SUPPLEMENT TO “OPTIMAL CHANGE POINT DETECTION AND LOCALIZATION IN SPARSE DYNAMIC NETWORKS”

BY DAREN WANG*, YI YU† AND ALESSANDRO RINALDO‡

University of Chicago *, University of Warwick † and Carnegie Mellon University ‡

S.1. Proofs of Lemmas 3 and 4. In this subsection, we provide proofs of Lemma 1 in Section 2 and Lemma 2 in Section 3.1, which provide the minimax lower bounds for detection and localization respectively. In addition, Lemma S.1 is used in the proofs of Lemmas 1 and 2.

**Lemma S.1.** Let $\Theta \in \mathbb{R}^{n \times n}$ such that $\Theta_{ij} = \rho$ for all $1 \leq i, j \leq n$, where $0 < \rho < 1/2$. Let $A$ be an adjacency matrix of an inhomogeneous Bernoulli network with independent edges such that $\mathbb{E}(A) = \Theta$. For any $v_b, v_c \in [-\sqrt{\rho}, \sqrt{\rho}]^n$, let $B$ and $C$ be adjacency matrices of inhomogeneous Bernoulli networks with independent edges such that $\mathbb{E}(B) = v_b v_b^\top + \Theta$ and $\mathbb{E}(C) = v_c v_c^\top + \Theta$. Let $P_A, P_B, P_C$ be the distributions of $A, B$ and $C$. Then

$$
\mathbb{E}_{P_A} \left( \frac{dP_B}{dP_A} \frac{dP_C}{dP_A} \right) \leq \exp \left( \frac{(v_b^\top v_c)^2}{\rho(1 - \rho)} \right).
$$

Let $A' = A - \text{diag}(A)$, $B' = B - \text{diag}(B)$ and $C' = C - \text{diag}(C)$. Then

$$
\mathbb{E}_{P_{A'}} \left( \frac{dP_{B'}}{dP_{A'}} \frac{dP_{C'}}{dP_{A'}} \right) \leq \exp \left( \frac{(v_b^\top v_c)^2}{\rho(1 - \rho)} \right).
$$

**Proof.** Let $\Gamma = v_b v_b^\top$ and $\Lambda = v_c v_c^\top$.

$$
\mathbb{E}_{P_A} \left( \frac{dP_B}{dP_A} \frac{dP_C}{dP_A} \right) = \prod_{1 \leq i, j \leq n} \left( \frac{(\Gamma_{ij} + \rho)(\Lambda_{ij} + \rho)}{\rho} + \frac{(1 - \Gamma_{ij} - \rho)(1 - \Lambda_{ij} - \rho)}{(1 - \rho)} \right)
$$

$$
= \prod_{1 \leq i, j \leq n} \left( 1 + \frac{\Gamma_{ij} \Lambda_{ij}}{\rho(1 - \rho)} \right) \leq \prod_{1 \leq i, j \leq n} \exp \left( \frac{\Gamma_{ij} \Lambda_{ij}}{\rho(1 - \rho)} \right) = \exp \left( \frac{(v_b^\top v_c)^2}{\rho(1 - \rho)} \right).
$$

Note that

$$
\mathbb{E}_{P_{A'}} \left( \frac{dP_{B'}}{dP_{A'}} \frac{dP_{C'}}{dP_{A'}} \right) = \prod_{i \neq j} \left( \frac{(\Gamma_{ij} + \rho)(\Lambda_{ij} + \rho)}{\rho} + \frac{(1 - \Gamma_{ij} - \rho)(1 - \Lambda_{ij} - \rho)}{(1 - \rho)} \right)
$$

$$
= \prod_{i \neq j} \left( 1 + \frac{\Gamma_{ij} \Lambda_{ij}}{\rho(1 - \rho)} \right) \leq \prod_{1 \leq i, j \leq n} \left( 1 + \frac{\Gamma_{ij} \Lambda_{ij}}{\rho(1 - \rho)} \right) \leq \prod_{1 \leq i, j \leq n} \exp \left( \frac{\Gamma_{ij} \Lambda_{ij}}{\rho(1 - \rho)} \right) = \exp \left( \frac{(v_b^\top v_c)^2}{\rho(1 - \rho)} \right),
$$

where the fist inequality follows from the observation that $\Gamma_{ii} = (v_b)_i^2 \geq 0$ and $\Lambda_{ii} = (v_c)_i^2 \geq 0$. □

**Remark S.1.** Let $\Theta_{ij} = \rho + (vv^\top)_{ij}$, where $v \in \{ \pm \sqrt{\kappa_{ij}} \}^n$, $0 < \rho < 1/2$ and $0 < \kappa_{ij} < 1$, then the community labels can be decided according to the vector sign(v). More precisely let

$$
C_1 = \{ i : v_i > 0 \} \text{ and } C_2 = \{ i : v_i < 0 \}.
$$

The probability within $C_1$ or $C_2$ is $\rho(1 + \kappa_0)$. The probability between $C_1$ and $C_2$ is $\rho(1 - \kappa_0)$. 1
**Proof of Lemma 1.** Without loss of generality, suppose that $L = 4^{-1} T \log^2(T)$ is an integer. For $l \in \{1, \ldots, L\}$, $v \in \{1, -1\}^n$, let $\tilde{P}^l$ be the joint distribution of a collection of independent adjacency matrices $\{A(t)\}_{t=1}^T$ such that

$$\mathbb{E}\{(A(t))_{ij}\} = 1/2, \quad i, j \in \{1, \ldots, n\}, \quad t \in \{1, \ldots, T\} \setminus \{(l-1) \log^2(T) + 1, \ldots, l \log^2(T)\},$$

and

$$\mathbb{E}\{(A(t))_{ij}\} = 1/2 + \frac{2^{1/3} \xi^{1/2} (vv^\top)_{ij}}{n^{1/2}}, \quad i, j \in \{1, \ldots, n\}, \quad t \in \{(l-1) \log^2(T) + 1, \ldots, l \log^2(T)\}.$$

Let $\tilde{Q}^l = \tilde{P}^{4L-l}$,

$$\tilde{P} = \frac{1}{L} \sum_{l=1}^L \tilde{P}^l \quad \text{and} \quad \tilde{Q} = \frac{1}{L} \sum_{l=1}^L \tilde{Q}^l.$$

Note that for each $l \in \{1, \ldots, L\}$, $\tilde{P}^l$ has two change points and $\Delta = \log^2(T)$. Furthermore,

$$\rho = 1/2 \quad \text{and} \quad \kappa_0 = \frac{2^{1/2} \xi^{1/2}}{n^{1/2}}.$$

As a result,

$$\kappa_0 \sqrt{n \rho \Delta} = 2 \xi^{1/2} \log(T),$$

which implies that $\tilde{P}^l \in \mathcal{P}$ and therefore $\tilde{Q}^l \in \mathcal{P}$ for all $l$.

It follows from Le Cam’s lemma that

$$\inf_{\tilde{P} \in \mathcal{P}} \sup_{\tilde{Q} \in \mathcal{P}} \mathbb{E}_P(H(\tilde{Q}, \eta(P))) \geq \frac{T}{4} \{1 - d_{TV}(\tilde{P}, \tilde{Q})\}.$$

Let $P^l$ be a finite-dimensional distribution of $\tilde{P}^l$ consisting of only the first $T/2$ time points and $P_0$ be the joint distribution of a collection of independent adjacency matrices $\{B(t)\}_{t=1}^T$ such that

$$\mathbb{E}\{(B(t))_{ij}\} = 1/2, \quad i, j \in \{1, \ldots, n\}, \quad t \in \{1, \ldots, T/2\}.$$

Due to the symmetry, we have

$$d_{TV}(\tilde{P}, \tilde{Q}) \leq 2d_{TV}(P, P_0),$$

where

$$P = \frac{4}{L} \sum_{l=1}^L P^l.$$

Since $d_{TV}(\cdot, \cdot) \leq \sqrt{\chi^2(\cdot, \cdot)}$, it suffices to bound $\chi^2(P, P_0)$. We have

$$\chi^2(P, P_0) = \left(\frac{4}{L}\right)^2 \left[ \sum_{l=1}^L \mathbb{E}_{P_0} \left( \frac{dP^l}{dP_0} \right) \left( \frac{dP^l}{dP_0} \right)^\top + (L/4)(L/4 - 1) \right] - 1$$

$$\leq \left(\frac{4}{L}\right)^2 \left[ L \exp(8 \xi n \log(T)) + (L/4)(L/4 - 1) \right] - 1$$

$$= (4/L) \left\{ \exp(8 \xi n \log(T)) - 1 \right\} = 4 \log(T)(T^{8\xi n} - T^{-1}),$$

where the inequality follows from Lemma S.1. Therefore, one can set $n$ to be fixed, and for any $0 < \zeta < 1/8n$, there exists a sufficiently large $T(\zeta)$ such that $\log(T)(T^{8\zeta n} - T^{-1}) = 64^{-1}$ and that concludes the proof. \hfill \Box
Proof of Lemma 2. Let \( \Theta(1), \Theta(2) \in \mathbb{R}^{n \times n} \) be such that for all \( i, j = 1, \ldots, n \), \( \Theta_{ij}(1) = \rho/2 \) and that \( \Theta_{ij}(2) = \rho/2 + \kappa_0 \rho \). Since \( \kappa_0 \leq 1/2 \), it holds that \( \|\Theta(2)\|_\infty \leq \rho \).

