Supplementary Material for "Testing and estimation for clustered signals"

Section S1 provides proofs for some results in Sections 2 and 3, while Section S2 presents some additional simulation studies and real data analysis.

S1 Proofs of Main Results

S1.1 Preliminary Lemmas

We shall first provide Lemma S1.1, a probability inequality. The constant in $\lesssim$ of Lemma S1.1 only depends on υ. Lemma S1.1(i) follows from [S9] and it implies (ii), which is a version of Nagaev's inequality [S6]. Recall Condition 2.3 for sub-Gaussian random variables. Lemma S1.1(iii) follows from Ottaviani's inequality for maximal partial sums and the concentration inequality for sums of sub-Gaussian variables. Lemma S1.1(S4) follows from Ottaviani's inequality and the Hanson-Wright inequality (cf. [S8]).

Lemma S1.1. *Let $X_1, \ldots, X_n$ be independent random variables with $E(X_i) = 0$, $E(X_i^2) = \sigma_i^2 < \infty$; let $S_j = \sum_{i=1}^j X_i$. (i) ([S9]) There exists i.i.d. $N(0,1)$ random variables η_i with $G_j = \sum_{i=1}^j \sigma_i \eta_i$ such that for all $t > 0$,*

$$P(\max_{1\le i\le n} |S_i - G_i| \ge t) \lesssim \sum_{i=1}^n \mathcal{M}_\upsilon(X_i/t), \tag{S1}$$

where $\mathcal{M}_v(\cdot)$, $v > 2$, is the truncated moment defined in (2.18). (ii) Let $b_j = E(S_j^2) = \sum_{i=1}^j \sigma_i^2$. Then for all $t > 0$,

$$P(\max_{1\le i\le n} |S_i| \ge t) \lesssim \sum_{i=1}^n \mathcal{M}_v(X_i/t) + \exp(-t^2/(3b_n)). \tag{S2}$$

(iii) Assume that X_i are σ_i^2-sub-Gaussian. Then for all $t > 0$,

$$P(\max_{1\le i\le n} |S_i| \ge 2t) \le 4\exp(-t^2/(2b_n)), \tag{S3}$$

and there exists absolute constants $c_1, c_2, c_3 > 0$ such that

$$P(\max_{1\le i\le n} |\sum_{j=1}^i (X_j^2 - \sigma_j^2)| \ge t) \le c_1[\exp(-c_2 t^2/b_n) + \exp(-c_3 t/\max_{j\le n} \sigma_j)]. \tag{S4}$$

S1.2 Proof of Theorem 2.2

Let $\mathcal{S}_j^\diamond = \{i \in \mathbb{Z} : \tau_j + k \le i \le \tau_{j+1} - 1 - k\}$ and $\mathcal{B}_j^\diamond = \{i \in \mathbb{Z} : \tau_j - k \le i < \tau_j + k\}$. Hence $\mathcal{S}^\diamond = \mathcal{S}_1^\diamond \cup \mathcal{S}_3^\diamond \cup \ldots$ is the interior signal set, $\mathcal{N}^\diamond = \mathcal{S}_0^\diamond \cup \mathcal{S}_2^\diamond \cup \ldots$ is the interior non-signal set and $\mathcal{B}^\diamond = \mathcal{B}_1^\diamond \cup \mathcal{B}_2^\diamond \cup \ldots$ is the boundary set. Under Condition 2.2, we can decompose $\{1, \ldots, p\} = \mathcal{S}^\diamond \cup \mathcal{N}^\diamond \cup \mathcal{B}^\diamond$. Let event

$$\mathcal{A}_1 = \{\max_{k\le i\le p} |L_i - EL_i| \ge \gamma\} = \{\max_{k\le i\le p} |S_i - S_{i-k}| \ge k\gamma\},$$

where $S_i = \sum_{h=1}^i Z_h$. To prove Theorem 2.2(ii), we shall use Lemma S1.1(ii), while proof of Theorem 2.2(i) requires Lemma S1.1(iii). Since the two cases can be dealt with similarly, we only show the arguments for the former. By Lemma S1.1(ii), we have

$$P(\max_{1\le i\le k} |S_i| \ge k\gamma) \lesssim k\mathcal{M}_v(Z_1/(k\gamma)) + e^{-c_3 k\gamma^2/\sigma^2}, \tag{S5}$$

where c_3 and the constant in $\lesssim$ only depend on θ. By (S5) and the Bonferroni inequality,

$$\begin{aligned} P(\mathcal{A}_1) &\leq P(\max_{k\leq i\leq p}|S_i - S_{\lfloor i/k\rfloor k}| \geq \frac{k\gamma}{2}) + P(\max_{k\leq i\leq p}|S_i - S_{\lceil i/k\rceil k}| \geq \frac{k\gamma}{2}) \\ &\lesssim p\mathcal{M}_v(Z_1/(k\gamma)) + pk^{-1}e^{-c_4 k\gamma^2/\sigma^2}, \end{aligned} \tag{S6}$$

where $c_4 = c_3/4$. Under event $\mathcal{A}_1^c$, by Condition 2.2 and since $d > 2\gamma$, we have $Q_i = 0$ if $i \in \mathcal{N}^\diamond$ and $Q_i = 2$ if $i \in \mathcal{S}^\diamond$. Similarly as (S6), we have

$$P(\max_{k\leq i\leq 3k}|S_i - S_{i-k}| \geq k\delta) \lesssim k\mathcal{M}_v(Z_1/(k\delta)) + e^{-c_4 k\delta^2/\sigma^2} =: \varpi. \tag{S7}$$

Hence for event $\mathcal{A}_2 = \{\max_{i\in\mathcal{B}^\diamond}|L_i - EL_i| \geq \delta\}$, we similarly have

$$P(\mathcal{A}_2) \lesssim l\varpi. \tag{S8}$$

Let $m = \lfloor 2k\delta/d\rfloor$. If $\tau_1 - k \leq i \leq \tau_1 - m - 1$, then $E(R_i - R_{\tau_1}) \leq k^{-1}(i-\tau_1)d < -2\delta$. By (S7), we have $P(\max_{k\leq i\leq 3k}|R_i - ER_i| \geq \delta) \lesssim \varpi$ and

$$P\left[\max_{\tau_1-k\leq i\leq\tau_1-m-1} R_i > R_{\tau_1}\right] \lesssim \varpi. \tag{S9}$$

Let $g = \tau_1 + m$. If $g+1 \leq i \leq \tau_1 + k$, then $(i-\tau_1)d > 2k\delta$ and $EL_i > 2\delta$. Again by (S7),

$$P\left[\min_{g+1\leq i\leq\tau_1+k} L_i \leq \delta\right] \lesssim \varpi. \tag{S10}$$

On event $\mathcal{A} = \mathcal{A}_1^c \cap \mathcal{A}_2^c$, we have $\hat{\tau}_1 \in \mathcal{B}_1^\diamond$. Thus by (S9) and (S10),

$$P(\mathcal{A} \cap \{|\hat{\tau}_1 - \tau_1| > m\}) \lesssim \varpi. \tag{S11}$$

Similar inequalities hold for $\hat{\tau}_j - \tau_j$, $2 \le j \le l$. By (S6) and (S8), (2.20) follows from (S11) via the Bonferroni inequality. $\square$

S1.3 Proof of Theorem 2.3

We should prove Theorem 2.3 in the main paper together with the following Theorem S1.1 which concerns polynomial-tailed Z_j.

