Detecting multiple change points: the PULSE criterion

[1] *Renmin University of China,* [2] *Xi'an Jiaotong University,*

[3] *Beijing Normal University,* [4] *Hong Kong Baptist University*

Supplementary Material

This Supplementary material includes the lemmas and technical proofs of the Theorems.

S1 Proof

S1.1 Proof of Theorem 1

We give two lemmas first.

Lemma 1. *Assume that* $X_i - EX_i$ *are independent identically distributed random variables and* $\frac{n^{1/4}\log n}{\sqrt{\alpha}_n} \to 0$. *The second finite monments exists. Then we have*

$$\Pr\{\max_{1\le i\le n}\Big||\tilde{D}_n(i)| - |D(i)|\Big| > \tau_n\} = o(1) \tag{S1.1}$$

where $\tau_n = O(\sqrt{\frac{\log n}{\alpha_n}})$.

Proof of Lemma 1 We first rewrite $\tilde{D}_n(i)$ as a sum of independent variables:

$$\begin{aligned}\tilde{D}_n(i) &= \frac{1}{\alpha_n^2}\Big\{\sum_{j=i}^{i+\alpha_n-1}(j-i+1)X_j + \sum_{i+\alpha_n}^{i+2\alpha_n-1}(3\alpha_n-2j+2i-2)X_j \\ &\quad + \sum_{i+2\alpha_n}^{i+3\alpha_n-1}(3\alpha_n-j+i-1)X_j\Big\}. \qquad \text{(S1.2)}\end{aligned}$$

Then the variance of $\tilde{D}_n(i)$ equals, for a constant $C>0$:

$$\begin{aligned}&\mathrm{Var}\Big\{\frac{1}{\alpha_n^2}(\sum_{j=i}^{i+\alpha_n-1}(j-i+1)X_j + \sum_{j=i+\alpha_n}^{i+2\alpha_n-2}(2i+3\alpha_n-2j-2)X_j \\ &+ \sum_{i+2\alpha_n-1}^{i+3\alpha_n-2}(i+3\alpha_n-j-1)X_j)\Big\} \\ =\ & \frac{Var(X_1)}{\alpha_n^4}(\sum_{i=1}^{\alpha_n}2\cdot i^2+\sum_{h=1}^{\alpha_n}(3\alpha_n-2h)^2) := \frac{C^2}{\alpha_n} = \sigma_n^2. \qquad \text{(S1.3)}\end{aligned}$$

It is obvious that the variance of $\tilde{D}(i)$ is then free of the index i with $\sigma_n = C/\sqrt{\alpha}_n$. In addition, as $\tilde{D}_n(i)$ is a weighted sum of $\{X_i\}_{i=1}^n$, we then further rewrite it. Define a weight function $w_n(t,j)$ as denoting $[nt]$ as the largest integer that is smaller or equal to $[nt]$,

$$\begin{aligned}w_n(t,j) = \mathrm{I}\{[nt]\le j\le [nt]+\alpha_n-1\}\frac{(j-[nt]+1)}{\alpha_n^2} \\ +\mathrm{I}\{[nt]+\alpha_n\le j\le [nt]+2\alpha_n-1\}\frac{(3\alpha_n-2j+2[nt]-2)}{\alpha_n^2} \\ +\mathrm{I}\{[nt]+2\alpha_n\le j\le [nt]+3\alpha_n-1\}\frac{(3\alpha_n-j+[nt]-1)}{\alpha_n^2},\end{aligned}$$

where $\mathrm{I}\{B\}$ denotes indicator function of set B. As for evert i there exists

$t_i \in (0,1)$ such that $i = [nt_i]$, we have

$$w_n(t_i, j) = \mathrm{I}\{[nt_i] \le j \le [nt_i] + \alpha_n - 1\}\frac{(j-i+1)}{\alpha_n^2} + \mathrm{I}\{[nt_i] + \alpha_n \le j \le [nt_i] + 2\alpha_n - 1\}\frac{(3\alpha_n - 2j + 2[nt_i] - 2)}{\alpha_n^2} + \mathrm{I}\{[nt_i] + 2\alpha_n \le j \le [nt_i] + 3\alpha_n - 1\}\frac{(3\alpha_n - j + [nt_i] - 1)}{\alpha_n^2}. \tag{S1.4}$$

$\tilde{D}_n(i)$ can then be rewritten as $\tilde{D}_n(i) = \sum_{j=1}^{n} w_n(t_i, j)X_j$. Then $\tilde{D}_n(i) - \tilde{D}(i) = \sum_{j=1}^{n} w_n(t_i, j)(X_j - E(X_j))$. Thus we have

$$\frac{\tilde{D}_n(i) - \tilde{D}(i)}{\sigma_n} = \sum_{j=1}^{n} \frac{w_n(t_i, j)}{\sigma_n}(X_j - E(X_j))$$