For \( \delta > 0 \) to be chosen later, let \( P_1^\delta \) be the joint distribution of a collection of independent adjacency matrices \( \{A(t)\}_{t=1}^T \) such that
\[
\mathbb{E}(A(t)) = \begin{cases} 
\Theta(1), & \text{if } t \leq T/2 + \delta, \\
\Theta(2), & \text{if } t > T/2 + \delta. 
\end{cases}
\]

Let \( P_2^\delta \) be the joint distribution of a collection of independent adjacency matrices \( \{B(t)\}_{t=1}^T \) such that
\[
\mathbb{E}(B(t)) = \begin{cases} 
\Theta(1), & \text{if } t \leq T/2, \\
\Theta(2), & \text{if } t > T/2. 
\end{cases}
\]

Then we have,
\[
2d_{TV}^2(P_1, P_2) \leq KL(P_1, P_2)
\]
\[
= \delta n^2 \left( \frac{\rho/2 + \kappa_0 \rho}{\rho/2} \log \left( \frac{\rho/2 + \kappa_0 \rho}{\rho/2} \right) + (1 - \rho/2 - \kappa_0 \rho) \log \left( \frac{1 - \rho/2 - \kappa_0 \rho}{1 - \rho/2} \right) \right)
\]
\[
\leq \delta n^2 \left( \frac{\rho/2 + \kappa_0 \rho}{\rho/2} \right)^2 + (1 - \rho/2 - \kappa_0 \rho) \frac{-\kappa_0 \rho}{1 - \rho/2}
\]
\[
= \delta n^2 \left( \kappa_0 \rho + 2\kappa_0^2 \rho - \kappa_0 \rho + \kappa_0^2 \rho^2 (1 - \rho/2)^{-1} \right) \leq 4\delta \kappa_0^2 n^2 \rho = 4\delta \kappa_0^2 n^2 \rho.
\]

Since
\[
\inf_{\hat{\eta}} \sup_{\eta \in \mathcal{P}} \mathbb{E}_P(\|\hat{\eta} - \eta\|) \geq \delta(1 - d_{TV}(P_1, P_2)),
\]

taking \( \delta = \frac{1}{8\kappa_0^2 n^2 \rho} \), we have
\[
\inf_{\hat{\eta}} \sup_{\eta \in \mathcal{P}} \mathbb{E}_P(\|\hat{\eta} - \eta\|) \geq \frac{1}{16\kappa_0^2 n^2 \rho}.
\]

\( \square \)

S.2. Proofs of technical results used in Theorem 1. Throughout this section, for notational convenience we set \( p = n^2 \) and assume \( p \sqrt{n} \geq \log(p) \). We admit the discrepancy with (2) – where we require \( \rho n \geq \log(n) \). This will only affect the constants.

Observe that in Section 2, no additional structure is imposed on the adjacency matrix. In addition, for two matrices \( A, B \in \mathbb{R}^{n \times n} \), we have
\[
(A, B) = \{\text{vec}(A)\}^\top \text{vec}(B),
\]
where \( \text{vec}(\cdot) \) is the vectorized version of a matrix by stacking the columns thereof. It, therefore, suffices to view \( A \) as a sparse Bernoulli vector with \( p = n^2 \) entries. The assumptions below are vector versions of Assumption 1. We include them here for brevity.

Assumption S.1. Let \( X(1), \ldots, X(T) \in \mathbb{R}^p \) be independent random vectors with independent Bernoulli entires. Suppose that the \( i \)th coordinate \( X_i(t) \) of \( X(t) \) satisfies \( \mathbb{E}(X_i(t)) = \mu_i(t) \) and that
\[
\max_{1 \leq i \leq T} \|\mu(t)\|_\infty \leq \rho.
\]
Note that in fact if $A$ is an adjacency matrix of an inhomogeneous Bernoulli network defined in Definition 1, then due to symmetry, there are in fact $p = n(n - 1)/2$ independent entries. In this section, for notational simplicity, we let $p = n^2$ which has the same order as $n(n - 1)/2$.

**Assumption S.2.** Let $\{\eta_k\}_{k=0}^{K+1} \subset \{1, \ldots, T + 1\}$ be a collection of change points, such that $1 = \eta_0 < \eta_1 < \ldots < \eta_K \leq T < \eta_{K+1} = T + 1$ and, for $t = 2, \ldots, T$, 

$$
\mu(t) \neq \mu(t - 1) \text{ if and only if } t \in \{\eta_1, \ldots, \eta_K\}.
$$

Assume the spacing $\Delta$ satisfy that

$$
\Delta := \min_{k=1, \ldots, K+1} \{\eta_k - \eta_{k-1}\} \leq T,
$$

and the normalized jump size $\kappa_0$ satisfies

$$
\inf_{k=1, \ldots, K} \| \mu(\eta_k) - \mu(\eta_k - 1) \| = \inf_{k=1, \ldots, K} \kappa_k \geq \kappa_0 \sqrt{p} > 0.
$$

**S.2.1. Probability bounds.** In this subsection, our task is to provide a probability bound for the event $A(s, e, t)$ defined in (24) to hold. The result is formally stated in Lemma S.4, and necessary technical details are provided in Lemmas S.2 and S.3.

Suppose $\{w_t\}_{t=1}^T \subset \mathbb{R}$ satisfies

$$
\sum_{t=1}^T w_t^2 = 1.
$$

**Lemma S.2.** Suppose Assumption S.1 holds. Let $v \in \mathbb{R}^p$ and $\{w_t\}_{t=1}^T \subset \mathbb{R}$ satisfy (1). Then for any $\varepsilon > 0$, we have

$$
P\left( \left| \sum_{i=1}^p v_i \sum_{t=1}^T w_t (X_i(t) - \mu_i(t)) \right| \geq \varepsilon \right) \leq \exp\left( - \frac{3/2 \varepsilon^2}{3 \rho \|v\|^2_2 + \varepsilon \max_{i=1}^p |v_i| \max_{t=1}^T |w_t|} \right).
$$

**Proof.** Observe that

$$
\mathbb{E}\left( \sum_{i=1}^p v_i \sum_{t=1}^T w_t (X_i(t) - \mu_i(t)) \right)^2 = \sum_{i=1}^p \sum_{t=1}^T v_i^2 w_t^2 \mathbb{E}(X_i(t) - \mu_i(t))^2 \leq \rho \|v\|^2_2,
$$

due to the independence assumption and the fact that $\sum_{t=1}^T w_t^2 = 1$, and that

$$
\max_{t=1, \ldots, T} |w_t v_i (X_i(t) - \mu_i(t))| \leq \frac{p}{\max_{i=1}^p |v_i|} \max_{t=1}^T |w_t|,
$$

since $X_i(t)$ is a Bernoulli random variable with mean $\mu_i(t)$. The desired result follows from Bernstein inequality.

**Lemma S.3.** Assume that the collection $\{Y(t)\}_{t=1}^T$ satisfies Assumption S.1. Let $v = \sum_{t=1}^T w_t (Y(t) - \mu(t)) \in \mathbb{R}^p$. Then there exists $C > 0$ depending on $c > 0$ such that

$$
P\left( \max_{1 \leq i \leq p} |v_i| \geq C \sqrt{\log(p) \vee \log(T)} \right) \leq T^{-c},
$$

and

$$
P\left( \|v\| \geq C \sqrt{\log(p) \vee \log(T) + \sqrt{p}} \right) \leq T^{-c},
$$
**Proof.** For the first part observe that it follows from Lemma 5.9 in Vershynin (2010) that there exists some absolute constant $C_1 > 0$ such that

$$
\|v_i\|_{\psi_2}^2 \leq C_1 \sum_{i=1}^{T} w_i^2 \|Y_i(t) - \mu_i(t)\|_{\psi_2}^2 \leq 2C_1,
$$

where $\| \cdot \|_{\psi_2}$ is the Orlicz norm (e.g. Definition 5.7 in Vershynin, 2010), and the second inequality follows from $\|Y_i(t) - \mu_i(t)\|_{\psi_2}^2 \leq 2$ and $\sum_{i=1}^{T} w_i^2 = 1$. Therefore for each $i = 1, \ldots, p$, $v_i$ is sub-Gaussian and there exist a constant $c > 0$ and a large enough $C > 0$ depending on $c$ and $C_1$ such that

$$
\mathbb{P} \left( v_i \geq C \sqrt{\log(p) \vee \log(T)} \right) \leq (p \vee T)^{-c},
$$

Since

$$
p(p \vee T)^{-c} \leq \begin{cases} T^{-c}, & p \leq T, \\ p^{-c} \leq T^{-c}, & p \geq T, \end{cases}
$$

the desired result follows from a union bound argument.

For the second part, define $F(x_1, \ldots, x_p) = \|x\|$ and $G_i(y_1, \ldots, y_l) = \sum_{t=1}^{T} w_i(y_t - \mu_i(t))$, $i = 1, \ldots, p$. Since for all $i$, both $F$ and $G_i$ are one Lipschitz function, $\|v\|$ is a one Lipschitz function of \{Y_i(t)\}_{i=1}^{p} \sum_{t=1}^{T}$. It follows from the proof of Corollary 4 in Samson (2000) that, for any $\varepsilon > 0$,

$$
\mathbb{P} (\|v\| > \mathbb{E}\|v\| + \varepsilon) \leq \exp \left( -\frac{\varepsilon^2}{2} \right).
$$

Since $\mathbb{E}\|v\| \leq \sqrt{ \sum_{i=1}^{p} \mathbb{E}(v_i^2)} \leq \sqrt{mp}$, the desired results follows by by taking $\varepsilon = C \sqrt{\log(p) \vee \log(T)}$.