Theorem S1.1. *Let* $\gamma^\dagger = (k^{-1}\log p)^{1/2} + k^{-1}p^{2/\theta}$ *and* $\delta^\dagger = (k^{-1}\log p)^{1/2} + k^{-1}(lk)^{2/\theta}$. *Assume* $\theta > 4$, *Conditions 2.1, 2.4, and* $\gamma^\dagger = o(d^2)$. *Let* $\gamma = c_1\gamma^\dagger$ *and* $\delta = c_2\delta^\dagger$, *where* c_1 *and* c_2 *are sufficiently large constants. Then there exists a constant* $c > 0$ *independent of* k *and* p *such that*

$$P\left[\hat{l} = l, \max_{j\le l} |\hat{\tau}_j - \tau_j| \le \frac{ck\delta}{d^2}\right] \to 1. \tag{S12}$$

Proof of Theorem S1.1. The argument is similar to the one of Theorem 2.2. Recall the latter proof for $\mathcal{S}^\diamond$, $\mathcal{B}^\diamond$ and $\mathcal{N}^\diamond$. Let $\zeta_j = Z_j^2 - \sigma^2$ and $H_j = \sum_{i=1}^j \zeta_i$. Let $\upsilon > \theta/2$ and recall (2.15) for κ. Similarly as (S5),

$$P(\max_{1\le i\le k} |H_i| \ge k\gamma) \lesssim k\mathcal{M}_\upsilon(\zeta_1/(k\gamma)) + e^{-c_3k\gamma^2/\kappa^2}, \tag{S13}$$

which by the Bonferroni inequality imply that there exists a constant $c_1 > 0$ such that

$$P(\mathcal{A}_1) \to 0, \text{ where } \mathcal{A}_1 = \left\{\max_{i\in\mathcal{N}^\diamond} |L_i^\dagger| \vee |R_i^\dagger| \ge c_1\gamma^\dagger\right\}. \tag{S14}$$

Hence under $\mathcal{A}_1^c$, we have $Q_i^\dagger = 0$ if $i \in \mathcal{N}^\diamond$. Now let $i \in \mathcal{S}^\diamond$. Then for j with $i \le j \le i+k-1$, $|\mu_j| \ge d$. By Lemma S1.1(ii), let $t = \sum_{j=i}^{i+k-1} \mu_j^2$,

$$P(\sum_{j=i}^{i+k-1} (\mu_j^2 + 4\mu_j Z_j) < 0) \lesssim \sum_{j=i}^{i+k-1} \mathcal{M}_v(\mu_j Z_j/t) + \exp(-c_4 t). \tag{S15}$$

By Conditions 2.3, since $i \in \mathcal{S}^\diamond$ and $\gamma^\dagger = o(d^2)$, $t \ge kd^2 \to \infty$. For $i \le j \le i+k-1$, since $2t \ge t + \mu_j^2 \ge 2t^{1/2}|\mu_j|$, we have $\varpi := \min_{i \le j \le i+k-1} t/|\mu_j| \ge t^{1/2}$. $h := \sup_{y \ge \varpi} y^\theta \mathcal{M}_v(Z_1/y) \to 0$. Hence

$$\sum_{j=i}^{i+k-1} \mathcal{M}_v(\mu_j Z_j/t) \le \sum_{j=i}^{i+k-1} \frac{|\mu_j|^\theta h}{t^\theta} \le \frac{h}{t^{\theta/2}} \le \frac{h}{(kd^2)^{\theta/2}}. \tag{S16}$$

By (S15) and the Bonferroni inequality, since $\gamma = o(d^2)$, we have

$$P(\min_{i \in \mathcal{S}^\diamond} \sum_{j=i}^{i+k-1} (\mu_j^2 + 4\mu_j Z_j) < 0) \le p[\frac{h}{(kd^2)^{\theta/2}} + \exp(-c_4 kd^2)] \to 0 \tag{S17}$$

by the definition of $\gamma^\dagger$. Again by $\gamma = o(d^2)$, $X_i^2 - \sigma^2 = \mu_i^2/2 + (\mu_i^2/2 + 2\mu_i Z_i) + Z_i^2 - \sigma^2$, (S13) and (S17), we have

$$P(\mathcal{A}_2) \to 1, \text{ where } \mathcal{A}_2 = \left\{ \min_{i \in \mathcal{S}^\diamond} L_i - \sigma^2 > \gamma \right\}. \tag{S18}$$

Under $\mathcal{A}_2$, if $i \in \mathcal{S}^\diamond$, we have $Q_i^\dagger = 2$. Let $\mathcal{A} = \mathcal{A}_1^c \cap \mathcal{A}_2$.

Let $m = \lfloor ck\delta/d^2 \rfloor$. For $\tau_1 - k \le i \le \tau_1 - m$, we have

$$k(R_i^\dagger - R_{\tau_1}^\dagger) = \sum_{f=i+1}^{\tau_1} (X_f^2 - X_{k+f}^2) = Q_i - \pi_i - U_i, \tag{S19}$$

where $Q_i = \sum_{f=i+1}^{\tau_1} (Z_f^2 - Z_{k+f}^2)$, $U_i = \sum_{f=i+1}^{\tau_1} (\mu_{k+f}^2/2 + 2\mu_{k+f} Z_{k+f})$ and $\pi_i = \sum_{f=i+1}^{\tau_1} \mu_{k+f}^2/2$.

By (S13), there exists a constant $c_5 > 0$ such that

$$P\left[\max_{\tau_1-k\le i\le\tau_1-1}|Q_i|\ge k\delta\right]\lesssim k\mathcal{M}_v(\zeta_1/(k\delta))+e^{-c_3k\delta^2/\kappa^2}=o(l^{-1}) \tag{S20}$$

by letting $\delta = c_2\delta^\dagger$, where $c_2 > 0$ is a sufficiently large constant. To see the second relation in (S20), we use the fact that $y^{\theta/2}\mathcal{M}_v(\zeta_1/y)\to 0$ as $y\to\infty$. We shall now deal with the term U_i. Similarly as (S15), let $t_i=\sum_{f=i+1}^{\tau_1}\mu_{k+f}^2$, $\tau_1-k\le i\le\tau_1-m$. Then $\min_{i+1\le f\le\tau_1} t_i/|\mu_{k+f}|\ge t_i^{1/2}\ge m^{1/2}d$. Let $h_1:=\sup_{y\ge m^{1/2}d} y^\theta\mathcal{M}_v(Z_1/y)\to 0$. Similarly as (S17),

$$\begin{aligned}P(\min_{\tau_1-k\le i\le\tau_1-m}U_i<0) &\le \sum_{i=\tau_1-k}^{\tau_1-m}[\frac{h_1}{((\tau_1-i)d^2)^{\theta/2}}+\exp(-c_4(\tau_1-i)d^2)]\\ &\lesssim \frac{h_1}{m^{\theta/2-1}d^\theta}+\exp(-c_4md^2)=o(l^{-1}).\end{aligned} \tag{S21}$$

The last relation in (S21) is due to the fact that $l=o(\gamma^\dagger(k\delta^\dagger)^{\theta/2-1})$ and $\gamma^\dagger=o(d^2)$. By (S19), (S20) and (S21), we obtain

$$P[\max_{\tau_1-k\le i\le\tau_1-m-1}k(R_i^\dagger-R_{\tau_1}^\dagger)\ge 0]=o(l^{-1}). \tag{S22}$$

For $L_i^\dagger$ with $\tau_1+m+1\le i\le\tau_1+k$, we similarly have

$$P[\min_{\tau_1+m+1\le i\le\tau_1+k}L_i\le\delta]=o(l^{-1}). \tag{S23}$$

By (S22) and (S23), $P(\mathcal{A}\cap\{|\hat\tau_1-\tau_1|\le m\})=o(l^{-1})$. The latter inequality also holds for $\hat\tau_f-\tau_f$ with $2\le f\le l$. Hence (S12) follows from (S14) and (S18). □

Proof of Theorem 2.3 in the main paper. Theorem 2.3 assumes sub-Gaussian Z_j. We can use the arguments in the proof of Theorem S1.1, with the following modifications: by Lemma S1.1(iii), the term $k\mathcal{M}_v(\zeta_1/(k\gamma))$ in (S13) can be replaced by $\exp(-c_4k\gamma)$; the term $\mathcal{M}_v(\mu_jZ_j/t)$ in (S15) vanishes; the term $k\mathcal{M}_v(\zeta_1/(k\delta))$ in (S20) can be replaced by $\exp(-c_5k\delta)$, and the polynomial term $h_1/(m^{\theta/2-1}d^\theta)$ in (S21) vanishes. Then Theorem 2.3 follows from the arguments in the proof of Theorem S1.1.