Let $\tilde{w}_n(t_i, j) = \frac{w_n(t_i,j)}{\sigma_n}$, $Y_n(t_i) = \tilde{D}_n(i) - \tilde{D}(i)/\sigma_n$ and $e_j = X_j - E(X_j)$. Then we have that

$$Y_n(t_i) = \sum_{j=1}^{n} \tilde{w}_n(t_i, j)e_j, \tag{S1.5}$$

where $\tilde{w}_n(t_i, j)$ can be seen as a special case of Equation (18) in Wu and Zhao (2007). In addition, define $\Omega_n(t_i) = |\tilde{w}_n(t_i, 1)| + \sum_{j=2}^{n} |\tilde{w}_n(t_i, j) - \tilde{w}_n(t_i, j-1)|$ and $\Omega_n = \max_{1 \le i \le n}\{\Omega_n(t_i)\}$. Some elementary calculations lead to

$$\Omega_n(t_i) = \frac{4\alpha_n + 3}{\alpha_n^2 \sigma_n}. \tag{S1.6}$$

As $\Omega_n(t_i)$ is free of i and then $\Omega_n = \frac{4\alpha_n+3}{\alpha_n^2\sigma_n}$. The application of Theorem 3 in Wu (2007) and Equation (6) in Wu and Zhao (2007) suggest that there exists a Gaussian process below with the standard Brownian motion $\mathbb{B}(\cdot)$,

$$Y_n^*(t_i) = \sum_{j=1}^{n} \tilde{w}_n(t_i, j)\sqrt{\mathrm{Var}(X_1)}\{\mathbb{B}(j) - \mathbb{B}(j-1)\} \tag{S1.7}$$

such that almost surely for all i

$$|Y_n(t_i) - Y_n^*(t_i)| \le o(\Omega_n(t_i) n^{1/4} \log n), \tag{S1.8}$$

and then

$$\max_{1\le i\le n} |Y_n(t_i) - Y_n^*(t_i)| = o(\Omega_n n^{1/4} \log n). \tag{S1.9}$$

This yields that almost surely

$$\begin{aligned}\max_{1\le i\le n} |Y_n(t_i)| &= \max_{1\le i\le n} |Y_n(t_i) - Y_n^*(t_i) + Y_n^*(t_i)| \\ &\le \max_{1\le i\le n} |Y_n^*(t_i)| + \max_{1\le i\le n} |Y_n(t_i) - Y_n^*(t_i)| \\ &\le \max_{1\le i\le n} |Y_n^*(t_i)| + o(\Omega_n n^{1/4} \log n),\end{aligned} \tag{S1.10}$$

and

$$\begin{aligned}\max_{1\le i\le n} \left| |\tilde{D}_n(i)| - |\tilde{D}(i)| \right| / \sigma_n \le \max_{1\le i\le n} \left| \tilde{D}_n(i) - \tilde{D}(i) \right| / \sigma_n \\ = \max_{1\le i\le n} \left| Y_n(t_i) \right| \\ \le \max_{1\le i\le n} |Y_n^*(t_i)| + o(\Omega_n n^{1/4} \log n).\end{aligned} \tag{S1.11}$$

Due to the fact $\sigma_n = O(1/\sqrt{\alpha_n})$ and the result in (S1.6), we can see that $\Omega_n = \frac{4\alpha_n+3}{\alpha_n^2 \sigma_n} = O(1/\sqrt{\alpha_n})$. By the condition $\frac{n^{1/4}\log n}{\sqrt{\alpha_n}} \to 0$, we have for any τ_n

$$\begin{aligned}\Pr\{\max_{1\le i\le n}\Big||\tilde{D}_n(i)|-|\tilde{D}(i)|\Big|>\tau_n\} &= \Pr\{\max_{1\le i\le n}\Big||\tilde{D}_n(i)|-|\tilde{D}(i)|\Big|/\sigma_n>\tau_n/\sigma_n\}\\ &\le \Pr\{\max_{1\le i\le n}\Big|Y_n^*(t_i)\Big|+o(\frac{n^{1/4}\log n}{\sqrt{\alpha}_n})>\tau_n/\sigma_n\}\\ &\le \Pr\{\max_{1\le i\le n}\Big|Y_n^*(t_i)\Big|+1>\tau_n/\sigma_n\}.\end{aligned} \tag{S1.12}$$

From (S1.7), we have

$$\mathrm{Var}(Y_n^*(i))=\frac{\mathrm{Var}(X_1)}{\sigma_n^2\alpha_n^4}(\sum_{i=1}^{\alpha_n}2\cdot i^2+\sum_{h=1}^{\alpha_n}(3\alpha_n-2h)^2)=1. \tag{S1.13}$$

In other words, $Y_n^*(t_i)$ follows the standard normal distribution, and thus, with an application of Proposition 2.1.2 in Roman (2017), we have, for large τ_n/σ_n,

$$\begin{aligned}\Pr\{\max_{1\le i\le n}\Big|Y_n^*(t_i)\Big|+1>\tau_n/\sigma_n\} \le& n\max_{1\le i\le n}\Pr\{\Big|Y_n^*(t_i)\Big|+1>\tau_n/\sigma_n\}\\ &= n\Pr\{\Big|Y_n^*(t_1)\Big|>\tau_n/\sigma_n-1\}\\ &\le n/(\frac{\tau_n}{\sigma_n}-1)\exp\{-\frac{1}{2}(\frac{\tau_n}{\sigma_n}-1)^2\}.\end{aligned} \tag{S1.14}$$

Taking $\tau_n/\sigma_n=\sqrt{2\log n}+1$, we have as $n\to\infty$

$$\begin{aligned}n/(\frac{\tau_n}{\sigma_n}-1)\exp\left\{-\frac{1}{2}(\frac{\tau_n}{\sigma_n}-1)^2\right\} &= \exp\left\{\log n-\log\sqrt{2\log n}-\log n\right\}\\ &= \sqrt{\frac{1}{2\log n}}\to 0.\end{aligned}$$

That is when $\tau_n=\sigma_n(\sqrt{2\log n}+1)=O(\sqrt{\log\alpha_n}/\sqrt{\alpha_n})$ and $n\to\infty$, we have

$$\Pr\{\max_{1\leq i\leq n}\left||\tilde{D}_n(i)|-|\tilde{D}(i)|\right|>\tau_n\}\leq\sqrt{\frac{1}{2\log n}}\to 0. \tag{S1.15}$$

This means that $\max_{1\leq i\leq n}\left||\tilde{D}_n(i)|-|\tilde{D}(i)|\right|=O_p(\sqrt{\frac{\log n}{\alpha_n}})$. We complete the proof of Lemma 1.

