Supplementary Material for "LAWS: A Locally Adaptive Weighting and Screening Approach To Spatial Multiple Testing"

We first prove the main results in Section A. Additional explanations for the variance-bias tradeoff in choosing τ and Equation (3.6) are provided in Sections B and C, respectively.

A Proofs of Main Results

$w(s) = \frac{\pi(s)}{1-\pi(s)}$

A.1 Proof of Theorem 1

$Q_w(t) = \frac{\sum_{s\in S}(1-\pi(s))w(s)t}{\sum_{s\in S}(1-\pi(s))w(s)t + \sum_{s\in S}\pi(s)G_1(w(s)t|s)}$

代入即证.

Proof: Note that

$$Q_w(t) = \frac{\sum_{s\in\mathbb{S}}(1-\pi(s))t}{\sum_{s\in\mathbb{S}}(1-\pi(s))t + \frac{\sum_{s\in\mathbb{S}}(1-\pi(s))}{\sum_{s\in\mathbb{S}}\pi(s)}\sum_{s\in\mathbb{S}}[\pi(s)G_1\{w(s)t|s\}]},$$

and

$$Q_1(t) = \frac{\sum_{s\in\mathbb{S}}(1-\pi(s))t}{\sum_{s\in\mathbb{S}}(1-\pi(s))t + \sum_{s\in\mathbb{S}}[\pi(s)G_1\{t|s\}]}$$

Because $t \to G_1(t|s)$ is concave and $x \to G_1(t/x|s)$ is convex for $\min_{s\in\mathbb{S}} w^{-1}(s) \le x \le \max_{s\in\mathbb{S}} w^{-1}(s)$, together with the condition that $\frac{\sum_{s\in\mathbb{S}}\pi(s)}{\sum_{s\in\mathbb{S}}(1-\pi(s))} \le 1$, we have

$\sum(1-\pi(s)) \geq \sum\pi(s).$

$\sum\pi(s)G_1(w(s)t) \geq$

$$\begin{aligned}
\frac{\sum_{s\in\mathbb{S}}(1-\pi(s))}{\sum_{s\in\mathbb{S}}\pi(s)}\sum_{s\in\mathbb{S}}\pi(s)G_1\{w(s)t|s\} &= \frac{\sum_{s\in\mathbb{S}}(1-\pi(s))}{\sum_{s\in\mathbb{S}}\pi(s)}\sum_{s\in\mathbb{S}}\pi(s)G_1\{t/w^{-1}(s)|s\} \\
&\ge \sum_{s\in\mathbb{S}}(1-\pi(s))G_1\left(\frac{t}{\sum_{s\in\mathbb{S}}\left(\frac{\pi(s)}{\sum_{s\in\mathbb{S}}\pi(s)}w^{-1}(s)\right)}|s\right) \\
&= \sum_{s\in\mathbb{S}}(1-\pi(s))G_1\left(\frac{\sum_{s\in\mathbb{S}}\pi(s)}{\sum_{s\in\mathbb{S}}(1-\pi(s))}t|s\right) \\
&\ge \sum_{s\in\mathbb{S}}\pi(s)G_1(t|s).
\end{aligned}$$

凸函数的权.

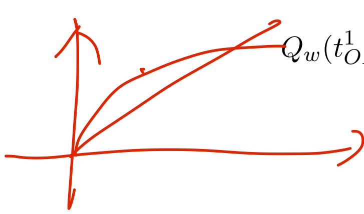
$x \to G_1(\frac{t}{x}|s)$ 函数凸性.

$t \to G_1(t|s)$ 凹性.

$\geq \sum(1-\pi(s)) \frac{\sum\pi(s)}{\sum(1-\pi(s))} G_1(\frac{\sum\pi(s)}{\sum(1-\pi(s))}t|$

Thus, take $t = t^1_{OR}$, we have

$\frac{\pi(s)}{1-\pi(s)} G_1(\sum$



$$Q_w(t^1_{OR}) \le Q_1(t^1_{OR}) \le \alpha.$$

由上确界定义

Hence, we have $t_{OR}^{w} \geq t_{OR}^{1}$, which yields that

由Ψ的单调性

$$\Psi_w(t_{OR}^{w}) \geq \Psi_w(t_{OR}^{1}) \geq \Psi_1(t_{OR}^{1}).$$

$\sum \pi(s) G_1\left(\frac{\pi(s)}{1-\pi(s)} t\right) \geq \sum \pi(s) G_1(t)$

This concludes the proof of Theorem 1. ∎

A.2 Proof of Proposition 1

Proof: Note that, by the definition of $v_h(s, s')$,

$$1 - \hat{\pi}^{\tau}(s) = \frac{\sum_{s' \in \mathcal{T}_\tau} v_h(s, s')}{(1-\tau)\sum_{s' \in \mathbb{S}} v_h(s, s')} = \frac{\sum_{s' \in \mathcal{T}_\tau} K_h(s - s')}{(1-\tau)\sum_{s' \in \mathbb{S}} K_h(s - s')}.$$

