Multiple Testing of Local Extrema for Detection of Structural Breaks in Piecewise Linear Models

Zhibing He, Dan Cheng and Yunpeng Zhao

Southeast University, Arizona State University and Colorado State University

Supplementary Material

S1. Characteristics of the Type II change point v_j with large $|q_j|$

A Type II change point v_j with large $|q_j|$ behaves similarly to a Type I change point; its derivatives exhibit similar characteristics (see Figure 1 in the Supplementary Materials). Specifically, a Type II change point with q_j large enough such that $q_j > c/\gamma$ cannot generate a local extremum in the first derivative $\mu'_\gamma(t)$; and the corresponding second derivative $\mu''_\gamma(t)$ has only one local extremum around the change point v_j. Since $a_j \neq 0$ and $q_j > c/\gamma$, one has $v_j + \gamma^2 q_j > v_j + c\gamma$, implying no local extremum in the first derivative $\mu'_\gamma(t)$

Corresponding author: Dan Cheng. Email: chengdan@asu.edu

(see Lemma 2). Additionally,

$$v_j + \frac{1}{2}\left(\gamma^2 q_j + \gamma\sqrt{\gamma^2 q_j^2 + 4}\right) > v_j + c\gamma, \quad v_j + \frac{1}{2}\left(\gamma^2 q_j - \gamma\sqrt{\gamma^2 q_j^2 + 4}\right) \approx v_j. \tag{S1.1}$$

Hence, there is only one local extremum around v_j in the second derivative $\mu''_\gamma(t)$ (see Lemma 2). Given this behavior, it is reasonable to interpret a Type II change point with large $|q_j|$ as a special case that essentially behaves like Type I change points ($q_j = \infty$ as $a_j = 0$).

Intuitively, for a Type II change point v_j, if the $|q_j|$ is large enough such that $q_j > c/\gamma$, it indicates that the slope change dominates the jump size; thereafter, this Type II change point behaves similarly to a Type I change point. In this paper, we do not consider the case of large $|q_j|$. However, as discussed above, this case can be essentially treated as the case of Type I change points.

Figure 1 shows a Type II change point v_j with large $|q_j|$ behaves similarly to a Type I change point, its derivatives exhibit similar characteristics. Specifically, a Type II change point with q_j large enough such that $q_j > c/\gamma$ cannot generate a local extremum in the first derivative $\mu'_\gamma(t)$; and the second derivative $\mu''_\gamma(t)$ has only one local extremum around the change point v_j. Hence, this special Type II change point can be essentially treated as a Type I change point.

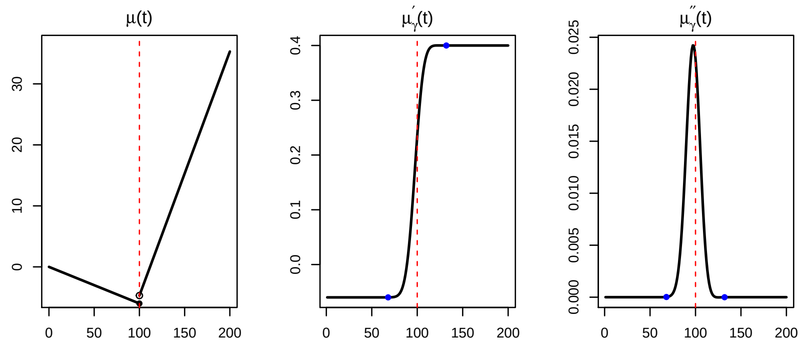

Figure 1: Characteristics of the Type II change point v_j with large $|q_j|$. In the left plot of signal $\mu(t)$, the slope change dominates the jump size, giving a large $|q_j|$. The middle plot of the first derivative μ'_γ shows no local extremum, and the right plot of the second derivative $\mu''_\gamma(t)$ shows a local maximum around the change point v_j.

S2. Proofs in Section 2

Proof of Lemma 1. Recall that $w_\gamma(t)$ defined in equation (3) is a truncated Gaussian kernel with support $[-c\gamma, c\gamma]$ and bandwidth γ. For $t \in (v_j - c\gamma, v_j + c\gamma)$, we have

$$
\begin{aligned}
\mu_\gamma(t) &= w_\gamma(t) * \mu(t) = \int_{t-c\gamma}^{t+c\gamma} w_\gamma(t-s)\mu(s)ds \\
&= \int_{t-c\gamma}^{v_j} \frac{1}{\gamma}\phi(\frac{t-s}{\gamma})(c_j + k_j s)ds + \int_{v_j}^{t+c\gamma} \frac{1}{\gamma}\phi(\frac{t-s}{\gamma})(c_{j+1} + k_{j+1}s)ds \\
&= \int_{-c}^{\frac{v_j-t}{\gamma}} \phi(x)(c_j + k_j t + k_j\gamma x)dx + \int_{\frac{v_j-t}{\gamma}}^{c} \phi(x)(c_{j+1} + k_{j+1}t + k_{j+1}\gamma x)dx \\
&= [c_j + k_j t - (c_{j+1} + k_{j+1}t)]\Phi(\frac{v_j - t}{\gamma}) + [c_j + k_j t + (c_{j+1} + k_{j+1}t)]\Phi(c) - (c_j + k_j t) \\
&\quad + (k_j - k_{j+1})\gamma\phi(c) + (k_{j+1} - k_j)\gamma\phi(\frac{v_j - t}{\gamma}).
\end{aligned} \tag{S2.2}
$$

Similarly, we obtain that, for $t \in (v_j + c\gamma, v_{j+1} - c\gamma)$,

$$
\begin{aligned}
\mu_\gamma(t) = w_\gamma(t) * \mu(t) &= \int_{t-c\gamma}^{t+c\gamma} \frac{1}{\gamma}\phi(\frac{t-s}{\gamma})(c_{j+1} + k_{j+1}s)ds \\
&= (c_{j+1} + k_{j+1}t)[2\Phi(c) - 1];
\end{aligned} \tag{S2.3}
$$

and for $t \in (v_{j-1} + c\gamma, v_j - c\gamma)$,

$$
\begin{aligned}
\mu_\gamma(t) = w_\gamma(t) * \mu(t) &= \int_{t-c\gamma}^{t+c\gamma} \frac{1}{\gamma}\phi(\frac{t-s}{\gamma})(c_j + k_j s)ds \\
&= (c_j + k_j t)[2\Phi(c) - 1].
\end{aligned} \tag{S2.4}
$$

Taking the first and second derivatives of $\mu_\gamma(t)$ in (S2.2), (S2.3) and (S2.4) respectively, we obtain

$$
\mu'_\gamma(t) = \begin{cases}
\frac{a_j}{\gamma}\phi(\frac{v_j-t}{\gamma}) + (k_j - k_{j+1})\Phi(\frac{v_j-t}{\gamma}) + (k_j + k_{j+1})\Phi(c) - k_j & t \in (v_j - c\gamma, v_j + c\gamma), \\
k_j[2\Phi(c) - 1] & t \in (v_{j-1} + c\gamma, v_j - c\gamma), \\
k_{j+1}[2\Phi(c) - 1] & t \in (v_j + c\gamma, v_{j+1} - c\gamma),
\end{cases}
$$

and

$$\mu''_\gamma(t) = \begin{cases} \frac{a_j(v_j-t)+(k_{j+1}-k_j)\gamma^2}{\gamma^3}\phi(\frac{v_j-t}{\gamma}) & t \in (v_j - c\gamma, v_j + c\gamma), \\ 0 & \text{otherwise}. \end{cases}$$

□

Proof of Lemma 2. Note that, if $a_j = 0$, then $\mu'_\gamma(t)$ is monotone in $(v_j - c\gamma, v_j + c\gamma)$ and hence there is no local extremum. If $a_j \neq 0$, by letting $\mu''_\gamma(t) = 0$, we see that the local extremum of $\mu'_\gamma(t)$ is achieved at $t = v_j + \gamma^2 q_j$. Similarly, solving the equation

$$\mu^{(3)}_\gamma(t) = [\frac{a_j}{\gamma^3} + \frac{a_j(v_j-t)+(k_{j+1}-k_j)\gamma^2}{\gamma^3}\frac{v_j-t}{\gamma^2}]\phi(\frac{v_j-t}{\gamma}) = 0,$$

we obtain that the local extremum of $\mu''_\gamma(t)$ is achieved at

$$t = \begin{cases} v_j + \frac{\gamma^2 q_j \pm \gamma\sqrt{4+q_j^2\gamma^2}}{2} & a_j \neq 0, \\ v_j & a_j = 0. \end{cases}$$

□

S3. Peak Height Distributions for $z'_\gamma(t)$ and $z''_\gamma(t)$

Proof of Lemma 3 and Proposition 2. Due to the stationarity of $z_\gamma(t)$, without loss of generality, we only consider the case when $t = 0$. The variances of $z^{(d)}_\gamma(0)$, $d = 1, \ldots, 4$, are calculated as follows.

$$\mathrm{Var}(z_\gamma'(0)) = \int_{\mathbb{R}} \frac{s^2}{\xi^6}\phi^2\left(\frac{s}{\xi}\right) ds = \frac{1}{\xi^6}\frac{\xi}{2\sqrt{\pi}}\int_{\mathbb{R}} s^2 \frac{\sqrt{2}}{\xi}\phi\left(\frac{s}{\xi/\sqrt{2}}\right) ds$$
$$= \frac{1}{\xi^6}\frac{\xi}{2\sqrt{\pi}}\frac{\xi^2}{2} = \frac{1}{4\sqrt{\pi}\xi^3};$$

$$\mathrm{Var}(z_\gamma''(0)) = \int_{\mathbb{R}} \frac{1}{\xi^6}\phi^2\left(\frac{s}{\xi}\right) ds - 2\int_{\mathbb{R}} \frac{s^2}{\xi^8}\phi^2\left(\frac{s}{\xi}\right) ds + \int_{\mathbb{R}} \frac{s^4}{\xi^{10}}\phi^2\left(\frac{s}{\xi}\right) ds$$
$$= \frac{1}{\xi^4}\frac{1}{2\sqrt{\pi}\xi} - \frac{2}{\xi^2}\frac{1}{4\sqrt{\pi}\xi^3} + \frac{1}{\xi^{10}}\frac{\xi}{2\sqrt{\pi}}\int_{\mathbb{R}} s^4 \frac{\xi}{\sqrt{2}}\phi\left(\frac{s}{\xi/\sqrt{2}}\right) ds$$
$$= \frac{1}{\xi^{10}}\frac{\xi}{2\sqrt{\pi}}\frac{3\xi^4}{4} = \frac{3}{8\sqrt{\pi}\xi^5};$$

$$\mathrm{Var}(z_\gamma^{(3)}(0)) = 9\int_{\mathbb{R}} \frac{s^2}{\xi^{10}}\phi^2\left(\frac{s}{\xi}\right) ds - 6\int_{\mathbb{R}} \frac{s^4}{\xi^{12}}\phi^2\left(\frac{s}{\xi}\right) ds + \int_{\mathbb{R}} \frac{s^6}{\xi^{14}}\phi^2\left(\frac{s}{\xi}\right) ds$$
$$= \frac{9}{\xi^4}\frac{1}{4\sqrt{\pi}\xi^3} - \frac{6}{\xi^2}\frac{3}{8\sqrt{\pi}\xi^5} + \frac{1}{\xi^{14}}\frac{\xi}{2\sqrt{\pi}}\int_{\mathbb{R}} s^6 \frac{\sqrt{2}}{\xi}\phi\left(\frac{s}{\xi/\sqrt{2}}\right) ds$$
$$= \frac{1}{\xi^{14}}\frac{\xi}{2\sqrt{\pi}}\frac{15\xi^6}{8} = \frac{15}{16\sqrt{\pi}\xi^7};$$

$$\mathrm{Var}(z_\gamma^{(4)}(0)) = 9\int_{\mathbb{R}} \frac{1}{\xi^{10}}\phi^2\left(\frac{s}{\xi}\right) ds - 36\int_{\mathbb{R}} \frac{s^2}{\xi^{12}}\phi^2\left(\frac{s}{\xi}\right) ds + 42\int_{\mathbb{R}} \frac{s^4}{\xi^{14}}\phi^2\left(\frac{s}{\xi}\right) ds$$
$$- 12\int_{\mathbb{R}} \frac{s^6}{\xi^{16}}\phi^2\left(\frac{s}{\xi}\right) ds + \int_{\mathbb{R}} \frac{s^8}{\xi^{18}}\phi^2\left(\frac{s}{\xi}\right) ds$$
$$= \frac{1}{\xi^9}\left(9 \times \frac{1}{2\sqrt{\pi}} - 36 \times \frac{1}{4\sqrt{\pi}} + 42 \times \frac{3}{8\sqrt{\pi}} - 12 \times \frac{15}{16\sqrt{\pi}} + \frac{1}{2\sqrt{\pi}} \times \frac{105}{16}\right)$$
$$= \frac{105}{32\sqrt{\pi}\xi^9}.$$