**Lemma S.4.** Let \{X(t)\}_{t=1}^{T} and \{Y(t)\}_{t=1}^{T} be two independent copies, both of which satisfying Assumption S.1. Suppose in addition that

$$
\rho \sqrt{p} \geq \log(p).
$$

For \{w_i\}_{t=1}^{T} satisfying $\sum_{t=1}^{T} w_i^2 = 1$, let $\bar{X} = \sum_{t=1}^{T} w_t X(t)$, $\bar{Y} = \sum_{t=1}^{T} w_t Y(t)$ and $\bar{\mu} = \sum_{t=1}^{T} w_t \mu(t)$. There exists $C_\beta > 0$ depending on $c$ and $c_T$ such that

$$
\mathbb{P} \left( \left| \sum_{i=1}^{p} \bar{X}_i \bar{Y}_i - \sum_{i=1}^{p} \bar{\mu}_i^2 \right| \geq C_\beta \log(T) \left( \|\bar{\mu}\| + \log^{1/2}(T) \rho \sqrt{p} \right) \right) \leq 6T^{-cr} + 2T^{-c},
$$

where $C_\beta > \max\{4c_T/3, \sqrt{3c_T(C+1)^2 + C^2}\}$, and $C, c$ are from Lemma S.2.

**Proof.** Note that $\sum_{i=1}^{p} \bar{X}_i \bar{Y}_i - \sum_{i=1}^{p} \bar{\mu}_i^2 = I + II + III$, where

$$
I = \sum_{i=1}^{p} (\bar{X}_i - \bar{\mu}_i)(\bar{Y}_i - \bar{\mu}_i), \quad II = \sum_{i=1}^{p} (\bar{X}_i - \bar{\mu}_i)\bar{\mu}_i \quad \text{and} \quad III = \sum_{i=1}^{p} (\bar{Y}_i - \bar{\mu}_i)\bar{\mu}_i.
$$

It suffices to bound $I$ and $II$, due to the fact that \{X(t)\}_{t=1}^{T} and \{Y(t)\}_{t=1}^{T} are iid.

As for $I$, for any $i = 1, \ldots, p$, let $v_i = \sum_{t=1}^{T} w_t (Y_i(t) - \mu_i(t))$. Conditional on \{Y(t)\}_{t=1}^{T}, it follows from Lemma S.2 that for any $\varepsilon > 0$, we have

$$
\mathbb{P}_{X|Y} \left( \left| \sum_{i=1}^{p} v_i \sum_{t=1}^{T} w_t (X_i(t) - \mu_i(t)) \right| \geq \varepsilon \right) \leq 2 \exp \left( -\frac{3/2\varepsilon^2}{3\rho \|v\|^2 + \varepsilon \max_i |v_i|} \right),
$$
due to the fact that $\max_t |w_t| \leq 1$. By Lemma S.3, there exist $C, c > 0$ such that
\[
P_Y \left( \max_{i=1,\ldots,p} |v_i| \geq C\sqrt{\log(p) \lor \log(T)} \right) \leq T^{-c},
\]
and that
\[
P_Y \left( \|v\| \geq C\sqrt{\log(p) \lor \log(T)} + \sqrt{p\rho} \right) \leq T^{-c}.
\]
Thus for any $\varepsilon > 0$, it holds that
\[
P_{X,Y} \left( \left| \sum_{i=1}^{p} v_i \sum_{t=1}^{T} w_t(X_i(t) - \mu_i(t)) \right| \geq \varepsilon \right)
\leq 2\exp \left( -\frac{3/2\varepsilon^2}{3\rho(C\sqrt{\log(p) \lor \log(T)} + \sqrt{p\rho})^2 + C\varepsilon\sqrt{\log(p) \lor \log(T)}} \right) + 2T^{-c}.
\]
Since $\rho\sqrt{\rho} \geq \log(p)$, by taking $\varepsilon = C''\rho\sqrt{\rho}\log^{3/2}(T)$ for sufficiently large
\[
C'' \geq \sqrt{3c_T(C + 1)^2 + C^2},
\]
it holds that
\[
P(|I| \geq C''\rho\sqrt{\rho}\log^{3/2}(T)) \leq 2T^{-c_T} + 2T^{-c}.
\]
Observe that $III$ is identically distributed as $II$. For $II$, observe that for $\varepsilon > 0$, it follows from Lemma S.2,
\[
P \left( \left| \sum_{i=1}^{p} \mu_i \sum_{t=1}^{T} w_t(X_i(t) - \mu_i(t)) \right| \geq \varepsilon \right) \leq 2\exp \left( -\frac{3/2\varepsilon^2}{3\rho\|\bar{\mu}\|^2 + \varepsilon \max_t |\mu_i| \max_t |w_t|} \right).
\]
Let $\varepsilon = C'\|\bar{\mu}\| \log(T), \text{ with } C' > 4c_T/3,$
\[
3\rho\|\bar{\mu}\|^2 + \varepsilon \max_t |\mu_i| \max_t |w_t| \leq 3\rho\|\bar{\mu}\|^2 + \varepsilon \rho \leq 3\|\bar{\mu}\|^2 + \varepsilon \leq 3/(2c_T)\varepsilon^2/\log(T).
\]
Therefore $P(|II| \geq C'\|\bar{\mu}\| \log(T)) \leq 2T^{-c_T}$.
\]
\[
S.2.2. Localization. \text{ This is the key lemma used in the proof of Theorem 1 to localize the change points. We deliberately present this lemma with seemingly low-level conditions, in order for us to directly check the conditions in the proof of Theorem 1.}
\]
\[
\text{LEMMA S.5. Assume } \{X_i\}_{i=1}^{T} \text{ and } \{Y_i\}_{i=1}^{T} \text{ be two independent copies } \mathbb{E}(X_i) = \mathbb{E}(Y_i) = \mu(t) \text{ such that Assumption S.2 holds.}
\]
\[
\text{Let } [s_0, e_0] \text{ be any interval with } e_0 - s_0 \leq C_R \Delta \text{ and containing at least one change point } \eta_r \text{ such that}
\]
\[
\eta_{r-1} \leq s \leq \eta_r \leq \ldots \leq \eta_{r+q} \leq e \leq \eta_{r+q+1}, \quad q \geq 0
\]
\[
\text{and that } \min\{s_0 - \eta_r, e_0 - \eta_{r+q}\} \geq \Delta/2. \text{ Denote } \kappa^{s,e}_{\text{max}} = \max\{\kappa_p : \min\{\eta_p - s_0, e_0 - \eta_p\} \geq \Delta/16\}.
\]
\[
\text{Consider any generic } [s, e] \subset [s_0, e_0] \text{ such that } [s, e] \text{ contains at least one change point. Let } b \in \arg\max_{s \leq t \leq e}(\bar{X}^{s,e}(t), \bar{Y}^{s,e}(t)). \text{ For some } c_4 > 0 \text{ and } \lambda > 0, \text{ suppose that}
\]
\[
(b) \qquad (\bar{X}^{s,e}(b), \bar{Y}^{s,e}(b)) \geq c_4(\kappa^{s,e}_{\text{max}})^2 \Delta
\]
\[
(3) \qquad \sup_{s < t < e} \| (\bar{X}^{s,e}(t), \bar{Y}^{s,e}(t)) - \|\bar{\mu}^{s,e}(t)\|^2 \| \leq \lambda
\]
If there exists a sufficiently small absolute constant $c_5 > 0$ satisfying
\[
c_5 < \min \left\{ \frac{c_3}{2C_R^2 + 2c_3}, \frac{1}{2 + 32C_R^2 \min\{1/4, 1/2 - 2c_3\}} \right\}
\]
with $c_3$ defined in Lemma S.13, such that
\[
\lambda \leq c_5 \max_{s < t < e} \|\tilde{\mu}^{s,e}(t)\|_2^2
\]
then there exists a change point $\eta_k \in (s, e)$ such that
\[
\min\{e - \eta_k, \eta_k - s\} > \Delta/4, \quad |\eta_k - b| \leq \frac{C_3 \Delta \lambda}{\|\tilde{\mu}^{s,e}(\eta_k)\|_2^2} \quad \text{and} \quad \|\tilde{\mu}^{s,e}(\eta_k)\|_2^2 \geq (1 - 2c_5) \max_{s < t < e} \|\tilde{\mu}^{s,e}(t)\|_2^2,
\]
where $C_3 = 2C_R^2 / \min\{1/4, 1/2 - 2c_3\}$.

**Proof.** For any $t \in \{s+1, \ldots, e-1\}$, denote $\tilde{Z}^{s,e}(t) = (\tilde{X}^{s,e}(t), \tilde{Y}^{s,e}(t))$. It follows from Proposition S.1 that without loss of generality, we can assume $\|\tilde{\mu}^{s,e}(t)\|^2$ is locally decreasing at $b$. Observe that this implies that there exists a change point $\eta_k \in [s, b]$, since otherwise $\|\tilde{\mu}^{s,e}(t)\|_2^2$ is increasing on $[s, b]$ as a consequence of Lemma S.16. Therefore, we have
\[
s \leq \eta_k \leq b \leq e.
\]
Observe that
\[
\tilde{Z}^{s,e}(b) \geq \max_{s < t < e} \|\tilde{\mu}^{s,e}(t)\|_2^2 - \lambda \geq c_5^{-1}(1 - c_5)\lambda,
\]
which follows from (3) and (4), and
\[
\|\tilde{\mu}^{s,e}(b)\|_2^2 \geq \tilde{Z}^{s,e}(b) - \lambda \geq \max_{s < t < e} \|\tilde{\mu}^{s,e}(t)\|_2^2 - 2\lambda \geq c_5^{-1}(1 - 2c_5)\lambda,
\]
which follows from (5). We, consequently, have
\[
\|\tilde{\mu}^{s,e}(b)\|_2^2 \geq \tilde{Z}^{s,e}(b) - \lambda \geq (1 - c_5(1 - c_5)^{-1})\tilde{Z}^{s,e}(b) > \tilde{Z}^{s,e}(b)/2 \geq (c_4/2)(\kappa_{\max}^{s,e})^2 \Delta.
\]
where the second inequality follows from (5) and the last inequality follows from (2).