Thus, uniformly for all $s \in \mathbb{S}$, we have

$$\begin{aligned} \mathbb{E}\Big(\sum_{s' \in \mathcal{T}_\tau} K_h(s - s')\Big) &= \mathbb{E}\Big(\sum_{s' \in \mathbb{S}} [K_h(s - s') I\{p(s') > \tau\}]\Big) \\ &= \sum_{s' \in \mathbb{S}} [K_h(s - s') \mathbb{P}\{p(s') > \tau\}]. \end{aligned}$$

Under the conditions of the proposition, the eigenvalues of the Hessian matrix $\mathbb{P}^{(2)}(p(s) > \tau) \in \mathbb{R}^{d \times d}$ are bounded from above and below, namely, we have $-C \leq \lambda_{\min}(s) \leq \lambda_{\max}(s) \leq C$, for all $s \in \mathbb{S}$. Furthermore, $\pi^{\tau}(s)$ is bounded, and thus the entries of the first derivative vector are also bounded. Hence, by multivariate Taylor expansion, we have

$$\begin{aligned} &\int_{\mathcal{S}} K_h(s - s') \mathbb{P}(p(s') > \tau) ds' \\ &\quad = \mathbb{P}(p(s) > \tau) \int_{\mathcal{S}} K_h(s - s') ds' + v^{\intercal} \int_{\mathcal{S}} \frac{s' - s}{h} K\Big(\frac{s - s'}{h}\Big) ds' \\ &\qquad + \frac{h^2}{2} O(1) \int_{\mathbb{R}^d} x^{\intercal} x K(x) dx + o(h^2), \end{aligned}$$

where $v = (v_1, \ldots, v_d)^{\intercal}$ satisfies $v_j = O(1)$ for $j = 1, \ldots, d$. It follows that, as $\mathbb{S} \to \mathcal{S}$ in the summation $\sum_{s' \in \mathbb{S}}$,

$$\mathbb{E}(1 - \hat{\pi}^{\tau}(s)) \to \frac{\int_{\mathcal{S}} K_h(s - s') \mathbb{P}(p(s') > \tau) ds'}{(1-\tau) \int_{\mathcal{S}} K_h(s - s') ds'}.$$

Thus, we have, uniformly for all $s \in \mathbb{S}$, there exists some constant $c > 0$,

$$
\begin{aligned}
[\mathbb{E}\{\hat{\pi}^{\tau}(s)\} - \pi^{\tau}(s)]^2 \quad \le \quad & c\Big(v^{\intercal}\int_{\mathcal{S}} \frac{s'-s}{h} K(\frac{s-s'}{h})ds' \Big/ \int_{\mathcal{S}} K_h(s-s')ds'\Big)^2 \\
& + ch^4\Big(\int x^{\intercal} x K(x)dx \Big/ \int_{\mathcal{S}} K_h(s-s')ds'\Big)^2. \qquad \text{(A.2)}
\end{aligned}
$$

We next show the variance term. Note that, for some constant $c' > 1$,

$$
\begin{aligned}
\mathsf{Var}\Big(\sum_{s' \in \mathcal{T}_{\tau}} K_h(s-s')\Big) \quad &= \quad \mathsf{Var}\Big(\sum_{s' \in \mathbb{S}}[K_h(s-s')I\{p(s') > \tau\}]\Big) \\
&\le \quad c' \sum_{s' \in \mathbb{S}}[K_h^2(s-s')\mathbb{P}\{p(s') > \tau\}(1-\mathbb{P}\{p(s') > \tau\})].
\end{aligned}
$$