Hence, the parameters of the peak height distributions for $z'_\gamma(t)$ and $z''_\gamma(t)$ are given respectively by

$$\sigma_1^2 = \mathrm{Var}(z'_\gamma(0)) = \frac{1}{4\sqrt{\pi}\xi^3}, \quad \eta_1 = \frac{\mathrm{Var}(z''_\gamma(0))}{\sqrt{\mathrm{Var}(z'_\gamma(0))\mathrm{Var}(z^{(3)}_\gamma(0))}} = \frac{\sqrt{3}}{\sqrt{5}};$$
$$\sigma_2^2 = \mathrm{Var}(z''_\gamma(0)) = \frac{3}{8\sqrt{\pi}\xi^5}; \quad \eta_2 = \frac{\mathrm{Var}(z^{(3)}_\gamma(0))}{\sqrt{\mathrm{Var}(z''_\gamma(0))\mathrm{Var}(z^{(4)}_\gamma(0))}} = \frac{\sqrt{5}}{\sqrt{7}}.$$

This leads to the desired peak height distributions in Proposition 2. □

S4. FDR Control and Power Consistency for Type I Change Points

Proof of Theorem 1. Note that Type I change points are detected using the second derivative $y''_\gamma(t) = \mu''_\gamma(t) + z''_\gamma(t)$, where a change point v_j becomes a local extremum in $\mu''_\gamma(t)$ precisely at v_j by Proposition 1 and $z''_\gamma(t)$ is a smooth Gaussian process. Moreover, $\mu'(t)$ is piecewise constant with size of jump $|k_{j+1} - k_j|$ at v_j. Thus the detection of Type I change points using the second derivative is essentially equivalent to the detection of change points in piecewise constant signals (Cheng et al. 2020). Under condition (C1), by substituting the original signal μ with μ' and replacing the differential smoothed noise z'_γ with z''_γ, we see that Theorem 1 follows from Theorems 3.1 and 3.3 in Cheng et al. (2020). □

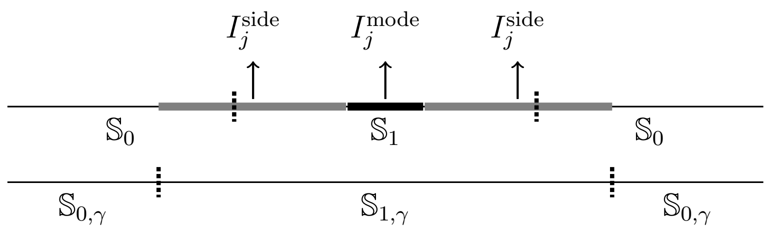

Figure 2: Illustration of the *signal region* $\mathbb{S}_1$ (S_j in this example), *null region* $\mathbb{S}_0$, *smoothed signal region* $\mathbb{S}_{1,\gamma}$ ($S_{j,\gamma}$ in this example) and *smoothed null region* $S_{0,\gamma}$. I_j^{mode} and I_j^{side} are a partition of $S_{j,\gamma}$ such that $I_j^{\text{mode}} \subset S_j \subset S_{j,\gamma}$.

S5. Supporting Results for FDR Control and Power Consistency for Type II Change Points

To prove Theorem 2 and Theorem 3, we require the following lemmas in this section. It is important to note that the lemmas provided below are based on Type II change points detection, where $a_j \neq 0$.

Recall that *signal region* is defined as $\mathbb{S}_1 = \cup_{j=1}^J S_j = \cup_{j=1}^J (v_j - b, v_j + b)$, and the smoothed signal is generated by convolving the signal and a kernel function with support $[-c\gamma, c\gamma]$. Hence, *smoothed signal region* is $\mathbb{S}_{1,\gamma} = \cup_{j=1}^J S_{j,\gamma} = \cup_{j=1}^J (v_j - c\gamma, v_j + c\gamma)$, where $S_{j,\gamma}$ is the smoothed signal region of the change point v_j (see Figure 2).

Lemma S1. *Let* $I_j^{\text{side}} \cup I_j^{\text{mode}} = (v_j - c\gamma, v_j + c\gamma) = S_{j,\gamma}$ *be a partition of* $S_{j,\gamma}$*, where* $I_j^{\text{mode}} := \{t \in U(L) : |t - v_j| \leq \delta\}$ *and* $0 < \delta < \min\{b, \gamma\}$ *such that*

$I_j^{\text{mode}} \subset S_j$. *If* $q = \sup_j |q_j|$ *is sufficiently small, then*

(1). $M_\gamma = \inf_j M_{j,\gamma} > 0$ *is fixed, where* $M_{j,\gamma} = \inf_{I_j^{\text{mode}}} \frac{|\mu'_\gamma(t) - k(t)|}{|a_j|}$ *and* $k(t)$ *is the slope at* t.

(2). $C_\gamma = \inf_j C_{j,\gamma} > 0$ *is fixed, where* $C_{j,\gamma} = \frac{1}{|a_j|} \inf_{I_j^{\text{side}}} |\mu''_\gamma(t)|$.

(3). $D_\gamma = \inf_j D_{j,\gamma} > 0$ *is fixed, where* $D_{j,\gamma} = \frac{1}{|a_j|} \inf_{I_j^{\text{mode}}} |\mu_\gamma^{(3)}(t)|$.

Proof. (1). It follows from the arguments of Lemma 1 that

$$\begin{aligned}\frac{\mu'_\gamma(t) - k(t)}{a_j} &= \frac{1}{a_j}[\frac{a_j}{\gamma}\phi(\frac{v_j - t}{\gamma}) - (k_{j+1} - k_j)\Phi(\frac{v_j - t}{\gamma}) + k_{j+1}] - \frac{k(t)}{a_j} \\ &= \frac{1}{\gamma}\phi(\frac{v_j - t}{\gamma}) - q_j\Phi(\frac{v_j - t}{\gamma}) + \frac{k_{j+1} - k(t)}{a_j} \\ &= \frac{1}{\gamma}\phi(\frac{v_j - t}{\gamma}) - q_j\Phi(\frac{v_j - t}{\gamma}) + q_j I\{k(t) = k_j\}.\end{aligned}$$

For $t \in I_j^{\text{mode}}$, $(k_{j+1} - k(t))/a_j = 0$ or q_j as $k(t) = k_j$ or k_{j+1}. If q is sufficiently small, then $M_{j,\gamma} > 0$ and it follows that $M_\gamma = \inf M_{j,\gamma} > 0$.

(2). Note that $t \in I_j^{\text{side}} \subset S_{j,\gamma}$ implies $\delta < |v_j - t| < c\gamma$. By Lemma 1,

$$C_{j,\gamma} = \frac{1}{|a_j|} \inf_{I_j^{\text{side}}} |\mu''_\gamma(t)| = \inf_{I_j^{\text{side}}} |\frac{(v_j - t) + (k_{j+1} - k_j)\gamma^2/a_j}{\gamma^3}|\phi(\frac{v_j - t}{\gamma}).$$

As $k_{j+1} - k_j = 0$ and $|v_j - t| > 0$ for $t \in I_j^{\text{side}}$, then $C_{j,\gamma} > 0$ and $C_\gamma = \inf_j C_{j,\gamma} > 0$.

(3). For $t \in I_j^{\text{mode}}$, it follows from the proof of Lemma 2 that

$$D_{j,\gamma} = \frac{1}{|a_j|} \inf_{I_j^{\text{mode}}} |\mu_\gamma^{(3)}(t)| = \inf_{I_j^{\text{mode}}} |\frac{1}{\gamma^3} + \frac{(v_j - t)^2 + q_j(v_j - t)\gamma^2}{\gamma^5}|\phi(\frac{v_j - t}{\gamma}).$$

If q is sufficiently small for $t \in I_j^{\text{mode}}$, then $D_{j,\gamma} > 0$ and $D_\gamma = \inf_j D_{j,\gamma} > 0$. $\square$

To simplify the notations below, we regard $\tilde{T}_{\text{II}} \cap I_j^{\text{side}}$ as $\tilde{T}_{\text{II}}^{+} \cap I_j^{\text{side}}$ if $j \in \mathcal{I}_{\text{II}}^{+}$ and as $\tilde{T}_{\text{II}}^{-} \cap I_j^{\text{side}}$ if $j \in \mathcal{I}_{\text{II}}^{-}$, respectively. The same notation applies accordingly when $\tilde{T}_{\text{II}}$ is replaced with $\tilde{T}_{\text{II}}(u)$ or when I_j^{side} is replaced with I_j^{mode}.

Lemma S2. *For a Type* II *change point v_j, suppose that $q = \sup_j |q_j|$ is sufficiently small, then $\mu'_\gamma(t)$ has a local extremum at $t = v_j + \gamma^2 q_j \in I_j^{\text{mode}}$. Let $\sigma_1 = sd(z'_\gamma(t))$, $\sigma_2 = sd(z''_\gamma(t))$ and $\sigma_3 = sd(z_\gamma^{(3)}(t))$. Then*

(1). $P(\#\{t \in \tilde{T}_{\text{II}} \cap I_j^{\text{side}}\} = 0) \geq 1 - \exp(-\frac{a_j^2 C_{j,\gamma}^2}{2\sigma_2^2})$.

(2). $P(\#\{t \in \tilde{T}_{\text{II}} \cap I_j^{\text{mode}}\} = 1) \geq -1 + 2\Phi(\frac{|a_j| C_{j,\gamma}}{\sigma_2}) - \exp(-\frac{a_j^2 D_{j,\gamma}^2}{2\sigma_3^2})$.

(3). $P(\#\{t \in \tilde{T}_{\text{II}}(u) \cap I_j^{\text{mode}}\} = 1) \geq 1 - \exp(-\frac{a_j^2 D_{j,\gamma}^2}{2\sigma_3^2}) - \Phi(\frac{u - |a_j| M_{j,\gamma}}{\sigma_1})$ *for any fixed u.*

Proof. (1). As the probability that there are no local extrema of $y'_\gamma(t)$ in I_j^{side} is greater than the probability that $|y''_\gamma(t)| > 0$ for all $t \in I_j^{\mathrm{side}}$, we have

$$\begin{aligned}
P(\#\{t \in \tilde{T}_{\mathrm{II}} \cap I_j^{\mathrm{side}}\} = 0) &\geq P(\inf_{I_j^{\mathrm{side}}} |y''_\gamma(t)| > 0) \\
&\geq P(\sup_{I_j^{\mathrm{side}}} |z''_\gamma(t)| < \inf_{I_j^{\mathrm{side}}} |\mu''_\gamma(t)|) \\
&= 1 - P(\sup_{I_j^{\mathrm{side}}} |z''_\gamma(t)| > \inf_{I_j^{\mathrm{side}}} |\mu''_\gamma(t)|) \\
&\geq 1 - \exp(-\frac{a_j^2 C_{j,\gamma}^2}{2\sigma_2^2}),
\end{aligned} \tag{S5.5}$$

where the last line holds due to the Borell-TIS inequality.

(2). Consider here $j \in \mathcal{I}_{\mathrm{II}}^+$ since the case of $j \in \mathcal{I}_{\mathrm{II}}^-$ is similar. The probability that $y'_\gamma(t)$ has no local maximum in I_j^{mode} is less than the probability that $y''_\gamma(v_j - \delta) \leq 0$ or $y''_\gamma(v_j + \delta) \geq 0$. Thus, the probability of no local maximum in I_j^{mode} is bounded above by

$$\begin{aligned}
P(\#\{t \in \tilde{T}_{\mathrm{II}}^+ \cap I_j^{\mathrm{mode}}\} = 0) &\leq P(y''_\gamma(v_j - \delta) \leq 0 \cup y''_\gamma(v_j + \delta) \geq 0) \\
&\leq P(y''_\gamma(v_j - \delta) \leq 0) + P(y''_\gamma(v_j + \delta) \geq 0) \\
&= \Phi(-\frac{\mu''_\gamma(v_j - \delta)}{\sigma_2}) + \Phi(\frac{\mu''_\gamma(v_j + \delta)}{\sigma_2}) \\
&= 1 - \Phi(\frac{\mu''_\gamma(v_j - \delta)}{\sigma_2}) + 1 - \Phi(-\frac{\mu''_\gamma(v_j + \delta)}{\sigma_2}) \\
&\leq 2 - 2\Phi(\frac{|a_j| C_{j,\gamma}}{\sigma_2}),
\end{aligned}$$

where the last line holds because $\mu''_\gamma(v_j - \delta) \geq |a_j| C_{j,\gamma} > 0$ and $-\mu''_\gamma(v_j + \delta) \geq |a_j| C_{j,\gamma} > 0$.