For the consistency of the estimated change points defined in the criterion, we first give the detailed computation of $\tilde{D}(i)$. It is easy to see that

$$|\tilde{D}(i)|=\begin{cases}0, & z_{k-1}+\alpha_n\leq i\leq z_k-2\alpha_n;\\ \frac{1+\cdots+(i-(z_k-2\alpha_n))}{\alpha_n^2}\beta_k, & z_k-2\alpha_n<i\leq z_k-\alpha_n;\\ \frac{[(i-(z_k-\alpha_n-1))+\cdots+\alpha_n]+[(\alpha_n-1)+\cdots+(\alpha_n-(i-(z_k-\alpha_n)))]}{\alpha_n^2}\beta_k, & z_k-\alpha_n<i\leq z_k-\frac{\alpha_n}{2};\\ \frac{[(z_k-i+2)+\cdots+\alpha_n]+[(\alpha_n-1)+\cdots+(\alpha_n-(z_k-i+1))]}{\alpha_n^2}\beta_k, & z_k-\frac{\alpha_n}{2}<i\leq z_k;\\ \frac{1+\cdots+((z_k+\alpha_n)-i+1)}{\alpha_n^2}\beta_k, & z_k<i\leq z_k+\alpha_n;\\ 0, & z_k+\alpha_n<i\leq z_{k+1}-2\alpha_n.\end{cases}$$

From this formula, we have a more detailed calculation that will be used in the proof of Lemma 2 and Theorem 2.1:

$$|\tilde{D}(i)| = \begin{cases} 0, & z_{k-1}+\alpha_n \le i \le z_k - 2\alpha_n; \\ \frac{(i-z_k+2\alpha_n+1)\cdot(i-z_k+2\alpha_n)}{\alpha_n^2}\beta_k, & z_k - 2\alpha_n < i \le z_k - \alpha_n; \\ \frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k, & z_k-\alpha_n < i < z_k - \frac{\alpha_n}{2} - B_n; \\ (\frac{3}{4} - \frac{B_n^2}{\alpha_n^2} + \frac{B_n}{\alpha_n^2})\beta_k, & i = z_k - \frac{\alpha_n}{2} - B_n; \\ \frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k, & z_k - \frac{\alpha_n}{2} - B_n < i < z_k - \frac{\alpha_n}{2}; \\ \frac{3}{4}\beta_k, & i = z_k - \frac{1}{2}\alpha_n; \\ \frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k, & z_k - \frac{\alpha_n}{2} < i < z_k - \frac{\alpha_n}{2} + B_n; \\ (\frac{3}{4} - \frac{B_n^2}{\alpha_n^2} + \frac{B_n}{\alpha_n^2})\beta_k, & i = z_k - \frac{\alpha_n}{2} + B_n; \\ \frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k, & z_k - \frac{\alpha_n}{2} + B_n < i \le z_k; \\ \frac{(-i+z_k+\alpha_n+2)(-i+1+\alpha_n+z_k)}{\alpha_n^2}\beta_k, & z_k < i \le z_k + \alpha_n; \\ 0, & z_k + \alpha_n < i \le z_{k+1} - 2\alpha_n. \end{cases} \tag{S1.16}$$

We can then know that when $z_{k-1} - 2\alpha_n \le i \le z_k - \frac{\alpha_n}{2}$, $|\tilde{D}(i)|$ monotonically increases with i while when $z_k - \frac{\alpha_n}{2} \le i \le z_k + \alpha_n$, $|\tilde{D}(i)|$ monotonically decreases.

Similarly, we can derive $T(i) = \frac{|\tilde{D}(i)|+c_n}{|\tilde{D}(i+\frac{3\alpha_n}{2})|+c_n}$ as:

$$T(i) = \begin{cases} \frac{0+c_n}{0+c_n} = 1, & z_{k-1}+\alpha_n \le i \le z_k - \frac{7}{2}\alpha_n, \\ \frac{0+c_n}{\frac{(i-z_k+2\alpha_n+1)\cdot(i-z_k+2\alpha_n)}{\alpha_n^2}\beta_k+c_n}, & z_k-\frac{7}{2}\alpha_n < i \le z_k-\frac{5}{2}\alpha_n, \\ \frac{0+c_n}{\frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k+c_n}, & z_k-\frac{5}{2}\alpha_n < i < z_k-2\alpha_n-B_n, \\ \frac{0+c_n}{(\frac{3}{4}-\frac{B_n^2}{\alpha_n^2}+\frac{B_n}{\alpha_n^2})\beta_k+c_n}, & i = z_k-2\alpha_n-B_n, \\ \frac{0+c_n}{\frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k+c_n}, & z_k-2\alpha_n-B_n < i < z_k-2\alpha_n, \\ \frac{0+c_n}{\frac{3}{4}\beta_k+c_n}, & i = z_k-2\alpha_n, \\ \frac{\frac{(i-z_k+2\alpha_n+1)\cdot(i-z_k+2\alpha_n)}{\alpha_n^2}\beta_k+c_n}{\frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k+c_n}, & z_k-2\alpha_n < i \le z_k-2\alpha_n+B_n, \\ \frac{\frac{B_n(B_n+1)}{\alpha_n^2}+c_n}{(\frac{3}{4}-\frac{B_n^2}{\alpha_n^2}+\frac{B_n}{\alpha_n^2})\beta_k+c_n}, & i = z_k-2\alpha_n+B_n, \\ \frac{\frac{(i-z_k+\frac{1}{2}\alpha_n+1)\cdot(i-z_k+\frac{1}{2}\alpha_n)}{\alpha_n^2}\beta_k+c_n}{\frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k+c_n}, & z_k-2\alpha_n+B_n < i \le z_k-\frac{3}{2}\alpha_n, \\ \frac{\frac{(i-z_k+\frac{1}{2}\alpha_n+1)\cdot(i-z_k+\frac{1}{2}\alpha_n)}{\alpha_n^2}\beta_k+c_n}{\frac{(-i+z_k+\alpha_n+1)\cdot(-i+z_k+\alpha_n+2)}{\alpha_n^2}\beta_k+c_n}, & z_k-\frac{3}{2}\alpha_n < i \le z_k-\alpha_n, \\ \frac{\frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k+c_n}{\frac{((z_k+\alpha_n)-i+1)((z_k+\alpha_n)-i+2)\beta_k}{\alpha_n^2}+c_n}, & z_k-\alpha_n < i \le z_k-\frac{1}{2}\alpha_n, \\ \frac{\frac{-i^2-\alpha_n i+2iz_k-i+z_k-z_k^2+\alpha_n z_k+\frac{1}{2}(\alpha_n^2-\alpha_n)}{\alpha_n^2}\beta_k+c_n}{0+c_n}, & z_k-\frac{1}{2}\alpha_n < i \le z_k, \\ \frac{\frac{((z_k+\frac{3}{2}\alpha_n)-i+1)((z_k+\frac{3}{2}\alpha_n)-i+2)}{\alpha_n^2}+c_n}{0+c_n}, & z_k < i \le z_k+\alpha_n, \\ \frac{0+c_n}{0+c_n} = 1, & z_k+\alpha_n < i \le z_{k+1}-\frac{7}{2}\alpha_n. \end{cases} \tag{S1.17}$$

We now give another lemma and its proof.

Lemma 2. *Assume that $X_i - EX_i$ are independent identically distributed* *random variables, we could define* $A^d = \{i : T(i) < d\}$ *and* $A_n^d = \{i : T_n(i) \leq d\}$ *for any* $0 < d < 1$. *We have for any* d_1, d_2 *and* d_3 *with* $0 < d_3 < d_1 < d_2 < 1$.

$$\Pr\{A_n^{d_1} \subseteq A^{d_2}\} \to 1 \quad \Pr\{A^{d_3} \subseteq A_n^{d_1}\} \to 1. \tag{S1.18}$$

Further, for any $k = 1, \ldots, K$ *the intervals* (m_k, M_k) *are disjoint and each contains only one local minimizer* $z_k - 3\alpha_n/2$ *of* $T(i)$. *Further, for any* d *with* $0 < d < 1$,

$$\max_{i \in A_n^d} |T_n(i) - T(i)| = o_p(1). \tag{S1.19}$$

Proof of Lemma 2 To prove this lemma, we first analyse the properties of $T_n(i) = \frac{\tilde{D}_n(i)+c_n}{\tilde{D}_n(i+\frac{3}{2}\alpha_n)+c_n}$ around the point $z_k - 2\alpha_n$ where z_k is the change point. Write it as

$$\begin{aligned} T_n(i) &= \frac{|\tilde{D}_n(i)| + c_n}{|\tilde{D}_n(i + \frac{3}{2}\alpha_n)| + c_n} \\ &= \frac{|\tilde{D}_n(i)| - |\tilde{D}(i)| + |\tilde{D}(i)| + c_n}{|\tilde{D}_n(i)| - |\tilde{D}(i + \frac{3}{2}\alpha_n)| + |\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n} \\ &= \frac{O_p(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + |\tilde{D}(i)| + c_n}{O_p(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + |\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n} \end{aligned} \tag{S1.20}$$

For the flat parts in the sequence with $|\tilde{D}(i)| = 0$ for all i, we have

$$T_n(i) = \frac{O_p(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + 0 + c_n}{O_p(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + 0 + c_n} = o_p(1). \tag{S1.21}$$

When a change point appears, we have that, from (S1.16) and the discussion right below it, for $\forall i \in [z_k - \frac{7}{2}\alpha_n, z_k - 2\alpha_n]$, $|\tilde{D}(i)| = 0$, $|\tilde{D}(i + \frac{3}{2}\alpha_n)|$

monotonically increases and at $i = z_k - 2\alpha_n$, we have

$$T_n(z_k - 2\alpha_n) = \frac{|\tilde{D}_n(z_k - 2\alpha_n)| + c_n}{|\tilde{D}_n(z_k - \frac{1}{2}\alpha_n)| + c_n} = \frac{O_p(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + 0 + c_n}{O_p(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + \frac{3}{4}\beta_k + c_n} = o_p(1). \tag{S1.22}$$