Since $s \leq \eta_k \leq b \leq e$ and $\|\tilde{\mu}^{s,e}(t)\|_2^2$ is locally decreasing at $b$, by Proposition S.1, $\|\tilde{\mu}^{s,e}(t)\|_2^2$ is decreasing within $[\eta_k, b]$. Therefore
\[
\|\tilde{\mu}^{s,e}(\eta_k)\|_2^2 \geq \|\tilde{\mu}^{s,e}(b)\|_2^2.
\]
Equation (8) combining with (6) gives
\[
\|\tilde{\mu}^{s,e}(\eta_k)\|_2^2 \geq (1 - 2c_5) \max_{s < t < e} \|\tilde{\mu}^{s,e}(t)\|_2^2.
\]

**Step 1.** In this step, it will be shown that $\min\{\eta_k - s, e - \eta_k\} \geq \min\{1, c_4\}\Delta/16$.

Suppose $\eta_k$ is the only change point in $(s, e)$. It must hold that $\min\{\eta_k - s, e - \eta_k\} \geq \min\{1, c_4\}\Delta/16$, otherwise by Lemma S.15,
\[
\|\tilde{\mu}^{s,e}(\eta_k)\|_2^2 = \frac{(\eta_k - s)(e - \eta_k)}{e - s} \kappa_2^2 \leq \frac{c_4}{16} \kappa_2^2 \Delta \leq \frac{c_4}{2}(\kappa_{\max}^{s,e})^2 \Delta,
\]
which contradicts (7).
Suppose \([s, e]\) contains at least two change points. For the sake of contradiction, suppose \(\min\{\eta_k - s, e - \eta_k\} < \min\{1, c_4\} \Delta / 16\). Reversing the time series if necessary, it suffices to consider

\[
\eta_k - s < \min\{1, c_4\} \Delta / 16.
\]

Observe that (9) implies that \(\eta_k\) is the first change point in \([s, e]\). Therefore

\[
\|\tilde{\mu}^{s,e}(\eta_k)\|^2 \leq \frac{1}{8}\|\tilde{\mu}^{s,e}(\eta_{k+1})\|^2 + 4\kappa^2(\eta_k - s) \leq \frac{1}{8} \max_{s < t < e} \|\tilde{\mu}^{s,e}(t)\|^2 + \frac{c_4}{4} \kappa^2 \Delta
\]

\[
\leq \frac{1}{8}(1 - 2c_5)^{-1}\|\tilde{\mu}^{s,e}(b)\|^2 + \frac{1}{2}\|\tilde{\mu}^{s,e}(b)\|^2 < \|\tilde{\mu}^{s,e}(b)\|_2,
\]

where the first inequality follows from Lemma S.17 and (9), the second inequality follows from (9), the third inequality follows from (6) and (7), and the fourth inequality follows from \(c_5 < 3/8\). This contradicts (8).

**Step 2.** In order to apply Lemma S.13, it suffices to check that (17) for \(\tilde{\mu}^{s,e}(t)\). Observe that

\[
\max_{s < t < e} \|\tilde{\mu}^{s,e}(t)\|^2 - \|\tilde{\mu}^{s,e}(\eta_k)\|^2 \leq 2c_5 \max_{s < t < e} \|\tilde{\mu}^{s,e}(t)\|^2 \leq 2c_5(1 - 2c_5)^{-1}\|\tilde{\mu}^{s,e}(\eta_k)\|^2
\]

\[
\leq \frac{2c_5 c_R^2}{c_3(1 - 2c_5)} c3\|\tilde{\mu}^{s,e}(\eta_k)\|^2 \Delta^2(e - s)^{-2} \leq c_3\|\tilde{\mu}^{s,e}(\eta_k)\|^2 \Delta^2(e - s)^{-2},
\]

where \(c_3\) is defined as in (17), the first and the second inequality follow from (6), the third inequality follows from \(e - s \leq C_R \Delta\) and the last inequality hold for sufficiently small

\[
0 < c_5 < \frac{c_3}{2C_R^2 + 2c_3}.
\]

Let \(c\) be defined in Lemma S.13. Since \(e - s \leq C_R \Delta\),

\[
\frac{2\lambda(e - s)^2}{c\Delta\|\tilde{\mu}^{s,e}(\eta_k)\|^2} \leq \frac{2c_2 c_R^2}{cc_5^2 (1 - c_5) \lambda} \leq \frac{\lambda}{\Delta} < 1/16,
\]

where the first inequality follows from (6) and the last inequality holds for sufficiently small \(c_5\) satisfying

\[
c_5 < \frac{1}{2 + 32C_R^2 / c}.
\]

By Lemma S.13 if \(d\) is chosen such that

\[
d - \eta_k = \frac{2\lambda(e - s)^2}{c\Delta\|\tilde{\mu}^{s,e}(\eta_k)\|^2} \leq \Delta / 16,
\]

and that

\[
\|\tilde{\mu}^{s,e}(\eta_k)\|^2 - \|\tilde{\mu}^{s,e}(d)\|^2 > c\|\tilde{\mu}^{s,e}(\eta_k)\|^2\|d - \eta_k\|\Delta(e - s)^{-2} \geq 2\lambda,
\]

where the first inequality follows from Lemma S.13 and the second inequality follows from (10).

For the sake of contradiction, suppose \(b \geq d\). Then

\[
\|\tilde{\mu}^{s,e}(b)\|^2 < \|\tilde{\mu}^{s,e}(d)\|^2 < \|\tilde{\mu}^{s,e}(\eta_k)\|^2 - 2\lambda \leq \max_{s < t < e} \|\tilde{\mu}^{s,e}(t)\|^2 - 2\lambda \leq \max_{s < t < e} \tilde{Z}(t) + \lambda - 2\lambda = \tilde{Z}(b) - \lambda,
\]
where the first inequality follows from Proposition S.1, which ensures that \( \| \tilde{\mu}^{s,e}(t) \|^2 \) is decreasing on \([\eta_k, b]\) and \( d \in [\eta_k, b] \), the second inequality follows from (11). This is a contradiction to (3). Thus \( b \leq d \) and so

\[
0 \leq b - \eta_k \leq d - \eta_k \leq \frac{2\lambda(e - s)^2}{c\Delta \| \tilde{\mu}^{s,e}(\eta_k) \|^2} \leq \frac{2C^2_R}{c} \frac{\Delta \lambda}{\| \tilde{\mu}^{s,e}(\eta_k) \|^2}
\]

where the third inequality follows from \( e - s \leq C_R \Delta \).

\[ \square \]

S.3. Proofs of technical results used in Theorems 2 and 3.

S.3.1. Matrix estimation. We first establish some results concerning matrix estimation.

**Lemma S.6.** 1. Let \( \{A(t)\}_{t=1}^T \) be a collection of independent matrices with independent Bernoulli entries satisfying

\[
\max_{1 \leq t \leq T} \| \mathbb{E} A(t) \|_\infty \leq \rho,
\]

with \( n\rho \geq \log(n) \). Let \( \{w(t)\}_{t=1}^T \subset \mathbb{R} \) be a collection of scalars such that \( \sum_{t=1}^T w(t)^2 = 1 \) and \( \sum_{t=1}^T w(t) = 0 \). Then there exists an absolute constant \( C > 32 \times 2^{1/2} \varepsilon^2 \) such that

\[
\mathbb{P} \left( \left\| \sum_{t=1}^T w(t)A(t) - \mathbb{E} \left( \sum_{t=1}^T w(t)A(t) \right) \right\|_{\text{op}} \geq C \sqrt{n\rho} + \varepsilon \right) \leq \exp(-\varepsilon^2/2).
\]

2. If \( \{A(t)\}_{t=1}^T \) are symmetric matrices, then (12) still holds.

**Proof.** Observe that the conclusion in 2 is a consequence of that in 1, as if \( A(t) \) is symmetric, then \( A(t) = A'(t) + A''(t) \), where \( A'(t) \) is the upper diagonal matrix of \( A(t) \) including the diagonal and \( A''(t) \) is the lower diagonal matrix of \( A(t) \). Therefore the conclusion in 2 follows by applying the conclusion in 1 to \( A'(t) \) and \( A''(t) \). In the rest of this proof, we will only consider 1.

Let \( B(t) = A(t) - \mathbb{E}(A(t)) \) and \( \tilde{B} = \sum_{t=1}^T w(t)B(t) \). The function

\[
H(B(1), \ldots, B(T)) = \left\| \sum_{t=1}^T w(t)B(t) \right\|_{\text{op}} = \left\| \tilde{B} \right\|_{\text{op}}
\]

is one-Lipschitz, therefore by Corollary 4 in Samson (2000), one has for any \( \varepsilon > 0 \),

\[
\mathbb{P} \left( \left\| \sum_{t=1}^T w(t)B(t) \right\|_{\text{op}} \geq \mathbb{E} \left( \left\| \sum_{t=1}^T w(t)B(t) \right\|_{\text{op}} \right) + \varepsilon \right) \leq \exp(-\varepsilon^2/2).
\]

To complete the argument, it suffices to bound \( \mathbb{E} \left( \left\| \sum_{t=1}^T w(t)B(t) \right\|_{\text{op}} \right) \). By Lemma S.7 and for all \( t \in \{1, \ldots, T\} \), the entries of \( w(t)B(t) \) are bounded on \([-\rho, 1]\), there exists a collection of random matrices \( \{Z(t)\}_{t=1}^T \subset \mathbb{R}^{n \times n} \) such that \( \mathbb{E}(Z|B) = B \), where \( Z = (Z(1), \ldots, Z(T)) \) and
\[ B = (B(1), \ldots, B(T)), \text{ and that } (Y_t)_{ij} = (1 - \rho)(Z(t))_{ij} + \rho \text{ are mutually independent Bernoulli random variables with parameter } \rho. \text{ Denote } G(B) = \| \sum_{t=1}^{T} w(t)B(t) \|_{op}. \text{ Then}
\]

\[
\mathbb{E} \left( \left\| \sum_{t=1}^{T} w(t)B(t) \right\|_{op} \right) = \mathbb{E} (G(B)) = \mathbb{E} (G(\mathbb{E}(Z|B))) \leq \mathbb{E} (G(Z)|B) = \mathbb{E} \left( \left\| \sum_{t=1}^{T} w(t)Z(t) \right\|_{op} \right)
\]

\[
= \frac{1}{1 - \rho} \mathbb{E} \left( \left\| \sum_{t=1}^{T} w(t)Y(t) \right\|_{op} \right),
\]

where \( G \) being convex is used in the inequality and \( \sum_{t=1}^{T} w(t) = 0 \) is used in the last equality. Since the entries of \( \sum_{t=1}^{T} w(t)Y(t) \) are independent and identically distributed, by Lemma S.8,

\[
\mathbb{E} \left( \left\| \sum_{t=1}^{T} w(t)B(t) \right\|_{op} \right) \leq C \sqrt{m},
\]

where \( C > 32 \times 2^{1/4}e^2 \).