Thus, as $\mathbb{S} \to \mathcal{S}$ in the summation $\sum_{s' \in \mathbb{S}}$, we have,

$$
\begin{aligned}
\mathsf{Var}(1-\hat{\pi}^{\tau}(s)) \quad &\le \quad c''|\mathbb{S}|^{-1} \frac{\int_{\mathcal{S}} K_h^2(s-s')\mathbb{P}\{p(s') > \tau\}(1-\mathbb{P}\{p(s') > \tau\})ds'}{[(1-\tau)\int_{\mathcal{S}} K_h(s-s')ds']^2} \\
&\le \quad c''|\mathbb{S}h|^{-1}(1-\tau)^{-2}\int K^2(x)dx \Big/ \Big(\int_{\mathcal{S}} K_h(s-s')ds'\Big)^2 \\
&\le \quad c'''|\mathbb{S}h|^{-1}(1-\tau)^{-2} \Big/ \Big(\int_{\mathcal{S}} K_h(s-s')ds'\Big)^2, \qquad \text{(A.3)}
\end{aligned}
$$

for some constants $c'', c''' > 0$, where the last inequality comes from the fact that $K(\cdot)$ is positive and bounded. Combining (A.2) and (A.3) and letting $|\mathbb{S}| \to \infty$ and thus $h \to 0$, by the conditions that the domain $\mathcal{S}$ is finite, $\int_{\mathbb{R}^d} K(t)dt = 1$, $\int_{\mathbb{R}^d} tK(t)dt = 0$ and $\int_{\mathbb{R}^d} t^{\intercal} tK(t)dt < \infty$, Proposition 1 is proved. ∎

A.3 Proof of Theorem 2

Proof: We first introduce the following procedure in order to prove the theorem.

Procedure 1 *Calculate the weights as*

$$
\tilde{w}(s) = \left\{\sum_{s \in \mathbb{S}} \frac{\pi^{\tau}(s)}{1-\pi^{\tau}(s)}\right\}^{-1} \frac{m\pi^{\tau}(s)}{1-\pi^{\tau}(s)}, \ s \in \mathbb{S} \qquad \text{(A.4)}
$$

权重不一样.

For hypothesis $H_0(s) : \theta(s) = 0$, define adjusted p-values as $\tilde{p}^w(s) = \min\{p(s)/\tilde{w}(s), 1\}$. Apply the BH procedure at level α to all adjusted p-values.

The following lemma develops the theoretical property of Procedure 1 for each realization of $\{\theta(s), s \in \mathbb{S}\}$.

Lemma 1 *Under Conditions (A3) and (A5), and assume that, there exists a sufficiently small constant* $\xi > 0$*, such that* $\pi^\tau(s) \in [\xi, 1-\xi]$*, then we have*

$$\overline{\lim_{m\to\infty}} FDR_{\text{Procedure 1}} \le \alpha, \quad \textit{and} \quad \lim_{m\to\infty} \mathbb{P}(FDP_{\text{Procedure 1}} \le \alpha + \epsilon) = 1.$$

for any $\epsilon > 0$.

We remark here that, from the proof of Lemma 1, by replacing $\pi^\tau(s)$ by $\hat{\pi}(s)$ in Procedure 1, Lemma 1 still holds under the conditions of Proposition 1.

Now we prove Theorem 2. We first note that, by the proof of Lemma 3 in Xia et al. (2019), Algorithm 1 is equivalent to the following algorithm.

Algorithm 2 An equivalent LAWS Procedure

1: Estimate the FDP by

↙ 唯分布变化，

$$\widehat{\text{FDP}}_{\hat{w}}(t) = \frac{\sum_{s\in\mathbb{S}} \hat{\pi}(s)t}{\max\{\sum_{s\in\mathbb{S}} I(p^{\hat{w}}(s) \le t), 1\}} \tag{A.5}$$

2: Obtain the data-driven threshold $\hat{t}_w = \sup_t \left\{t : \widehat{\text{FDP}}_{\hat{w}}(t) \le \alpha\right\}$.

3: Reject $H_0(s)$ if $p^{\hat{w}}(s) \le \hat{t}_w$.

Thus, in the oracle case when $\{\pi(s), s \in \mathbb{S}\}$ are known, the corresponding oracle FDP and its conservative estimator can be equivalently written by

本应是 $(1-\pi(s))t$. 由于

$$\text{FDP}_w(t) = \frac{\sum_{s\in\mathcal{H}_0} I(p^w(s) \le t)}{\max\{R_w(t), 1\}}, \quad \widehat{\text{FDP}}_w(t) = \frac{\sum_{s\in\mathbb{S}} \pi^\tau(s)t}{\max\{R_w(t), 1\}}, \tag{A.6}$$

$w(s) = \frac{\pi(s)}{1-\pi(s)} \to \pi(s)t$

where $R_w(t) = \sum_{s\in\mathbb{S}} I(p^w(s) \le t)$ denotes the oracle total number of rejections with the threshold t, and $w(s) = \frac{\pi^\tau(s)}{1-\pi^\tau(s)}$. Define the oracle threshold $t_w = \sup_t \left\{t : \widehat{\text{FDP}}_w(t) \le \alpha\right\}$. The oracle decision rule $\boldsymbol{\delta}^w(t)$ is then equivalent to reject $H_0(s)$ if $p^w(s) \le t_w$, and we have that $\text{FDP}_w(t_w) = \text{FDP}[\boldsymbol{\delta}^w\{p^w_{(k^w)}\}]$.