S1.4 Proof of Theorem 2.4

Let $p' = p - m + 1$ and define the empirical distribution function

$$\hat{G}(v) = \frac{1}{p'}\sum_{i=1}^{p'} \mathbf{1}_{\hat{\sigma}_i^2 \le v}. \tag{S24}$$

Then $\hat{G}(\hat{\sigma}^2_{(k)}) = k/p'$. Let $G(v) = E\hat{G}(v)$. Since $\hat{\sigma}_i^2$ are m-dependent, we have $\mathrm{cov}(\mathbf{1}_{\hat{\sigma}_i^2 \le v}, \mathbf{1}_{\hat{\sigma}_j^2 \le v}) = 0$ if $|j-i| \ge m$. Hence we have

$$\hat{G}(v) - G(v) \to 0 \text{ in probability} \tag{S25}$$

in view of $m/p' \to 0$ and

$$E[\hat{G}(v) - G(v)]^2 = \frac{1}{p'^2}\sum_{i,j=1}^{p'} \mathrm{cov}(\mathbf{1}_{\hat{\sigma}_i^2 \le v}, \mathbf{1}_{\hat{\sigma}_j^2 \le v}) \le \frac{2m}{p'}. \tag{S26}$$

Let $W_i = \sum_{j=1}^{i}(Z_j^2 - \sigma^2)$, $y = m\gamma_p$ and $v > \theta/2$. By Lemma S1.1(ii),

$$P(\max_{i \le m} |W_i| \ge y) \lesssim m\mathcal{M}_v(\frac{Z_1^2 - \sigma^2}{y}) + \exp(-C_3 m\gamma_p^2) = o(\frac{m}{p}), \tag{S27}$$

where constant C_3 only depends on θ and κ. The second relation in (S27) holds in view of (2.22) and $y \to \infty$. Let $V_1 = m^{-1}\sum_{i=1}^{m} Z_i^2$. Note that, if $i < \tau_1 - m$ or $\tau_l \le i < \tau_{l+1} - m$ for even l, $\hat{\sigma}_i^2$ and V_1 have the same distribution. Hence, by Assumption 1, for sufficiently large p, we have $G(\sigma^2 + \gamma_p) > 1/2$, which implies

$$P(\hat{\sigma}^2_{(k/2)} \le \sigma^2 + \gamma_p) \to 1. \tag{S28}$$

Therefore, to prove (2.23), by (S25), it remained to show that

$$P(\hat{\sigma}^2_{(1)} \geq \sigma^2 - c\gamma_p) = P(\min_{i\leq p'} \hat{\sigma}^2_i \geq \sigma^2 - c\gamma_p) \to 1 \tag{S29}$$

holds for some constant $c > 0$. By (S27), and the Bonferroni inequality, since $\max_{i\leq m} |\sum_{j=i}^{i+m-1} Z_j^2 - m\sigma^2| \leq 2\max_{i\leq 2m} |W_i|$, we have

$$P(\max_{i\leq p'} |m^{-1} \sum_{j=i}^{i+m-1} Z_j^2 - \sigma^2| \geq c_\theta \gamma_p) \to 0, \tag{S30}$$

where the constant c_θ only depends on θ. Since $X_i^2 = \mu_i^2 + 2\mu_i Z_i + Z_i^2$, to prove (S29), it remains to verify that

$$P(\min_{i\leq p'} \sum_{j=i}^{i+m-1} (\mu_j^2 + \gamma_p + 2\mu_j Z_j) < 0) \to 0. \tag{S31}$$

As in (S15), (S16), (S17), let $T = \sum_{j=i}^{i+m-1} \mu_j Z_j$, $\varrho_l = \sum_{j=i}^{i+m-1} |\mu_j|^l$, $u = \sum_{j=i}^{i+m-1}(\mu_j^2 + \gamma_p) = \varrho_2 + m\gamma_p$. By Lemma S1.1(ii), there exists a positive constant C_4, only depending on θ such that

$$P(2T \leq -u) \leq \sum_{j=i}^{i+m-1} \mathcal{M}_v(\mu_j Z_j/u) + \exp(-C_3 \frac{u^2}{\varrho_2}), \tag{S32}$$

where $\varphi = \min_j u/|\mu_j|$. Since $(\mu_j^2 + m\gamma_p)/|\mu_j| \geq 2(m\gamma_p)^{1/2}$, we have $\varphi \geq 2p^{1/q} \to \infty$. Also $2u^q \geq \varrho_2^q + (m\gamma_p)^q \geq \varrho_q^2 + (m\gamma_p)^q \geq 2\varrho_q p$ and $u^2 \geq 2\varrho_2 p^{2/q}$. So $(u/\mu_j)^\theta \mathcal{M}_v(\mu_j Z_j/u) \to 0$. Hence by (S32) we obtain $P(2T \leq -u) = o(p^{-1})$ holds uniformly in i. Thus (S31) follows.

To deal with the sub-Gaussian case with Condition 2.3, instead of using (S27), the Nagaev inequality Lemma S1.1(ii), we can apply the Bernstein inequality

$$P(\max_{i\leq m} |W_i| \geq y) \lesssim \exp(-C_2 y) + \exp(-C_3 y^2/m). \tag{S33}$$

With elementary calculations, we can similarly show that $\hat{\sigma}^2_{(k)} = \sigma^2 + O_P(\gamma_p)$ with $y = m\gamma_p$ and the new $\gamma_p = (m^{-1}\log p)^{1/2}$. Details are omitted. $\square$

S1.4.1 Proof of Corollary 2.2

Corollary 2.2 can be proved by using the same arguments for Theorem 2.4. Let $U_1 = m^{-1}\sum_{i=1}^m (Z_i - Z_{i-1})^4$. Note that the differences $Z_i - Z_{i-1}$ are 1-dependent. We can have a concentration inequality for $P(|U_1 - \nu| \geq \phi_p)$, by considering separately even and odd indices i. Then we apply the following version of Bernstein's inequality: if Y_i are i.i.d. with mean 0, variance σ^2 and $E(\exp(t|Y_i|)) < \infty$ for some $t > 0$, then $P(\max_{j\leq n}|\sum_{i=1}^j (Y_i^2 - \sigma^2)| \geq u) \leq c_1\exp(-c_2\sqrt{u}) + c_1\exp(-c_3u^2/n)$ for some constants $c_1, c_2, c_3 > 0$. Other arguments are similar so we omit the details.

S1.5 Proof of Theorem 3.1

We shall first prove (3.6). For $1 \leq g \leq np$, let $g = n(a-1) + i$, where $i = i_g, a = a_g \in \mathbb{N}$ and $1 \leq i \leq n$. Write $S_g = \sum_{b=1}^{a-1}\sum_{j=1}^n Y_{jb} + \sum_{j=1}^i Y_{ja}$. Let $v_g = \mathrm{var}(S_g) = n(\sigma_1^2 + \ldots + \sigma_{a-1}^2) + i\sigma_a^2$. Note that $S_{an} - S_{bn} = n(\hat{\mu}_{b+1} + \ldots + \hat{\mu}_a)$. Since Y_{ij} are independent with mean 0, under Condition 3.1, by Lemma S1.1(i) with $\upsilon = \theta$, we have

$$P\left[\max_{g\leq np}|S_g - IB(v_g)| \geq c_1 u\right] \leq \sum_{a=1}^p\sum_{i=1}^n E|\frac{Y_{ia}}{u}|^\theta \leq npu^{-\theta}K_\theta^\theta,$$

where the constant c_1 only depends on θ. So (3.6) follows.