**Lemma S.7.** Let \( X \in [-\rho, 1] \) be a centered Bernoulli random variable. Then there exists a random variable \( Y \) such that

- \( \mathbb{E}(Y|X) = X \), and
- \( (1 - \rho)Y + \rho \) is a Bernoulli random variable with parameter \( \rho \).

**Proof.** The proof is taken from the proof of Lemma 2 in Tomozei and Massoulié (2014), by letting

\[
Y = 1 - \mathbb{1}\{X \leq (1 + \rho)U - \rho\},
\]

where \( U \) is a Uniform[0,1] random variable independent with \( X \).

**Lemma S.8.** Let \( \{A(t)\}_{t=1}^{T} \) be a collection of independent adjacency matrices whose entries are independent Bernoulli random variables with parameter \( \rho \) satisfying with \( n\rho \geq c_2 \log(n) \), \( c_2 > 4 \), and let \( B_t = A_t \) - \( E(A_t) \). Suppose \( \{w_t\}_{t=1}^{T} \subset \mathbb{R} \) be a collection of scalar such that \( \sum_{t=1}^{T} w_t^2 = 1 \). Then there exists an absolute constant \( C > 32 \times 2^{1/4}e^2 \) such that

\[
\mathbb{E} \left( \left\| \sum_{t=1}^{T} w_tB_t \right\|_{op} \right) \leq C \sqrt{n\rho}.
\]

**Proof.** Let \( \widetilde{B} = \sum_{t=1}^{T} w_tB(t) \). To bound \( \mathbb{E}(\|\widetilde{B}\|_{op}) \), since the entries of \( \widetilde{B} \) are independent and identically distributed with \( \mathbb{E}(\widetilde{B}) = 0 \), by Corollary 2.2 in Segnier (2000), one has

\[
\mathbb{E} \left( \|\widetilde{B}\|_{op} \right) \leq C_1 \mathbb{E} \left( \max_{1 \leq i \leq n} \|\widetilde{B}_{is}\| \right),
\]

where \( C_1 = 16 \times 2^{1/4}e^2 \). For any \( i \in \{1, \ldots, n\} \), since \( \|\widetilde{B}_{is}\| \) is one-Lipschitz convex function, by Corollary 4 in Samson (2000), it holds that for any \( \varepsilon > 0 \),

\[
\mathbb{P} \left( \|\widetilde{B}_{is}\| \geq \mathbb{E}\|\widetilde{B}_{is}\| + \varepsilon \right) \leq \exp(-\varepsilon^2/2).
\]
Since
\[(\mathbb{E}||\tilde{B}_{1*}||^2)^2 \leq \mathbb{E}(||\tilde{B}_{1*}||^2) = \sum_{t=1}^{T} w_t^2 \mathbb{E}(||(B(t))_{1*}||^2) + \sum_{s \neq t} w_s w_t \mathbb{E}(B_s, B_t) = \sum_{t=1}^{T} w_t^2 n \rho (1 - \rho) \leq n \rho ,\]
one has
\[(13) \quad \mathbb{P}\left( ||\tilde{B}_{1*}|| \geq \sqrt{n \rho} + \varepsilon \right) \leq \exp(-\varepsilon^2 / 2).\]

Using the above display, it follows that
\[
\mathbb{E}\left( \max_{1 \leq i \leq n} ||\tilde{B}_{1*}|| \right) = \int_0^\infty \mathbb{P}\left( \max_{1 \leq i \leq n} ||\tilde{B}_{1*}|| \geq \varepsilon \right) d\varepsilon \leq \int_0^{2\sqrt{n \rho}} 1 d\varepsilon + \int_{2\sqrt{n \rho}}^\infty n \mathbb{P}(||\tilde{B}_{1*}|| \geq \varepsilon) d\varepsilon
\]
\[
= 2\sqrt{n \rho} + \int_{2\sqrt{n \rho}}^\infty n \mathbb{P}(||\tilde{B}_{1*}|| \geq \varepsilon + \sqrt{n \rho}) d\varepsilon \leq 2\sqrt{n \rho} + \int_{\sqrt{n \rho}}^\infty n \exp(-\varepsilon^2 / 2) d\varepsilon
\]
\[
\leq 2\sqrt{n \rho} + \frac{1}{\sqrt{n \rho}} \int_{\sqrt{n \rho}}^\infty n \varepsilon \exp(-\varepsilon^2 / 2) d\varepsilon \leq 2\sqrt{n \rho} + n^{1-c_2 / 2} \frac{1}{\sqrt{c_2 \log(n)}} < C_2 \sqrt{n \rho},
\]
where \(C_2 > 2\), the first inequality follows from the observation that \(||\tilde{B}_{1*}||\) are identically distributed, the second inequality follows from \((13)\) and the last inequality follows from \(\rho n \geq c_2 \log(n), c_2 > 2\).

\(\square\)

Lemmas S.9 and S.10 are from Lemma 1 in Xu (2018).

**Lemma S.9.** Let \(A, B \in \mathbb{R}^{n \times n}\) be two symmetric matrices with \(\|A - B\|_{\text{op}} < \tau / (1 + \delta), \tau > 0\). Then for a fixed \(\delta < 1\), we have
\[
||\text{USVT}(A, \tau, \infty) - B||^2_F \leq 16 \min_{0 \leq r \leq n} \left\{ r \tau^2 + (1 + \delta)^2 \delta^{-2} \sum_{i=r+1}^{n} \lambda_i^2(B) \right\}.
\]

**Lemma S.10.** Let \(A\) and \(B\) be defined as in Lemma S.9, and that \(\|B\|_{\infty} \leq \tau\), then
\[
||\text{USVT}(A, \tau, \tau') - B||^2_F \leq 16 \min_{0 \leq r \leq n} \left\{ r \tau^2 + (1 + \delta)^2 \delta^{-2} \sum_{i=r+1}^{n} \lambda_i^2(B) \right\}.
\]

S.3.2. Proofs of technical results used in Theorem 3.

**Lemma S.11.** Suppose \(A, \Gamma \in \mathbb{R}^{n \times n}\) are symmetric matrices, satisfying that the entries of \(A\) are Bernoulli random variables, \(||\Gamma||_{\infty} \leq \rho\) and \(||A - (\Gamma - \text{diag}(\Gamma))||_{\text{op}} \leq (1 + \delta)\tau\). Then
\[
||\text{USVT}(A, \tau, \infty) - \Gamma||^2_F \leq 16 \min_{0 \leq r \leq n} \left\{ r \tau^2 + 2(1 + \delta)^2 \delta^{-2} \sum_{i=r+1}^{n} \lambda_i \right\} + 32(1 + \delta)^2 \delta^{-2} ||\text{diag}(\Gamma)||^2_F,
\]
where \(\{\lambda_i\}_{i=1}^{n}\) are the eigenvalues of \(\Gamma\) ordered in decreasing absolute values.
Proof. Let \( \lambda_i^r \) be the eigenvalues of \( \Gamma - \text{diag}(\Gamma) \) ordered in absolute value, \( \lambda_i \) be the eigenvalues of \( \Gamma \) ordered in absolute value and \( v_i \) be the eigenvectors of \( \Lambda \). Observe that for any orthonormal basis \( \{u_i\}_{i=1}^n \) and any \( r = 1, \ldots, n-1 \),

\[
\sum_{i=r+1}^n (\lambda_i^r)^2 \leq \sum_{i=r+1}^n u_i^T (\Gamma - \text{diag}(\Gamma))^2 u_i.
\]

By Lemma S.9, one has

\[
\| \text{USVT}(A, \tau, \infty) - (\Gamma - \text{diag}(\Gamma)) \|_F^2 \leq 16 \min_{0 \leq r \leq n} \left\{ r \tau^2 + (1 + \delta)^2 \delta^{-2} \sum_{i=r+1}^n (\lambda_i^r)^2 \right\}.
\]

For any \( r = 1, \ldots, n \),

\[
\sum_{i=r+1}^n (\lambda_i^r)^2 \leq \sum_{i=r+1}^n v_i^T (\Gamma - \text{diag}(\Gamma))^2 v_i - v_i^T \Gamma^2 v_i + \sum_{i=r+1}^n \lambda_i^2
\]

\[
= \sum_{i=r+1}^n v_i^T (-2\Gamma \text{diag}(\Gamma) + \text{diag}(\Lambda)^2) v_i + \sum_{i=r+1}^n \lambda_i^2 \leq \sum_{i=r+1}^n \| \Gamma v_i \|_2^2 + 2v_i^T \text{diag}(\Gamma)^2 v_i + \sum_{i=r+1}^n \lambda_i^2
\]

\[
\leq 2 \sum_{i=r+1}^n \lambda_i^2 + 2\| \text{diag}(\Gamma) \|_F^2,
\]

which leads to the desired results. \( \square \)

S.4. Properties of the population CUSUM statistics. Recall that in Definition 1 we introduced a general version of CUSUM statistics, which can be applied to various types of data. In Sections S.4.1 and S.4.2, we apply Definition 1 to vectors and scalars respectively.