Based on the definition of t_w, to prove Theorem 2, it is enough of show that, uniformly for all $t \ge t_w$,

趋于 $c\pi(s)t$ 即可.

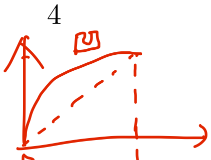
$$\left|\frac{\sum_{s\in\mathbb{S}}[I(p^w(s) \le t, \theta(s) = 0) - c\pi(s)t]}{\sum_{s\in\mathbb{S}} \pi(s)t}\right| \to 0$$

$\pi^\tau(s) - \pi(s) \le 0$. 低估.

且 $\left|\frac{\pi^\tau(s) - \pi(s)}{\pi(s)}\right| = c$.

$C = \frac{1-G_1(\tau(s))}{1-\tau} \le 1$.

$G_1(\tau(s)) \ge \tau$



in probability, for some constant $0 < c \leq 1$.

Note that the weights in the oracle procedure and the weights in Procedure 1 are proportional. Thus the adjusted p-values $\{p^w(s), i = 1, \ldots, m\}$ and $\{\tilde{p}^w(s), i = 1, \ldots, m\}$ have exactly the same order, with $\tilde{p}^w(s) = p^w(s) m^{-1} \sum_{s \in \mathbb{S}} \frac{\pi^\tau(s)}{1-\pi^\tau(s)}$. Also note that Procedure 1 is equivalent to find

成比例关系.

Xia (2019) 的引理4, 做伸缩(加权)后p值在H0下的pdf做相应的变化.

BH方法:

$$\tilde{t}_w = \sup_t \{ t : \frac{mt}{\max\{\sum_{s \in \mathbb{S}} I(\tilde{p}^w(s) \leq t), 1\}} \leq \alpha \},$$

and reject the hypothesis $H_0(s)$ if $\tilde{p}^w(s) < \tilde{t}_w$. By letting $\tilde{t} = m(\sum_{s \in \mathbb{S}} \frac{\pi^\tau(s)}{1-\pi^\tau(s)})^{-1} t$, we have that

$$\frac{mt}{\max\{\sum_{s \in \mathbb{S}} I(\tilde{p}^w(s) \leq t), 1\}} = \frac{\sum_{s \in \mathbb{S}} \frac{\pi^\tau(s)}{1-\pi^\tau(s)} \tilde{t}}{\max\{\sum_{s \in \mathbb{S}} I(p^w(s) \leq \tilde{t}), 1\}},$$

and thus the threshold for $\tilde{p}^w(s)$ in Procedure 1 is

$$\tilde{t}_w = \sup_{\tilde{t}} \{ \tilde{t} : \frac{\sum_{s \in \mathbb{S}} \frac{\pi^\tau(s)}{1-\pi^\tau(s)} \tilde{t}}{\max\{\sum_{s \in \mathbb{S}} I(p^w(s) \leq \tilde{t}), 1\}} \leq \alpha \} \cdot m^{-1} \sum_{s \in \mathbb{S}} \frac{\pi^\tau(s)}{1-\pi^\tau(s)}.$$

Hence the threshold for $p^w(s)$ in Procedure 1 is

$tw = \sup\{t : \frac{\sum \pi(s) t}{\quad} \leq \alpha\}$

$$t_w^1 = \sup_t \{ t : \frac{\sum_{s \in \mathbb{S}} \frac{\pi^\tau(s)}{1-\pi^\tau(s)} t}{\max\{\sum_{s \in \mathbb{S}} I(p^w(s) \leq t), 1\}} \leq \alpha \}.$$

分子更大, $\therefore t_w^1 \leq t_w$.