We now prove (3.7). Let the index set $\mathcal{A} = \{(i,a) : 1 \leq i \leq n, 1 \leq a \leq p\}$. For $\alpha = (i,a) \in \mathcal{A}$, let $X_{\alpha j} = \psi_{\alpha j}Y_\alpha$, $0 \leq j \leq p-k$, where $\psi_{\alpha j} = \mathbf{1}_{j+1\leq a\leq j+k}(nv_j)^{-1/2}$. Let vector

$\psi_\alpha = (\psi_{\alpha,0}, \ldots, \psi_{\alpha,p-k})^T$ and $\psi_{\alpha *} = \max_{0\le j\le p-k} \psi_{\alpha j}$. Observe that

$$\max_{0\le j\le p-k} R_j^\star = \max_{0\le j\le p-k} \sum_{\alpha\in\mathcal{A}} Y_\alpha \psi_{\alpha j} \tag{S34}$$

and $Y_\alpha, \alpha \in \mathcal{A}$, are independent random variables. We now apply Theorem 2.1 in [S2]. To this end, note that $\psi_{\alpha j} \le (nk)^{-1/2}/\sigma_*$ and $\psi_{\alpha j} = 0$ if $j \ge a$ or $a > j + k$,

$$L := \max_{0\le j\le p-k} \sum_{\alpha\in\mathcal{A}} E|Y_\alpha \psi_\alpha|^3 \le knK_3^3((nk)^{1/2}\sigma_*)^{-3}. \tag{S35}$$

Using the inequality $E(|Z|^3 \mathbf{1}_{|Z|\ge A}) \le A^{3-\theta} E(|Z|^\theta)$ since $\theta > 3$, we have

$$M(\phi) := \sum_{\alpha\in\mathcal{A}} E|Y_\alpha \psi_{\alpha *}|^3 \mathbf{1}_{|Y_\alpha \psi_{\alpha *}|\ge (4\phi \log(p-k+1))^{-1}} \le \frac{pn(4\phi\log p)^{\theta-3} K_\theta^\theta}{((nk)^{1/2}\sigma_*)^\theta} =: M^\circ(\phi). \tag{S36}$$

Note that $E(R_j^2) = 1$ for all $0 \le j \le p-k$. By Theorem 2.1 in [S2], there exists an absolute constant $c > 0$ such that

$$\rho^\star \le c \min_{\ell \ge L} [(\ell^2 \log^7 p)^{1/6} + \ell^{-1} M((\ell^2 \log^4 p)^{-1/6})]. \tag{S37}$$

Choose $\ell = \ell^\circ$ such that $(\ell^2 \log^7 p)^{1/6} = \ell^{-1} M^\circ((\ell^2 \log^4 p)^{-1/6})$. Then (3.7) follows from elementary manipulations by considering two cases $\ell^\circ > L$ and $\ell^\circ \le L$ separately. In the latter case, the right hand size of (S37) is minimized at $\ell = L$. □

S1.6 Proof of Proposition 3.1

We first consider the case $\theta > 4$. Assume that $\mu_j = 0$ for all $j \le p$. Let $\bar{Y}_{\cdot j} = n^{-1} \sum_{i=1}^n Y_{ij}$,

$$S_j = \sum_{l=1}^j W_j, \text{ where } W_j = \sum_{i=1}^n (Y_{ij} - \bar{Y}_{\cdot j})^2 - (n-1)\sigma_j^2.$$

By Burkholder's inequality, $\|\bar{Y}_{\cdot j}\|_\theta \lesssim n^{-1/2}\|Y_{ij}\|_\theta$ and $\|\sum_{i=1}^n Y_{ij}^2 - n\sigma_j^2\|_{\theta/2} \lesssim n^{1/2}\|Y_{ij}\|_\theta^2$. By Condition 3.1,

$$E(|W_j|^{\theta/2}) \lesssim nK_\theta^\theta \text{ and } E(|W_j|^2) \lesssim nK_4^4$$

Since Y_{ij} are independent, by Lemma S1.1(ii), for $u > 0$,

$$P(\max_{j\le k}|S_j| \ge u) \lesssim \frac{nkK_\theta^\theta}{u^{\theta/2}} + \exp(-c_3\frac{u^2}{nkK_2^2}), \tag{S38}$$

where c_3 only depends on θ. Define the oscillation

$$\Upsilon = \Upsilon(k) = \max_{1\le j,h\le p:\ |j-h|\le k} |S_j - S_h|. \tag{S39}$$

By the triangle inequality, we have

$$\begin{aligned} P(\Upsilon \ge u) &\le \sum_{g=0}^{\lfloor p/k \rfloor} P(\max_{1\le j\le k} |S_{j+gk} - S_{gk}| \ge u/3) \\ &\lesssim \frac{p}{k}\frac{nkK_\theta^\theta}{u^{\theta/2}} + \frac{p}{k}\exp(-c_7\frac{u^2}{nkK_2^2}), \end{aligned} \tag{S40}$$

implying (3.11). If $2 < \theta \le 4$, instead of (S38), we have by the Markov and the Burkholder inequalities that

$$P(\max_{j\le k}|S_j| \ge u) \le \frac{E(\max_{j\le k}|S_j|^{\theta/2})}{u^{\theta/2}} \le c_4\frac{nkK_\theta^\theta}{u^{\theta/2}}. \tag{S41}$$

By the argument in (S40), $P(\Upsilon \ge u) \lesssim npK_\theta^\theta u^{-\theta/2}$. So (3.12) follows. $\square$

S1.7 Proof of Theorem 3.2

Recall (S39) for Υ, (3.3) for W_j, and (3.14) for W_j^*. Let E^* be the conditional expectation given $Y = (Y_1, \ldots, Y_n)$. Define coefficients $f_{j,l} = 0$ if $l \le j$ or $l > j+k$, and $f_{j,l} = 1$ if $j < l \le j+k$. Then $W_j^* = \sum_{l=0}^{p-k} f_{j,l}\hat{\sigma}_l \eta_l$, and for all $k \le j, j' \le p-k$, we have

$$|E^*(W_j^* W_{j'}^*) - E(W_j W_{j'})| = |\sum_{l=0}^{p-k} f_{j,l} f_{j',l}(\hat{\sigma}_l^2 - \sigma_l^2)| \le \frac{\Upsilon}{n}. \tag{S42}$$

By Condition 3.1, (S42) and the triangle inequality, there exist constants $c_1, c_2 > 0$ only depending on σ_*, σ^* such that

$$\begin{aligned} |\gamma_{j,j'}(\hat{\sigma}) - \gamma_{j,j'}(\sigma)| &\le |\gamma_{j,j'}(\hat{\sigma}) - \frac{E^*(W_j^* W_{j'}^*)}{v_j^{1/2} v_{j'}^{1/2}}| + \frac{\Upsilon/n}{v_j^{1/2} v_{j'}^{1/2}} \\ &\le |\gamma_{j,j'}(\hat{\sigma})||1 - \frac{\hat{v}_j^{1/2} \hat{v}_{j'}^{1/2}}{v_j^{1/2} v_{j'}^{1/2}}| + c_1 \frac{\Upsilon}{nk} \le c_2 \frac{\Upsilon}{nk}. \end{aligned}$$

Applying inequality (S40) with $u = nkt/c_2$, letting $t = ct_*$ with a sufficiently large constant c, we obtain (3.15) after elementary manipulations in view of Theorem 4.1 in [S2] (see also Theorem 3.1 in [S1]). By Theorem 3.1, under (3.7), $\rho^* = o(1)$, implying that the right hand side of (3.15) is $o(1)$ via elementary manipulations. $\square$

S1.8 Proof of Theorem 3.3

We should prove Theorem 3.3 in the main paper together with the following Theorem S1.2 which concerns polynomial-tailed Z_j.