As we discussed before, for any $i \in [z_k - 2\alpha_n, z_k - \frac{1}{2}\alpha_n]$, $|\tilde{D}(i)|$ monotonically increases, and $|\tilde{D}(i + \frac{3}{2}\alpha_n)|$ monotonically decreases, then $T_n(i)$ uniformly converges to the monotonically increasing $T(i)$ and

$$T_n(z_k - \frac{1}{2}\alpha_n) = \frac{|\tilde{D}_n(z_k - \frac{1}{2}\alpha_n)| + c_n}{|\tilde{D}_n(z_k + \alpha_n)| + c_n} = \frac{O_p(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + \frac{3}{4}\beta_k + c_n}{O_p(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + 0 + c_n} \xrightarrow{P} \infty. \tag{S1.23}$$

Step 1 To prove the subset equations in (S1.18) and the uniform convergence in (S1.19). Define $A^{d_2} = \{i : T(i) < d_2\}$ and $A_n^{d_1} = \{i : T_n(i) < d_1\}$ where $d_1 < d_2$. Recall the decomposition of (S1.20). By the definition of $A_n^{d_1}$, we have for all $i \in A_n^{d_1}$, we have $T_n(i) \le d_1$. Then,

$$o_p(c_n) + |\tilde{D}(i)| + c_n \le d_1(o_p(c_n) + |\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n).$$

That is,

$$|\tilde{D}(i)| + c_n \le d_1(|\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n) + o_p(c_n).$$

We can get, uniformly over all i, in probability, for large n

$$\begin{aligned} T(i) &= \frac{|\tilde{D}(i)| + c_n}{|\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n} \\ &\le d_1 + o(1) < d_2. \end{aligned} \tag{S1.24}$$

In other words, with a probability going to one, $A_n^{d_1} \subseteq A^{d_2} = \{i : T(i) < d_2\}$. We can similarly prove that with a probability tending to one, $A^{d_3} \subseteq A_n^{d_1}$ for d_3 with $d_3 < d_1 < 1$.

Step 2. To prove that for any $k = 1, \ldots, K$ the intervals (m_k, M_k) are disjoint and each contains only one local minimizer $z_k - 2\alpha_n$ of $T(i)$. Consider a value d with $d > 0.5$. Let $\tilde{m}_k$ and $\tilde{M}_k$ satisfy the following conditions:

$$T(\tilde{m}_k - 1) \geq d, \quad T(\tilde{m}_k) < d,$$

$$T(\tilde{M}_k) < d, \quad T(\tilde{M}_k + 1) \geq d.$$

Denote the interval $(\tilde{m}_k, \tilde{M}_k)$. From the previous proof, we can easily derive that in probability, $(m_k, M_k) \subseteq (\tilde{m}_k, \tilde{M}_k)$. Further, from the properties, we also know that all $(\tilde{m}_k, \tilde{M}_k)$ are contained in A^d and disjoint, also each interval contains only one local minimizer $z_k - 2\alpha_n$ of $T(i)$. When we choose a value d with $0 < d < 0.5$ we can derive that in probability, $(\tilde{m}_k, \tilde{M}_k) \subseteq (m_k, M_k)$. Similarly, we also know that all $(\tilde{m}_k, \tilde{M}_k)$ are contained in A^d and disjoint, also each interval contains only one local minimizer $z_k - 2\alpha_n$ of $T(i)$. These two properties imply that in probability (m_k, M_k) are contained in $A_n^{0.5}$ and disjoint, also each interval contains only one local minimizer $z_k - 2\alpha_n$ of $T(i)$.

Step 3. To prove the weak convergence of $T_n(i)$ to $T(i)$ over the set

$A_n^{d_1}$. As in probability $A_n^{d_1} \subseteq A^{d_2}$ such that $T(i) \le d_2 < 1$, we consider a large set to derive the uniform convergence. For any $i \in A^{d_2}$, we have, uniformly,

$$\begin{aligned}
T_n(i) - T(i) &= \frac{|\tilde{D}_n(i)| + c_n}{|\tilde{D}_n(i + \frac{3}{2}\alpha_n)| + c_n} - \frac{|\tilde{D}(i)| + c_n}{|\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n} \\
&= \frac{|(D_n(i)| + c_n)(|\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n) - (|\tilde{D}(i)| + c_n)(|\tilde{D}_n(i + \frac{3}{2}\alpha_n)| + c_n)}{(|\tilde{D}_n(i + \frac{3}{2}\alpha_n)| + c_n)(|\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n)} \\
&= \{\frac{[(|\tilde{D}_n(i)| - |\tilde{D}(i))||(\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n)]}{(|\tilde{D}_n(i + \frac{3}{2}\alpha_n)| + c_n)(|\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n)} \\
&- \frac{[(|\tilde{D}_n(i + \frac{3}{2}\alpha_n)| - |\tilde{D}(i + \frac{3}{2}\alpha_n)|)(|\tilde{D}(i)| + c_n)]}{(|\tilde{D}_n(i + \frac{3}{2}\alpha_n)| + c_n)(|\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n)}\} \\
&= \{\frac{[o_p(c_n)(|\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n)] - [o_p(c_n)(|\tilde{D}(i)| + c_n)]}{(|\tilde{D}_n(i + \frac{3}{2}\alpha_n)| + c_n)(|\tilde{D}(i + \frac{3}{2}\alpha_n)| + c_n)}\} \\
&= \frac{o_p(c_n)}{o_p(c_n) + (|\tilde{D}(i + \frac{3\alpha_n}{2})| + c_n)} - o_p(c_n)T(i) \\
&= \frac{o_p(c_n)}{c_n} - o_p(c_n) = o_p(1).
\end{aligned}$$

Thus $\max_{i \in A^{d_2}} |T_n(i) - T(i)| = o_p(1)$. The proof is finished.