On the other hand, the probability that $y'_\gamma(t)$ has at least two local maxima in I_j^{mode} is less than the probability that $y_\gamma^{(3)}(t) > 0$ for some $t \in I_j^{\mathrm{mode}}$ and $y_\gamma^{(3)}(t) < 0$ for some other $t \in I_j^{\mathrm{mode}}$. Thus, for $t \in \tilde{T}_{\mathrm{II}}^{+} \cap I_j^{\mathrm{mode}}$,

$$\begin{aligned}
P(\#\{t \in \tilde{T}_{\mathrm{II}}^{+} \cap I_j^{\mathrm{mode}}\} \geq 2) &\leq P(\sup y_\gamma^{(3)}(t) > 0 \cap \inf y_\gamma^{(3)}(t) < 0) \\
&\leq P(\sup y_\gamma^{(3)}(t) > 0) \wedge P(\inf y_\gamma^{(3)}(t) < 0) \\
&\leq P(\sup z_\gamma^{(3)}(t) > \inf -\mu_\gamma^{(3)}(t)) \wedge P(\sup z_\gamma^{(3)}(t) > \inf \mu_\gamma^{(3)}(t)) \\
&\leq \exp(-\frac{a_j^2 D_{j,\gamma}^2}{2\sigma_3^2}),
\end{aligned}$$

The last line holds due to Lemma S1.

The probability that $y'_\gamma(t)$ has only one local maximum in I_j^{mode} is calculated as

$$\begin{aligned}
&P(\#\{t \in \tilde{T}_{\mathrm{II}}^{+} \cap I_j^{\mathrm{mode}}\} = 1) \\
&= 1 - P(\#\{t \in \tilde{T}_{\mathrm{II}}^{+} \cap I_j^{\mathrm{mode}}\} = 0) - P(\#\{t \in \tilde{T}_{\mathrm{II}}^{+} \cap I_j^{\mathrm{mode}}\} \geq 2) \qquad \text{(S5.6)} \\
&\geq -1 + 2\Phi(\frac{|a_j| C_{j,\gamma}}{\sigma_2}) - \exp(-\frac{a_j^2 D_{j,\gamma}^2}{2\sigma_3^2}).
\end{aligned}$$

(3). Consider first $j \in \mathcal{I}_{\mathrm{II}}^{+}$. Since the probability that at least two local maxima of $y'_\gamma(t)$ in I_j^{mode} exceed $u + k(t)$ is less than the probability that $y'_\gamma(t)$ has at least two maxima in I_j^{mode}, then

$$\begin{aligned}
&P(\#\{t \in \tilde{T}_{\mathrm{II}}^{+} \cap I_j^{\mathrm{mode}} : y'_\gamma(t) - k(t) > u\} \geq 2) \\
&\leq P(\#\{t \in \tilde{T}_{\mathrm{II}}^{+} \cap I_j^{\mathrm{mode}}\} \geq 2) \leq \exp(-\frac{a_j^2 D_{j,\gamma}^2}{2\sigma_3^2}).
\end{aligned}$$

Similarly, $P(\#\{t \in \tilde{T}_{\mathrm{II}}^{-} \cap I_j^{\mathrm{mode}} : y'_\gamma(t) - k(t) < -u\} \geq 2) \leq \exp(-\frac{a_j^2 D_{j,\gamma}^2}{2\sigma_3^2})$.

On the other hand, since the probability that no local maxima of $y'_\gamma(t)$ in I_j^{mode} exceeding $u + k(t)$ is less than the probability that $y'_\gamma(t) - k(t) \le u$ anywhere in I_j^{mode}, then

$$\begin{aligned} & P(\#\{t \in \tilde{T}_{\text{II}}^+ \cap I_j^{\text{mode}} : y'_\gamma(t) - k(t) > u\} = 0) \\ & \le P(y'_\gamma(t) - k(t) \le u,\ \forall t \in I_j^{\text{mode}}) \\ & = P(z'_\gamma(t) + \mu'_\gamma(t) - k(t) \le u,\ \forall t \in I_j^{\text{mode}}) \\ & \le \Phi(\frac{u - a_j M_{j,\gamma}}{\sigma_1}) = \Phi(\frac{u - |a_j| M_{j,\gamma}}{\sigma_1}), \end{aligned}$$

where the last line holds since $a_j > 0$ for $j \in \mathcal{I}_{\text{II}}^+$.

Therefore, the probability that only one local maximum of $y'_\gamma(t)$ in I_j^{mode} exceeds $u + k(t)$ is

$$\begin{aligned} & P(\#\{t \in \tilde{T}_{\text{II}}^+ \cap I_j^{\text{mode}} : y'_\gamma(t) - k(t) > u\} = 1) \\ & = 1 - P(\#\{t \in \tilde{T}_{\text{II}}^+ \cap I_j^{\text{mode}} : y'_\gamma(t) - k(t) > u\} = 0) \\ & \quad - P(\#\{t \in \tilde{T}_{\text{II}}^+ \cap I_j^{\text{mode}} : y'_\gamma(t) - k(t) > u\} \ge 2) \\ & \ge 1 - \Phi(\frac{u - |a_j| M_{j,\gamma}}{\sigma_1}) - \exp(-\frac{a_j^2 D_{j,\gamma}^2}{2\sigma_3^2}). \end{aligned} \tag{S5.7}$$

The case of $j \in \mathcal{I}_{\text{II}}^-$ can be proved similarly. □

Lemma S3. *Let $\mathbb{T}_\gamma = \mathbb{S}_{1,\gamma} \setminus \mathbb{S}_1 = \cup_{j=1}^J (S_{j,\gamma} \setminus S_j)$ be the transition region. Let $W_{\text{II},\gamma} = \#\{t \in \tilde{T}_{\text{II}} \cap \mathbb{S}_{1,\gamma}\}$ and $W_{\text{II},\gamma}(u) = \#\{t \in \tilde{T}_{\text{II}}(u) \cap \mathbb{S}_{1,\gamma}\}$, where $W_{\text{II},\gamma}(u) = W_{\text{II},\gamma}^+(u) + W_{\text{II},\gamma}^-(u) = \#\{t \in \tilde{T}_{\text{II}}^+(u) \cap \mathbb{S}_{1,\gamma}\} + \#\{t \in \tilde{T}_{\text{II}}^-(u) \cap \mathbb{S}_{1,\gamma}\}$. Then, under condition $(C2)$, there exists $\eta > 0$ such that*

---

(1). $P(\#\{t \in \tilde{T}_{\text{II}} \cap \mathbb{T}_\gamma\} \geq 1) = o(\exp(-\eta a^2))$*;*

(2). $P(W_{\text{II},\gamma} = J) = 1 - o(\exp(-\eta a^2))$*;*

(3). $P(W_{\text{II},\gamma}(u) = J) = 1 - o(\exp(-\eta a^2))$*;*

(4). $W_{\text{II},\gamma}/L = A + o_p(1)$*;*

(5). $W_{\text{II},\gamma}(u)/W_{\text{II},\gamma} = 1 + o_p(1)$.

Proof. (1). Note that $I_j^{\text{mode}} \subset S_j$ implies $T_{j,\gamma} \subset I_j^{\text{side}}$. Applying Lemma S2, we obtain

$$\begin{aligned}
&P(\#\{t \in \tilde{T}_{\text{II}} \cap \mathbb{T}_\gamma\} \geq 1) \leq P(\#\{t \in \tilde{T}_{\text{II}} \cap (\cup_{j=1}^J I_j^{\text{side}})\} \geq 1) = P(\cup_{j=1}^J \#\{t \in \tilde{T}_{\text{II}} \cap I_j^{\text{side}}\} \geq 1) \\
&\quad \leq \sum_{j=1}^J [1 - P(\#\{t \in \tilde{T}_{\text{II}} \cap I_j^{\text{side}}\} = 0)] \leq \sum_{j=1}^J \exp(-\frac{a_j^2 C_{j,\gamma}^2}{2\sigma_2^2}) \leq \sum_{j=1}^J \exp(-\frac{a^2 C_\gamma^2}{2\sigma_2^2}) \\
&\quad = \frac{J}{L} L \exp(-\frac{a^2 C_\gamma^2}{2\sigma_2^2}) = o(\exp(-\eta a^2)),
\end{aligned}$$

where $0 < \eta < C_\gamma^2/(2\sigma_2^2)$, and the last line holds because $J/L \to A$ and

$$L \exp(-\frac{a^2 C_\gamma^2}{2\sigma_2^2}) = \exp\{a^2(\frac{\log L}{a^2} - \frac{C_\gamma^2}{2\sigma_2^2})\} = o(\exp(-\eta a^2)) \text{ as } \frac{\log L}{a^2} \to 0.$$

(2). Applying Lemma S2, we have that there exists $\eta > 0$ such that

$$
\begin{aligned}
P(W_{\mathrm{II},\gamma} = J) &= P(\#\{t \in \tilde{T}_{\mathrm{II}} \cap \mathbb{S}_{1,\gamma}\} = J) \\
&\geq P[\cap_{j=1}^{J}(\#\{t \in \tilde{T}_{\mathrm{II}} \cap I_j^{\mathrm{mode}}\} = 1 \cap \#\{t \in \tilde{T}_{\mathrm{II}} \cap I_j^{\mathrm{side}}\} = 0)] \\
&\geq 1 - \sum_{j=1}^{J}[1 - P(\#\{t \in \tilde{T}_{\mathrm{II}} \cap I_j^{\mathrm{mode}}\} = 1 \cap \#\{t \in \tilde{T}_{\mathrm{II}} \cap I_j^{\mathrm{side}}\} = 0)] \\
&\geq 1 - \sum_{j=1}^{J}[2 - P(\#\{t \in \tilde{T}_{\mathrm{II}} \cap I_j^{\mathrm{mode}}\} = 1) - P(\#\{t \in \tilde{T}_{\mathrm{II}} \cap I_j^{\mathrm{side}}\} = 0)] \\
&\geq 1 - \sum_{j=1}^{J}[2 - 2\Phi(\frac{|a_j|C_{j,\gamma}}{\sigma_2}) + \exp(-\frac{a_j^2 D_{j,\gamma}^2}{2\sigma_3^2}) + \exp(-\frac{a_j^2 C_{j,\gamma}^2}{2\sigma_2^2})] \\
&\geq 1 - \frac{J}{L}\{2L[1 - \Phi(\frac{aC_\gamma}{\sigma_2})] + L\exp(-\frac{a^2 C_\gamma^2}{2\sigma_2^2}) + L\exp(-\frac{a^2 D_\gamma^2}{2\sigma_3^2})\} \\
&= 1 - o(\exp(-\eta a^2)),
\end{aligned}
$$

where the last line holds because $L[1 - \Phi(Ka)] \leq L\phi(Ka)/(Ka)$ for any $K > 0$.

(3). Denote by $B_{j,0}$ and $B_{j,1}$ the events $\#\{t \in \tilde{T}_{\mathrm{II}} \cap I_j^{\mathrm{side}}\} = 0$ and $\#\{t \in \tilde{T}_{\mathrm{II}}(u) \cap I_j^{\mathrm{mode}}\} = 1$, respectively. Then, by Lemma S2, there exists $\eta > 0$ such that

$$P(W_{\mathrm{II},\gamma}(u)=J)=P(W^{+}_{\mathrm{II},\gamma}(u)+W^{-}_{\mathrm{II},\gamma}(u)=J)\geq P(\cap_{j=1}^{J}(B_{j,0}\cap B_{j,1}))$$
$$\geq 1-\sum_{j=1}^{J}[1-P(B_{j,0}\cap B_{j,1})]$$
$$\geq 1-\sum_{j=1}^{J}[1-(P(B_{j,0})+P(B_{j,1})-1)]$$
$$\geq 1-\sum_{j=1}^{J}[\exp(-a_j^2\frac{D_{j,\gamma}^2}{2\sigma_3^2})+\Phi(\frac{u-|a_j|M_{j,\gamma}}{\sigma_1})+\exp(-a_j^2\frac{C_{j,\gamma}^2}{2\sigma_2^2})]$$
$$\geq 1-\sum_{j=1}^{J}[\exp(-a^2\frac{D_{\gamma}^2}{2\sigma_3^2})+\Phi(\frac{u-aM_{\gamma}}{\sigma_1})+\exp(-a^2\frac{C_{\gamma}^2}{2\sigma_2^2})]=1-o(\exp(-\eta a^2)).$$

(4). It follows from part (2) that

$$\frac{W_{\mathrm{II},\gamma}}{L}=\frac{W_{\mathrm{II},\gamma}}{J}\frac{J}{L}=A(1+o_p(1))=A+o_p(1).$$

(5). Similarly to the proof of part (4), it follows from part (3) that $W_{\mathrm{II},\gamma}(u)/L=A+o_p(1)$. Then

$$W_{\mathrm{II},\gamma}(u)/W_{\mathrm{II},\gamma}=(W_{\mathrm{II},\gamma}(u)/L)/(W_{\mathrm{II},\gamma}/L)=1+o_p(1).$$

Remark S1. *The proof of FDR control and power consistency mainly relies on Lemma S3. While we have shown that the statements in Lemma S3 hold under the asymptotic condition C2, we now demonstrate that this condition is stricter than necessary.*

Proposition 1. *As $L\to\infty$ and $a\to\infty$, for any $K>0$, if $K-\frac{\log L}{a^2}>\eta>0$,* *then*

$$L\exp(-Ka^2) = \exp\{a^2[\frac{\log L}{a^2} - K]\} = o(\exp(-\eta a^2)).$$

The above Proposition 1 is straightforward to verify. By applying Proposition 1, to ensure all statements in Lemma S3 hold, it suffices to set $K = \min\{\frac{C_\gamma^2}{2\sigma_2^2}, \frac{D_\gamma^2}{2\sigma_3^2}\}$ *and require that* $\frac{\log L}{a^2} < K$. *This implies that the asymptotic condition* $\log L/a^2 \to 0$ *in condition C2 is overly restrictive and can be relaxed to* $\frac{\log L}{a^2} < \min\{\frac{C_\gamma^2}{2\sigma_2^2}, \frac{D_\gamma^2}{2\sigma_3^2}\}$*, which depends on the variance of the second and third derivatives of* $z_\gamma(t)$.

