n_1} \hat{Z}_k(\boldsymbol{s}) - \vartheta(\boldsymbol{s}) \sum_{k=n_1+1}^{n} \hat{Z}_k(\boldsymbol{s})}{\mathsf{Var}\left\{\sum_{k=1}^{n_1} \hat{Z}_k(\boldsymbol{s}) - \vartheta(\boldsymbol{s}) \sum_{k=n_1+1}^{n} \hat{Z}_k(\boldsymbol{s})\right\}^{1/2}},$$

where $\hat{Z}_k(\boldsymbol{s}) = Z_k(\boldsymbol{s})\mathsf{I}\left\{|Z_k(\boldsymbol{s})| \le \tau_n\right\} - \mathsf{E}\left[Z_k(\boldsymbol{s})\mathsf{I}\left\{|Z_k(\boldsymbol{s})| \le \tau_n\right\}\right]$ and $\tau_n = C\log(m \vee n)$ for some constnat $C > 0$.

Let $\hat{V}_{\tilde{w}}(\boldsymbol{s}) = G^{-1}\{G(|\hat{V}(\boldsymbol{s})|)/\tilde{w}(\boldsymbol{s}, U(\boldsymbol{s}))\}$. By the fact $G\left(t + O\{(\log m)^{-2}\}\right)/G(t) = 1+O\left\{(\log m)^{-2/3}\right\}$ uniformly in $0 \le t \le a_0(\log m)^{1/2}$ for any constant $a_0 > 0$. Then, similar to the proof of Proposition 1, we can choose τ_n, such that for some constant C_5 specified later,

$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\left|V_{\tilde{w}}(\boldsymbol{s}) - \hat{V}_{\tilde{w}}(\boldsymbol{s})\right| \ge (\log m)^{-2}\right\} = O(m^{-C_5}), \tag{S2.23}$$
$$\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_0}\left|W(\boldsymbol{s}) - \hat{W}(\boldsymbol{s})\right| \ge (\log m)^{-2}\right\} = O(m^{-C_5}), \tag{S2.24}$$

and that for any $(\boldsymbol{s}, \boldsymbol{l}) \in \mathcal{S}_{03}$ defined in Lemma 2,

$$\begin{aligned}&\left\|\mathsf{Cov}\{\hat{V}_{\tilde{w}}(\boldsymbol{s}), \hat{V}_{\tilde{w}}(\boldsymbol{l}), \hat{W}(\boldsymbol{s}), \hat{W}(\boldsymbol{l})\} - \mathsf{Cov}\{V_{\tilde{w}}(\boldsymbol{s}), V_{\tilde{w}}(\boldsymbol{l}), W(\boldsymbol{s}), W(\boldsymbol{l})\}\right\|_2\\ &\quad = O\left\{(\log m)^{-2-\gamma}\right\}.\end{aligned} \tag{S2.25}$$

Following the proof of Proposition 1, let $\boldsymbol{N} = (N_1, N_2, N_3, N_4)$ be a multivariate normal vector with mean zero and covariance matrix $\boldsymbol{\Sigma} = \mathsf{Cov}\{\hat{V}_{\tilde{w}}(\boldsymbol{s}), \hat{V}_{\tilde{w}}(\boldsymbol{l}), \hat{W}(\boldsymbol{s}), \hat{W}(\boldsymbol{l})\}$ for $(\boldsymbol{s}, \boldsymbol{l}) \in \mathcal{S}_{03}$. Then by Theorem 1.1 in Zaitsev (1987), we have that,

$$\begin{aligned}&\Pr_{H_0}\left\{|\hat{V}_{\tilde{w}}(\boldsymbol{s})| \ge t, |\hat{V}_{\tilde{w}}(\boldsymbol{l})| \ge t, |\hat{W}(\boldsymbol{s}) - \tilde{u}_{\boldsymbol{s}}| \le \epsilon, |\hat{W}(\boldsymbol{l}) - \tilde{u}_{\boldsymbol{l}}| \le \epsilon\right\}\\ &\le \Pr\left\{\min\left(|N_1|, |N_2|\right) \ge t - \epsilon_n, \max\left(|N_3 - \tilde{u}_{\boldsymbol{s}}|, |N_4 - \tilde{u}_{\boldsymbol{l}}|\right) \le \epsilon + \epsilon_n\right\} + c_1\exp\left\{\frac{n^{1/2}\epsilon_n}{c_2\tau_n}\right\},\end{aligned}$$

and

$$\Pr_{H_0}\left\{|\hat{V}_{\tilde{w}}(\boldsymbol{s})|\geq t, |\hat{V}_{\tilde{w}}(\boldsymbol{l})|\geq t, |\hat{W}(\boldsymbol{s})-\tilde{u}_{\boldsymbol{s}}|\leq \epsilon, |\hat{W}(\boldsymbol{l})-\tilde{u}_{\boldsymbol{l}}|\leq \epsilon\right\}$$
$$\geq \Pr\left\{\min\left(|N_1|,|N_2|\right)\geq t+\epsilon_n, \max\left(|N_3-\tilde{u}_{\boldsymbol{s}}|,|N_4-\tilde{u}_{\boldsymbol{l}}|\right)\leq \epsilon-\epsilon_n\right\} - c_1\exp\left\{\frac{n^{1/2}\epsilon_n}{c_2\tau_n}\right\},$$

where c_1, c_2 are positive constants, and $\epsilon, \epsilon_n \to 0$ are to be specified later. Let $\mathcal{A} = \{\boldsymbol{N} : \min(|N_1|,|N_2|) \geq t_1, \max(|N_3-\tilde{u}_{\boldsymbol{s}}|,|N_4-\tilde{u}_{\boldsymbol{l}}|) \leq t_2\}$ for any $t_1, t_2 \in \mathbb{R}$. Furthermore, by (S2.25) and Conditions (C5) and (C6), we have $\|\boldsymbol{\Sigma}-\boldsymbol{I}\|_2 = O\{(\log m)^{-2-\gamma}\}$. Since $\Pr\left(\|\boldsymbol{N}\|_2^2 > C\log m\right) \leq 4\Pr\left(N_1^2 > C\log m/4\right)$, for any constant $C>0$, we have that,

$$\begin{aligned}\int_{\mathcal{A},\|\boldsymbol{N}\|_2^2>C\log m}\exp\left(-\frac{1}{2}\boldsymbol{N}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{N}\right)d\boldsymbol{N} \leq& c_1\int_{N_1^2>C\log m/4}\exp\left(-\frac{1}{2}N_1^2\right)dN_1\\ \leq& c_2(\log m)^{-1/2}\exp\left(-C\log m/8\right)\\ =& O\left\{(\log m)^{-1/2}m^{-C/8}\right\}, \qquad \text{(S2.26)}\end{aligned}$$

for some constants $c_1, c_2 > 0$. Then by $\|\boldsymbol{\Sigma}^{-1}-\boldsymbol{I}\|_2 \leq \|\boldsymbol{\Sigma}^{-1}\|_2\|\boldsymbol{\Sigma}-\boldsymbol{I}\|_2 = O\left\{(\log m)^{-2-\gamma}\right\}$ and the fact that $\boldsymbol{x}^\intercal\boldsymbol{A}\boldsymbol{x} \leq \|\boldsymbol{A}\|_2\|\boldsymbol{x}\|_2^2$, we have that,

$$\begin{aligned}&\int_{\mathcal{A},\|\boldsymbol{N}\|_2^2\leq C\log m}\exp\left(-\frac{1}{2}\boldsymbol{N}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{N}\right)d\boldsymbol{N}\\ =&\int_{\mathcal{A},\|\boldsymbol{N}\|_2^2\leq C\log m}\exp\left\{-\frac{1}{2}\boldsymbol{N}^\intercal(\boldsymbol{\Sigma}^{-1}-\boldsymbol{I})\boldsymbol{N}-\frac{1}{2}\|\boldsymbol{N}\|_2^2\right\}d\boldsymbol{N}\\ =&\left[1+O\{(\log m)^{-1-\gamma}\}\right]\int_{\mathcal{A},\|\boldsymbol{N}\|_2^2\leq C\log m}\exp\left(-\frac{1}{2}\|\boldsymbol{N}\|_2^2\right)d\boldsymbol{N}\\ =&\left[1+O\{(\log m)^{-1-\gamma}\}\right]\int_{\mathcal{A}}\exp\left(-\frac{1}{2}\|\boldsymbol{N}\|_2^2\right)d\boldsymbol{N}\\ &-\left[1+O\{(\log m)^{-1-\gamma}\}\right]\int_{\mathcal{A},\|\boldsymbol{N}\|_2^2>C\log m}\exp\left(-\frac{1}{2}\|\boldsymbol{N}\|_2^2\right)d\boldsymbol{N}\\ =&\left[1+O\{(\log m)^{-1-\gamma}\}\right]\int_{\mathcal{A}}\exp\left(-\frac{1}{2}\|\boldsymbol{N}\|_2^2\right)d\boldsymbol{N}+O\left\{(\log m)^{-1/2}m^{-C/8}\right\}. \qquad \text{(S2.27)}\end{aligned}$$