Comparing the definition of t_w and t_w^1, it is easily to see that $t_w^1 \leq t_w$. By the proofs of Lemma 1, we have that, the threshold for the z-values corresponding to threshold of p-values: t_w^1, is no larger than t_m as defined in the proof of Lemma 1. We further learn from the proofs of Lemma 1 that, for every realization of $\{\theta(s), s \in \mathbb{S}\}$, uniformly for all $t \geq t_w^1$, as $m \to \infty$,

$$\left| \frac{\sum_{\theta(s)=0} [I(p^w(s) \leq t) - \mathbb{P}(p^w(s) \leq t | \theta(s) = 0)]}{\sum_{\theta(s)=0} \mathbb{P}(p^w(s) \leq t | \theta(s) = 0)} \right| \to 0$$

in probability. Note that, by Condition (A4),

$$\left| \frac{\sum_{\theta(s)=0} \mathbb{P}(p^w(s) \leq t | \theta(s) = 0) - \sum_{s \in \mathbb{S}} \mathbb{P}(p^w(s) \leq t, \theta(s) = 0)}{\sum_{s \in \mathbb{S}} \mathbb{P}(p^w(s) \leq t, \theta(s) = 0)} \right| \to 0$$

$P(p \leq t | \theta(s) = 0) = \frac{P(p \leq t, \theta(s) = 0)}{P(\theta(s) = 0)}$.

in probability, where $\{\theta(s), s \in \mathbb{S}\}$ in the above equation represent random variables as modeled in (2.3). Thus we have

大数定律.

$$\left| \frac{\sum_{s\in\mathbb{S}}[I(p^w(s) \le t, \theta(s) = 0) - \mathbb{P}(p^w(s) \le t, \theta(s) = 0)]}{\sum_{s\in\mathbb{S}} \mathbb{P}(p^w(s) \le t, \theta(s) = 0)} \right| \to 0$$

in probability. Note that

$P(\frac{p}{w} \le t | \theta = 0) = P(p \le wt | \theta = 0)$

$$\begin{aligned} \sum_{s\in\mathbb{S}} \mathbb{P}(p^w(s) \le t, \theta(s) = 0) &= \sum_{s\in\mathbb{S}} \mathbb{P}(p^w(s) \le t | \theta(s) = 0)\mathbb{P}(\theta(s) = 0) \\ &= \sum_{s\in\mathbb{S}} \left[\frac{\pi^\tau(s)}{1 - \pi^\tau(s)} t(1 - \pi(s)) \right] \le \sum_{s\in\mathbb{S}} \pi^\tau(s)t. \end{aligned}$$

调权后会变.

$\because \pi^\tau(s) \le \pi(s)$.

Thus we have that, uniformly for all $t \ge t_w^1$, there exists a constant $0 < c \le 1$, as $m \to \infty$,

$$\left| \frac{\sum_{s\in\mathbb{S}}[I(p^w(s) \le t, \theta(s) = 0) - c\pi(s)t]}{\sum_{s\in\mathbb{S}} \pi(s)t} \right| \to 0$$

in probability. This concludes the proof of Theorem 2. ∎

A.4 Proof of Theorem 3

Proof: As shown in the proof of Theorem 2, the data-driven decision rule $\boldsymbol{\delta}^{\hat{w}}(t)$ is equivalent to reject $H_0(s)$ if $p^{\hat{w}}(s) \le t_{\hat{w}}$, and we have that $\text{FDP}_{\hat{w}}(t_{\hat{w}}) = \text{FDP}[\boldsymbol{\delta}^{\hat{w}}\{p^{\hat{w}}_{(k^{\hat{w}})}\}]$.

Similarly as the proof of Theorem 2, based on the proofs of Lemma 1 and Condition (A4), we have that

$$\left| \frac{\sum_{s\in\mathbb{S}}(I(p^{\hat{w}}(s) \le t, \theta(s) = 0) - \mathbb{P}(p^{\hat{w}}(s) \le t, \theta(s) = 0))}{\sum_{s\in\mathbb{S}} \mathbb{P}(p^{\hat{w}}(s) \le t, \theta(s) = 0)} \right| \to 0$$

in probability, uniformly for all $t \ge t_w^1$, where t_w^1 is defined as in the proof of Theorem 2, by replacing $\pi^\tau(s)$ by $\hat{\pi}(s)$ in Procedure 1. Recall that $\pi^\tau(s) = 1 - \frac{\mathbb{P}(p(s) > \tau)}{1 - \tau}$. and that the bias of $1 - \pi^\tau(s)$ is always non-negative. By Proposition 1 and the fact that $\pi^\tau(s) \in [\xi, 1 - \xi]$ for some sufficiently small constant $\xi > 0$,

$$\sum_{s\in\mathbb{S}} \mathbb{P}(p^{\hat{w}}(s) \le t, \theta(s) = 0) = \sum_{s\in\mathbb{S}} \mathbb{P}(p^{\hat{w}}(s) \le t | \theta(s) = 0)\mathbb{P}(\theta(s) = 0)$$