S.4.1. Vector CUSUM.

Assumption S.3. Let \( \{V(t)\}_{t=1}^T \subset \mathbb{R}^p \). Assume there exists \( \{\nu_m\}_{m=0}^M \subset \{1, \ldots, T\} \) such that

\[
1 = \nu_0 < \nu_1 < \ldots < \nu_M \leq T < \nu_{M+1} = T + 1
\]

and, for \( t = 2, \ldots, T \),

\[
V(t) \neq V(t-1) \text{ if and only if } t \in \{\nu_1, \ldots, \nu_M\}.
\]

Let \( \inf_{m=1, \ldots, M} \| V(\nu_m) - V(\nu_m - 1) \| = \inf_{m=1, \ldots, M} \kappa_m \geq \kappa = \kappa_0 \sqrt{p} \).

The results in this subsection are used in the proofs of the main theorems. Below, \( \{V(t)\}_{t=1}^T \) corresponds to \( \{\mu(t)\}_{t=1}^T \) as defined in Assumption 1, and \( \kappa = \kappa_0 \sqrt{p} \) (see Assumption 1). For brevity, we introduce new notation in this subsection such that it is self-contained within this subsection.

For \( 0 \leq s < t < c \leq T \), denote the CUSUM statistics

\[
\tilde{V}_{s,c}(t) = \sum_{r=s+1}^t \frac{t - r}{(e - s)(e - t)} V(r) - \sum_{r=t+1}^c \frac{r - s}{(e - s)(e - t)} V(r).
\]

For simplicity denote \( \tilde{V}(t) = \tilde{V}_{0,T}(t) \). It is desired to show that this vector version CUSUM statistics have the same properties as the univariate CUSUM statistic.
Remark S.2. The CUSUM statistic defined in (14) is translational invariant. In other words, let \( W \in \mathbb{R}^p \) and \( U(t) = V(t) + W \) for all \( t \), then
\[
\tilde{V}(t) = \tilde{U}(t).
\]
Consequently it can be assumed that \( \sum_{t=1}^{T} V(t) = 0 \), and
\[
(15) \quad \tilde{V}(t) = \left( \frac{1}{T} \sum_{r=1}^{t} V(r) - \frac{t}{T} \sum_{r=1}^{T} V(r) \right) / \sqrt{\frac{t(T-t)}{T}} = \left( \frac{1}{T} \sum_{r=1}^{T} V(r) \right) / \sqrt{\frac{t(T-t)}{T}}.
\]

Proposition S.1. The quantity \( ||\tilde{V}(t)||^2 \) is maximized at the change points. For \( t \in [\nu_{m-1}, \nu_m] \), \( ||\tilde{V}(t)||^2 \) is either monotone or decreases and then increases.

Proof. Let \( t \in (\nu_{m-1}, \nu_m) \). By Equation (2.7) of Lemma 2.2 in Venkatraman (1992), for every \( j = 1, \ldots, p \), \( \tilde{V}_j(t) \) can be continuously extended to the function
\[
f_j(x) = \frac{a_j - b_j x}{x(1-x)},
\]
where \( x = t/T \), \( a_j \) and \( b_j \) are defined similarly as in Lemma 2.2 in Venkatraman (1992). Thus it suffices to show that for \( x \in (c, d) \) where \( 0 \leq c \leq d \leq 1 \), the function
\[
f(x) = \sum_{j=1}^{p} \frac{(a_j - b_j x)^2}{x(1-x)}
\]
is maximized at either \( c \) or \( d \).

Let
\[
f'(x) = \sum_{j=1}^{n} \frac{-2a_j x - b_j x - a_j (b_j x - a_j)}{(x-1)^2 x^3} = \frac{g(x)}{(x-1)^2 x^2}.
\]
The desired result follows if \( f'(x) \) is either nonpositive, or nonnegative or that there exists \( x_0 \in (0, 1) \) such that
\[
f'(x) \begin{cases} \leq 0 & \text{when } x \leq x_0 \ \\
 \geq 0 & \text{when } x \geq x_0 \end{cases}
\]
Since \( (x-1)^2 x \geq 0 \) for all \( x \in (0, 1) \). Observe that \( g \) is quadratic and that \( g(0) = -\sum_{i=1}^{n} a_i^2 \leq 0 \) and \( g(1) = (b_i x - a_i)^2 \geq 0 \). Therefore \( g(x) \) can have at most one root in \((c, d)\). If \( g(x) \) has no root in \((c, d)\), then \( g(x) \) is either positive or negative. If \( g(x) \) has a root \( x_0 \in (c, d) \), then (16) holds. \( \square \)

Lemma S.12. Suppose there exists a change point \( \nu \in (0, T) \) such that any other change point \( \nu' \) within \((0, T)\) satisfies \( \min\{|\nu' - \nu|\} \geq \Delta \). Then
\[
\max_{0 < \iota < T} ||\tilde{V}(t)||^2 \geq \frac{||V(\nu) - V(\nu + 1)||^2 \Delta^2}{48T}.
\]

Proof. Denote \( \kappa = ||V(\nu) - V(\nu + 1)|| \).
Step 1. Let

\[ I_1 = \left\{ i : \left| \sum_{r=1}^{\nu} V_i(r) \right| \geq \Delta |V_i(\nu) - V_i(\nu + 1)|/4 \right\}, \]

\[ I_2 = \left\{ i : \left| \sum_{r=1}^{\nu} V_i(r) \right| \geq \Delta |V_i(\nu) - V_i(\nu + 1)|/4 \right\}, \]

\[ I_3 = \left\{ i : \left| \sum_{r=1}^{\nu + \Delta} V_i(r) \right| \geq \Delta |V_i(\nu) - V_i(\nu + 1)|/4 \right\}, \]

Then by Lemma S.18, \( I_1 \cup I_2 \cup I_3 = \{1, \ldots, p\} \). We have

\[ \sum_{l=1}^{3} \left\{ \sum_{i \in I_l} (V_i(\nu) - V_i(\nu + 1))^2 \right\} \geq \sum_{i=1}^{p} (V_i(\nu) - V_i(\nu + 1))^2 = \kappa^2, \]

which implies that

\[ \max_{l=1,2,3} \left\{ \sum_{i \in I_l} (V_i(\nu) - V_i(\nu + 1))^2 \right\} \geq \kappa^2/3. \]

Without loss of generality, suppose \( \sum_{i \in I_1} (V_i(\nu) - V_i(\nu + 1))^2 \geq \kappa^2/3 \). Then

\[ \max_{1 < t < T} \left\| \mathbf{\tilde{V}}(t) \right\|^2 \geq \left\| \mathbf{\tilde{V}}(\nu - \Delta) \right\|^2 = \frac{T}{(\nu - \Delta)(T - (\nu - \Delta))} \left[ \sum_{r=1}^{\nu - \Delta} V(r) \right]^2 \]

\[ \geq \frac{1}{T} \sum_{i \in I_1} \left( \sum_{r=1}^{\nu - \Delta} V_i(r) \right)^2 \geq \frac{1}{T} \sum_{i \in I_1} \left( \sum_{r=1}^{\Delta} |V_i(\nu) - V_i(\nu + 1)|/4 \right)^2 \]

\[ \geq \frac{\Delta^2}{48T} \kappa^2, \]

where the first equality follows from (15) and the second last inequality follows from the definition of \( I_1 \).

\[ \tag{15} \]

**Lemma S.13.** Let \([s, e] \subset [0, T]\) be any generic interval containing a change point \( \nu \) satisfying

\[ \min\{\nu - s, e - \nu\} \geq c_1 \Delta. \]

If

\[ \left\| \mathbf{\tilde{V}}^{s,e}(\nu) \right\|^2 \geq \kappa^2 \Delta^2(e - s)^{-1}, \]

and there exists a sufficient small absolute constant \( c_3 > 0 \) such that

\[ \tag{17} \max_{1 < t < T} \left\| \mathbf{\tilde{V}}^{s,e}(t) \right\|^2 - \left\| \mathbf{\tilde{V}}^{s,e}(\nu) \right\|^2 \leq c_3 \left\| \mathbf{\tilde{V}}^{s,e}(\nu) \right\|^2 \Delta^2(e - s)^{-2}, \]

then there exists an absolute constant \( c, c_1 > 0 \) such that \( d \in [s, e] \) satisfying \( |d - \nu| \leq c_1 \Delta/16 \), and

\[ \left\| \mathbf{\tilde{V}}^{s,e}(\nu) \right\|^2 - \left\| \mathbf{\tilde{V}}^{s,e}(d) \right\|^2 \geq c \left\| \mathbf{\tilde{V}}^{s,e}(\nu) \right\|^2 |\nu - d|\Delta(e - s)^{-2}, \]

where \( c = \min\{c_1, 1/2 - 2c_3\} \).
Proof. Denote \( \tilde{V}^{s,c}(t) = \tilde{V}(t) \) and \( l = d - \nu \). It suffices consider the case of \( l \geq 0 \), as the case of \( l \leq 0 \) follows by reversing the time series. Let \( \nu' > \nu \) be the next change point. Then either \( \nu' = e \) which means that \( \nu \) is the last change point, or \( \nu' < T \) which indicates that \( \nu \) is not the last change point.