Proof of Theorem 1. We consider the first part in the theorem. By Lemma 2, in probability $z_k - 2\alpha_n \in (\tilde{m}_k, \tilde{M}_k) \subseteq A^d$ implies that $z_k - 2\alpha_n \in (m_k, M_k) \subseteq A_n^{0.5} \subseteq A^d$. Thus uniformly over $1 \le k \le K$ in probability, we have

$$\tilde{m}_k \le z_k - 2\alpha_n \le \tilde{M}_k. \qquad \text{(S1.25)}$$

At the population level with $T(i)$'s, by the uniqueness of $z_k - 2\alpha_n$ in the interval (m_k, M_k), searching for $z_k - 2\alpha_n$ in (m_k, M_k) is equivalent to searching for $z_k - 2\alpha_n$ in the non-random $(\tilde{m}_k, \tilde{M}_k)$ in probability.

Write $\hat{z}_k - 2\alpha_n$ as the local minimizer of $T_n(i)$'s in the interval $(m_k, M_k) \subseteq (\tilde{m}_k, \tilde{M}_k) \subseteq A^d$. Recall that by Lemma 2 $\max_{i \in A^{d_2}} |T_n(i) - T(i)| = o_p(1)$. We can then work on each interval (m_k, M_k). For any k with $1 \leq k \leq K$, from (S1.17), $T(z_k - 2\alpha_n)$ is the only local minimum and by the definition of $\hat{z}_k - 2\alpha_n$, $T_n(i) \geq T_n(\hat{z}_k - 2\alpha_n)$ in the interval in probability. From (S1.22) and (S1.23), we have that, as $|\tilde{D}(z_k - 2\alpha_n)| = 0$,

$$|\tilde{D}_n(z_k - 2\alpha_n)| = O_p(\sqrt{\log n}/\sqrt{\alpha_n}) = o_p(c_n) \tag{S1.26}$$

and, as $|\tilde{D}(z_k - \frac{1}{2}\alpha_n)| = 3\beta_k/4$,

$$|\tilde{D}_n(z_k - \frac{1}{2}\alpha_n)| - 3\beta_k/4 = O_p(\sqrt{\log n}/\sqrt{\alpha_n}) = o_p(c_n). \tag{S1.27}$$

Further, from the calculation of $T(i)$ before, we can see that letting $B_n = \alpha_n(\log \alpha_n)^{-1/5}$, for any $j = O(B_n)$

$$|\tilde{D}(z_k - 2\alpha_n \pm j)| = O(c_n). \tag{S1.28}$$

To prove that $\hat{z}_k/z_k - 1 = o_p(1)$, we only need to prove that $|\hat{z}_k - z_k| = O_p(B_n)$. To this end, applying the strictly decreasing and increasing monotonicity of $T(i)$ on the two sides of $z_k - 2\alpha_n$ respectively, and the uniform convergence of $T_n(i)$ to $T(i)$ in probability in the set $A_n^{0.5}$, we only need to show that $T_n(z_k - 2\alpha_n \pm B_n) - T_n(z_k - 2\alpha_n) > 0$ in probability.

Consider $T_n(z_k - 2\alpha_n - B_n)$ first. Note that

$$T_n(z_k - 2\alpha_n - B_n) = \frac{0 + c_n + o_p(c_n)}{(\frac{3}{4} - \frac{B_n^2}{\alpha_n^2} + \frac{B_n}{\alpha_n^2})\beta_k + c_n + o_p(c_n)}. \tag{S1.29}$$

Let $b_{n1} = (\frac{B_n^2}{\alpha_n^2} - \frac{B_n}{\alpha_n^2})\beta_k$. To simplify the notations, in the following all derivations are in probability. We can derive that

$$\begin{aligned}
& T_n(z_k - 2\alpha_n - B_n) - T_n(z_k - 2\alpha_n) \\
&= \frac{c_n + O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}})}{O(\frac{1}{\sqrt{\alpha_n}}) + c_n + \frac{3}{4}\beta_k - b_n} - \frac{c_n + O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}})}{O(\frac{1}{\sqrt{\alpha_n}}) + c_n + \frac{3}{4}\beta_k} \\
&:= \frac{c_n + a_{n2}}{\beta_{n2} - b_{n1}} - \frac{c_n + a_{n1}}{\beta_{n1}} \\
&= \frac{(a_{n2} + c_n)\beta_{n1} - (a_{n1} + c_n)(\beta_{n2} - b_{n1})}{\beta_{n1}(\beta_{n2} - b_{n1})} \\
&= \frac{(a_{n1} + c_n)(\beta_{n1} - \beta_{n2}) + (a_{n2} - a_{n1})\beta_{n1} + (a_{n1} + c_n)b_{n1}}{\beta_{n1}(\beta_{n2} - b_{n1})} \\
&= \frac{(a_{n1} + c_n)O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}})\beta_{n1} + (a_{n1} + c_n)b_{n1}}{\beta_{n1}(\beta_{n2} - b_{n1})} \\
&= \frac{((a_{n1} + c_n)b_{n1})[O(\frac{\sqrt{\log n}}{b_{n1}\sqrt{\alpha_n}}) + O(\frac{\sqrt{\log n}}{(a_{n1}+c_n)b_{n1}\sqrt{\alpha_n}})\beta_{n1} + 1]}{\beta_{n1}(\beta_{n2} - b_{n1})}.
\end{aligned} \tag{S1.30}$$