Theorem S1.2. *Assume Conditions 2.2, 2.5 and 3.1 and $2\sigma^*\gamma \le d\sqrt{nk}$. Let $m = \lfloor 2k^{1/2}\delta\sigma^* n^{-1/2} d^{-1} \rfloor$. Then*

$$1 - P\left[\hat{l} = l, \max_{j \le l} |\hat{\tau}_j - \tau_j| \le m\right] \lesssim \frac{p}{k}\left[\frac{knK_\theta^\theta}{(kn)^{\theta/2}\gamma^\theta} + \exp(-c_1\gamma^2)\right]$$

$$+l[\frac{knK_\theta^\theta}{(kn)^{\theta/2}\delta^\theta} + \exp(-c_2\delta^2)], \tag{S43}$$

where the constant in $\lesssim$ *and* $c_1, c_2 > 0$ *are independent of* k, d, n *and* p.

Proof of Theorem S1.2. We shall use the argument in the proof in Theorem 2.2. Let $S_l = \sum_{j=1}^l \sum_{i=1}^n Z_{ij}$. Similarly as (S5), by Lemma S1.1(ii),

$$P[\max_{1\le l\le k} |S_l| \ge (nk)^{1/2}\sigma_*\gamma] \lesssim \frac{knK_\theta^\theta}{(kn)^{\theta/2}(\sigma_*\gamma)^\theta} + \exp(-c_3(\gamma\sigma_*/\sigma^*)^2). \tag{S44}$$

As (S7), the above inequality implies

$$P(\max_{k\le l\le 3k} |S_l - S_{l-k}| \ge (nk)^{1/2}\sigma_*\delta) \lesssim \frac{knK_\theta^\theta}{(kn)^{\theta/2}\delta^\theta} + \exp(-c_4\delta^2) =: \varpi. \tag{S45}$$

Let event $\mathcal{A}_1 = \{\max_{j\le p-k} |R_j^\star - ER_j^\star| \ge \gamma\}$. By (S44) and Condition 3.1,

$$P(\mathcal{A}_1) \lesssim \frac{p}{k}[\frac{knK_\theta^\theta}{(kn)^{\theta/2}\gamma^\theta} + \exp(-c\gamma^2)]. \tag{S46}$$

Recall the proof of Theorem 2.2 for the definitions of $\mathcal{B}^\diamond$, $\mathcal{N}^\diamond$ and $\mathcal{S}^\diamond$. If $j \in \mathcal{S}^\diamond$, since $2\sigma^*\gamma \le d\sqrt{nk}$, $ER_j^\star > 2\gamma$. Hence under $\mathcal{A}_1^c$, we have $Q_j^\star = 0$ if $j \in \mathcal{N}^\diamond$ and $Q_j^\star = 2$ if $j \in \mathcal{S}^\diamond$. Let event $\mathcal{A}_2 = \{\max_{j\in\mathcal{B}^\diamond} |R_j^\star - ER_j^\star| \ge \delta\}$ and $\mathcal{A} = \mathcal{A}_1^c \cap \mathcal{A}_2^c$. Similarly as (S8), we have

$$P(\mathcal{A}_2) \lesssim l\varpi. \tag{S47}$$

If $\tau_1 + m + 1 \le j \le \tau_1 + k$, then $EL_j^\star > 2\delta$. Similarly as (S10), by (S45),

$$P\left[\min_{\tau_1+m+1\le j\le\tau_1+k} L_j^\star \le \delta\right] \lesssim \varpi. \tag{S48}$$

If $\tau_1 - k \le j \le \tau_1 - m - 1$, then $E(R_j^\flat - R_{\tau_1}^\flat) \le k^{-1/2}(j-\tau_1)n^{1/2}d < -2\delta$. So by (S45)

$$P\left[\max_{\tau_1-k\le j\le \tau_1-m-1} R_j^\flat > R_{\tau_1}^\flat\right] \lesssim \varpi. \tag{S49}$$

Thus by the argument in the proof of Theorem 2.2, (S43) follows. □

Proof of Theorem 3.3. In Theorem 3.3 Z_j are sub-Gaussian. By Lemma S1.1(iii), the terms $knK_\theta^\theta/((kn)^{\theta/2}(\sigma_*\gamma)^\theta)$ in (S44) and $knK_\theta^\theta/((kn)^{\theta/2}\gamma^\theta)$ in (S46) vanish. Then the claim in Theorem 3.3 follows from the arguments in the proof of Theorem S1.2. □

Based on Theorem S1.2, let $\gamma^\flat = (\log p)^{1/2} + (nk)^{-1/2}(np)^{1/\theta}$, $\delta^\flat = (\log p)^{1/2} + (kn)^{1/\theta-1/2}l^{1/\theta}$; let $\gamma = C_1\gamma^\flat$, $\delta = C_2\delta^\flat$, where $C_1, C_2 > 0$ are constants. Then the right hand side of (S43) can be arbitrarily small by letting C_1, C_2 sufficiently large. Theorem S1.2 implies that we can have exact recovery with probability $P[\hat{l} = l, \max_{j\le l}|\hat{\tau}_j - \tau_j| = 0]$ going to 1 if

$$\frac{k^{1/2}}{n^{1/2}}\frac{(\log p)^{1/2} + (kn)^{1/\theta-1/2}l^{1/\theta}}{d} \to 0.$$

S1.9 Proof of Theorem 3.4.

Note that the bound in (3.20) of Theorem 3.4 is the same as (2.10) in Theorem 2.1. By Proposition S1.1, $\|((n^2-n)/2)^{1/2}W_j\|_\theta = O(1)$. With the latter, Theorem 3.4 can be proved along similar lines as the argument in Theorem 3.1. Details are omitted.

Proposition S1.1. *Assume that $U_1, \ldots, U_n$ are i.i.d. with mean 0 and $K_\theta := \|U_i\|_\theta < \infty$, $\theta > 2$. Then*

$$\|\sum_{1\le i<i'\le n} U_i U_{i'}\|_\theta \lesssim nK_2^2 + n^{2/\theta}K_\theta^2, \tag{S50}$$

where the constant in $\lesssim$ only depends on θ.

Proof of Proposition S1.1. Let U_i, U'_j, $i, j \in \mathbb{Z}$, be i.i.d. random variables. By the decoupling equality (see for example [S4]), we have

$$\begin{aligned} \| \sum_{1\le i<i'\le n} U_i U_{i'} \|_\theta &\lesssim \| \sum_{1\le i<i'\le n} U_i U'_{i'} \|_\theta \\ &\le \| \sum_{1\le i,i'\le n} U_i U'_{i'} \|_\theta + \| \sum_{1\le i\le n} U_i U'_i \|_\theta \end{aligned}$$

By the Rosenthal inequality, we have $\| \sum_{1\le i\le n} U_i \|_\theta \lesssim n^{1/2} K_2 + n^{1/\theta} K_\theta$ and $\| \sum_{1\le i\le n} U_i U'_i \|_\theta \lesssim n^{1/2} K_2^2 + n^{1/\theta} K_\theta^2$. Hence Proposition S1.1 follows. □

S1.10 Proof of Theorem 3.5

We should prove Theorem 3.5 in the main paper together with the following Theorem S1.3 which concerns polynomial-tailed Z_j.