Case 1. Suppose \( \nu' = T \). Let \( i = \nu - s \) and \( h = e - \nu \). For any \( u \in \{1, \ldots, p\} \), by Case 1 in Lemma 2.6 of Venkatraman (1992), it holds that

\[
\tilde{V}_u(\nu) = \frac{a_u \sqrt{i + h}}{\sqrt{ih}}, \quad \tilde{V}_u(\nu + l) = \frac{h - l}{h} \frac{a_u \sqrt{i + h}}{\sqrt{(i + l)(h - l)}}.
\]

Thus

\[
\tilde{V}_u(\nu)^2 - \tilde{V}_u(\nu + l)^2 = \frac{l(a_u^2(i + h) + h + i)}{h(i + l)} = \frac{l(h + i)}{h(i + l)} \tilde{V}_u(\nu)^2.
\]

So

\[
\|\tilde{V}(\nu)\|^2 - \|\tilde{V}(\nu + l)\|^2 = \frac{l(h + i)}{h(i + l)} \|\tilde{V}(\nu)\|^2 \geq \frac{l(e - s)}{(e - s)^2} \|\tilde{V}(\nu)\|^2 \geq \frac{c_1 l \Delta}{(e - s)^2} \|\tilde{V}(\nu)\|^2.
\]

Case 2. Suppose \( \nu' < e \). Let \( i = \nu - s \), \( h = \Delta / 2 \) and \( j = e - \nu - h \). Let \( l \leq h / 2 \). For any \( u \in \{1, \ldots, p\} \), by Case 2 in Lemma 2.6 of Venkatraman (1992),

\[
\tilde{V}_u(\nu) = \frac{a_u \sqrt{i + h}}{\sqrt{ih}}, \quad \tilde{V}_u(\nu + h) = \frac{(a_u + h \theta) \sqrt{i + j + h}}{\sqrt{(i + h)j}} \quad \text{and} \quad \tilde{V}_u(\nu + l) = \frac{(a_u + l \theta) \sqrt{i + j + h}}{\sqrt{(i + l)(j + h - l)}},
\]

where \( \theta \) is the solution of

\[
\tilde{V}_u^2(\nu + h) - \tilde{V}_u^2(\nu) = \frac{(a_u + h \theta)^2(i + j + h)}{(i + h)j} - \frac{a_u^2(i + h)}{ih}.
\]

Denote \( B = \|\tilde{V}(\nu + h)\|^2 - \|\tilde{V}(\nu)\|^2 \) and \( B_u = \tilde{V}_u(\nu + h)^2 - \tilde{V}_u(\nu)^2 \). Thus by (17),

\[
(18) \quad B \leq c_3 \|\tilde{V}^{s,c}(\nu)\|_2^2 \Delta^2 (e - s)^{-2}.
\]

Then by Lemma S.14,

\[
\|\tilde{V}(\nu)\|^2 - \|\tilde{V}(\nu + l)\|^2 = \sum_{u=1}^{p} \left\{ \tilde{V}_u(\nu)^2 - \tilde{V}_u(\nu + l)^2 \right\}
\]

\[
\geq \sum_{u=1}^{p} \left\{ \tilde{V}_u(\nu)^2 \frac{(hl - l^2)}{(i + l)(j + h - l)} - B_u \frac{l(i + h)j}{h(i + l)(j + h - l)} \right\}
\]

\[
\geq \frac{\|\tilde{V}(\nu)\|_2^2 \Delta}{2(e - s)^2} - 2B \frac{l}{\Delta} \geq (1/2 - 2c_3) \frac{\|\tilde{V}(\nu)\|_2^2 \Delta}{(e - s)^2},
\]

where the last inequality follows from (18). □

Lemma S.14. Denote

\[
\Theta_{\nu} = \frac{a \sqrt{i + j + h}}{\sqrt{i(j + h)}}, \quad \Theta_{\nu + h} = \frac{(a + h \theta) \sqrt{i + j + h}}{\sqrt{(i + h)j}} \quad \text{and} \quad \Theta_{\nu + l} = \frac{(a + l \theta) \sqrt{i + j + h}}{\sqrt{(i + l)(j + h - l)}}.
\]

Then

\[
\Theta_{\nu}^2 - \Theta_{\nu + l}^2 \geq \frac{\Theta_{\nu}^2 \frac{(hl - l^2)}{(i + l)(j + h - l)} - (\Theta_{\nu + h}^2 - \Theta_{\nu}^2)}{15} \frac{l(i + h)j}{h(i + l)(j + h - l)}.
\]
PROOF. Observe that
\[ \Theta_\nu^2 \Theta_{\nu+l}^2 = \frac{a^2(i + j + h)}{i(j + h)} - \frac{(a + l\theta)^2(i + j + h)}{(i + l)(j + h - l)} \]
\[ = \frac{a^2(i + j + h)}{i(j + h)(i + l)(j + h - l)}((i + l)(j + h - l) - i(j + h)) - \frac{(2l\theta a + l^2\theta^2)(i + j + h)}{(i + l)(j + h - l)} \]
\[ = \frac{a^2(i + j + h)}{i(j + h)(i + l)(j + h - l)}(-il + lj + lh - l^2) - \frac{(2l\theta a + l^2\theta^2)(i + j + h)}{(i + l)(j + h - l)}. \]

To bound the term \(2l\theta a + l^2\theta^2\), let \(b = \Theta_{\nu+h}^2 - \Theta_\nu^2\). Then
\[ b = \frac{(a + h\theta)^2(i + j + h)}{(i + h)j} - \frac{a^2(i + j + h)}{i(j + h)}. \]
Therefore
\[ \frac{bij(i + h)(j + h)}{i + j + h} = \frac{(a^2 + 2h\theta a + h^2\theta^2)i(j + h) - a^2(i + h)j,} \]
which gives
\[ 2h\theta a + h^2\theta^2 = \frac{bij(i + h)}{i + j + h} + \frac{a^2(j - i)h}{i(j + h)}. \]
Therefore
\[ 2l\theta a + l^2\theta^2 \leq 2l\theta a + lh\theta^2 = \frac{l}{h}(2l\theta a + h^2\theta^2) = \frac{l}{h} \left( \frac{bij(i + h)}{i + j + h} + \frac{a^2(j - i)h}{i(j + h)} \right) \]
which implies that
\[ \Theta_\nu^2 \Theta_{\nu+l}^2 = \frac{a^2(i + j + h)}{i(j + h)(i + l)(j + h - l)}(-il + lj + lh - l^2) - \frac{(2l\theta a + l^2\theta^2)(i + j + h)}{(i + l)(j + h - l)} \]
\[ \leq \frac{a^2(i + j + h)}{i(j + h)(i + l)(j + h - l)}(-il + lj + lh - l^2) \]
\[ = \frac{l}{h} \left( \frac{bij(i + h)}{i + j + h} + \frac{a^2(j - i)h}{i(j + h)} \right) (i + j + h) \]
\[ = \frac{a^2(i + j + h)}{i(j + h)(i + l)(j + h - l)}(-il + lj + lh - l^2) \]
\[ = \frac{l}{h}(i + l)(j + h - l) - \frac{a^2(i + j + h)}{i(j + h)(i + l)(j + h - l)}(j - i)l \]
\[ = \frac{a^2(i + j + h)}{i(j + h)(i + l)(j + h - l)}(lh - l^2) - \frac{l(i + h)j}{h(i + l)(j + h - l)}. \]

which concludes the proof.

\[ \square \]

**Lemma S.15.** Suppose \([s, e]\) contains one and only one change point \(\eta_k\), then
\[ \| V_{s,e}(t) \| = \begin{cases} \frac{t-s}{e-s}(e - \eta_k)^2 \| V(\eta) - V(\eta + 1) \|^2, & t \leq \eta_k, \\ \frac{e-t}{e-s}(\eta_k - s)^2 \| V(\eta) - V(\eta + 1) \|^2, & t \geq \eta_k. \end{cases} \]
Proof. This is a straightforward result from the definitions. \hfill \square

Lemma S.16. Let \( \eta_1 \) be the first change point in \( \{1, \ldots, T\} \). Then for any \( 1 \leq t \leq \eta_1 \),
\[
\| \tilde{V}_{0,T}(t) \|^2 = \frac{t(T - \eta_1)}{\eta_1(t - \eta_1)} \| \tilde{V}_{0,T}(\eta_1) \|^2.
\]

Proof. This is a direct consequence of Lemma S.19. \hfill \square

Lemma S.17. Let \([s, e]\) contain two or more change points such that
\[
\eta_{r-1} \leq s < \eta_r < \ldots < \eta_{r+q} \leq e \leq \eta_{r+q+1}, \quad q \geq 1.
\]
If \( \eta_r - s \leq c\Delta \) for some \( c \leq 1/4 \) and \( \eta_{r+1} - \eta_r \geq \Delta \), then
\[
\| \tilde{V}_{s,e}(\eta_r) \|^2 \leq 2c\| \tilde{V}_{s,e}(\eta_{r+1}) \|^2 + 4\kappa_r^2(\eta_r - s).
\]
If there are two and only two change points, then
\[
\max_{s < t < e} \| \tilde{V}_{s,e}(t) \|^2 \leq (e - \eta_{r+1})\kappa_{r+1}^2 + (\eta_r - s)\kappa_r^2.
\]

Proof. This follows from a similar calculation as in Lemma S.20. \hfill \square

S.4.2. Univariate CUSUM.