$$
\begin{aligned}
&= (1+o(1))\sum_{s\in\mathbb{S}}\left\{\frac{\pi^\tau(s)}{1-\pi^\tau(s)}t(1-\pi(s))\right\}\\
&\le (1+o(1))\sum_{s\in\mathbb{S}}\left\{\frac{\pi^\tau(s)}{1-\pi^\tau(s)}t(1-\pi^\tau(s))\right\}\\
&= (1+o(1))\sum_{s\in\mathbb{S}}\pi^\tau(s)t,
\end{aligned}
$$

where $o(1)$ in the above equations are in the limit of $\mathbb{S}\to\mathcal{S}$. Thus, based on Proposition 1, we have that, uniformly for all $t\ge t_w^1$, there exists a constant $0<c\le 1$ such that

$$
\left|\frac{\sum_{s\in\mathbb{S}}(I(p^{\hat{w}}(s)\le t,\theta(s)=0)-c\hat{\pi}(s)t)}{\sum_{s\in\mathbb{S}}c\hat{\pi}(s)t}\right|\to 0
$$

in probability. Namely, the data-driven procedure provides a more conservative FDR and FDP control asymptotically. This concludes the proof of Theorem 3.

A.5 Proof of Lemma 1

We now prove Lemma 1 below. We let $p^w(s)=\tilde{p}^w(s)$ in the proof of this lemma for notation simplicity. We arrange $\{s\in\mathbb{S}\}$ in any order $\{s_1,\dots,s_m\}$ and note that $\sum_{\theta(s_i)=0}\mathbb{P}(z_i^w\ge t)$ is equivalent to $\sum_{\theta(s_i)=0}\mathbb{P}(z_i^w\ge t|\theta(s_i)=0)$ and $\sum_{i=1,\dots,m}\mathbb{P}(z_i^w\ge t,\theta(s_i)=0)$ for a given realization of the hypotheses, where $z_i^w=\Phi^{-1}(1-p^w(s_i)/2)$, for $i=1,\dots,m$. We first show that by applying BH to the weighted p-values controls FDR exactly under the independence of z_i^w. We then prove that, in the dependent case, it performance asymptotically the same as case when z_i^w are independent.

Let $t_m=(2\log m-2\log\log m)^{1/2}$. By Condition (A5), we have

$$
\sum_{\theta(s_i)=1} I\{|z_i|\ge(c\log m)^{1/2+\rho/4}\}\ge\{1/(\pi^{1/2}\alpha)+\delta\}(\log m)^{1/2},
$$

with probability going to one, for some constant $c>0$. Recall that we have $w(s)=\left\{\sum_{s\in\mathbb{S}}\frac{\pi^\tau(s)}{1-\pi^\tau(s)}\right\}^{-1}\frac{m\pi^\tau(s)}{1-\pi^\tau(s)}$ and $\pi^\tau(s)\in[\xi,1-\xi]$. Thus, for those indices $i\in\mathcal{H}_1$ (equivalently $\theta(s_i)=1$) such that $|z_i|\ge(c\log m)^{1/2+\rho/4}$, we have

$$
p^w(s_i)=p(s_i)/w(s_i)\le(1-\Phi((c\log m)^{1/2+\rho/4}))/w(s_i)=o(m^{-M}),
$$

for any constant $M > 0$. Thus we have

$$\sum_{1\le i\le m} I\{z_i^w \ge (2\log m)^{1/2}\} \ge \{1/(\pi^{1/2}\alpha)+\delta\}(\log m)^{1/2},$$

with probability going to one. Hence, with probability tending to one, we have

$$\frac{2m}{\sum_{1\le i\le m} I\{z_i^w \ge (2\log m)^{1/2}\}} \le 2m\{1/(\pi^{1/2}\alpha)+\delta\}^{-1}(\log m)^{-1/2}.$$

Because $1-\Phi(t_m) \sim 1/\{(2\pi)^{1/2}t_m\}\exp(-t_m^2/2)$, it suffices to show that, uniformly in $0 \le t \le t_m$, there exists a constant $0 < c \le 1$, such that

$$\left|\frac{\sum_{\theta(s_i)=0} I(z_i^w \ge t) - cm_0G(t)}{cm_0G(t)}\right| \to 0, \tag{A.7}$$

in probability, where $G(t) = 2(1-\Phi(t))$.