Combining (S2.26) and (S2.27), by $|\boldsymbol{\Sigma}| = \left(1+O\left\{(\log m)^{-2-\gamma}\right\}\right)^4|\boldsymbol{I}|$, with the density function of multivariate normal variable, we can let C be sufficiently large, such that

$$\Pr\left\{\min\left(|N_1|, |N_2|\right) \geq t_1, \max\left(|N_3 - \tilde{u}_{\boldsymbol{s}}|, |N_4 - \tilde{u}_{\boldsymbol{l}}|\right) \leq t_2\right\}$$
$$= \left[1 + O\{(\log m)^{-1-\gamma}\}\right] G^2(t_1) \left\{\Phi(\tilde{u}_{\boldsymbol{s}} + t_2) - \Phi(\tilde{u}_{\boldsymbol{s}} - t_2)\right\} \left\{\Phi(\tilde{u}_{\boldsymbol{l}} + t_2) - \Phi(\tilde{u}_{\boldsymbol{l}} - t_2)\right\}.$$

Then, combined with (S2.23) and (S2.24), we obtain that,

$$\Pr_{H_0}\left\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t, |V_{\tilde{w}}(\boldsymbol{l})| \geq t, |W(\boldsymbol{s}) - \tilde{u}_{\boldsymbol{s}}| \leq \epsilon, |W(\boldsymbol{l}) - \tilde{u}_{\boldsymbol{l}}| \leq \epsilon\right\}$$
$$\leq \left[1 + O\{(\log m)^{-1-\gamma}\}\right] G^2\left\{t - \epsilon_n\right\} \cdot 2\left\{\epsilon + \epsilon_n\right\}^2 \phi(u_{\boldsymbol{s}1})\phi(u_{\boldsymbol{l}1}) + c_1 \exp\left\{\frac{n^{1/2}\epsilon_n}{c_2\tau_n}\right\} + O(m^{-C_5}),$$

and

$$\Pr_{H_0}\left\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t, |V_{\tilde{w}}(\boldsymbol{l})| \geq t, |W(\boldsymbol{s}) - \tilde{u}_{\boldsymbol{s}}| \leq \epsilon, |W(\boldsymbol{l}) - \tilde{u}_{\boldsymbol{l}}| \leq \epsilon\right\}$$
$$\geq \left[1 + O\{(\log m)^{-1-\gamma}\}\right] G^2\left\{t + \epsilon_n\right\} \cdot 2\left\{\epsilon - \epsilon_n\right\}^2 \phi(u_{\boldsymbol{s}2})\phi(u_{\boldsymbol{l}2}) - c_1 \exp\left\{\frac{n^{1/2}\epsilon_n}{c_2\tau_n}\right\} + O(m^{-C_5}),$$

where $u_{\boldsymbol{b}1} = \arg\max \phi(\breve{u})$ for $\breve{u} \in [\tilde{u}_{\boldsymbol{b}} - \epsilon - \epsilon_n, \tilde{u}_{\boldsymbol{b}} + \epsilon + \epsilon_n]$, and $u_{\boldsymbol{b}2} = \arg\min \phi(\breve{u})$, for $\breve{u} \in [\tilde{u}_{\boldsymbol{b}} - \epsilon + \epsilon_n, \tilde{u}_{\boldsymbol{b}} + \epsilon - \epsilon_n]$, for $\boldsymbol{b} = \boldsymbol{s}, \boldsymbol{l}$ respectively. Similarly, we obtain that,

$$\Pr_{H_0}\left\{|W(\boldsymbol{s}) - \tilde{u}_{\boldsymbol{s}}| \leq \epsilon, |W(\boldsymbol{l}) - \tilde{u}_{\boldsymbol{l}}| \leq \epsilon\right\}$$
$$\leq \left[1 + O\{(\log m)^{-1-\gamma}\}\right] 2\left\{\epsilon + \epsilon_n\right\}^2 \phi(u_{\boldsymbol{s}1})\phi(u_{\boldsymbol{l}1}) + c_1 \exp\left\{\frac{n^{1/2}\epsilon_n}{c_2\tau_n}\right\} + O(m^{-C_5}),$$

and

$$\Pr_{H_0}\left\{|W(\boldsymbol{s}) - \tilde{u}_{\boldsymbol{s}}| \leq \epsilon, |W(\boldsymbol{l}) - \tilde{u}_{\boldsymbol{l}}| \leq \epsilon\right\}$$
$$\geq \left[1 + O\{(\log m)^{-1-\gamma}\}\right] 2\left\{\epsilon - \epsilon_n\right\}^2 \phi(u_{\boldsymbol{s}2})\phi(u_{\boldsymbol{l}2}) - c_1 \exp\left\{\frac{n^{1/2}\epsilon_n}{c_2\tau_n}\right\} + O(m^{-C_5}).$$

Letting $\epsilon = \epsilon_n = (\log m)^{-2}$, then under $\log m = o(n^{1/8})$, by $|u_{\boldsymbol{b}1} - u_{\boldsymbol{b}2}| = O\left\{(\log m)^{-2}\right\}$ for $\boldsymbol{b} = \boldsymbol{s}, \boldsymbol{l}$, we can let C_5 sufficiently large such that,

$$\Pr_{H_0}\left\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t, |V_{\tilde{w}}(\boldsymbol{l})| \geq t | W(\boldsymbol{s}), W(\boldsymbol{l})\right\}$$
$$\leq \lim_{\epsilon \to 0} \frac{\left[1 + O\{(\log m)^{-1-\gamma}\}\right] G^2\left\{t - \epsilon_n\right\} \cdot 2\left\{\epsilon + \epsilon_n\right\}^2 \phi(u_{\boldsymbol{s}1})\phi(u_{\boldsymbol{l}1}) + O(m^{-C_5})}{2\left\{\epsilon - \epsilon_n\right\}^2 \phi(u_{\boldsymbol{s}2})\phi(u_{\boldsymbol{l}2}) + O(m^{-C_5})}$$
$$= \left[1 + O\{(\log m)^{-1-\gamma_1}\}\right] G^2(t),$$

and

$$
\begin{aligned}
&\Pr_{H_0}\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t, |V_{\tilde{w}}(\boldsymbol{l})| \geq t | W(\boldsymbol{s}), W(\boldsymbol{l})\} \\
&\geq \lim_{\epsilon \to 0} \frac{[1 + O\{(\log m)^{-1-\gamma}\}]\, G^2\{t + \epsilon_n\} \cdot 2\{\epsilon - \epsilon_n\}^2 \phi(u_{\boldsymbol{s}2})\phi(u_{\boldsymbol{l}2}) + O(m^{-C_5})}{2\{\epsilon + \epsilon_n\}^2 \phi(u_{\boldsymbol{s}1})\phi(u_{\boldsymbol{l}1}) + O(m^{-C_5})} \\
&= \left[1 + O\{(\log m)^{-1-\gamma_1}\}\right] G^2(t),
\end{aligned}
$$

uniformly for $0 \leq t \leq (2\log m)^{1/2}$, where $\gamma_1 = \min(\gamma, 1/2)$. Hence we have

$$
\begin{aligned}
&\mathsf{E}_{W(\boldsymbol{s}),W(\boldsymbol{l})}\left[\Pr_{H_0}\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t, |V_{\tilde{w}}(\boldsymbol{l})| \geq t | W(\boldsymbol{s}), W(\boldsymbol{l})\} \,\middle|\, \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] \\
&= \left[1 + O\{(\log m)^{-1-\gamma_1}\}\right] G^2(t),
\end{aligned}
$$

uniformly for $0 \leq t \leq (2\log m)^{1/2}$ and $(\boldsymbol{s}, \boldsymbol{l}) \in \mathcal{S}_{03}$.