Assumption S.4. Let \( \{f(t)\}_{t=1}^T \subset \mathbb{R} \). Assume there exists a sequence \( \{\nu_m\}_{m=0}^M \subset \{1, \ldots, T\} \) such that \( 1 = \nu_0 < \nu_1 < \ldots < \nu_M \leq T < \nu_{K+1} = T + 1 \) and, for \( t = 2, \ldots, T \),
\[
f(t) \neq f(t-1) \quad \text{if and only if} \quad t \in \{\nu_1, \ldots, \nu_M\}.
\]
We set
\[
|f(\nu_m) - f(\nu_{m-1})| = \kappa_m \geq \kappa.
\]

For the same reasons as we described after Assumption S.3, in this subsection we use a self-contained notation system, and one can interpret \( \kappa = \kappa_0\rho \) as we used in Assumption 1.

Lemma S.18. Suppose \( \nu_m \) is a change point of \( \{f(t)\}_{t=1}^T \) such that \( \min_{m' \neq m} \{\nu_m - \nu_{m'}\} \geq \Delta \). Then
\[
\max \left\{ \left| \sum_{r=1}^{\nu_m - \Delta} f(r) \right|, \quad \left| \sum_{r=1}^{\nu_m} f(r) \right|, \quad \left| \sum_{r=1}^{\nu_m + \Delta} f(r) \right| \right\} \geq \Delta |f(\nu_m) - f(\nu_m + 1)|/4.
\]

Proof. For simplicity denote \( \nu_m = \nu \). Observe that
\[
\max\{|f(\nu)|, |f(\nu + 1)|\} \geq |f(\nu) - f(\nu + 1)|/2.
\]
Thus
\[
\max \left\{ \left| \sum_{r=\nu - \Delta}^{\nu} f(r) \right|, \quad \left| \sum_{r=\nu + \Delta}^{\nu} f(r) \right| \right\} \geq \Delta |f(\nu) - f(\nu + 1)|/2.
\]
Since

\[
(21) \quad \left| \sum_{r=r-\Delta}^{\nu} f(r) \right| \leq \left| \sum_{r=1}^{\nu-\Delta} f(r) \right| + \left| \sum_{r=1}^{\nu} f(r) \right| \quad \text{and} \quad \left| \sum_{r=r+1}^{\nu+\Delta} f(r) \right| \leq \left| \sum_{r=1}^{\nu} f(r) \right| + \left| \sum_{r=1}^{\nu+\Delta} f(r) \right| ,
\]

we have that (20) and (21) directly imply (19).

**Lemma S.19.** Let \( \eta_1 \) be the first change point in \( \{2, \ldots, T\} \). Then for any \( 1 \leq t < \eta_1 \),

\[
\bar{f}_{t-1} = \frac{t(T - \eta_1)}{\eta_1(T - t)} \tilde{f}_{t-1} .
\]

**Proof.** Without loss of generality assume \( \sum_{t=1}^{T} f_t = 0 \). Thus \( \eta_1 f_1 = \sum_{t=1}^{\eta_1} f_t = -\sum_{t=\eta_1+1}^{T} f_t \). As a result, for any \( 1 \leq t < \eta_1 \),

\[
\bar{f}_{t-1} = \sqrt{\frac{T - t}{T t}} \sum_{i=1}^{t} f_i - \sqrt{\frac{t}{T(T - t)}} \sum_{i=t+1}^{T} f_i \\
= \sqrt{\frac{T - t}{T t}} t f_1 - \sqrt{\frac{t}{T(T - t)}} \left( (\eta_1 - t) f_1 + \sum_{i=\eta_1+1}^{T} f_i \right) \\
= \sqrt{\frac{T - t}{T t}} t f_1 - \sqrt{\frac{t}{T(T - t)}} \{ (\eta_1 - t) f_1 - \eta_1 f_1 \} = \frac{(T - t)\sqrt{t} + t\sqrt{t}}{\sqrt{T(T - t)}} f_1 = \frac{T t}{T - t} f_1.
\]

**Remark S.3.** If there exists \( b \in [1, \eta_1] \) such that \( \bar{f}_{b} > 0 \), then by Lemma S.19, \( \bar{f}_{\eta_1} > 0 \). Since for \( t \in [1, \eta_1] \),\( \frac{t(T - \eta_1)}{\eta_1(T - t)} \) is an increasing function of \( t \), this also implies \( \bar{f}_{t-1} > 0 \) is increasing within \([1, \eta_1]\), as a function of \( t \).

**Lemma S.20.** Let \( [s, e] \) contain two or more change points such that

\( \eta_{r-1} \leq s \leq \eta_r \leq \ldots \leq \eta_{r+q} \leq e \leq \eta_{r+q+1} , \quad q \geq 1. \)

If \( \eta_r - s \leq c_1^2 \Delta \) for some \( c_1 \leq 1/4 \) and \( \eta_{r+1} - \eta_r \geq \Delta \), then

\[
|\tilde{f}_{\eta_r}^{s,e}| \leq c_1 |\tilde{f}_{\eta_{r+1}}^{s,e}| + 2\kappa_r \eta_r - s.
\]

If \( [s, e] \) contains two and only two change points \( \eta_r \) and \( \eta_{r+1} \), then

\[
\max_{s \leq t \leq e} |\tilde{f}_t^{s,e}| \leq \sqrt{e - \eta_{r+1} \kappa_{r+1} + \eta_r - s \kappa_r}.
\]

**Proof.** Consider the sequence \( \{g_t\}_{t=s+1}^{e} \) be such that

\[
g_t = \begin{cases} 
    f_{\eta_{r+1}}, & \text{if } s + 1 \leq t \leq \eta_r, \\
    f_t, & \text{if } \eta_r + 1 \leq t \leq e.
\end{cases}
\]
For any $t \geq \eta_r + 1$,

$$f_t^{s,e} - g_t^{s,e} = \sqrt{\frac{e - t}{(e - s)(t - s)}} \left( \sum_{i=s+1}^{\eta_r} f_{\eta_r} - \sum_{i=\eta_r+1}^{\eta_r} g_{\eta_r} - \sum_{i=\eta_r+1}^{t} g_{\eta_r+1} \right)$$

$$= \sqrt{\frac{e - t}{(e - s)(t - s)}} (\eta_r - s)(f_{\eta_r+1} - f_{\eta_r}) \leq \sqrt{\eta_r - s} \kappa_r.$$

Thus

$$|f_{\eta_r} - g_{\eta_r}| + \sqrt{\eta_r - s} \kappa_r \leq \sqrt{\frac{(\eta_r - s)(e - \eta_r+1)}{(\eta_r+1 - s)(e - \eta_r)} |g_{\eta_r+1} - g_{\eta_r}|} + \sqrt{\eta_r - s} \kappa_r$$

$$\leq \sqrt{\frac{c^2 \Delta}{\Delta} |g_{\eta_r+1} - g_{\eta_r}|} + \sqrt{\eta_r - s} \kappa_r \leq c_1|f_{\eta_r+1}^{s,e}| + 2\sqrt{\eta_r - s} \kappa_r,$$

where the first inequality follows from Lemma S.19 and the observation that the first change point of $g_t$ in $[s,e]$ is $\eta_{r+1}$.

If there are two and only two change points, then

$$\max_{s < t < e} |f_t^{s,e}| = \max \left\{|f_{\eta_r}^{s,e}|, |f_{\eta_r+1}^{s,e}| \right\} \leq \max_{s < t < e} |g_t^{s,e}| + \sqrt{\eta_r - s} \kappa_r \leq \sqrt{e - \eta_r+1} \kappa_{r+1} + \sqrt{\eta_r - s} \kappa_r.$$

\[\square\]

\textbf{S.5. Additional lemmas.}

\textbf{Lemma S.21.} Suppose $x > 0$ and that $x^2 + bx - c \geq 0$ where $b,c > 0$ and that

$$b \leq \sqrt{c}/4.$$

Then $x \geq 7\sqrt{c}/8$.

\textbf{Proof.} We have either $x \geq \frac{-b + \sqrt{b^2 + 4c}}{2}$ or $x \leq \frac{-b - \sqrt{b^2 + 4c}}{2}$. Since $x,b,c > 0$ and $b \leq \sqrt{c}/4$, we have

$$x \geq \frac{-b + \sqrt{b^2 + 4c}}{2} \geq 7\sqrt{c}/8.$$

\[\square\]

We independently select at random from $\{1, \ldots, T\}$ two sequences $\{\alpha_m\}_{m=1}^{M_1}, \{\beta_m\}_{m=1}^{M_1}$, then we keep the pairs which satisfy $\beta_m - \alpha_m \leq C_R \Delta$, with $C_R \geq 3/2$. For notational simplicity, we label them as $\{\alpha_m\}_{m=1}^{M}, \{\beta_m\}_{m=1}^{M}$. Let

$$(22) \quad \mathcal{M} = \bigcap_{k=1}^{K} \{ \alpha_m \in \mathcal{S}_k, \beta_m \in \mathcal{E}_k, \text{for some } m \in \{1, \ldots, M\} \},$$

where $\mathcal{S}_k = [\eta_k - 3\Delta/4, \eta_k - \Delta/2]$ and $\mathcal{E}_k = [\eta_k + \Delta/2, \eta_k + 3\Delta/4], k = 1, \ldots, K$. In the lemma below, we give a lower bound on the probability of $\mathcal{M}$. 

19
Lemma S.22. For the event \( \mathcal{M} \) defined in (22), we have
\[
\mathbb{P}(\mathcal{M}) \geq 1 - \exp \left\{ \log \left( \frac{T}{\Delta} \right) - \frac{M\Delta}{4C_R T} \right\}.
\]

Proof. Since the number of change points are bounded by \( T/\Delta \),
\[
\mathbb{P}(\mathcal{M}^c) \leq \sum_{k=1}^{K} \prod_{m=1}^{M} \left\{ 1 - \mathbb{P}(\alpha_m \in S_k, \beta_m \in E_k) \right\} \leq K(1 - \Delta/(4C_R T))^M \leq (T/\Delta)(1 - \Delta/(4C_R T))^M.
\]

References.