When $(a_{n1} + c_n)b_{n1}\sqrt{\alpha}_n/\sqrt{\log n} \to \infty$, and $b_{n1}\sqrt{\alpha}_n/\sqrt{\log n} \to \infty$, we then have for large n, the value in the brackets is larger than a positive constant and then the numerator is positive as $c_n\sqrt{\alpha}_n/\sqrt{\log n} \to \infty$ and $c_n > 0$ such that $a_{n1} + c_n = c_n(1 + \frac{a_{n1}}{c_n}) = c_n(1 + O(\frac{\sqrt{\log n}}{\sqrt{\alpha}_n c_n})) > 0$ and $(a_{n1} + c_n)b_{n1} > 0$. We then have $T_n(z_k - 2\alpha_n - B_n) - T_n(z_k - 2\alpha_n) > 0$ when $b_{n1} \cdot c_n \cdot \sqrt{\alpha}_n = \frac{B_n^2}{\alpha_n^2} \cdot c_n \cdot \sqrt{\alpha}_n/\sqrt{\log n} > \frac{B_n^2}{\alpha_n^2} \cdot \sqrt{\log \alpha_n} \to \infty$.

For $i = z_k - 2\alpha_n + B_n$, we have

$$T_n(z_k - 2\alpha_n + B_n) = \frac{\frac{B_n(B_n+1)}{\alpha_n^2}\beta_k + c_n + o_p(c_n)}{(\frac{3}{4} - \frac{B_n^2}{\alpha_n^2} + \frac{B_n}{\alpha_n^2})\beta_k + c_n + o_p(c_n)}. \tag{S1.31}$$

Let $b_{n2} = \frac{B_n(B_n+1)}{\alpha_n^2}\beta_k$. We similarly have, in probability,

$$\begin{aligned}
&T_n(z_k - 2\alpha_n + B_n) - T_n(z_k - 2\alpha_n)\\
&= \frac{c_n + O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + b_{n2}}{c_n + O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + \frac{3}{4}\beta_k - b_{n1}} - \frac{c_n + O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}})}{O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + c_n + \frac{3}{4}\beta_k}\\
&=: \frac{c_n + a_{n3} + b_{n2}}{\beta_{n3} - b_{n1}} - \frac{c_n + a_{n1}}{\beta_{n1}}\\
&= \frac{(c_n + a_{n3} + b_{n2})\beta_{n1} - (a_{n1} + c_n)\beta_{n3} + (a_{n1} + c_n)b_{n1}}{\beta_{n1}(\beta_{n3} - b_{n1})}\\
&= \frac{(a_{n3} - a_{n1})\beta_{n1} + (a_{n1} + c_n)(\beta_{n1} - \beta_{n3}) + (a_{n1} + c_n)b_{n1} + b_{n2}\beta_{n1}}{\beta_{n1}(\beta_{n3} - b_{n1})}\\
&\geq \frac{O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}})\beta_{n1} + (a_{n1} + c_n)O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + b_{n2}\beta_{n1}}{\beta_{n1}(\beta_{n3} - b_{n1})}\\
&= \frac{b_{n2}[O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}b_{n2}})\beta_{n1} + (a_{n1} + c_n)O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}b_{n2}}) + \beta_{n1}]}{\beta_{n1}(\beta_{n3} - b_{n1})}
\end{aligned} \tag{S1.32}$$

The inequality is due to $(a_{n1}+c_n)b_{n1} > 0$. Thus as long as $b_{n2}\cdot\sqrt{\alpha_n}/\sqrt{\log n} > B_n^2\alpha_n^{-3/2}/\sqrt{\log n} \to \infty$, the first term in the brackets converges to zero. Note that a_{n1} and c_n both tend to zero. The second term converges to zero. As $\beta_{n1} = O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}}) + c_n + \frac{3}{4}\beta_k$, in which $O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}})$ and c_n go to zero, β_{n1} then tends to β_k and thus β_{n1} is larger than zero for large n. Therefore, $(O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}b_{n2}})\beta_{n1} + (a_{n1} + c_n)O(\frac{\sqrt{\log n}}{\sqrt{\alpha_n}b_{n2}}) + \beta_{n1})$ is greater than zero. The whole numerator and then the difference is larger than zero such that $T_n(z_k - 2\alpha_n + B_n) - T_n(z_k - 2\alpha_n) > 0$. Altogether, when $B_n^2 \cdot c_n \cdot \alpha_n^{-\frac{3}{2}}/\sqrt{\log n} \to \infty$, then

$$T_n(z_k - 2\alpha_n \pm B_n) - T_n(z_k - 2\alpha_n) > 0. \tag{S1.33}$$

As we argued before, $\hat{z}_k$ cannot be larger than $z_k \pm B_n$ in probability. Also, based on the definition in Lemma 2, we can get that $(z_k - 2\alpha_n - B_n, z_k - 2\alpha_n + B_n) \subset A_n^{d_1}$. That is

$$-B_n + z_k - 2\alpha_n \le \hat{z}_k - 2\alpha_n \le B_n + z_k - 2\alpha_n.$$

As $\frac{B_n}{\alpha_n} \to 0$

$$|\frac{\hat{z}_k - z_k}{\alpha_n}| \le \frac{B_n}{\alpha_n} \to 0$$

in probablity. In other words, for any $\epsilon > 0$, we have the uniform convergence over all $k \le K$: as $n \to \infty$

$$P(\max_{1 \le k \le K} |\frac{\hat{z}_k - z_k}{\alpha_n}| < \epsilon) \to 1 \tag{S1.34}$$

This proves that uniformly over all $k \le K$, $\hat{z}_k$ is a consistent estimator of z_k in the above sense. The proof of the first part of Theorem 1 is finished.