Meanwhile, we have

$$
\prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}} \mathsf{E}_{W(\boldsymbol{s}),W(\boldsymbol{l})}\left[\Pr_{H_0}\{|V_{\tilde{w}}(\boldsymbol{b})| \geq t | W(\boldsymbol{b})\} \,\middle|\, \{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}\right] = \left[1 + O\left\{(\log m)^{-3/2}\right\}\right] G^2(t).
$$

Then the desired result follows, which completes the proof of Lemma 3. □

S3 Additional Numerical Results and Discussions

S3.1 Additional 1D example

We consider an additional simulation with a 1D setting where there is no clear spatial pattern. More specifically, we consider $s = 1, \cdots, 5000$, set $\pi(s) = 0.7$ at 800 randomly sampled locations, and set $\pi(s) = 0.05$ for the rest of the locations. We first evaluate the accuracy of recovering $\pi(\boldsymbol{s}, U(\boldsymbol{s}))$. Figure S1 reports the result based on a single data replication. For this example, since there is no spatial pattern, LAWS performs very poorly. By contrast, the NAPA method performs much better. Next, we evaluate the empirical FDR and power. Figure S2 reports the empirical FDR and power of various testing methods based on 200 data replications. It is seen that LAWS suffers from some FDR inflation under this random signal

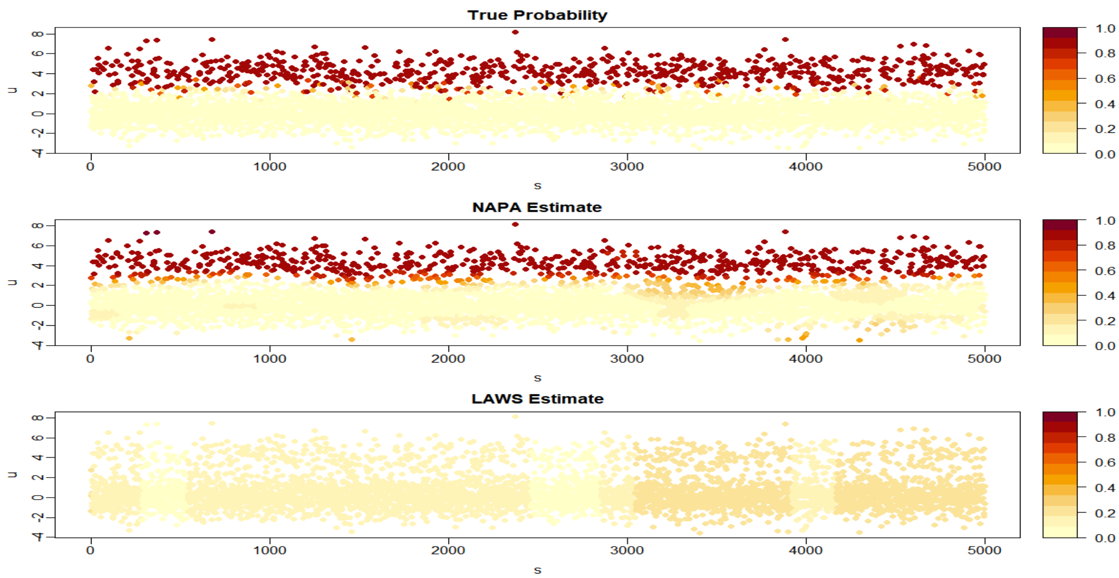

Figure S1: Estimation of the posterior non-null probability $\pi(\boldsymbol{s}, U(\boldsymbol{s}))$ for the additional 1D example. From top to bottom: the true probability, the estimated probability by NAPA, and the estimated probability by LAWS.

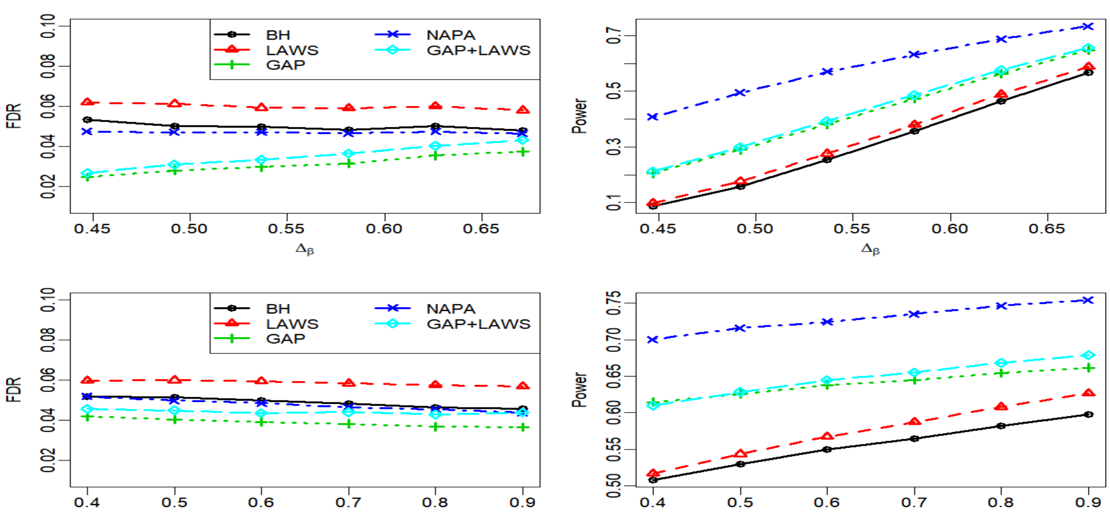

Figure S2: Empirical FDR and power for the additional 1D example. Top panels: varying Δ_β in scenario 1, and bottom panels: varying π in scenario 2. Five methods are compared: the proposed method (NAPA), the GAP method (Xia et al., 2020), the LAWS method (Cai et al., 2022), the simple combination of GAP and LAWS, and the BH method (Benjamini and Hochberg, 1995).

setting, and has little improvement over the BH procedure in most of the cases. Besides, the simple combination of GAP and LAWS has no apparent power advantage over GAP. By contrast, the proposed NAPA method enjoys the best power performance while having the FDR under control.

S3.2 Sensitivity analysis for bandwidth selection

We carry out a sensitivity analysis in the selection of the bandwidth matrix $\boldsymbol{H}$ in (5.8). There are three parameters involved in $\boldsymbol{H}$, i.e., $h_{\boldsymbol{s}}, h_U$ and ρ. Let $\hat{h}_{\boldsymbol{s}}, \hat{h}_U$ and $\hat{\rho}$ denote the selected parameters following the approach described in Section 5.1. We then vary one parameter while fixing the other two. For instance, we vary $h_{\boldsymbol{s}}$ as $k\hat{h}_{\boldsymbol{s}}$, with $k = \{0.6, 0.7, \ldots, 1.1, 1.2\}$, while fixing h_U and ρ at $\hat{h}_U$ and $\hat{\rho}$. Figure S3 reports the estimation result of the posterior non-null probability for a single data replication for the 2D example in Section 5.1. Figure S4 reports the empirical FDR and power of our testing method for the 1D example in Section 5.2.