We first consider the case when z_i^w are independent with each other. For $i = 1,\dots,m$, the ideal choice of the threshold t^o for z_i^w in order to control the FDR is that

$$t^o = \inf\{t \ge 0, \frac{\sum_{\theta(s_i)=0} I(z_i^w \ge t)}{\sum_{1\le i\le m} I(z_i^w \ge t)} \le \alpha\}.$$

It is easy to show that, under independence of z_i^w,

$$\left|\frac{\sum_{\theta(s_i)=0} I(z_i^w \ge t) - \sum_{\theta(s_i)=0}\mathbb{P}(z_i^w \ge t)}{\sum_{\theta(s_i)=0}\mathbb{P}(z_i^w \ge t)}\right| \to 0, \text{ as } m\to\infty.$$

Thus a good estimate of t^o can be written by

$$\hat{t}^o = \inf\{t \ge 0, \frac{\sum_{\theta(s_i)=0} \mathbb{P}(z_i^w \ge t)}{\sum_{1\le i\le m} I(z_i^w \ge t)} \le \alpha\}. \tag{A.8}$$

Since the spatial locations $\{s \in S\}$ does not change the null distribution of p-values, according to Theorem 1 in Genovese et al. (2006), and by the fact that the original p-values of the null hypotheses are uniformly distributed, the procedure by applying BH procedure on the weighted p-values controls the FDR at level $\alpha m_0/m$. That is, if

$$k = \max\{i : p^w_{(i)} \le i\alpha/m\},$$

where $p^w_{(1)} \le \cdots \le p^w_{(m)}$ are the ordered weighted p-values, and we reject all k hypotheses associated with $p^w_{(1)}, \ldots, p^w_{(k)}$, then we have

$$\mathbb{E}\left(\frac{\sum_{\theta(s)=0} I(p^w(s) \le p^w_{(k)})}{\max\{\sum_{1\le i\le m} I(p^w(s) \le p^w_{(k)}), 1\}}\right) \le \alpha m_0/m. \tag{A.9}$$

By the definition of z^w_i, it is equivalent to reject all k hypotheses with

$$\frac{2m(1-\Phi(z^w_{(i)}))}{i} \le \alpha.$$

That is to find

$$\hat{t} = \inf\{t \ge 0, \frac{2m(1-\Phi(t))}{\sum_{1\le i\le m} I(z^w_i \ge t)} \le \alpha\}, \tag{A.10}$$

and reject all hypotheses with $z^w_i \ge \hat{t}$. This yields that

$$\mathbb{E}\left(\frac{\sum_{\theta(s_i)=0} I(z^w_i \ge \hat{t})}{\sum_{1\le i\le m} I(z^w_i \ge \hat{t})}\right) \le \alpha m_0/m.$$

Hence, this procedure is more conservative than rejecting all hypotheses with $z^w_i \ge \hat{t}^o$ as defined in (A.8). Thus, there exists a constant $0 < c \le 1$, such that, uniformly in $0 \le t \le t_m$,

$$\left|\frac{\sum_{\theta(s_i)=0} \mathbb{P}(z^w_i \ge t) - cm_0 G(t)}{cm_0 G(t)}\right| \to 0. \tag{A.11}$$

Under Assumption (A1), that is, when the weighted z-values are weakly dependent with each other, by the proof of Theorem 1 in Xia et al. (2019) and the fact that

$$\Phi^{-1}\{1 - [1 - \Phi\{(c_1 \log m + c_2 \log\log m)^{1/2}\}]/w(s_i)\} = c_1 \log m + c_2 \log\log m + c_3,$$

for some constant c_1, c_2, c_3, we have,

$$\left| \frac{\sum_{\theta(s_i)=0} (I(z_i^w \geq t) - \mathbb{P}(z_i^w \geq t))}{\sum_{\theta(s_i)=0} \mathbb{P}(z_i^w \geq t)} \right| \to 0, \tag{A.12}$$

in probability, uniformly in $0 \leq t \leq t_m$. By (A.11) and (A.12), (A.7) is proved and thus Procedure 1 controls FDR and FDP asymptotically under dependency. This concludes the proof of Lemma 1. ∎

B Explanation of the bias-variance tradeoff in choosing τ

There is a bias-variance tradeoff in the choice of τ in the proposed estimator $\hat{\pi}^\tau(s)$. It is easy to see that a large τ will simultaneously reduce the bias (desirable) and decrease the sample size (undesirable).