Similarly, we can also relax the asymptotic condition for $\log L/k$ *in condition C1, such that it is bounded by a constant that depends on the third and fourth derivatives of* $z_\gamma(t)$.

□

S6. FDR Control and Power Consistency for Type II Change Points

S6.1 FDR control

Proof of Theorem 2 (FDR control). Recall that $V_{\text{II}}(u) = \#\{t \in \tilde{T}_{\text{II}}(u) \cap \mathbb{S}_0\}$ and $W_{\text{II},\gamma}(u) = \#\{t \in \tilde{T}_{\text{II}}(u) \cap \mathbb{S}_{1,\gamma}\}$. Let $W_{\text{II}}(u) = \#\{t \in \tilde{T}_{\text{II}}(u) \cap \mathbb{S}_1\}$ and $V_{\text{II},\gamma}(u) = \#\{t \in \tilde{T}_{\text{II}}(u) \cap \mathbb{S}_{0,\gamma}\}$. That is, $V_{\text{II}}(u)$ and $W_{\text{II}}(u)$ represent the number of local maxima above u plus local minima below $-u$ in the *null region* and *signal region* , respectively; and $V_{\mathrm{II},\gamma}(u)$ and $W_{\mathrm{II},\gamma}(u)$ represent the number of local maxima above u plus local minima below $-u$ in the *smoothed null region* and *smoothed signal region* , respectively. Accordingly, V_{II}, W_{II}, $V_{\mathrm{II},\gamma}$, and $W_{\mathrm{II},\gamma}$ are defined as the number of extrema in $\mathbb{S}_0$, $\mathbb{S}_1$, $\mathbb{S}_{0,\gamma}$, and $\mathbb{S}_{1,\gamma}$, respectively. Note that $R_{\mathrm{II}}(u) = \#\{t \in \tilde{T}_{\mathrm{II}}(u)\} = V_{\mathrm{II}}(u) + W_{\mathrm{II}}(u)$. By the definition of FDR for Type II change points detection, for any fixed u, we have

$$\mathrm{FDR}_{\mathrm{II}}(u) = E[\frac{V_{\mathrm{II}}(u)}{V_{\mathrm{II}}(u) + W_{\mathrm{II}}(u)}] = E[\frac{V_{\mathrm{II}}(u)/L}{V_{\mathrm{II}}(u)/L + W_{\mathrm{II}}(u)/L}]. \tag{S6.8}$$

Note that

$$P(V_{\mathrm{II},\gamma}(u) = V_{\mathrm{II}}(u)) = 1 - P(V_{\mathrm{II},\gamma}(u) \neq V_{\mathrm{II}}(u)) = 1 - P(\#\{t \in \tilde{T}_{\mathrm{II}}(u) \cap \mathbb{T}_\gamma\} \geq 1)$$
$$\geq 1 - P(\#\{t \in \tilde{T}_{\mathrm{II}} \cap \mathbb{T}_\gamma\} \geq 1) = 1 - o(\exp(-\eta a^2)).$$

Hence, $V_{\mathrm{II},\gamma}(u) = V_{\mathrm{II}}(u) + o_p(1)$. Similarly, we have $W_{\mathrm{II},\gamma}(u) = W_{\mathrm{II}}(u) + o_p(1)$.

Then it follows that

$$\begin{aligned}\frac{V_{\mathrm{II}}(u)}{L} &= \frac{V_{\mathrm{II}}(u)}{V_{\mathrm{II},\gamma}(u)}\frac{V_{\mathrm{II},\gamma}(u)}{L} = [1 + o_p(1)]\frac{[L - J(2c\gamma)]E[\tilde{m}_{z'_\gamma}(U(1), u)]}{L} \\ &= (1 - 2c\gamma A)E[\tilde{m}_{z'_\gamma}(U(1), u)] + o_p(1), \\ \frac{W_{\mathrm{II}}(u)}{L} &= \frac{W_{\mathrm{II}}(u)}{W_{\mathrm{II},\gamma}(u)}\frac{W_{\mathrm{II},\gamma}(u)}{L} = [1 + o_p(1)][A + o_p(1)] = A + o_p(1).\end{aligned}$$

Hence,

$$\frac{V_{\mathrm{II}}(u)/L}{V_{\mathrm{II}}(u)/L + W_{\mathrm{II}}(u)/L} = \frac{(1 - 2c\gamma A)E[\tilde{m}_{z'_\gamma}(U(1), u)]}{A + (1 - 2c\gamma A)E[\tilde{m}_{z'_\gamma}(U(1), u)]} + o_p(1) < 1. \tag{S6.9}$$

As $\frac{V_{\mathrm{II}}(u)}{V_{\mathrm{II}}(u)+W_{\mathrm{II}}(u)} \leq 1$, by the Dominated Convergence Theorem (DCT),

$$\begin{aligned}\lim E[\frac{V_{\mathrm{II}}(u)/L}{V_{\mathrm{II}}(u)/L + W_{\mathrm{II}}(u)/L}] &= E[\lim \frac{V_{\mathrm{II}}(u)/L}{V_{\mathrm{II}}(u)/L + W_{\mathrm{II}}(u)/L}] \\ &= \frac{(1-2c\gamma A)E[\tilde{m}_{z'_\gamma}(U(1),u)]}{A + (1-2c\gamma A)E[\tilde{m}_{z'_\gamma}(U(1),u)]}.\end{aligned}$$

That is,

$$\mathrm{FDR}_{\mathrm{II}}(u) \to \frac{(1-2c\gamma A)E[\tilde{m}_{z'_\gamma}(U(1),u)]}{A + (1-2c\gamma A)E[\tilde{m}_{z'_\gamma}(U(1),u)]}, \tag{S6.10}$$

completing Part (i) in Theorem 2 for the FDR control.

Next we will show the FDR control for random threshold $\tilde{u}_{\mathrm{II}}$. Let $\tilde{G}(u) = \#\{t \in \tilde{T}_{\mathrm{II}}(u)\}/\#\{t \in \tilde{T}_{\mathrm{II}}\}$ be the empirical marginal right cdf of $z'_\gamma(t)$ given $t \in \tilde{T}_{\mathrm{II}}$. By the Benjamini-Hochberg (BH) procedure, the BH threshold $\tilde{u}_{\mathrm{II}}$ satisfies $\alpha\tilde{G}(\tilde{u}_{\mathrm{II}}) = k\alpha/m_{\mathrm{II}} = F_{z'_\gamma}(\tilde{u}_{\mathrm{II}})$, so $\tilde{u}_{\mathrm{II}}$ is the largest u that solves the equation

$$\alpha\tilde{G}(u) = F_{z'_\gamma}(u). \tag{S6.11}$$

Note that

$$\tilde{G}(u) = \frac{V_{\mathrm{II},\gamma}(u) + W_{\mathrm{II},\gamma}(u)}{V_{\mathrm{II},\gamma} + W_{\mathrm{II},\gamma}} = \frac{V_{\mathrm{II},\gamma}(u)}{V_{\mathrm{II},\gamma}} \frac{V_{\mathrm{II},\gamma}}{V_{\mathrm{II},\gamma} + W_{\mathrm{II},\gamma}} + \frac{W_{\mathrm{II},\gamma}(u)}{W_{\mathrm{II},\gamma}} \frac{W_{\mathrm{II},\gamma}}{V_{\mathrm{II},\gamma} + W_{\mathrm{II},\gamma}},$$

By the weak law of large numbers or Lemma 8 in Schwartzman et al. (2011),

$$\frac{V_{\mathrm{II},\gamma}(u)/L}{V_{\mathrm{II},\gamma}/L} \xrightarrow{P} \frac{E[V_{\mathrm{II},\gamma}(u)]}{E[V_{\mathrm{II},\gamma}]} = F_{z'_\gamma}(u). \tag{S6.12}$$

In addition,

$$\frac{V_{\mathrm{II},\gamma}}{V_{\mathrm{II},\gamma}+W_{\mathrm{II},\gamma}}=\frac{V_{\mathrm{II},\gamma}/L}{V_{\mathrm{II},\gamma}/L+W_{\mathrm{II},\gamma}/L}\xrightarrow{P}\frac{E[\tilde{m}_{z'_\gamma}(U(1))](1-2c\gamma A)}{E[\tilde{m}_{z'_\gamma}(U(1))](1-2c\gamma A)+A},$$

and

$$\frac{W_{\mathrm{II},\gamma}}{V_{\mathrm{II},\gamma}+W_{\mathrm{II},\gamma}}=\frac{W_{\mathrm{II},\gamma}/L}{V_{\mathrm{II},\gamma}/L+W_{\mathrm{II},\gamma}/L}\xrightarrow{P}\frac{A}{E[\tilde{m}_{z'_\gamma}(U(1))](1-2c\gamma A)+A}.$$

Combining these with (S6.11) and Part (5) in Lemma S3, we obtain

$$\tilde{G}(u)\xrightarrow{P}F_{z'_\gamma}(u)\frac{E[\tilde{m}_{z'_\gamma}(U(1))](1-2c\gamma A)}{E[\tilde{m}_{z'_\gamma}(U(1))](1-2c\gamma A)+A}+\frac{A}{E[\tilde{m}_{z'_\gamma}(U(1))](1-2c\gamma A)+A}.$$

Plugging $\tilde{G}(u)$ by this limit in (S6.11) and solving for u gives the deterministic solution u^*_{II} such that

$$F_{z'_\gamma}(u^*_{\mathrm{II}})=\frac{\alpha A}{A+E[\tilde{m}_{z'_\gamma}(U(1))](1-2c\gamma A)(1-\alpha)}. \tag{S6.13}$$

Note that $\tilde{u}_{\mathrm{II}}$ is the solution of $\alpha\tilde{G}(u)=F_{z'_\gamma}(u)$ and u^*_{II} is the solution of $\lim\alpha\tilde{G}(u)=F_{z'_\gamma}(u)$, and since $F^{-1}_{z'_\gamma}(\cdot)$ is continuous and monotonic, we have $\tilde{u}_{\mathrm{II}}\xrightarrow{P}u^*_{\mathrm{II}}$. That is, for any $\epsilon>0$,

$$P(|\tilde{u}_{\mathrm{II}}-u^*_{\mathrm{II}}|>\epsilon)\to 0.$$

By the definition of FDR, for the random threshold $\tilde{u}_{\mathrm{II}}$ and $\epsilon>0$,

$$\begin{aligned}\mathrm{FDR}_{\mathrm{II,BH}}=\mathrm{FDR}_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})&=E[\frac{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})+W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}]\\&=E[\frac{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})+W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}\mathbb{1}(|\tilde{u}_{\mathrm{II}}-u^*_{\mathrm{II}}|\le\epsilon)]+E[\frac{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})+W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}\mathbb{1}(|\tilde{u}_{\mathrm{II}}-u^*_{\mathrm{II}}|>\epsilon)]\\&=E[\frac{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})+W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}\mathbb{1}(|\tilde{u}_{\mathrm{II}}-u^*_{\mathrm{II}}|\le\epsilon)]+o(1).\end{aligned} \tag{S6.14}$$