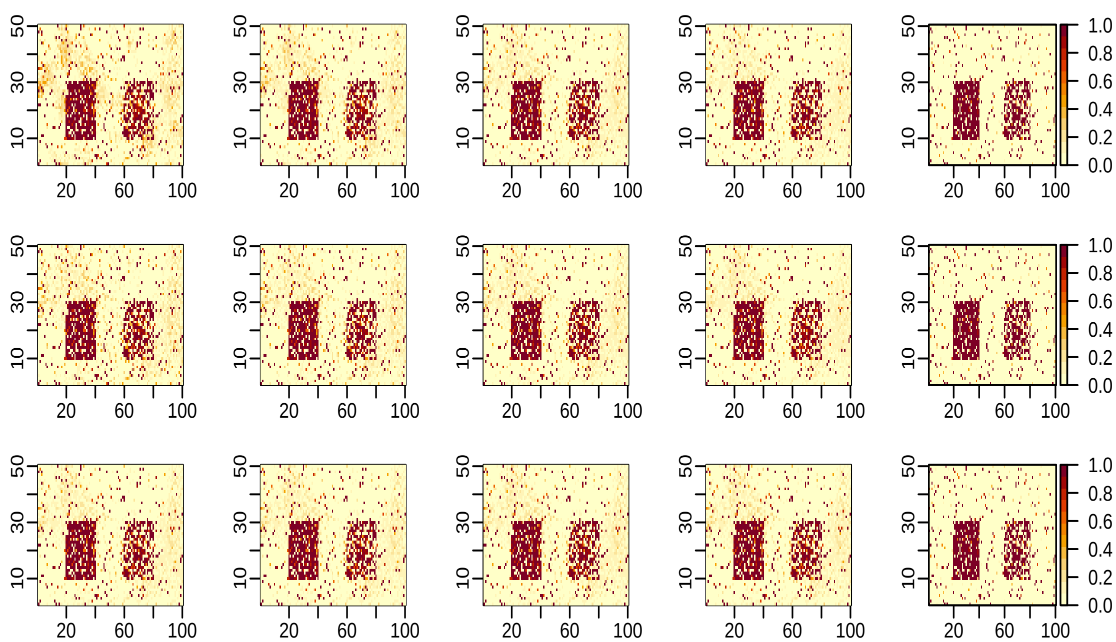

Figure S3: Sensitivity analysis: estimation of the posterior non-null probability for the 2D example. From top to bottom: we vary $h_{\boldsymbol{s}} = k\hat{h}_{\boldsymbol{s}}$, $h_U = k\hat{h}$, and $\rho = k\hat{\rho}$. From left to right: $k = 0.6, 0.8, 1.0, 1.2$, and the last column shows the true probability.

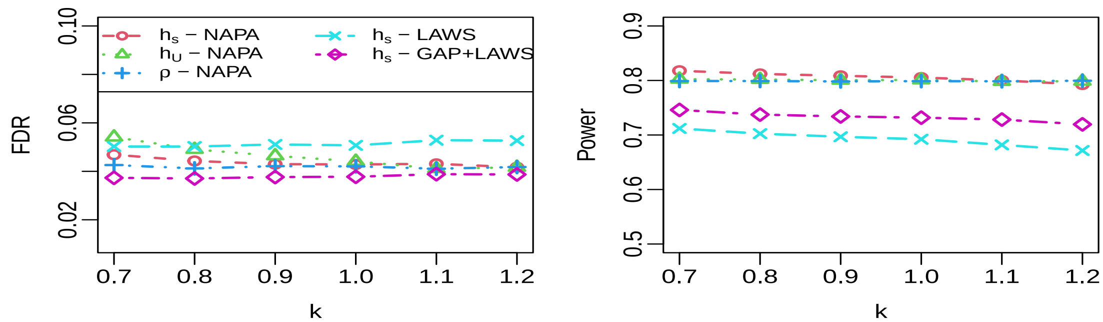

Figure S4: Sensitivity analysis: empirical FDR and power for the 1D example. We vary h_s, h_U and ρ respectively for NAPA, while vary h_s for LAWS and GAP+LAWS.

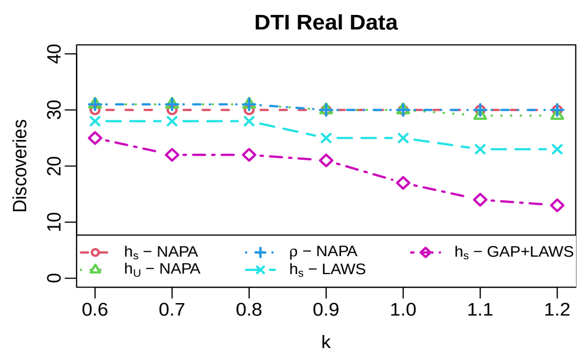

Figure S5: Sensitivity analysis: discoveries for the 1D DTI real data example.

We vary h_s, h_U and ρ respectively for NAPA, while vary h_s for LAWS and GAP+LAWS. All results are based on 200 data replications. Moreover, Figure S5 shows the discoveries of the 1D DTI real data example in Section 6.1 by NAPA, LAWS and GAP+LAWS, under varying bandwidth parameters. In all these plots, it is seen that our method achieves a relatively stable performance across a range of values of those parameters.

S3.3 Irregular domain and irregular lattice

Our method is generally applicable to spatial data with spatial smoothness, but usually regardless of the shape of the domain, or the specific lattice of the sampling observations. Next we consider some simulations examples with irregular domain and irregular lattice.

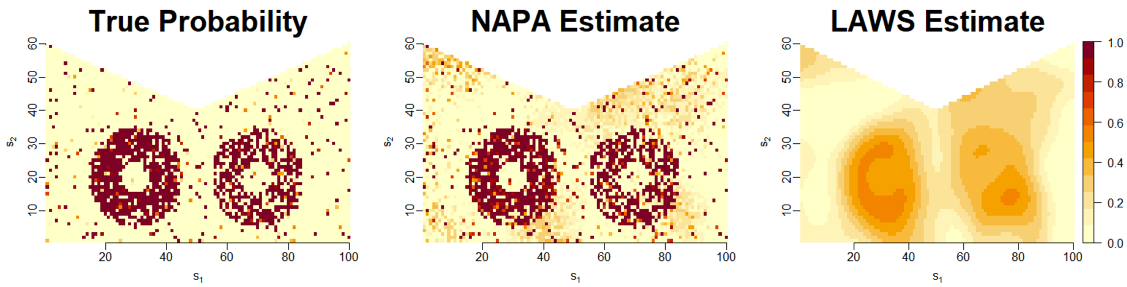

Figure S6: Irregular domain: estimation of the posterior non-null probability.

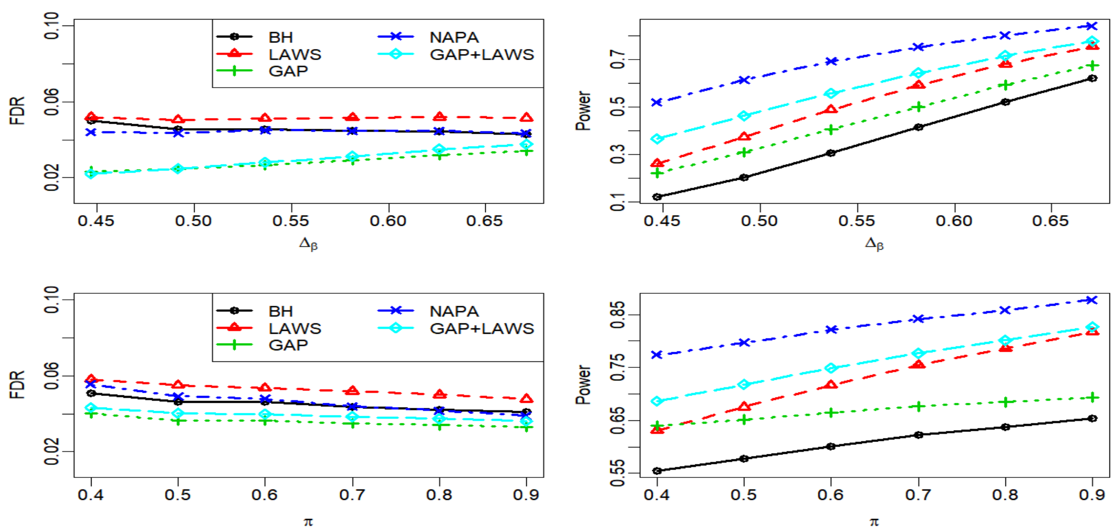

Figure S7: Irregular domain: empirical FDR and power.