Theorem S1.3. *Let $\theta > 2$. Assume either (i) n is bounded, $(|Z_{1j}|^\theta)_{j\ge 1}$ is uniformly integrable or (ii) $n \to \infty$. Let $\gamma^\natural = (\log p)^{1/2} + k^{-1/2} p^{2/\theta}$, $\delta^\natural = (\log p)^{1/2} + (lk)^{2/\theta} k^{-1/2}$; let $\gamma = c_1 \gamma^\natural$, $\delta = c_2 \delta^\natural$, where $c_1, c_2 > 0$ are sufficiently large constants. Assume Conditions 2.4, 3.1 and $\gamma^\natural = o(d^2 n \sqrt{k})$. Then there exists a constant $c > 0$ independent of n, k and p such that*

$$P\left[\hat{l} = l,\ \max_{j\le l} |\hat{\tau}_j - \tau_j| \le \frac{c k^{1/2} \delta^\natural}{n d^2}\right] \to 1. \tag{S51}$$

Proof of Theorem S1.3. Let $\bar{Z}_{\cdot j} = n^{-1} \sum_{i=1}^n Z_{ij}$ and write

$$W_j = \frac{2\xi_j}{n(n-1)} + \mu_j^2 + 2\mu_j \bar{Z}_{\cdot j}, \text{ where } \xi_j = \sum_{1\le i<i'\le n} Z_{ij} Z_{i'j}. \tag{S52}$$

By Proposition 3.2 and Condition 3.1, $E(|\xi_j|^\theta) \lesssim n^\theta (\sigma^*)^{2\theta} + n^2 K_\theta^\theta$. Let $T_h = \sum_{j=1}^h \xi_j$ and $y = c_4 \gamma^\natural n k^{1/2}$, where $c_4 > 0$ is a constant to be determined later. By Lemma S1.1(ii), there exists a constant $c_3 > 0$ such that

$$P(\max_{h\le k} |T_h| \ge y) \lesssim \sum_{j=1}^{k} E(|\xi_j/y|^\theta) + \exp(-\frac{c_3 y^2}{kn^2(\sigma^*)^4}) = O(k/p^2) \tag{S53}$$

by letting $c_4 = 2(\sigma^*)^2 c_3^{-1/2}$. By (S53), similarly as (S14), for some constant $c_1 > 0$,

$$P(\mathcal{A}_1) \to 0, \text{ where } \mathcal{A}_1 = \left\{ \max_{j\in\mathcal{N}^\diamond} |L_{j,4}| \vee |R_{j,4}| \ge c_1\gamma^\natural \right\}. \tag{S54}$$

Let $a \in \mathcal{S}^\diamond$, $a \le j \le a+k-1$ and $t = \sum_{j=a}^{a+k-1} \mu_j^2$. We first consider the case that $n \to \infty$. By Nagaev's inequality [S6] or Lemma S1.1(ii),

$$\begin{aligned} P(\sum_{j=a}^{a+k-1} (\mu_j^2 + 4\mu_j \bar{Z}_{\cdot j}) < 0) &\lesssim \frac{\sum_{j=a}^{a+k-1} n|\mu_j|^\theta K_\theta^\theta}{(\sum_{j=a}^{a+k-1} n|\mu_j|^2)^\theta} + \exp(-c_5 nt) \\ &\lesssim \frac{n^{1-\theta}}{(kd^2)^{\theta/2}} + \exp(-c_5 nkd^2). \end{aligned} \tag{S55}$$

By the condition $\gamma^\natural = o(d^2 n\sqrt{k})$, we have $n^{1-\theta}(kd^2)^{-\theta/2} = O(n^{1-\theta/2}p^{-1})$ and $\exp(-c_5 nkd^2) = o(p^{-2})$. Thus the right hand side of (S55) is of order $o(p^{-1})$ since $n \to \infty$. If n is bounded, using the uniform integrability of $(|Z_{1j}|^\theta)_{j\ge1}$, we shall show that the right hand side of (S55) is also of order $o(p^{-1})$. To this end, in view of (S15) and (S16), it suffices to show that

$$\lim_{u\to\infty} u^\theta \max_{j\ge1} \mathcal{M}_\upsilon(Z_{1j}/u) = 0. \tag{S56}$$

Clearly, $u^\theta \max_{j\ge1} E((Z_{1j}/u)^2 \mathbf{1}_{|Z_{1j}|\ge u}) \le E(|Z_{1j}|^\theta \mathbf{1}_{|Z_{1j}|\ge u}) \to 0$ as $u \to \infty$. Let $\ell = (1-\theta/\upsilon)/2$. Choose $K_u \in \mathbb{N}$ such that $2^{K_u} \le u^{1-\ell} < 2^{K_u+1}$. Then

$$\begin{aligned} E(|Z_{1j}|^\upsilon \mathbf{1}_{|Z_{1j}|\le u}) &\le E(|Z_{1j}|^\upsilon \mathbf{1}_{|Z_{1j}|\le u^\ell}) + \sum_{b=0}^{K_u} E(|Z_{1j}|^\upsilon \mathbf{1}_{2^{-b-1}u<|Z_{1j}|\le 2^{-b}u}) \\ &\le u^{\ell\upsilon} + \sum_{b=0}^{K_u} (2^b u)^{\upsilon-\theta} E(|Z_{1j}|^\theta \mathbf{1}_{2^{-K_u-1}u<|Z_{1j}|}) \end{aligned}$$

Hence $u^\theta \max_{j\geq 1} E((Z_{1j}/u)^v \mathbf{1}_{|Z_{1j}|\leq u}) \to 0$ as $u \to \infty$. Thus (S56) follows. Using the same argument in the proof of Theorem 2.3, we have (S51). The main difference is that ξ_j here has finite θth moment, while ζ_j therein only has finite $(\theta/2)$th moment. Details are omitted. $\square$

Proof of Theorem 3.5. Recall the proof of Theorem S1.3 for ξ_j. Let $Q_j = \xi_j/n$. By the Hanson-Wright inequality (cf. [S8]), there exists absolute constants $c_1, c_2 > 0$ such that $P(|Q_j| \geq u) \leq 2\exp(-c_1 u^2) + 2\exp(-c_2 u)$ for all $u > 0$. Hence Q_j is sub-exponential in the sense that $E\exp(t|Q_j|) < \infty$ for some $t > 0$. By Bernstein's inequality, for some positive constants c_3, c_4, c_5,

$$P(\max_{h\leq k} |\sum_{i=1}^{h} Q_i| \geq u) \leq c_3 \exp(-c_4 u) + c_3 \exp(-c_5 u^2/k). \tag{S57}$$

Following the arguments in the proof of Theorem S1.3, Theorem 3.5 follows by replacing the polynomial term $\sum_{j=1}^{k} E(|\xi_j/y|^\theta)$ in (S53) by the exponential term in (S57) and by removing the term $n^{1-\theta}/(kd^2)^{\theta/2}$ in (S55) in view of Lemma S1.1(iii).

S2 Additional Simulation Studies

In the main paper we presented one-sided test with one realization and compare it with [S10] in Section 4.1 and two-sided test with one realization in Section 4.2. In this Supplementary Material we shall present additional simulation studies. One-sided test with multiple realization is presented in Section S2.1. In Section S2.2, we examine two-sided test with multiple realization and compare it with [S3].