We now prove the second Part of Theorem 1. From the proof of the first part, we can see that we can consistently estimate all z_k for $1 \le k \le K$. Thus, clearly $\hat{K} = K$ with a probability going to one.

Now we prove the third part of Theorem 1. In the case with divergent K, along with the steps in the proof of Lemma 2 and of the first part of the theorem, we still have that $\max_k T_n(z_k - 2\alpha_n) \to 0$ in probability. That is, the local minima of $T_n(z_k - 2\alpha_n)$ can also converge to zero. The consistency can be proved almost the same as that for given K. Also $\hat{K} = K$ with a probability going to one in the divergent case. We then omit the details and finish the proof.

S1.2 Proof of Theorem 2

Denote the minimum change magnitude as $\beta_z = \min_{1\le k\le K_n} \beta_k$. β_z converges to 0 at the rate of $O((\log \alpha_n)^{-1/5})$ by the assumption.

From the proof of Lemma 2 and (S1.17), we have that, letting $B_n = \alpha_n(\log \alpha_n)^{-1/10}$, for any $j = O(B_n)$,

$$|\tilde{D}(z_k - 2\alpha_n \pm j)| = O(c_n). \tag{S1.35}$$

To this end, applying the strict monotonicity of $T(i)$, respectively, on the two sides of $z - 2\alpha_n$, and the uniform convergence of $T_n(i)$ to $T(i)$ in probability in the set $A_n^{0.5}$, we only need to show that $T_n(z_k - 2\alpha_n \pm B_n) - T_n(z_k - 2\alpha_n) > 0$ in probability. In other words, we only need to check, similarly as those in (S1.30) and (S1.32),

$$b_{n1} \cdot c_n \cdot \sqrt{\alpha}_n / \sqrt{\log n} \to \infty \tag{S1.36}$$

where $b_{n1} = (\frac{B_n^2}{\alpha_n^2} - \frac{B_n}{\alpha_n^2})\beta_z$. As $\beta_z = O((\log \alpha_n)^{-1/5})$ and $B_n = \alpha_n(\log \alpha_n)^{-1/10}$, we have the above convergence. Then

$$T_n(z_k - 2\alpha_n \pm B_n) - T_n(z - 2\alpha_n) > 0. \tag{S1.37}$$

Thus $z_k - B_n \leq \hat{z}_k \leq z_k + B_n$ in probability. As $\frac{B_n}{\alpha_n} \to 0$, we have uniformly over all $k \leq K$ in probability

$$|\frac{\hat{z}_k - z_k}{\alpha_n}| \leq \frac{B_n}{\alpha_n} \to 0.$$

The proof is finished.

S1.3 Proof of Theorem 3

We now prove the consistency of the estimators of the variance change points. From the criterion construction, the proof is very much similar to that for Theorem 1 as long as we pay attention to the rate of uniform convergence of $D_n(i)$ that is in this case the variance difference. Rather than only considering the first and second moment, we should take both second and forth moment into account. As the second monment of variable exists, there exists a constant C such that $E(X_j^2) \geq C$ for all j. Then we have that

$$
\begin{aligned}
\max_i |\tilde{D}_n(i) - \tilde{D}(i)| &= \max_i \log \frac{\sum_{j=1}^n w_n(t_i,j)X_j^2}{\sum_{j=1}^n w_n(t_i,j)E(X_j^2)} \\
&= \max_i \log \frac{\sum_{j=1}^n w_n(t_i,j)X_j^2 - \sum_{j=1}^n w_n(t_i,j)E(X_j^2) + \sum_{j=1}^n w_n(t_i,j)E(X_j^2)}{\sum_{j=1}^n w_n(t_i,j)E(X_j^2)} \\
&= \max_i \log(1 + \frac{\sum_{j=1}^n w_n(t_i,j)X_j^2 - \sum_{j=1}^n w_n(t_i,j)E(X_j^2)}{\sum_{j=1}^n w_n(t_i,j)E(X_j^2)}) \\
&\leq \max_i \frac{\sum_{j=1}^n w_n(t_i,j)X_j^2 - \sum_{j=1}^n w_n(t_i,j)E(X_j^2)}{\sum_{j=1}^n w_n(t_i,j)E(X_j^2)} \\
&\leq \max_i \frac{1}{C}(\sum_{j=1}^n w_n(t_i,j)X_j^2 - \sum_{j=1}^n w_n(t_i,j)E(X_j^2))
\end{aligned}
\tag{S1.38}
$$

For both of the mean and variance scenario, the number of variables that $\tilde{D}_n(i)$ involves is the same. As the forth finite moment exists, we have that the convergence rate $\max_i |\sum_{j=1}^n w_n(t_i,j)X_j^2 - \sum_{j=1}^n w_n(t_i,j)E(X_j^2)| = O_p(\sqrt{\frac{\log n}{\alpha_n}})$ And thus we have $\max_i |\tilde{D}_n(i) - \tilde{D}(i)| = O_p(\sqrt{\frac{\log n}{\alpha_n}})$. We then finish the proof without repeating the details that are used to prove Theorem 1.

Bibliography

Roman, V. (2007) High dimensional probability: an introduction with applications in data science *University of Michigen*

Wu, W. B. (2007) Strong invariance principles for dependent random varibles *The Annals of Probability*, **35**, 2294-2320

Wu, W. B. and Zhao, Z. (2007) Inference of trends in time series *Journal of the Royal Statistical Society: Series B*, **69**, 391-410