First, we simulate a 2D setting where the signals have donut shapes located on a plane with an absence of a triangle, as shown in Figure S6. We set $\{\boldsymbol{s} = (s_1, s_2)\}$ to be grid points on a 100×60 plane, with $0.4s_1 + s_2 \leq 60$ for $s_1 = 1, 2, \ldots, 50$, and $0.4s_1 - s_2 \geq -20$ for $s_1 = 51, 52, \ldots, 100$. We set $\pi(\boldsymbol{s}) = 0.8$ for the left donut where $\text{dist}\{(s_1, s_2), (30, 20)\} \in [5, 15]$, set $\pi(\boldsymbol{s}) = 0.6$ for the right donut where $\text{dist}\{(s_1, s_2), (70, 20)\} \in [5, 15]$, and $\text{dist}(\cdot, \cdot)$ is the Euclidean distance. Figure S6 shows the accuracy of recovering the posterior non-null probability based on a single data replication, and Figure S7 reports the empirical FDR and power of various testing methods based on 200 replications for the two scenarios in Section 5.2. It is seen that our method continues to perform well in this irregular domain setting.

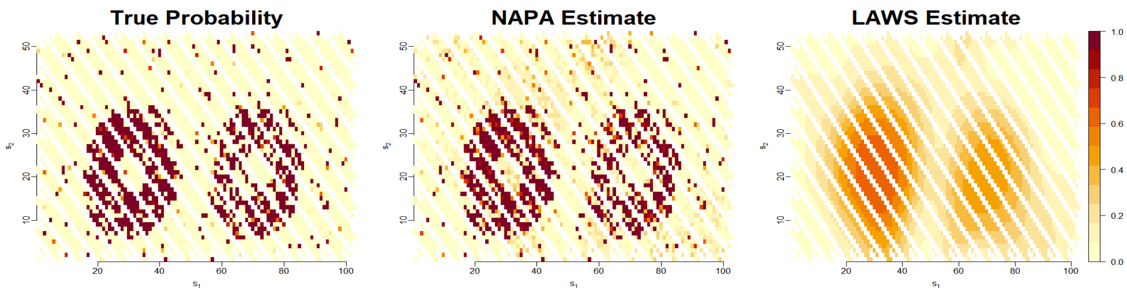

Figure S8: Irregular lattice: estimation of the posterior non-null probability.

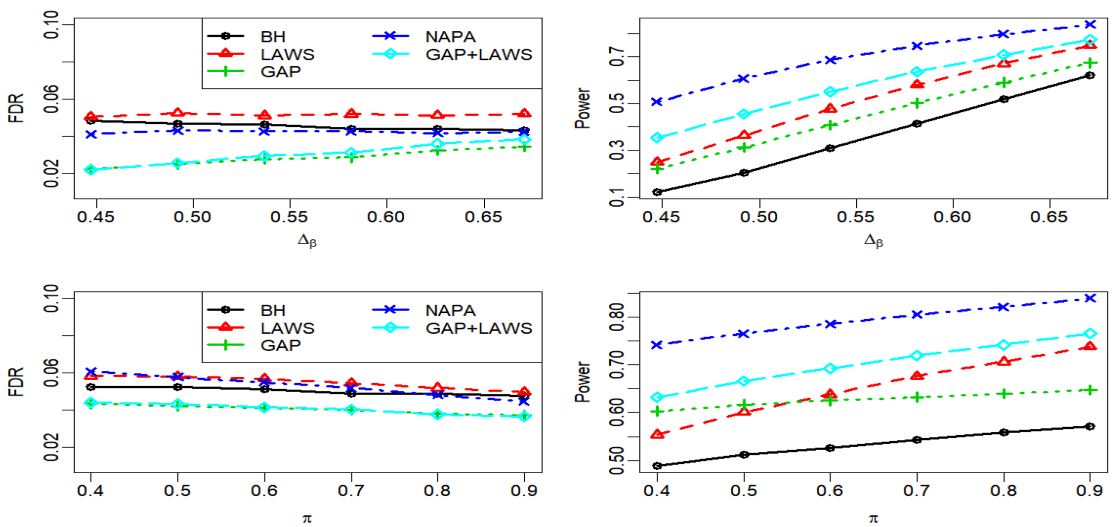

Figure S9: Irregular lattice: empirical FDR and power.

Next, we simulate a 2D setting where the signals have rough donut shapes and the lattice is irregular with a few diagonal strips missing, as shown in Figure S8. We set $\{\boldsymbol{s} = (s_{1a}, s_{2a})\}$ to be grid points on a 103×53 plane, where $s_{1a} = s'_{1a} + c_a$, $s_{2a} = s'_{2a} + c_a$ for $a = 1, \ldots, 5000$, $s'_{1a} = 1, 2, \ldots, 100$, $s'_{2a} = 1, 2, \ldots, 50$, and $(c_1, \ldots, c_{5000})$ is a vector that repeatedly replicates $(0, 1, 2, 3, 2, 1, 0)$. We set $\pi(\boldsymbol{s}) = 0.8$ for the left donut where $\text{dist}\{(s'_{1a}, s'_{2a}), (30, 20)\} \in [5, 15]$, and set $\pi(\boldsymbol{s}) = 0.6$ for the right donut where $\text{dist}\{(s'_{1a}, s'_{2a}), (70, 20)\} \in [5, 15]$. Figure S8 and Figure S9 report the results. It is seen again that our method performs well in this irregular lattice setting.

S3.4 Heavy-tailed distribution

In Section 5, we simulate the data from the normal distribution. We next consider an example that the data $\{Y_{i,1}(\boldsymbol{s})\}_{i=1}^{n_1}$ and $\{Y_{i,2}(\boldsymbol{s})\}_{i=1}^{n_2}$ are generated from the heavy-tailed t distribution,

$$Y_{i,1}(\boldsymbol{s}) \mid \theta(\boldsymbol{s}) \sim \{1-\theta(\boldsymbol{s})\}\, \mathrm{t}(\nu) \; + \; \theta(\boldsymbol{s})\, \{\beta_1(\boldsymbol{s}) + \mathrm{t}(\nu)\},$$
$$Y_{i,2}(\boldsymbol{s}) \mid \theta(\boldsymbol{s}) \sim \{1-\theta(\boldsymbol{s})\}\, 4\mathrm{t}(\nu) \; + \; \theta(\boldsymbol{s})\, \{\beta_1(\boldsymbol{s}) + \Delta_\beta + 4\mathrm{t}(\nu)\},$$

where $\theta(\boldsymbol{s}) \sim \text{Bernoulli}(1, \pi(\boldsymbol{s})), \beta_1(\boldsymbol{s}) = 1/\sqrt{20}, \Delta_\beta = 3/\sqrt{20}$ and $t(\nu)$ denotes a t distribution with the degree of freedom ν. We consider $\nu = \{3, 4, 5\}$ for the 2D example in Section 5.1. Figure S10 reports the estimation, and Figure S11 reports the testing results with varying Δ_β. It is seen that our method performs well under the heavy-tailed distribution.