Since both $V_{\rm II}(u)$ and $W_{\rm II}(u)$ are decreasing with respect to u, we have that, for L large enough,

$$
\begin{aligned}
&E[\frac{V_{\rm II}(\tilde{u}_{\rm II})}{V_{\rm II}(\tilde{u}_{\rm II})+W_{\rm II}(\tilde{u}_{\rm II})}\mathbb{1}(|\tilde{u}_{\rm II}-u^*_{\rm II}|\leq\epsilon)\geq E[\frac{V_{\rm II}(u^*_{\rm II}+\epsilon)}{V_{\rm II}(u^*_{\rm II}-\epsilon)+W_{\rm II}(u^*_{\rm II}-\epsilon)}\mathbb{1}(|\tilde{u}_{\rm II}-u^*_{\rm II}|\leq\epsilon)],\\
&E[\frac{V_{\rm II}(\tilde{u}_{\rm II})}{V_{\rm II}(\tilde{u}_{\rm II})+W_{\rm II}(\tilde{u}_{\rm II})}\mathbb{1}(|\tilde{u}_{\rm II}-u^*_{\rm II}|\leq\epsilon)\leq E[\frac{V_{\rm II}(u^*_{\rm II}-\epsilon)}{V_{\rm II}(u^*_{\rm II}+\epsilon)+W_{\rm II}(u^*_{\rm II}+\epsilon)}].
\end{aligned}
$$

Additionally, note that the expectation of the number of extrema exceeing level u is continuous in u by the Kac-Rice formula. Thus, similarly to (S6.9), we have that, for large L, as $\epsilon\to 0$,

$$
\begin{aligned}
&\left|E[\frac{V_{\rm II}(u^*_{\rm II}+\epsilon)}{V_{\rm II}(u^*_{\rm II}-\epsilon)+W_{\rm II}(u^*_{\rm II}-\epsilon)}\mathbb{1}(|\tilde{u}_{\rm II}-u^*_{\rm II}|\leq\epsilon)]-E[\frac{V_{\rm II}(u^*_{\rm II})}{V_{\rm II}(u^*_{\rm II})+W_{\rm II}(u^*_{\rm II})}]\right|\\
&\leq\left|E[\frac{V_{\rm II}(u^*_{\rm II}+\epsilon)}{V_{\rm II}(u^*_{\rm II}-\epsilon)+W_{\rm II}(u^*_{\rm II}-\epsilon)}\mathbb{1}(|\tilde{u}_{\rm II}-u^*_{\rm II}|\leq\epsilon)]-E[\frac{V_{\rm II}(u^*_{\rm II}+\epsilon)}{V_{\rm II}(u^*_{\rm II}-\epsilon)+W_{\rm II}(u^*_{\rm II}-\epsilon)}]\right|\\
&\quad+\left|E[\frac{V_{\rm II}(u^*_{\rm II}+\epsilon)/L}{V_{\rm II}(u^*_{\rm II}-\epsilon)/L+W_{\rm II}(u^*_{\rm II}-\epsilon)/L}]-E[\frac{V_{\rm II}(u^*_{\rm II})/L}{V_{\rm II}(u^*_{\rm II})/L+W_{\rm II}(u^*_{\rm II})/L}]\right|\\
&\to 0
\end{aligned}
$$

and

$$
\left|E[\frac{V_{\rm II}(u^*_{\rm II}-\epsilon)}{V_{\rm II}(u^*_{\rm II}+\epsilon)+W_{\rm II}(u^*_{\rm II}+\epsilon)}]-E[\frac{V_{\rm II}(u^*_{\rm II})}{V_{\rm II}(u^*_{\rm II})+W_{\rm II}(u^*_{\rm II})}]\right|\to 0,
$$

which implies

$$
E[\frac{V_{\rm II}(\tilde{u}_{\rm II})}{V_{\rm II}(\tilde{u}_{\rm II})+W_{\rm II}(\tilde{u}_{\rm II})}\mathbb{1}(|\tilde{u}_{\rm II}-u^*_{\rm II}|\leq\epsilon)]-E[\frac{V_{\rm II}(u^*_{\rm II})}{V_{\rm II}(u^*_{\rm II})+W_{\rm II}(u^*_{\rm II})}]\to 0. \tag{S6.15}
$$

Therefore, by (S6.14) and (S6.15), we obtain

$$\begin{aligned}
\lim \mathrm{FDR}_{\mathrm{II,BH}} &= \lim E[\frac{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}}) + W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})} \mathbb{1}(|\tilde{u}_{\mathrm{II}} - u^*_{\mathrm{II}}| \le \epsilon)] \\
&= \lim E[\frac{V_{\mathrm{II}}(u^*_{\mathrm{II}})}{V_{\mathrm{II}}(u^*_{\mathrm{II}}) + W_{\mathrm{II}}(u^*_{\mathrm{II}})}] \\
&= E[\lim \frac{V_{\mathrm{II}}(u^*_{\mathrm{II}})/L}{V_{\mathrm{II}}(u^*_{\mathrm{II}})/L + W_{\mathrm{II}}(u^*_{\mathrm{II}})/L}] \\
&= \frac{F_{z'_\gamma}(u^*_{\mathrm{II}}) E[\tilde{m}_{z'_\gamma}(U(1))](1 - 2c\gamma A)}{F_{z'_\gamma}(u^*_{\mathrm{II}}) E[\tilde{m}_{z'_\gamma}(U(1))](1 - 2c\gamma A) + A} \quad \text{(by (S6.9) and (S6.12))} \\
&= \alpha \frac{E[\tilde{m}_{z'_\gamma}(U(1))](1 - 2c\gamma A)}{E[\tilde{m}_{z'_\gamma}(U(1))](1 - 2c\gamma A) + A} \quad \text{(by (S6.13))}.
\end{aligned} \tag{S6.16}$$

□

S6.2 Power consistency

Proof of Theorem 2 (power consistency). By Lemma S2, for any fixed u,

$$\begin{aligned}
\mathrm{Power}_{\mathrm{II},j}(u) &= P(\#\{t \in \tilde{T}_{\mathrm{II}}(u) \cap S_j\} \ge 1) \\
&\ge P(\#\{t \in \tilde{T}_{\mathrm{II}}(u) \cap I_j^{\mathrm{mode}}\} \ge 1) \\
&\ge 1 - \exp(-\frac{a_j^2 D_{j,\gamma}^2}{2\sigma^3}) - \Phi(\frac{u - |a_j| M_{j,\gamma}}{\sigma_1}) \to 1.
\end{aligned}$$

Therefore,

$$\mathrm{Power}_{\mathrm{II}}(u) = \frac{1}{J} \sum_{j=1}^{J} \mathrm{Power}_{\mathrm{II},j}(u) \to 1.$$

For the random threshold $\tilde{u}_{\mathrm{II}}$ and arbitrary $\epsilon > 0$, we have

$$\begin{aligned}
P(\#\{t \in \tilde{T}_{\mathrm{II}}(\tilde{u}_{\mathrm{II}}) \cap S_j\} \ge 1) &= P(\#\{t \in \tilde{T}_{\mathrm{II}}(\tilde{u}_{\mathrm{II}}) \cap S_j\} \ge 1, |\tilde{u}_{\mathrm{II}} - u^*_{\mathrm{II}}| \le \epsilon) \\
&\quad + P(\#\{t \in \tilde{T}_{\mathrm{II}}(\tilde{u}_{\mathrm{II}}) \cap S_j\} \ge 1, |\tilde{u}_{\mathrm{II}} - u^*_{\mathrm{II}}| > \epsilon),
\end{aligned} \tag{S6.17}$$

where the last term is bounded above by $P(|\tilde{u}_{\mathrm{II}} - u^*_{\mathrm{II}}| > \epsilon) \to 0$ as $\epsilon \to 0$. Moreover,

$$
\begin{aligned}
&P(\#\{t \in \tilde{T}_{\mathrm{II}}(\tilde{u}_{\mathrm{II}}) \cap S_j\} \geq 1, |\tilde{u}_{\mathrm{II}} - u^*_{\mathrm{II}}| \leq \epsilon) \\
\geq\ & P(\#\{t \in \tilde{T}_{\mathrm{II}}(u^*_{\mathrm{II}} + \epsilon) \cap S_j\} \geq 1, |\tilde{u}_{\mathrm{II}} - u^*_{\mathrm{II}}| \leq \epsilon) \\
\geq\ & 1 - P(\#\{t \in \tilde{T}_{\mathrm{II}}(u^*_{\mathrm{II}} + \epsilon) \cap S_j\} = 0) - P(|\tilde{u}_{\mathrm{II}} - u^*_{\mathrm{II}}| > \epsilon) \to 1,
\end{aligned}
$$

where the last line holds because for any two events A and B, $P(A \cap B) = 1 - P(A^c \cup B^c) \geq 1 - P(A^c) - P(B^c)$. Therefore, by Lemma S2, $\mathrm{Power}_{\mathrm{II},j}(\tilde{u}_{\mathrm{II}}) \to 1$, and hence

$$
\mathrm{Power}_{\mathrm{II,BH}} = \mathrm{Power}_{\mathrm{II,BH}}(\tilde{u}_{\mathrm{II}}) = \frac{1}{J} \sum_{j=1}^{J} \mathrm{Power}_{\mathrm{II},j}(\tilde{u}_{\mathrm{II}}) \to 1.
$$

□

S7. FDR Control and Power Consistency for Mixture of Type I and Type II Change Points

S7.1 FDR control

Proof of Theorem 3 (FDR control). Recall that the FDR for mixture of Type I and Type II change points is defined as

$$
\mathrm{FDR}_{\mathrm{III}}(u_1, u_2) = E\left\{ \frac{V_{\mathrm{I}\backslash\mathrm{II}}(u_1) + V_{\mathrm{II}}(u_2)}{R_{\mathrm{I}\backslash\mathrm{II}}(u_1) + R_{\mathrm{II}}(u_2)} \right\},
$$

and that $\mathbb{S}_{1,\mathrm{I}\backslash\mathrm{II}} = \cup_{j=1}^{J_1}(v_j - b, v_j + b) \setminus \mathbb{S}_{\mathrm{II}}^*$, where v_j is a Type I change point and $\mathbb{S}_{\mathrm{II}}^* = \cup_{i=1}^{R_{\mathrm{II}}(u_2)}(\hat{v}_i - 2\gamma, \hat{v}_i + 2\gamma)$ for all significant $\hat{v}_i \in \tilde{T}_{\mathrm{II}}$. Let $\mathbb{S}_{1,\gamma,\mathrm{I}\backslash\mathrm{II}}$ and $\mathbb{S}_{0,\gamma,\mathrm{I}\backslash\mathrm{II}}$ be the smoothed signal region and smoothed null region of Type I change points, denoted by $\mathbb{S}_{1,\gamma,\mathrm{I}\backslash\mathrm{II}} = \cup_{j=1}^{J_1}(v_j - c\gamma, v_j + c\gamma) \setminus \mathbb{S}_{\mathrm{II}}^*$ and $\mathbb{S}_{0,\gamma,\mathrm{I}\backslash\mathrm{II}} = U(L) \setminus \mathbb{S}_{1,\gamma,\mathrm{I}\backslash\mathrm{II}}$.

Let $W_{\mathrm{I}\backslash\mathrm{II}}(u_1) = \#\{t \in \tilde{T}_{\mathrm{I}\backslash\mathrm{II}}(u_1) \cap \mathbb{S}_{1,\mathrm{I}\backslash\mathrm{II}}\}$ and $W_{\mathrm{I}\backslash\mathrm{II},\gamma}(u_1) = \#\{t \in \tilde{T}_{\mathrm{I}\backslash\mathrm{II}}(u_1) \cap \mathbb{S}_{1,\gamma,\mathrm{I}\backslash\mathrm{II}}\}$. Then $R_{\mathrm{I}\backslash\mathrm{II}}(u_1) = V_{\mathrm{I}\backslash\mathrm{II}}(u_1) + W_{\mathrm{I}\backslash\mathrm{II}}(u_1)$, and $\mathrm{FDR}_{\mathrm{III}}(u_1, u_2)$ can be rewritten as

$$\begin{aligned}
\mathrm{FDR}_{\mathrm{III}}(u_1, u_2) &= E[\frac{V_{\mathrm{I}\backslash\mathrm{II}}(u_1) + V_{\mathrm{II}}(u_2)}{V_{\mathrm{I}\backslash\mathrm{II}}(u_1) + W_{\mathrm{I}\backslash\mathrm{II}}(u_1) + V_{\mathrm{II}}(u_2) + W_{\mathrm{II}}(u_2)}] \\
&= E\{E[\frac{V_{\mathrm{I}\backslash\mathrm{II}}(u_1) + V_{\mathrm{II}}(u_2)}{V_{\mathrm{I}\backslash\mathrm{II}}(u_1) + W_{\mathrm{I}\backslash\mathrm{II}}(u_1) + V_{\mathrm{II}}(u_2) + W_{\mathrm{II}}(u_2)} | V_{\mathrm{II}}(u_2)]\}
\end{aligned}$$

Let $\Delta(u_2)$ be the length of $\mathbb{S}_{0,\gamma,\mathrm{I}\backslash\mathrm{II}}$. Then

$$L - 2c\gamma J_1 - 4\gamma R_{\mathrm{II}}(u_2) \leq \Delta(u_2) \leq L - 2c\gamma J_1,$$

where the left inequality holds because some of the falsely detected Type II change points may overlap with Type I change points.