S2.1 Simulation study 3

We now study performance of our testing and estimation procedures based on multiple realizations with heteroscedastic variances for the one-sided test. Consider the model

$$Y_{ij} = \mu_j + \sigma_j \epsilon_{ij}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, p, \tag{S1}$$

where μ_j follows the same configuration as in Table 1, $\sigma_j \sim \mathcal{U}(1, 2), j = 1, \ldots, p$ and $\epsilon_{ij} \sim N(0, 1), t(6)/1.5^{0.5}$ or $\mathrm{LP}(0, 1)/2^{0.5}, i = 1, \ldots, n; j = 1, \ldots, p$. We construct test statistic

$$\hat{R}_j = \frac{\sum_{l=j+1}^{j+k} \sqrt{n}\hat{\mu}_l}{(\sum_{l=j+1}^{j+k} \hat{\sigma}_l^2)^{1/2}}, \quad j = 0, \ldots, p-k,$$

where $\hat{\mu}_l = n^{-1} \sum_{i=1}^{n} Y_{il}$ and $\hat{\sigma}_l^2 = \sum_{i=1}^{n} (Y_{il} - \hat{\mu}_l)^2/(n-1)$ and let

$$R_j^* = \frac{\sum_{l=j+1}^{j+k} \sqrt{n}\hat{\mu}_l}{(\sum_{l=j+1}^{j+k} \sigma_l^2)^{1/2}}, \quad j = 0, \ldots, p-k.$$

Let $p = 600, 2000$ and 6000, sample size $n = 2$ and 5 and error distributions $N(0, 1), t(6)/1.5^{0.5}$ and $\mathrm{LP}(0, 1)/2^{0.5}$. We use $k = \lfloor p^{1/2} \rfloor$. We follow Algorithm 3.1. For $\hat{R}_j$, γ and δ are the 0.95th quantile of $\max_{0 \le j \le p-k} \hat{G}_j$ and $\max_{j \in W_1} \hat{G}_j$ respectively, where $\hat{G}_j = \sum_{l=j+1}^{j+k} \hat{\sigma}_l \eta_l / \hat{v}_j^{1/2}$ and η_l are i.i.d. $N(0, 1)$ that are independent of (Y_{ij}) and $\hat{v}_j = \sum_{l=j+1}^{j+k} \hat{\sigma}_l^2$. For R_j^*, γ and δ are the 0.95th quantile of $\max_{0 \le j \le p-k} G_j^*$ and $\max_{j \in W_1} G_j^*$, where $G_j^* = \sum_{l=j+1}^{j+k} \sigma_l \eta_l / v_j^{1/2}$ and $v_j = \sum_{l=j+1}^{j+k} \sigma_l^2$. From Table S1, we can see that the performance of using estimated variance is comparable as the one using the true variance and the detection accuracy improves with increased sample sizes and number of tests. The combined error rate, FDR, power and the difference between estimated break points and true break points are similar across different error terms.

Table S1: Summary statistics for one-sided test with multiple realization based on $1,000$ simulations. Underscore e (resp. underscore t) is computed based on $\hat{R}_j$ with the estimate $\hat{\sigma}_j^2$ (resp. the true σ_j^2).

p	CER_e	CER_t	FDR_e	FDR_t	Power_e	Power_t	$\hat{l}_e$	$\hat{l}_t$	Diff_e	Diff_t
$n = 2$										
$N(0,1)$										
600	0.0513	0.0513	0.0048	0.0037	0.75	0.74	2	2	5.49	5.43
2000	0.0294	0.0294	0.0019	0.0017	0.85	0.85	2	2	10.07	10.31
6000	0.0148	0.0149	0.0010	0.0009	0.93	0.93	2	2	18.14	17.76
$t(6)/1.5^{0.5}$										
600	0.0517	0.0517	0.0040	0.0040	0.75	0.74	2	2	5.42	5.52
2000	0.0293	0.0293	0.0022	0.0018	0.86	0.85	2	2	10.48	10.03
6000	0.0147	0.0147	0.0017	0.0015	0.93	0.93	2	2	14.84	14.73
$\mathrm{LP}(0,1)/2^{0.5}$										
600	0.0585	0.0584	0.0038	0.0026	0.71	0.70	2	2	5.30	5.34
2000	0.0290	0.0290	0.0022	0.0018	0.86	0.85	2	2	9.12	9.31
6000	0.0138	0.0138	0.0011	0.0011	0.93	0.93	2	2	13.48	13.68
$n = 5$										
$N(0,1)$										
600	0.0257	0.0257	0.0081	0.0083	0.88	0.88	2	2	3.69	3.63
2000	0.0114	0.0114	0.0046	0.0046	0.95	0.95	2	2	6.95	6.93
6000	0.0052	0.0052	0.0031	0.0031	0.98	0.98	2	2	15.66	15.63
$t(6)/1.5^{0.5}$										
600	0.0242	0.0242	0.0086	0.0077	0.89	0.88	2	2	3.58	3.57
2000	0.0117	0.0117	0.0048	0.0044	0.95	0.95	2	2	6.64	6.11
6000	0.0054	0.0054	0.0036	0.0035	0.98	0.98	2	2	15.25	15.21
$\mathrm{LP}(0,1)/2^{0.5}$										
600	0.0283	0.0283	0.0071	0.0067	0.87	0.86	2	2	3.09	3.07
2000	0.0123	0.0122	0.0050	0.0044	0.94	0.94	2	2	7.75	7.72
6000	0.0052	0.0052	0.0033	0.0032	0.98	0.98	2	2	12.86	12.80

S2.2 Simulation study 4

In this section, we study the empirical performance of the two-sided test with multiple realizations and compare it with the classical approach where individual self-normalized t-statistics are used (see for example [S3]). Data generation is the same as the one in simulation study 3 except that $\mu_j, j = 1, \ldots, p$ follow Table 5. We first compute the U-statistic W_j from (3.17). Our test statistic is

$$R^*_{j,4} = \sqrt{n(n-1)/2}\frac{W_{j+1} + \ldots + W_{j+k}}{(\hat{\omega}_{j+1} + \ldots \hat{\omega}_{j+k})^{1/2}},$$

where $\hat{\omega}_j, 0 \leq j \leq p - m$, are defined in (3.21) and are computed by (3.25). We use two ways to approximate the limiting distribution of $\max_{0\leq j\leq p-k} R^*_{j,4}$ by using (3.23) and (3.24), respectively. We compare their empirical performances with different number of tests ($p = 600, 2000$ and 6000), sample size ($n = 4$ and 10) and error terms ($N(0,1), t(6)/1.5^{0.5}$ and $\mathrm{LP}(0,1)/2^{0.5}$). Let $k = \lfloor p^{1/2} \rfloor$. We follow Algorithm 3.2. We compare these approximations and the method in [S3] at significance level 0.05 and summarize the results in Table S2.

Table S2 suggests that the proposed methods have smaller combined error rates and FDR and larger power across all scenarios and the performance improves with increased number of tests and sample size. When n is large, χ^2 approximation is better than normal approximation, as can be seen in Table S2 when $n = 10$. When n is small, their performances are comparable. The performance of the proposed methods improve with increased number of tests while the combined error rate based on [S3] does not change much as number of tests increases. Both approximations based on $G^\star_{j,4}$ and $G^\diamond_{j,4}$ identify 4 break points across different scenarios.