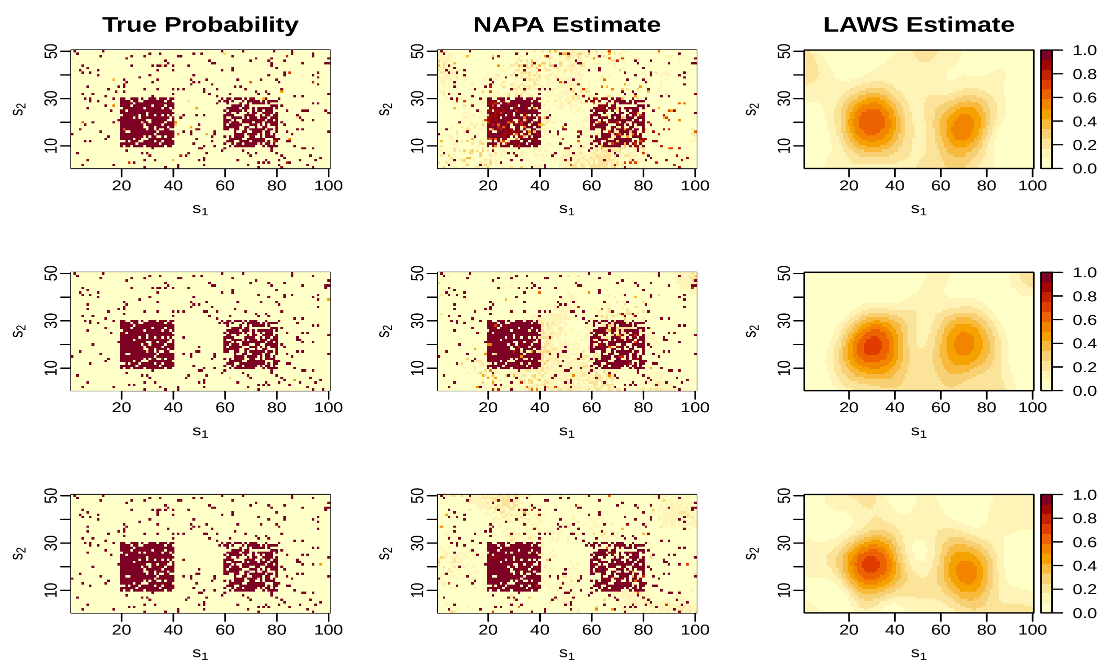

Figure S10: Heavy-tailed distribution: estimation of the posterior non-null probability. From top to bottom: $\nu = 3, 4, 5$.

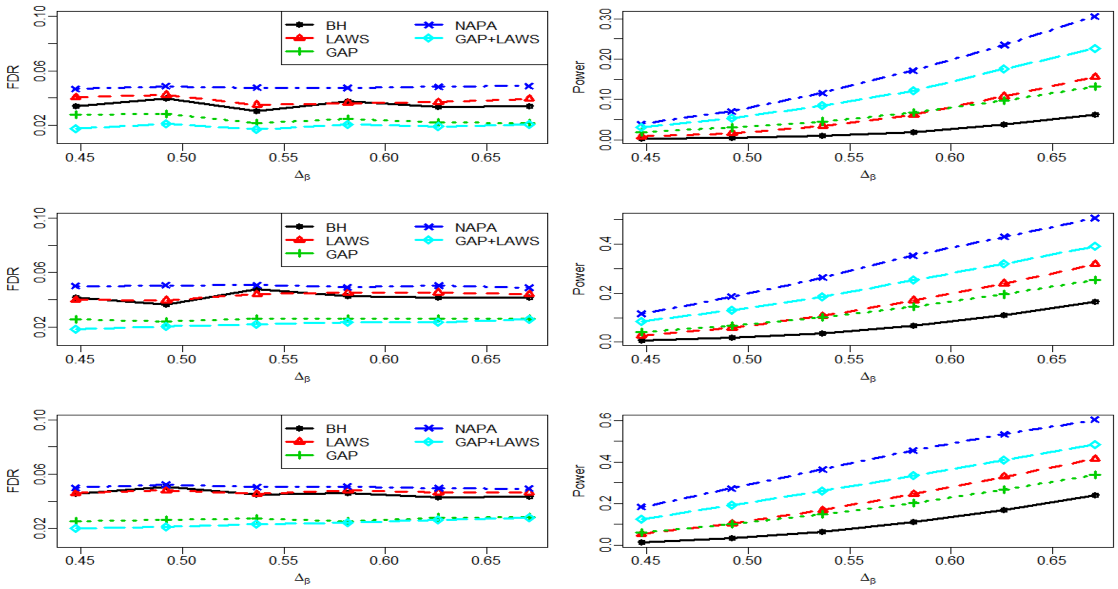

Figure S11: Heavy-tailed distribution: empirical FDR and power. From top to bottom: $\nu = 3, 4, 5$.

S3.5 Discussion of the conditions in Theorem 1

We give more discussion of the conditions in Theorem 1.

First, we assume $U(\boldsymbol{s})$ and $T(\boldsymbol{s})$ are independent under the alternative. In our oracle setting, we have $U(\boldsymbol{s})$ and $T(\boldsymbol{s})$ being independent under the null. Meanwhile, in our construction, it is relatively mild to obtain the asymptotic independence under the alternative too. More specifically, following (S1.19) and (S1.20) in the proof of Proposition 1, we can obtain that, for any constant $C > 0$, there exists some $b_m = o\{(\log m)^{-1/2}\}$, such that

$$\begin{aligned}
\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_1}\left|[T(\boldsymbol{s}) - \mu_T(\boldsymbol{s})] - \tilde{T}(\boldsymbol{s})\right| \geq b_m\right\} &= O(m^{-C}),\\
\Pr\left\{\max_{\boldsymbol{s}\in\mathcal{S}_1}\left|[U(\boldsymbol{s}) - \mu_U(\boldsymbol{s})] - \tilde{U}(\boldsymbol{s})\right| \geq b_m\right\} &= O(m^{-C}),
\end{aligned}$$

for $|T(\boldsymbol{s}) - \mu_T(\boldsymbol{s})| = O\{(\log m)^{1/2}\}, |U(\boldsymbol{s}) - \mu_U(\boldsymbol{s})| = O\{(\log m)^{1/2}\}$, where $\mu_T(\boldsymbol{s}) = [1 + o\{(\log m)^{-1}\}]\,\mathsf{E}\{T(\boldsymbol{s})\}$, $\mu_U(\boldsymbol{s}) = [1 + o\{(\log m)^{-1}\}]\,\mathsf{E}\{U(\boldsymbol{s})\}$, and $(\tilde{T}(\boldsymbol{s}), \tilde{U}(\boldsymbol{s}))$ are a

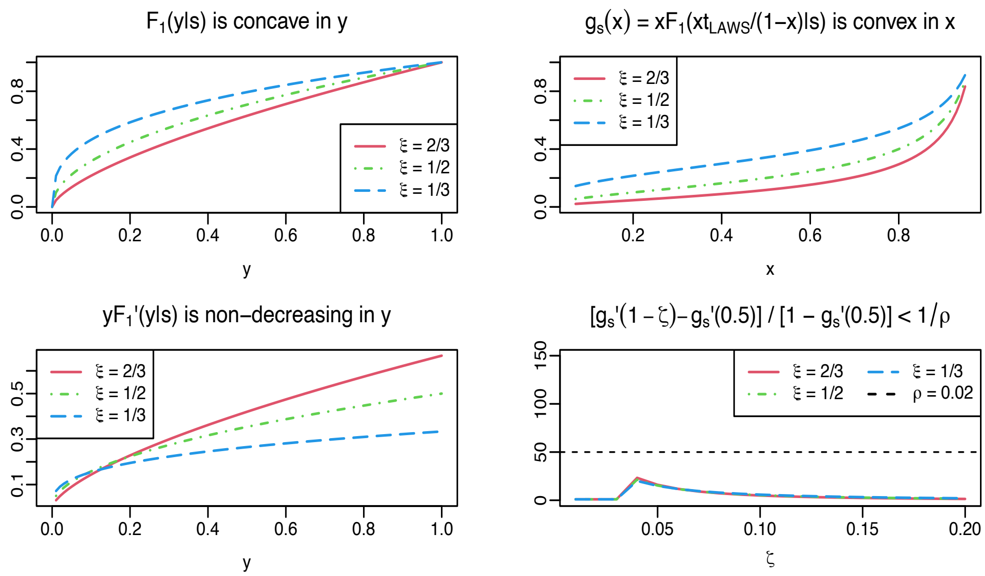

Figure S12: Verification of the four distributional conditions in Theorem 1.

pair of asymptotically independent normal variables. It then yields that

$$\mathsf{Cov}\left(T(\boldsymbol{s}), U(\boldsymbol{s})\right) \approx \mathsf{Cov}\left(\tilde{T}(\boldsymbol{s}) + \mu_T(\boldsymbol{s}), \tilde{U}(\boldsymbol{s}) + \mu_U(\boldsymbol{s})\right) \approx 0,$$

which indicates the asymptotic independence between $T(\boldsymbol{s})$ and $U(\boldsymbol{s})$ under the alternative.