Recall that for the detection of Type II change points in step 1 of Algorithm 3, we have

$$\begin{aligned}
&\frac{V_{\mathrm{II}}(u_2)}{L} \to E[\tilde{m}_{z'_\gamma}(U(1), u_2)](1 - 2c\gamma A_2), \\
&\frac{W_{\mathrm{II}}(u_2)}{L} \to A_2.
\end{aligned}$$

Since we assume that the distance between neighboring change points are large enough, and the detected Type II change points have been removed when detecting Type I change points, then $V_{\mathrm{I}\backslash\mathrm{II}}(u_1)$ and $W_{\mathrm{I}\backslash\mathrm{II}}(u_1)$ only depend on $V_{\mathrm{II}}(u_2)$,

$$\frac{V_{\mathrm{I}\backslash\mathrm{II}}(u_1)}{L}\big|V_{\mathrm{II}}(u_2) \sim E[\tilde{m}_{z''_\gamma}(U(1),u_1)]\Delta(u_2)/L,$$
$$\frac{W_{\mathrm{I}\backslash\mathrm{II}}(u_1)}{L}\big|V_{\mathrm{II}}(u_2) = \frac{W_{\mathrm{I}\backslash\mathrm{II}}(u_1)}{W_{\mathrm{I}\backslash\mathrm{II},\gamma}(u_1)}\frac{W_{\mathrm{I}\backslash\mathrm{II},\gamma}(u_1)}{L} = \frac{W_{\mathrm{I}\backslash\mathrm{II},\gamma}(u_1)}{L} + o_p(1).$$

Moreover, notice that the falsely detected Type II change points mainly result from random noise (due to FDR control), which are uniformly distributed over the null signal region of Type I change points. Hence in fact, $\cup_{j=1}^{J_1}(v_j - c\gamma, v_j + c\gamma) \cap \mathbb{S}_{\mathrm{II}}^*$ tends to an empty set. Thus $\frac{W_{\mathrm{I}\backslash\mathrm{II},\gamma}(u_1)}{L} \to A_1$ and

$$\begin{aligned}
&\frac{V_{\mathrm{I}\backslash\mathrm{II}}(u_1)+V_{\mathrm{II}}(u_2)}{V_{\mathrm{I}\backslash\mathrm{II}}(u_1)+W_{\mathrm{I}\backslash\mathrm{II}}(u_1)+V_{\mathrm{II}}(u_2)+W_{\mathrm{II}}(u_2)}\big|V_{\mathrm{II}}(u_2)\\
&= \frac{V_{\mathrm{I}\backslash\mathrm{II}}(u_1)/L+V_{\mathrm{II}}(u_2)/L}{V_{\mathrm{I}\backslash\mathrm{II}}(u_1)/L+W_{\mathrm{I}\backslash\mathrm{II}}(u_1)/L+V_{\mathrm{II}}(u_2)/L+W_{\mathrm{II}}(u_2)/L}\big|V_{\mathrm{II}}(u_2)\\
&\sim \frac{E[\tilde{m}_{z''_\gamma}(U(1),u_1)]\Delta(u_2)/L+V_{\mathrm{II}}(u_2)/L}{E[\tilde{m}_{z''_\gamma}(U(1),u_1)]\Delta(u_2)/L+V_{\mathrm{II}}(u_2)/L+A_1+A_2}\big|V_{\mathrm{II}}(u_2).
\end{aligned}$$

As $\Delta(u_2) \leq L - 2c\gamma J_1$, by DCT, we obtain

$$\begin{aligned}
\lim \mathrm{FDR}_{\mathrm{III}}(u_1,u_2) &= \lim E\{E[\frac{V_{\mathrm{I}\backslash\mathrm{II}}(u_1)+V_{\mathrm{II}}(u_2)}{V_{\mathrm{I}\backslash\mathrm{II}}(u_1)+W_{\mathrm{I}\backslash\mathrm{II}}(u_1)+V_{\mathrm{II}}(u_2)+W_{\mathrm{II}}(u_2)}\big|V_{\mathrm{II}}(u_2)]\}\\
&= E\{\lim E[\frac{V_{\mathrm{I}\backslash\mathrm{II}}(u_1)/L+V_{\mathrm{II}}(u_2)/L}{V_{\mathrm{I}\backslash\mathrm{II}}(u_1)/L+W_{\mathrm{I}\backslash\mathrm{II}}(u_1)/L+V_{\mathrm{II}}(u_2)/L+W_{\mathrm{II}}(u_2)/L}\big|V_{\mathrm{II}}(u_2)]\}\\
&\leq \frac{E[\tilde{m}_{z''_\gamma}(U(1),u_1)](1-2c\gamma A_1)+E[\tilde{m}_{z'_\gamma}(U(1),u_2)](1-2c\gamma A_2)}{E[\tilde{m}_{z''_\gamma}(U(1),u_1)](1-2c\gamma A_1)++E[\tilde{m}_{z'_\gamma}(U(1),u_2)](1-2c\gamma A_2)+A}.
\end{aligned} \tag{S7.18}$$

For the random thresholds $\tilde{u}_{\mathrm{I}}$ and $\tilde{u}_{\mathrm{II}}$,

$$\begin{aligned}\mathrm{FDR}_{\mathrm{III}}(\tilde{u}_{\mathrm{I}},\tilde{u}_{\mathrm{II}}) &= E[\frac{V_{\mathrm{I}\setminus\mathrm{II}}(\tilde{u}_{\mathrm{I}})+V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}{V_{\mathrm{I}\setminus\mathrm{II}}(\tilde{u}_{\mathrm{I}})+W_{\mathrm{I}\setminus\mathrm{II}}(\tilde{u}_{\mathrm{I}})+V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})+W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}] \\ &= E\{E\big[\frac{V_{\mathrm{I}\setminus\mathrm{II}}(\tilde{u}_{\mathrm{I}})+V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}{V_{\mathrm{I}\setminus\mathrm{II}}(\tilde{u}_{\mathrm{I}})+W_{\mathrm{I}\setminus\mathrm{II}}(\tilde{u}_{\mathrm{I}})+V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})+W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})}\big|V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})\big]\}.\end{aligned}$$

Similarly to (S6.16), we obtain that, as $\tilde{u}_{\mathrm{I}} \xrightarrow{P} u_{\mathrm{I}}^*$,

$$\mathrm{FDR}_{\mathrm{III}}(\tilde{u}_{\mathrm{I}},\tilde{u}_{\mathrm{II}}) \sim E\big[\frac{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})/L+F_{z''_\gamma}(u_{\mathrm{I}}^*)E[\tilde{m}_{z''_{\mathrm{I}}}(U(1))]\Delta(\tilde{u}_{\mathrm{II}})/L}{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})/L+W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})/L+F_{z''_\gamma}(u_{\mathrm{I}}^*)E[\tilde{m}_{z''_\gamma}(U(1))]\Delta(\tilde{u}_{\mathrm{II}})/L+A_1}\big],$$

where

$$\begin{aligned}&L-2c\gamma J_1-4\gamma R_{\mathrm{II}}(\tilde{u}_{\mathrm{II}}) \le \Delta(\tilde{u}_{\mathrm{II}}) \le L-2c\gamma J_1, \\ &F_{z''_\gamma}(u_{\mathrm{I}}^*) = \frac{\alpha A_1}{A_1+(1-\alpha)E[\tilde{m}_{z''_\gamma}(U(1))]\Delta(\tilde{u}_{\mathrm{II}})/L}.\end{aligned}$$

Additionally,

$$\begin{aligned}F_{z''_\gamma}(u_{\mathrm{I}}^*)\Delta(\tilde{u}_{\mathrm{II}})/L &= \frac{\alpha A_1\Delta(\tilde{u}_{\mathrm{II}})/L}{A_1+(1-\alpha)E[\tilde{m}_{z''_\gamma}(U(1))]\Delta(\tilde{u}_{\mathrm{II}})/L} \\ &\le \frac{\alpha A_1(1-2c\gamma A_1)}{A_1+(1-\alpha)E[\tilde{m}_{z''_\gamma}(U(1))](1-2c\gamma A_1)}.\end{aligned}$$

Therefore,

$$\begin{aligned}&\lim \mathrm{FDR}_{\mathrm{III}}(\tilde{u}_{\mathrm{I}},\tilde{u}_{\mathrm{II}}) \\ &= E\Big[\lim \frac{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})/L+F_{z''_\gamma}(u_{\mathrm{I}}^*)E[\tilde{m}_{z''_{\mathrm{I}}}(U(1))]\Delta(\tilde{u}_{\mathrm{II}})/L}{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})/L+W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})/L+F_{z''_\gamma}(u_{\mathrm{I}}^*)E[\tilde{m}_{z''_\gamma}(U(1))]\Delta(\tilde{u}_{\mathrm{II}})/L+\tilde{A}_1^*}\Big] \\ &\le E\Big[\lim \frac{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})/L+E[\tilde{m}_{z''_\gamma}(U(1))]\frac{\alpha A_1(1-2c\gamma A_1)}{A_1+(1-\alpha)E[\tilde{m}_{z''_\gamma}(U(1))](1-2c\gamma A_1)}}{V_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})/L+W_{\mathrm{II}}(\tilde{u}_{\mathrm{II}})/L+A_1+E[\tilde{m}_{z''_\gamma}(U(1))]\frac{\alpha A_1(1-2c\gamma A_1)}{A_1+(1-\alpha)E[\tilde{m}_{z''_\gamma}(U(1))](1-2c\gamma A_1)}}\Big] \\ &= E\Big[\lim \frac{V_{\mathrm{II}}(u_{\mathrm{II}}^*)/L+E[\tilde{m}_{z''_\gamma}(U(1))]\frac{\alpha A_1(1-2c\gamma A_1)}{A_1+(1-\alpha)E[\tilde{m}_{z''_\gamma}(U(1))](1-2c\gamma A_1)}}{V_{\mathrm{II}}(u_{\mathrm{II}}^*)/L+W_{\mathrm{II}}(u_{\mathrm{II}}^*)/L+A_1+E[\tilde{m}_{z''_\gamma}(U(1))]\frac{\alpha A_1(1-2c\gamma A_1)}{A_1+(1-\alpha)E[\tilde{m}_{z''_\gamma}(U(1))](1-2c\gamma A_1)}}\Big] \le \alpha,\end{aligned}$$

where the last line holds because by (S6.16),

$$E[\lim \frac{V_{\mathrm{II}}(u^*_{\mathrm{II}})/L}{V_{\mathrm{II}}(u^*_{\mathrm{II}})/L + A_2}] \le \alpha,$$

and on the other hand,

$$\frac{E[\tilde{m}_{z''_\gamma}(U(1))]\frac{\alpha A_1(1-2c\gamma A_1)}{A_1+(1-\alpha)E[\tilde{m}_{z''_\gamma}(U(1))](1-2c\gamma A_1)}}{A_1 + E[\tilde{m}_{z''_\gamma}(U(1))]\frac{\alpha A_1(1-2c\gamma A_1)}{A_1+(1-\alpha)E[\tilde{m}_{z''_\gamma}(U(1))](1-2c\gamma A_1)}}$$
$$= \alpha\frac{(1-2c\gamma A_1)E[\tilde{m}_{z''_\gamma}(U(1))]}{A_1 + (1-2c\gamma A_1)E[\tilde{m}_{z''_\gamma}(U(1))]} \le \alpha.$$

Here we have used an evident fact that $(a_1 + a_2)/(b_1 + b_2) \le \alpha$ if $a_1/b_1 \le \alpha$, $a_2/b_2 \le \alpha$ and $b_1 > 0, b_2 > 0$. □

S7.2 Power consistency

Proof of Theorem 3 (power consistency). Recall that if v_j is a Type II break, then $\mathrm{Power}_{\mathrm{II},j}(u_2) = P(\#\{t \in \tilde{T}_{\mathrm{II}}(u_2) \cap S_j\} \ge 1)$; and if v_j is a Type I break, then $\mathrm{Power}_{\mathrm{I},j}(u_1) = P(\#\{t \in \tilde{T}_{\mathrm{I}}(u_1) \cap S_j\} \ge 1)$. In the proofs of Theorems 1 and 2 it is shown that, for each j, both $\mathrm{Power}_{\mathrm{I},j}(u_1)$ and $\mathrm{Power}_{\mathrm{II},j}(u_2)$ tend to 1 as $L \to \infty$. Therefore, we obtain

$$\mathrm{Power}_{\mathrm{III}}(u_1, u_2) = \frac{1}{J}\sum_{j=1}^{J}\left[\mathrm{Power}_{\mathrm{I},j}(u_1)\mathbb{1}(v_j \text{ is Type I }) + \mathrm{Power}_{\mathrm{II},j}(u_2)\mathbb{1}(v_j \text{ is Type II })\right] \to 1.$$

Similarly to the proofs of power consistency in Theorems 1 and 2, for the random thresholds $\tilde{u}_{\mathrm{I}}$ and $\tilde{u}_{\mathrm{II}}$, we have that if v_j is a Type I break, then $\mathrm{Power}_{\mathrm{I},j}(\tilde{u}_{\mathrm{I}}) \to$

1; and if v_j is a Type II break, then $\text{Power}_{\text{II},j}(\tilde{u}_{\text{II}}) \to 1$. Thus,

$$\text{Power}_{\text{III,BH}} = \text{Power}_{\text{III}}(\tilde{u}_{\text{I}}, \tilde{u}_{\text{II}}) \to 1.$$

□

S8. Extra Simulation Studies

S8.1 FDR and Power versus Jump Size

To examine the robustness of the proposed mSTEM procedure under moderate or fixed jump sizes, we conducted additional simulations in which the jump magnitude a varied from 0.8 to 2.0, while all other parameters were fixed at $L = 1500$, $\gamma = b = 10$, and $\alpha = 0.05$. This setting directly targets the reviewer's concern regarding Condition (C2), which requires $a \to \infty$ asymptotically. The empirical False Discovery Rate (FDR) and power were computed from 1000 replications under the Type II change-point scenario.