To reduce the bias in the proposed estimator $\hat{\pi}^\tau(s)$, one needs to choose a relatively large τ to ensure the "purity" of $\mathcal{T}(\tau) = \{s \in \mathbb{S} : p(s) > \tau\}$, i.e. we wish to have a screening set where majority of the cases come from the null. Although the common choice of $\tau = 0.5$ suffices in many situations, we propose a new scheme to carefully calibrate τ that is adaptive to the observed data. Let τ be the threshold determined by the BH algorithm with $\alpha = 0.9$. Then roughly speaking, in the subset $\tilde{\mathcal{T}}(\tau) = \{s \in \mathbb{S} : p(s) \leq \tau\}$, 90% of the cases come from the null (e.g. the expected proportion of false positives made by BH). It is anticipated that in the remaining set $\mathcal{T}(\tau) = \{s \in \mathbb{S} : p(s) > \tau\}$, which is used to construct our estimator, an overwhelming proportion(more than 90%) of the cases should come from the null. In the simulation (say the 2D rectangle case), τ roughly ranges from 0.38 to 0.47 when we run BH algorithm at $\alpha = 0.9$. This data-driven scheme ensures the purity of the screening set while maintaining a larger sample size compared to the standard choice of $\tau = 0.5$.

C Explanation of Equation (3.6) in Section 3.3

In this section we justify the approximation in Equation (3.6). Denote $\boldsymbol{\delta}^v(t) = \{\delta^v(s,t) : s \in \mathcal{S}\}$ a class of testing rules, where $\delta^v(s,t) = \mathbb{I}\{p^v(s) \leq t\}$, $v(s)$ is the pre-specified weight and $p^v(s) = \min\left\{\frac{p(s)}{v(s)}, 1\right\}$. Denote by $V(t) = \sum_{s \in \mathcal{H}_0} \delta^v(s,t)$ and $R(t) = \sum_{s \in \mathbb{S}} \delta^v(s,t)$. We show that $Q^v(t)$ provides a good approximation to the actual FDR level under the following condition:

$$\mathsf{Var}\left[R(t)/\mathbb{E}\{R(t)\}\right] = o(1). \tag{C.13}$$

We remark here that for a fixed threshold $t > 0$ and $v(s) \in [\xi, 1-\xi]$ for a sufficiently small constant $\xi > 0$, Condition (C.13) is satisfied if $\mathsf{Var}\left\{\sum_{s\in\mathbb{S}} \mathbb{I}\{p(s) \le t\}/m\right\} = o(1)$. This is a weaker condition compared to (A2) in Section 4.1. Hence the condition can be fulfilled by both BH and LAWS under the general class of dependence structures being considered in this article.

Proposition 2 *Assume that Condition (C.13) holds, then we have*

$$FDR\{\boldsymbol{\delta}^v(t)\} = Q^v(t) + o(1) = \frac{\mathbb{E}\left\{\sum_{s\in\mathcal{H}_0} \delta^v(s,t)\right\}}{\mathbb{E}\left\{\sum_{s\in\mathcal{H}_0} \delta^v(s,t)\right\} + \mathbb{E}\left\{\sum_{s\in\mathcal{H}_1} \delta^v(s,t)\right\}} + o(1).$$

Proof of Proposition 2: Note that $R(t) = 0$ implies $V(t) = 0$, hence we have

$$\mathrm{FDR}\{\boldsymbol{\delta}^v(t)\} = \mathbb{E}\left[\frac{V(t)}{R(t)}\mathbb{I}\{R(t) > 0\}\right], \text{ and } Q^v(t) = \frac{\mathbb{E}\left[V(t)\mathbb{I}\{R(t) > 0\}\right]}{\mathbb{E}\left\{R(t)\right\}}.$$

It follows that

$$\begin{aligned} |\mathrm{FDR}\{\boldsymbol{\delta}^v(t)\} - Q^v(t)| &\le \mathbb{E}\left[\left|\frac{V(t)}{R(t)} - \frac{V(t)}{\mathbb{E}\left\{R(t)\right\}}\right| \mathbb{I}\{R(t) > 0\}\right] \\ &= \mathbb{E}\left[\left|\frac{V(t)}{R(t)}\frac{R(t) - \mathbb{E}\left\{R(t)\right\}}{\mathbb{E}\left\{R(t)\right\}}\right| \mathbb{I}\{R(t) > 0\}\right]. \end{aligned}$$

Note that $V(t)$ is always no larger than $R(t)$, we have

$$|\mathrm{FDR}\{\boldsymbol{\delta}^v(t)\} - Q^v(t)| \le \mathsf{Var}^{1/2}\left[R(t)/\mathbb{E}\{R(t)\}\right] = o(1),$$

which proves the proposition.