Next, we numerically verify the four distributional conditions in Theorem 1. Recall that, we require that $F_1(y|\boldsymbol{s})$ is concave in y; $g_{\boldsymbol{s}}(x)$ is convex in x for $x \leq 1/(1+t_{\text{LAWS}})$; $yF_1'(y|\boldsymbol{s})$ is non-decreasing in y; and $\{g_{\boldsymbol{s}}'(1-\zeta) - g_{\boldsymbol{s}}'(0.5)\}/\{1 - g_{\boldsymbol{s}}'(0.5)\} \leq 1/\varrho$, for some $0 < \varrho \leq 1$. We generate the alternative p-values from a Beta distribution, $\text{Beta}(\xi, 1)$. Such a p-value distribution is widely adopted in the literature; see for example, Sellke et al. (2001); Held and Ott (2018); Zhang and Chen (2022). We vary $\xi = \{2/3, 1/2, 1/3\}$. Figure S12 shows the results. The upper left panel shows the concavity of $F_1(y|\boldsymbol{s})$, the upper right panel shows the convexity of $g_{\boldsymbol{s}}(x)$ at $t_{\text{LAWS}} = 0.04$, the bottom left panel shows the non-decreasing pattern of $yF_1'(y|\boldsymbol{s})$, and the bottom right panel verifies that $\{g_{\boldsymbol{s}}'(1-\zeta) - g_{\boldsymbol{s}}'(0.5)\}/\{1 - g_{\boldsymbol{s}}'(0.5)\} \leq 1/\varrho$ for $\zeta \in (0, 0.2)$ and $\varrho = 0.02$.

S3.6 Discussion of Condition (C3)

We further discuss Condition (C3) from three aspects. First, we present a sufficient condition for (C3), and discuss how it can be relaxed when modifying the bandwidth matrix condition in (C4). Next, we verify that (C3) holds for some common spatial structures. Finally, we consider an example where (C3) may *not* hold, but our method still maintains a good performance.

We first note that, by our construction of $T(\boldsymbol{s})$ and $U(\boldsymbol{s})$, the correlation structure among $T(\boldsymbol{s})$ and $U(\boldsymbol{s})$ can be characterized by the correlation matrices of the influence functions,

$$\boldsymbol{R}_1 = \mathsf{Corr}(\boldsymbol{Z}_k) = (r_{\boldsymbol{s},\boldsymbol{l};1})_{m\times m}, \quad 1 \le k \le n_1,$$

$$\boldsymbol{R}_2 = \mathsf{Corr}(\boldsymbol{Z}_k) = (r_{\boldsymbol{s},\boldsymbol{l};2})_{m\times m}, \quad n_1 + 1 < k \le n_1 + n_2.$$

Consider a simple case where $U(\boldsymbol{s})$ is normally distributed, $\mathsf{Var}(Z_k)$'s are of the same order, and $c_0 m$ of the location-wise conditional variances in (C3) share the same order of magnitude for some $0 < c_0 \le 1$. Then a sufficient condition for (C3) is that both $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are s-sparse, i.e., there are at most s nonzero entries in each row, where s is of a constant order. We can further relax this sufficient condition. For instance, if we set the bandwidth matrix in Condition (C4) so that it satisfies $|\boldsymbol{H}| \gg (m\tilde{h})^{-1}$, then the order of s can be relaxed from a constant order to the order of $O(m^{1/2})$.

Next, we note that the above sufficient condition (where s is of a constant order) covers some common spatial structures, including the banded structure, where there are a fixed number of off-diagonal bands along the diagonal, the hub structure, where the nodes are partitioned into disjoint groups with a fixed number of nodes in each group, and the random structure, where the probability for the off-diagonal entries to be nonzero is of order $O(m^{-1})$.

Finally, we consider an example with a Toeplitz covariance structure, for which Condition (C3) may not hold. Specifically, we simulate the data $\{Y_{i,1}(\boldsymbol{s})\}_{i=1}^{n_1}$ and $\{Y_{i,2}(\boldsymbol{s})\}_{i=1}^{n_2}$ from

$$\mathbf{Y}_{i,1} \mid \boldsymbol{\theta} \sim (1-\boldsymbol{\theta})\,\text{Normal}(\mathbf{0}, \boldsymbol{\Sigma}) \; + \; \boldsymbol{\theta}\,\text{Normal}\big(\beta_1 \mathbf{1}_m, \boldsymbol{\Sigma}\big),$$

$$\mathbf{Y}_{i,2} \mid \boldsymbol{\theta} \sim (1-\boldsymbol{\theta})\,\text{Normal}(\mathbf{0}, 4\boldsymbol{\Sigma}) \; + \; \boldsymbol{\theta}\,\text{Normal}\big((\beta_1 + \Delta_\beta)\mathbf{1}_m, 4\boldsymbol{\Sigma}\big),$$

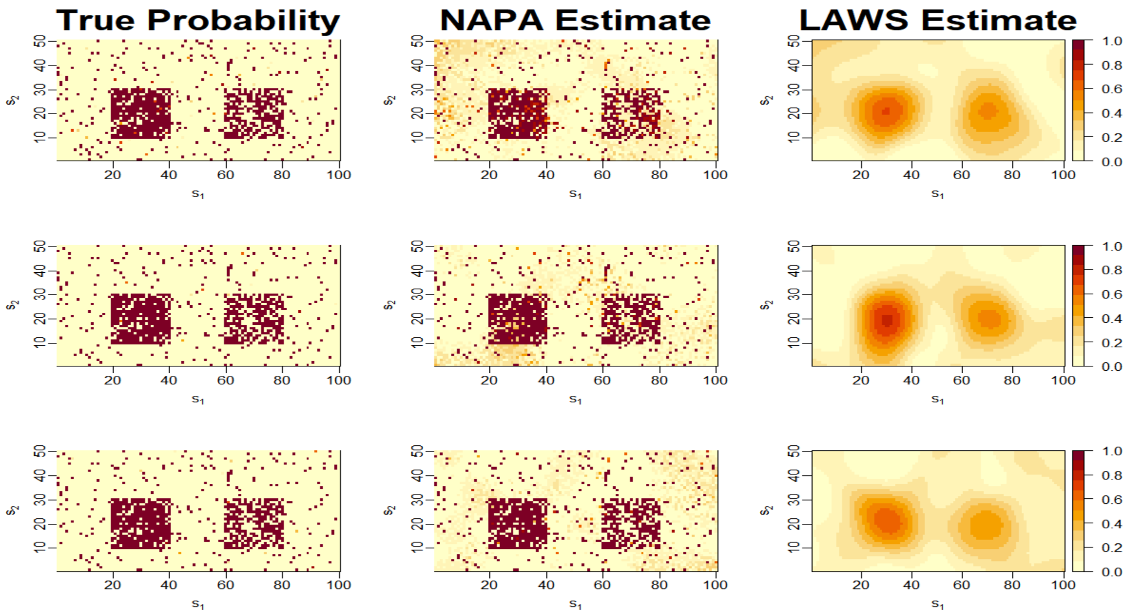

Figure S13: Robustness to Condition (C3): estimation of the posterior non-null probability under the Toeplitz structure. From top to bottom: $\rho = 0.4, 0.6, 0.8$.

where $\boldsymbol{\Sigma} = (\sigma_{i,j}), \sigma_{i,j} = \rho^{|i-j|}, 1 \leq i, j \leq m$, and $\mathbf{1}_m$ is a length-m vector with all ones. We vary $\rho = \{0.4, 0.6, 0.8\}$, and set $\beta_1 = 1/\sqrt{20}$, $\Delta_\beta = 3/\sqrt{20}$. Figure S13 reports the estimation result of the posterior non-null probability for a single data replication. It is seen that our NAPA method still outperforms LAWS and recovers the true posterior probability well.