The results are presented in Figure 3. Overall, the empirical FDR remains well controlled across all jump sizes, even when a is close to 1, indicating that the method is stable under moderate signal strengths. The empirical power increases with a, reaching nearly 1 when $a \geq 1.3$. These findings demonstrate that the asymptotic requirement in Condition (C2) is not restrictive in practice: mSTEM maintains valid FDR control and high detection power even when the jump size is moderate and does not diverge.

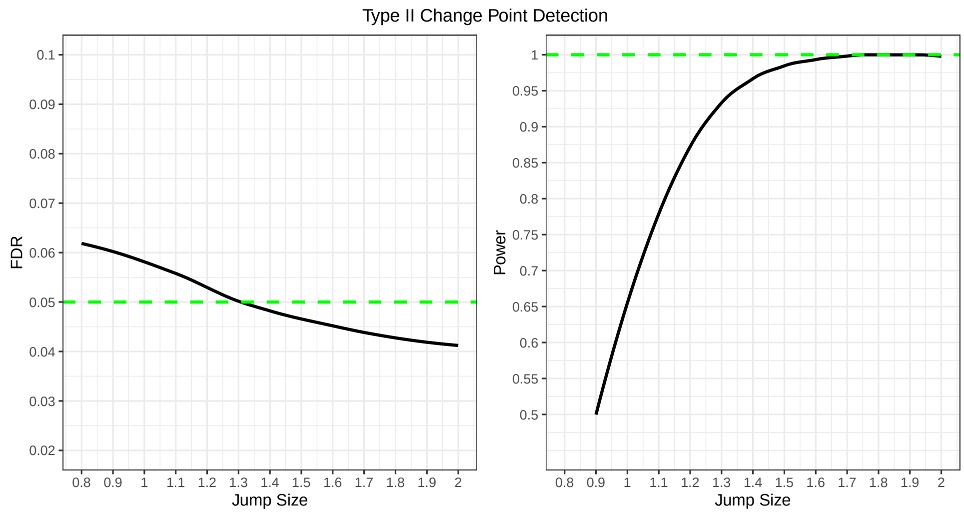

Figure 3: **Empirical FDR and power versus jump size.**

S8.2 FDR and Power versus Slope Change

To assess the robustness of the proposed method when Condition (C1) is relaxed, we generated signals with Type I change points where the size of the slope change varies from 0.1 to 1.0. For each configuration, the empirical FDR and power were computed over 1000 replications. The results are displayed in Figure 4.

Figure 4 shows that the empirical FDR remains below the nominal 0.05 level across all slope change magnitudes considered. Even when the slope change is as small as 0.3, the method continues to maintain valid FDR control. Meanwhile, the empirical power rapidly increases with the slope magnitude and reaches essentially 1 for slope changes above 0.3. These results indicate that the proposed mSTEM procedure is robust to small slope changes and performs reliably even when the theoretical spacing requirement in Condition (C1) may not be strictly satisfied for finite samples.

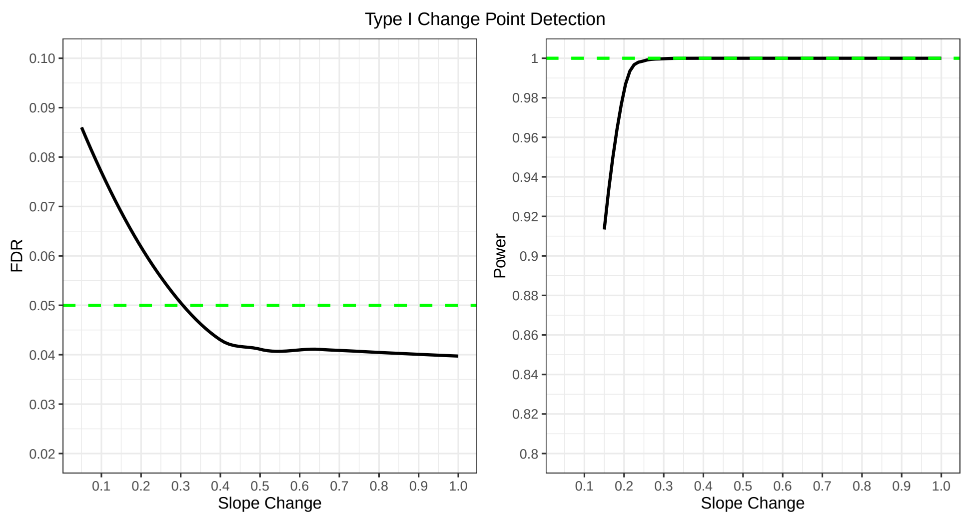

Figure 4: FDR and power versus slope change.

S8.3 Sensitivity of FDR and Power to the Bandwidth Parameter γ

We evaluate the robustness of the proposed mSTEM procedure with respect to the kernel bandwidth γ. The bandwidth determines the degree of local smoothing and affects both the effective signal-to-noise ratio and the separation between neighboring change points. We vary γ from 3 to 18 and examine empirical FDR and power across four settings: Type I change points, Type II (piecewise constant), Type II (piecewise linear), and the mixed Type I/II scenario. All results are based on 1000 replications with $L = 1500$, $b = 10$, and $\alpha = 0.05$.

Figure 5 shows that the proposed method is not robust to extreme bandwidth choices. When γ is too small, insufficient smoothing leads to a lower signal-to-noise ratio, resulting in reduced detection power. Conversely, when γ becomes too large, the smoothing windows overlap, and the FDR begins to increase while the power decreases due to a loss of local contrast. Across all scenarios, a moderate range of γ (approximately 8 – 12) achieves stable FDR control and near-perfect power; therefore, it provides a practical recommendation for real-data analysis.

S8.4 Robustness to Closely Spaced Change Points

To investigate the effect of closely spaced change points, we conducted additional experiments with minimal spacing reduced to $d = 50$, while keeping all other parameters the same as in Section 4.2. The noise was generated from a stationary ergodic Gaussian process with zero mean and variance $\nu = 1$, thereby introducing temporal correlation.

Figure 6 reports the empirical FDR and power across a range of bandwidth values $\gamma \in \{3, \ldots, 18\}$ for Type I, Type II (piecewise constant and piecewise

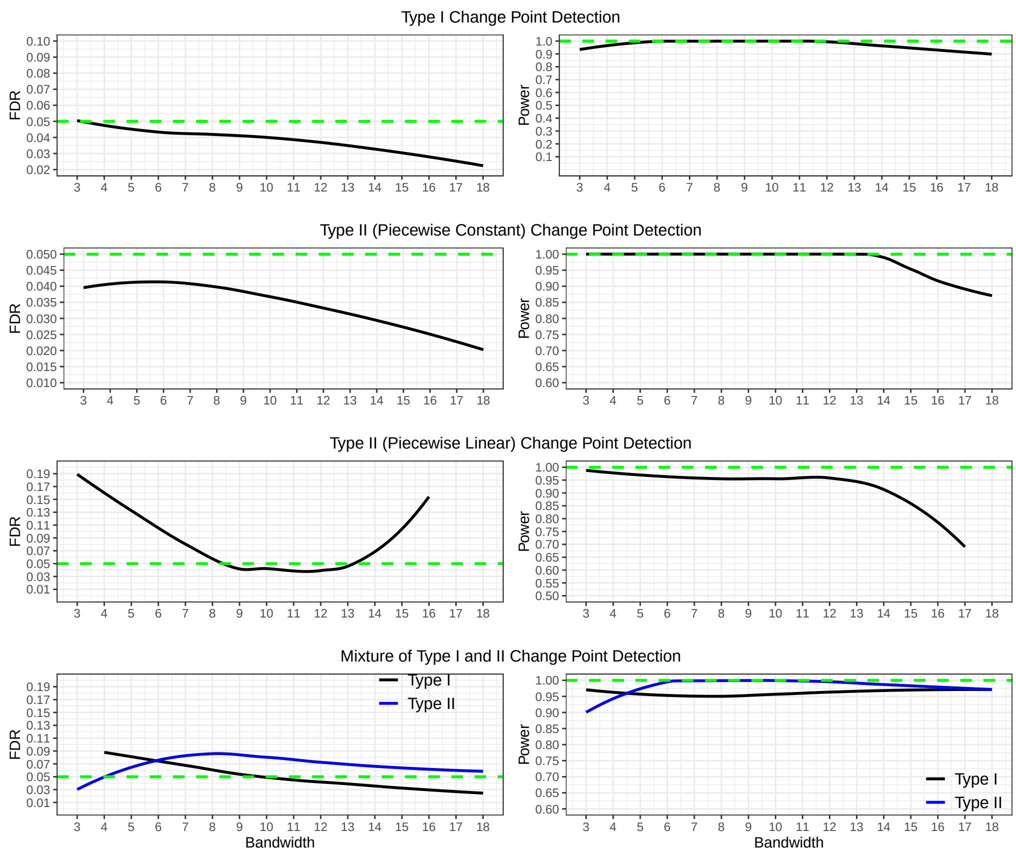

Figure 5: **Empirical FDR and power versus bandwidth γ.** Results are presented for four signal scenarios: (i) Type I change points, (ii) Type II (piecewise constant), (iii) Type II (piecewise linear), and (iv) the mixture of Type I and Type II change points. Each setting uses $L = 1500$, $b = 10$, and $\alpha = 0.05$, with averages taken over 1000 replications.

linear), and mixed scenarios. When the spacing d becomes smaller, a moderately smaller bandwidth is preferred, as a large γ would cause overlapping of smoothed neighborhoods and reduce detection accuracy. Consistent with this intuition, γ between 6 and 10 yields the best performance for $d = 50$, whereas the optimal range shifts to $\gamma = 8$–12 when $d = 150$ (Figure 5).

Importantly, the empirical FDR remains below the nominal level 0.05 for all settings, and the detection power stays above 0.85 even when structural changes are closely spaced. These results demonstrate that the proposed mSTEM method is robust to moderate violations of the spacing condition and performs well under correlated noise.

S8.5 Comparison with Baseline Methods under the Mixed Type I/II Scenario

To further evaluate performance, we compared the proposed mSTEM method with two representative change-point detection procedures: the Narrowest-Over-Threshold (NOT) and the Narrowest Significance Pursuit (NSP).

Table 1 shows the results in the mixed Type I/II scenario based on 1000 replications with $L = 1500$, $\gamma = 10$, $b = 10$, and $\alpha = 0.05$. Entries show the proportion of estimated change points within specified distance ranges from the truth, along with empirical FDR, power, and classification error (for mSTEM).