S2.3 Data analysis with one realization

We removed missing data and based our analysis on the resulting $p = 6233$ genes. We used $k = \lfloor p^{1/2} \rfloor = 78$ in the computation of $R^\circ_i, i = k, \ldots, p - k + 1$ and critical values γ and δ and $m = \lfloor p^{1/2} \rfloor = 78$ in the computation of $\hat{\sigma}^2_i, i = 1, \ldots, p - m + 1$. Critical values γ and δ were

Table S2: Summary statistics for two-sided test with multiple realization based on $1,000$ simulations. D represents difference between estimated break points and true break points. Underscore F is computed based on [S3], and underscore N (resp. underscore χ^2) is computed by distributional approximations based on $G^*_{j,4}$ (resp. $G^\diamond_{j,4}$).

p	CER_F	CER_N	CER_{χ^2}	FDR_F	FDR_N	FDR_{χ^2}	Power_F	Power_N	Power_{χ^2}	D_N	D_{χ^2}
$n = 4$											
$N(0,1)$											
600	0.20	0.07	0.09	0.53	0.04	0.03	0.11	0.66	0.55	16	18
2000	0.20	0.04	0.04	0.50	0.01	0.01	0.10	0.83	0.80	19	18
6000	0.20	0.02	0.02	0.50	0.01	0.005	0.08	0.91	0.90	34	34
$t(6)/1.5^{0.5}$											
600	0.19	0.08	0.11	0.44	0.03	0.02	0.12	0.60	0.48	15	24
2000	0.19	0.04	0.04	0.41	0.01	0.008	0.11	0.83	0.81	16	16
6000	0.19	0.02	0.02	0.42	0.008	0.007	0.10	0.90	0.89	44	34
$\mathrm{LP}(0,1)/2^{0.5}$											
600	0.18	0.07	0.09	0.29	0.03	0.03	0.16	0.67	0.57	10	13
2000	0.18	0.03	0.04	0.29	0.01	0.01	0.13	0.84	0.81	16	18
6000	0.18	0.02	0.02	0.27	0.005	0.005	0.12	0.90	0.89	51	33
$n = 10$											
$N(0,1)$											
600	0.17	0.03	0.03	0.08	0.04	0.04	0.15	0.89	0.89	15	13
2000	0.18	0.02	0.01	0.08	0.02	0.02	0.11	0.95	0.95	33	27
6000	0.18	0.01	0.01	0.07	0.01	0.01	0.08	0.97	0.97	49	47
$t(6)/1.5^{0.5}$											
600	0.17	0.03	0.03	0.07	0.04	0.04	0.18	0.88	0.89	17	14
2000	0.18	0.02	0.02	0.05	0.02	0.02	0.13	0.94	0.94	47	25
6000	0.18	0.01	0.01	0.04	0.01	0.01	0.11	0.97	0.98	51	47
$\mathrm{LP}(0,1)/2^{0.5}$											
600	0.16	0.03	0.03	0.03	0.03	0.03	0.18	0.89	0.90	21	13
2000	0.17	0.02	0.01	0.02	0.02	0.02	0.16	0.94	0.95	53	47
6000	0.18	0.01	0.01	0.02	0.01	0.01	0.12	0.97	0.97	52	48

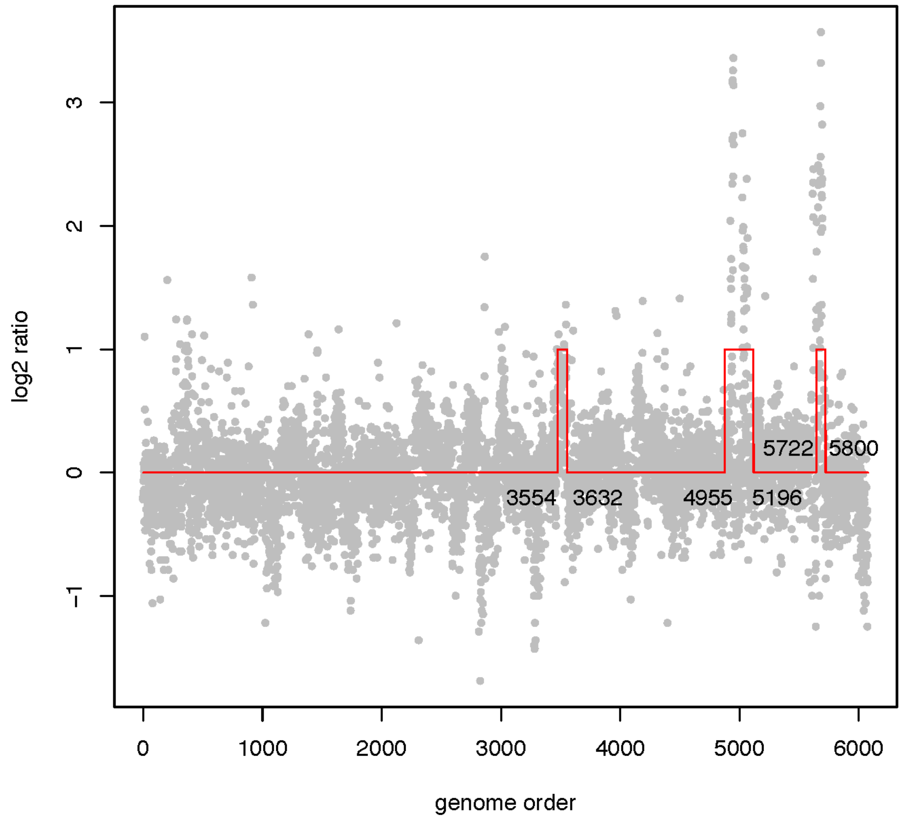

Figure S1: Testing results and detected break-points of amplification for breast cancer cell line BT474.

computed through the simulation-assisted approach. Specifically, γ and δ are empirical 0.95th quantiles of 10^4 independent realizations of $\hat{\sigma}\max_{1\le i\le p-k+1} G_i^\circ$, and $\hat{\sigma}\max_{i\in W_1} G_i^\circ$, respectively, where $G_i^\circ = k^{-1}\sum_{j=i}^{i+k-1}\eta_j, j \in \mathbb{Z}, \eta_j$ are i.i.d. $N(0,1)$ random variables and W_1 includes indices i such that the smoothed $\tilde{Q}_i^\circ = 1$. They are 0.4375 and 0.1204, respectively. We use the median of $\hat{\sigma}_i^2 = \sum_{j=i}^{i+m-1} X_j^2/m, i = 1, \ldots, p-m+1$ as our estimate of σ^2, which is 0.0897. We follow Algorithm 2.1 and present the results in Figure S2. It detects three clusters. Most amplifications on chromosomes 11, 17 and 20 are well known, as they have been identified by previous studies and in other breast cancer cell lines [S7, S5].

References

[S1] Chernozhukov, V., Chetverikov, D. and Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.*, **41** 2786–2819

[S2] Chernozhukov, V., Chetverikov, D. and Kato, K. (2017). Central limit theorems and bootstrap in high dimensions. *Ann. Probab.*, **45** 2309–2352

[S3] Fan, J., Hall, P. and Yao, Q. (2007). To how many simultaneous hypothesis tests can normal, student's t or bootstrap calibration be applied? *J. Amer. Statist. Assoc.*, **102** 1282–1288.

[S4] de la Pena, Victor (1992) Decoupling and Khintchine's Inequalities for U-Statistics, *Ann. Probab.*, **20** 1877–1892,

[S5] Lai, T.-L., Xing, H. and Zhang, N. (2008). Stochastic segmentation models for array-based comparative genomic hybridization data analysis. *Biostatistics*, **9** 290–307.

[S6] Nagaev, S. V. (1979) Large deviations of sums of independent random variables. *Ann. Probab.*, **7** 745–789,

[S7] Pollack, J. R., Perou, C. M., Alizadeh, A.A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D. and Brown, P. O. (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nature Genetics*, **23** 41–46.

[S8] Rudelson, M. and Vershynin, R. (2013). Hanson-Wright inequality and sub-gaussian concentration. *Electronic Communications in Probability*, **18**.

[S9] Sakhanenko, A. I. (2006) Estimates in the invariance principle in terms of truncated power moments. *Siberian Math. J.* **47** 1113–1127

[S10] Yao, Q. (1993) Tests for change-points with epidemic alternatives. *Biometrika*, **80** 179–191.