S3.7 Discussion of Condition (C6)

We further discuss Condition (C6) from three aspects. First, we explain why this condition is important for our setting. Next, we verify that (C6) holds for some common spatial structures. Finally, we revisit the example in Section S3.6, where (C6) does *not* hold, but our method still maintains a good performance.

We first note that, existing theory of multiple testing under dependency mainly considers the setting with one sequence of p-values, whereas NAPA needs to deal with two sequences.

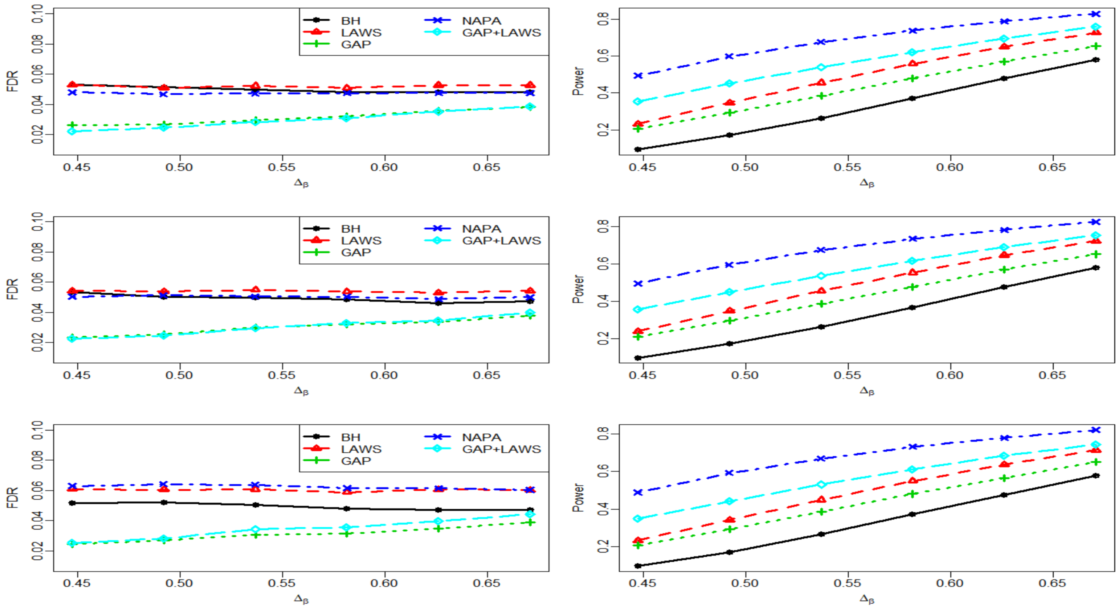

Figure S14: Robustness to Condition (C6): empirical FDR and power under the Toeplitz structure. From top to bottom: $\rho = 0.4, 0.6, 0.8$.

A key challenge in our theoretical analysis is to mitigate the effects of highly correlated pairs. Therefore, it requires a weak dependency condition that is stronger than those for the one sequence scenario. Specifically, for two pairs $(V_{\tilde{w}}(\boldsymbol{s}), W(\boldsymbol{s}))$ and $(V_{\tilde{w}}(\boldsymbol{l}), W(\boldsymbol{l}))$ that are highly correlated, where $(V_{\tilde{w}}(\boldsymbol{s}), V_{\tilde{w}}(\boldsymbol{l}))$ are the weighted approximations of the primary statistics $(T(\boldsymbol{s}), T(\boldsymbol{l}))$, and $(W(\boldsymbol{s}), W(\boldsymbol{l}))$ are the approximations of the auxiliary statistics $(U(\boldsymbol{s}), U(\boldsymbol{l}))$ as defined in Step 3.1 in the proof of Lemma 2, we need to show that

$$\frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{02}} \mathsf{E}_{W(\boldsymbol{s}),W(\boldsymbol{l})}\left[\Pr_{H_0}\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t, |V_{\tilde{w}}(\boldsymbol{l})| \geq t | \Omega\left(W(\boldsymbol{s})\right), \Omega\left(W(\boldsymbol{l})\right)\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{W(\boldsymbol{s})}\left[\Pr_{H_0}\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t | \Omega\left(W(\boldsymbol{s})\right)\}\right]\right\}^2}$$
$$- \frac{\sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathcal{S}_{02}} \prod_{\boldsymbol{b}=\boldsymbol{s},\boldsymbol{l}} \mathsf{E}_{W(\boldsymbol{b})}\left[\Pr_{H_0}\{|V_{\tilde{w}}(\boldsymbol{b})| \geq t | \Omega\left(W(\boldsymbol{b})\right)\}\right]}{\left\{\sum_{\boldsymbol{s}\in\mathcal{S}_0} \mathsf{E}_{W(\boldsymbol{s})}\left[\Pr\{|V_{\tilde{w}}(\boldsymbol{s})| \geq t | \Omega\left(W(\boldsymbol{s})\right)\}\right]\right\}^2} = O\left((\log m)^{-1-\gamma_1}\right),$$

for each realization of $\{\theta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}\}$, where $\Omega\left(W(\boldsymbol{s})\right)$ is some event relevant to $W(\boldsymbol{s})$, γ_1 is as defined in Step 3.2 in the proof of Lemma 2, and $\mathcal{S}_{02} = \{(\boldsymbol{s},\boldsymbol{l}) : \boldsymbol{s},\boldsymbol{l} \in \mathcal{S}_0, \boldsymbol{s} \in \Gamma_{\boldsymbol{l}}(\gamma), \text{ or } \boldsymbol{l} \in \Gamma_{\boldsymbol{s}}(\gamma)\}$. We should emphasize that, even if $T(\boldsymbol{s})$ and $U(\boldsymbol{s})$ are independent, the pairs $(T(\boldsymbol{s}), U(\boldsymbol{l}))$, $(T(\boldsymbol{s}), T(\boldsymbol{l}))$ and $(U(\boldsymbol{s}), U(\boldsymbol{l}))$ can still be highly dependent. This is significantly different from the existing literature where one only needs to deal with the correlations between one sequence of p-values.

Next, we note that Condition (C6) covers some common spatial structures, including the banded structure, the hub structure, and the random structure, similarly as for Condition (C3).

Finally, we revisit the example in Section S3.6. For this example, the Toeplitz covariance structure is exponentially decaying, and thus Condition (C6) does not hold. Figure S14 reports the empirical FDR and power of various testing methods based on 200 data replications. It is seen that our method still performs well, partly because Condition (C6) is sufficient but not necessary.

References

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B*, 57(1):289–300.

Cai, T. and Liu, W. (2011). Adaptive Thresholding for Sparse Covariance Matrix Estimation. *Journal of the American Statistical Association*, 106(494):672–684.

Cai, T. T., Sun, W., and Xia, Y. (2022). LAWS: A Locally Adaptive Weighting and Screening Approach to Spatial Multiple Testing. *Journal of the American Statistical Association*, 117:1370–1383.

Genovese, C. R., Roeder, K., and Wasserman, L. (2006). False discovery control with p-value weighting. *Biometrika*, 93(3):509–524.

Held, L. and Ott, M. (2018). On p-values and bayes factors. *Annual Review of Statistics and Its Application*, 5:393–419.

Sellke, T., Bayarri, M., and Berger, J. O. (2001). Calibration of ρ values for testing precise null hypotheses. *The American Statistician*, 55(1):62–71.

Xia, Y., Cai, T. T., and Sun, W. (2020). GAP: A General Framework for Information Pooling in Two-Sample Sparse Inference. *Journal of the American Statistical Association*, 115(531):1236–1250.

Zaitsev, A. Y. (1987). On the gaussian approximation of convolutions under multidimensional analogues of S.N. Bernstein's inequality conditions. *Probability Theory and Related Fields*, 74(4):535–566.

Zhang, X. and Chen, J. (2022). Covariate adaptive false discovery rate control with applications to omics-wide multiple testing. *Journal of the American Statistical Association*, 117(537):411–427.