S8.5 Comparison with Baseline Methods under the Mixed Type I/II Scenario

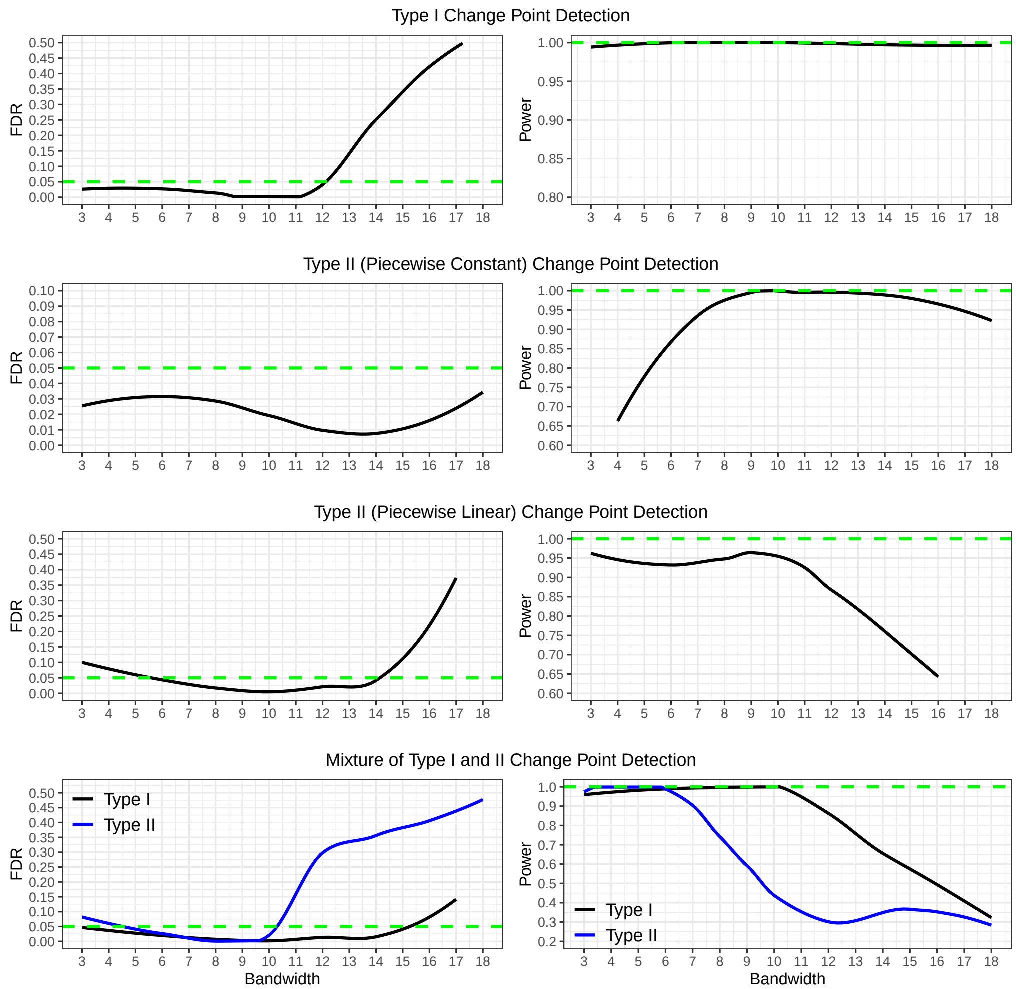

Figure 6: **Empirical FDR and power versus bandwidth γ under closely spaced change points ($d = 50$).**

S8.5 Comparison with Baseline Methods under the Mixed Type I/II Scenario

Table 1: **Empirical performance comparison under the mixed Type I/II change-point scenario.** Results are averaged over 1000 replications with $L = 1500$, $\gamma = 10$, $b = 10$, and $\alpha = 0.05$. The last column reports the empirical classification error for mSTEM.

Method	Type	$[0, \frac{1}{3}\gamma)$	$[\frac{1}{3}\gamma, \gamma)$	$[\gamma, 2\gamma)$	$[2\gamma, 4\gamma)$	$\geq 4\gamma$	FDR	Power	ClassError
mSTEM	Type I	0.8964	0.1083	0.0115	0.0086	0.0614	0.0741	0.9970	0.0575
	Type II	0.9666	0.0024	0.0013	0.0154	0.0349	0.0500	0.9690	
NOT		0.3403	0.1416	0.1136	0.2235	0.1200	0.4887	0.4819	–
NSP		0.1244	0.2099	0.3381	0.3203	0.1109	0.6971	0.3342	–

As shown in Table 1, the proposed mSTEM method achieves substantially lower FDR (0.05–0.07 compared with 0.49–0.70 for NOT and NSP) and markedly higher power (0.97–0.99 versus 0.33–0.48). Moreover, the average classification error for mSTEM is 5.8%, showing the accurate separation of Type I and Type II changes in mixed settings. These results demonstrate that mSTEM substantially outperforms existing methods in both detection accuracy and type identification.

S9. Extra Real Data Analysis

S9.1 Global Temperature Analysis Across Bandwidth

To examine the robustness of the proposed method in the global temperature application, we repeated the analysis over multiple bandwidth values, $\gamma \in \{4, 6, 8, 12\}$. Figure 7 displays the estimated piecewise-linear fits for each bandwidth choice.

Consistent with the simulation studies, the detection performance is sensitive when γ is chosen to be extremely small or extremely large. When $\gamma = 4$, the procedure generates too many small fluctuations and tends to over-segment the signal. Conversely, $\gamma = 12$ leads to excessive smoothing, resulting in the omission of moderate changes, including the 1971 shift often reported in climatology. The results for $\gamma = 6$ and $\gamma = 8$ are highly stable and nearly identical, both capturing the major warming trend and two primary structural shifts. These findings further confirm that moderate bandwidth values (e.g., $6 \leq \gamma \leq 8$) provide a good balance between sensitivity and robustness in real-data settings.

S9.2 Analysis of Stock price

Through analyzing the daily stock price of Host Hotels & Resorts, Inc. (HST) from January 1, 2018, to November 5, 2021, one can provide valuable insights into the company's stock performance and uncover any significant change points

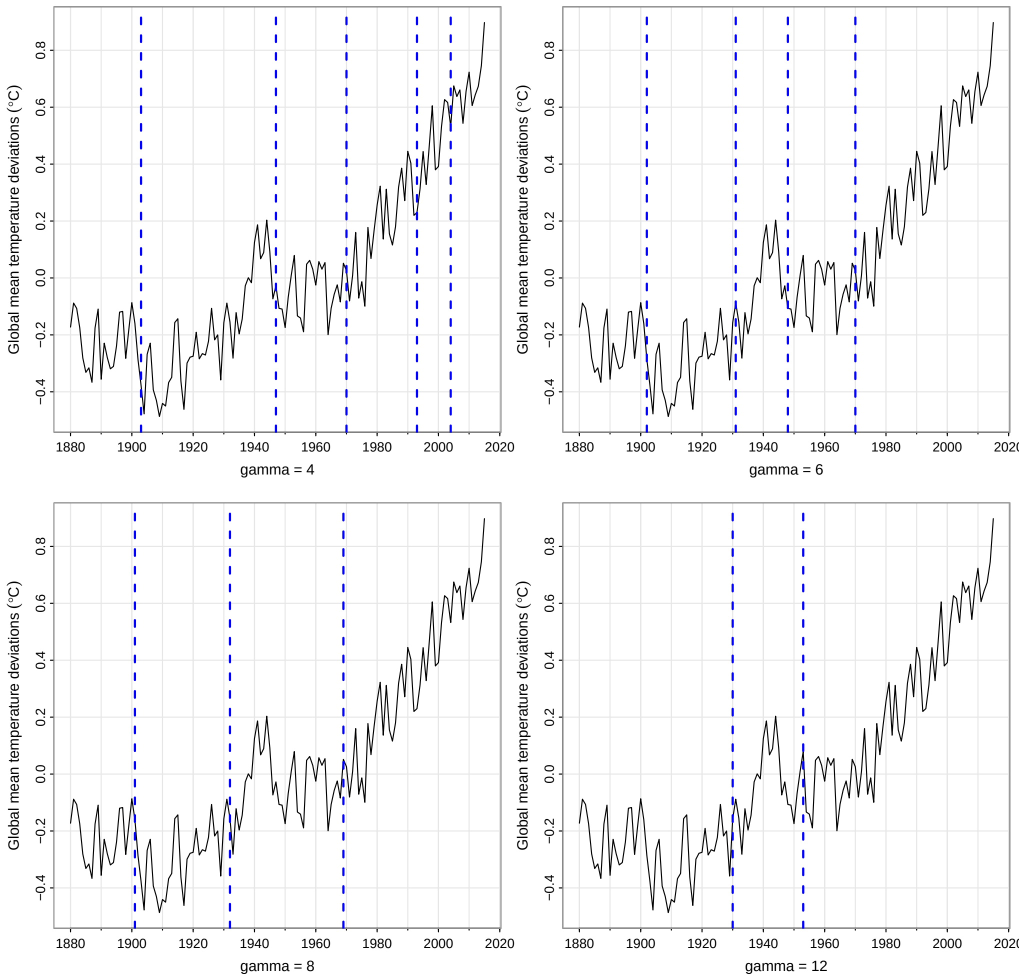

Figure 7: Global temperature analysis under different bandwidth values.

that may have occurred during this period. As the world's largest lodging and real estate investment trust (REIT), HST's stock price can be influenced by various factors, including market conditions, industry trends, company-specific events, and macroeconomic factors. Historical data for the HST stock price is available at Yahoo Finance.

The application of our method to detect the change points in the HST stock price allows for the identification of periods characterized by significant shifts in the stock's behavior or trends. These change points may correspond to specific events or factors that affect HST's stock price, such as earnings releases, mergers and acquisitions, changes in industry regulations, and market-moving news.

Figure 8 shows the results of change point detection for the HST stock price. It is seen that NOT and NSP, particularly NSP, tend to be sensitive to variations in the time series, resulting in the detection of numerous local extrema that can be attributed to noise. This leads to a higher False Discovery Rate (FDR). In contrast, our method provides detection results that are more interpretable. The detected change points align with significant events, such as the outbreak of the trade war between the USA and China in 2018 - 2019, which had a notable impact on the S&P 500 and other large-cap stocks. Similarly, the detected change points after 2020 correspond to the timeline of the Covid-19 outbreak, which significantly affected the global tourism industry.

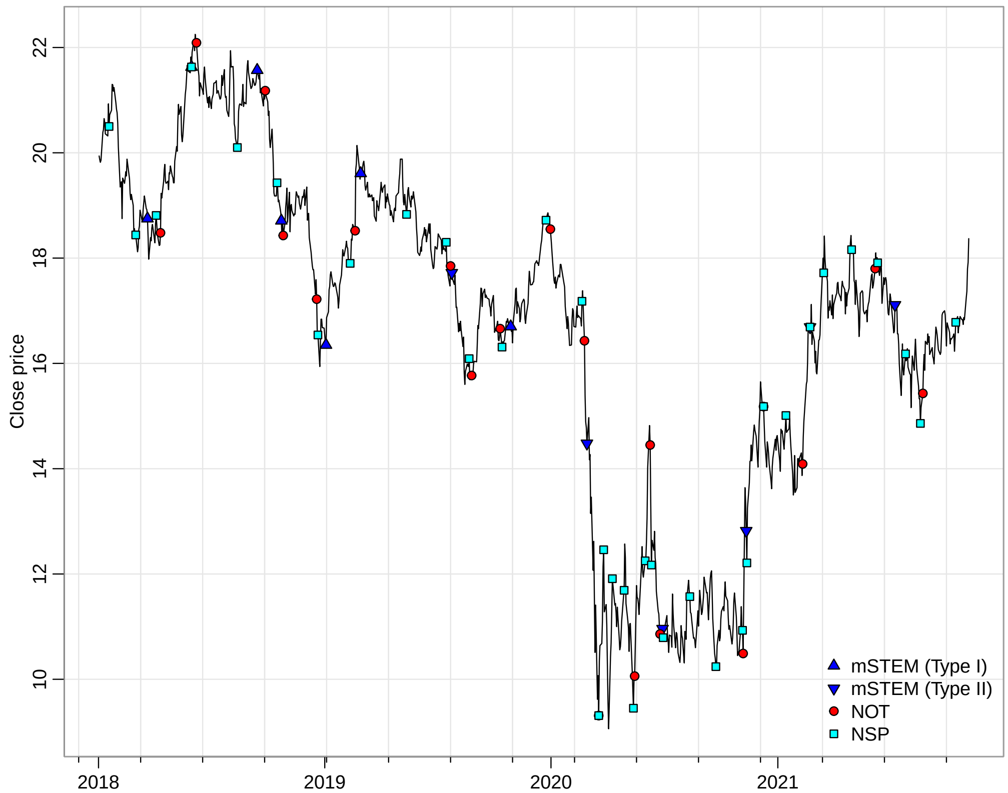

Figure 8: Change point detection of HST stock price.

References

Cheng, D., He, Z. & Schwartzman, A. (2020), ‘Multiple testing of local extrema for detection of change points’, *Electronic Journal of Statistics* **14**(2), 3705–3729.

Schwartzman, A., Gavrilov, Y. & Adler, R. J. (2011), ‘Multiple testing of local maxima for detection of peaks in 1d’, *Annals of statistics* **39**(6), 3290.