AIMS Mathematics, 11(5): 13149–13173.
DOI: 10.3934/math.2026542
Received: 04 March 2026
Revised: 04 April 2026
Accepted: 22 April 2026
Published: 12 May 2026

https://www.aimspress.com/journal/Math

Research article

Robust statistical inference for high-dimensional mean structural breaks under β-mixing dependence

Badr S. Alnssyan[1], **Abdelaziz Alsubie**[2], **and Javid Gani Dar**[3,*]

[1] Department of Management Information Systems, College of Business and Economics, Qassim University, Buraydah 51452, Saudi Arabia

[2] Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, Riyadh 11673, Saudi Arabia

[3] Department of Applied Sciences, Symbiosis Institute of Technology, Symbiosis International (Deemed) University, Lavale, Pune 412115, India

* **Correspondence:** Email: javinfo.stat@yahoo.co.in.

Abstract: We studied nonasymptotic inference for a single change-point in the mean of a high-dimensional time series under heavy-tailed marginals and temporal dependence. We developed a robust coordinatewise truncated cumulative sum (CUSUM) process on a trimmed candidate set and paired it with block self-normalization to adapt to an unknown long-run scale. Under β-mixing dependence and finite $(2 + \delta)$ moments, we derived explicit deviation bounds for the robust CUSUM process uniformly over candidate split points and coordinates. These bounds yield a finite-sample level-α test for the existence of a mean change, a finite-sample power guarantee under a separated alternative, and a localization guarantee for the argmax estimator with explicit dependence on $\log p$, the moment index, and the mixing profile. We also constructed a nonasymptotic confidence set for the change-point location by inverting a localized robust contrast, and we proved a finite-sample diameter bound for the resulting set. The proofs were explicit and included truncation bias control, block coupling under absolute regularity, and bounded-increment Bernstein arguments. A reproducible Monte Carlo study under heavy-tailed AR(1) dependence corroborated the finite-sample size control, power trends, localization behavior, and implementation trade-offs.

Keywords: robust change-point inference; high-dimensional mean shifts; heavy-tailed dependence; β-mixing processes; statistical model; blockwise self-normalization; nonasymptotic testing; finite-sample confidence regions; simulation

Mathematics Subject Classification: Primary 62G32, 62M10, 62H15; Secondary 60G10, 60E15, 62F03

1. Introduction

Change-point analysis provides statistical tools for detecting and localizing structural breaks in sequential data. Modern applications involve high-dimensional signals with dependence and departures from light-tailed models. Two persistent features in such data are heavy-tailed marginals and temporal dependence. Heavy tails increase sensitivity to outliers and inflate the variability of classical CUSUM statistics. Dependence modifies effective sample size and impacts threshold calibration.

This paper develops a nonasymptotic inference framework for a single mean change-point in a p-dimensional time series $X_1, \ldots, X_n \in \mathbb{R}^p$ under finite $(2 + \delta)$ moments and short-range dependence quantified by β-mixing. The target is finite-sample control with explicit dependence on (n, p, δ) and on a mixing profile. The method combines robustification through coordinatewise truncation, in the spirit of robust location estimation [5, 14], with self-normalization and recent high-dimensional change-point methodology [11, 20].

Classical univariate change-point theory was developed in [4, 8] and foundational works such as [1, 13, 17]. High-dimensional and panel procedures include sparsified binary segmentation and related multiscale methods [7, 10, 15], together with uniform tests in high dimension [6, 12]. Robust shift detection under dependence includes rank- and U-quantile-based methods [9]. The concentration tools used here rely on coupling and Bernstein-type bounds for weakly dependent sequences [2, 3, 16, 18, 19].

The present paper makes five contributions. First, it introduces a trimmed robust max-CUSUM process based on coordinatewise truncation and derives a uniform deviation inequality for the empirical robust contrast around its population counterpart under $(2 + \delta)$ moments and β-mixing. Second, it establishes a finite-sample level guarantee for a block self-normalized test under a block-scale nondegeneracy condition. Third, it proves a finite-sample power bound for a thresholded robust max-CUSUM detector under a separated alternative. Fourth, it derives a localization theorem for the robust argmax estimator using an explicit deterministic population margin on the trimmed set. Fifth, it constructs a finite-sample confidence set for the change-point location by inverting a localized robust contrast and proves an explicit diameter bound for that set.

The paper also now contains a reproducible numerical study that examines empirical size, finite-sample power across (γ, ℓ), localization accuracy across signal strengths, representative robust CUSUM trajectories, and runtime scaling relative to a wild Binary Segmentation (WBS)-style multiple-change extension.

Organization. Section 2 introduces the model, the trimmed candidate set, the robust statistics, and the assumptions. Section 3 collects the technical tools: block coupling, Bernstein inequalities, and truncation remainder bounds. Section 4 presents the main results and full proofs, including the new power theorem and the confidence-set diameter theorem. Section 5 reports the simulation study and implementation guidance. Section 6 contains the conclusion and future directions.

2. Model, estimands, and procedures

2.1. Single change-point in a high-dimensional mean

Let $X_1, \ldots, X_n \in \mathbb{R}^p$ be observed. We assume an unknown change location $\tau^\star \in \{1, \ldots, n-1\}$ such that

$$X_t = \begin{cases} \mu^{(1)} + \varepsilon_t, & 1 \le t \le \tau^\star, \\ \mu^{(2)} + \varepsilon_t, & \tau^\star + 1 \le t \le n, \end{cases} \tag{2.1}$$

where $\mu^{(1)}, \mu^{(2)} \in \mathbb{R}^p$ are deterministic and $(\varepsilon_t)_{t=1}^n$ is a strictly stationary, mean-zero, $\mathbb{R}^p$-valued process.

The mean shift is

$$\Delta := \mu^{(2)} - \mu^{(1)} \in \mathbb{R}^p.$$

The inferential tasks are:

- detect a change: $H_0 : \Delta = 0$ versus $H_1 : \Delta \neq 0$;
- estimate the change location $\tau^\star$;
- construct a confidence set for $\tau^\star$.

2.2. Heavy tails, dependence, and interiority

Assumption 2.1 (Moment condition)**.** *There exist $\delta > 0$ and $M_{2+\delta} \in (0, \infty)$ such that for every coordinate $j \in \{1, \dots, p\}$,*

$$\mathbb{E}\Big[|\varepsilon_{t,j}|^{2+\delta}\Big] \le M_{2+\delta} \qquad \text{for all } t.$$

Assumption 2.2 (β-mixing)**.** *Let $\mathcal{F}_a^b := \sigma(\varepsilon_t : a \le t \le b)$. The process (ε_t) is absolutely regular with coefficients*

$$\beta(k) := \sup_t \mathbb{E}\left[\sup_{B \in \mathcal{F}_{t+k}^\infty} \left| \mathbb{P}(B \mid \mathcal{F}_{-\infty}^t) - \mathbb{P}(B) \right| \right], \qquad k \ge 1.$$

We assume either geometric decay $\beta(k) \le c_0 e^{-c_1 k}$ for some $c_0, c_1 > 0$, or polynomial decay $\beta(k) \le c_0 k^{-a}$ for some $c_0 > 0$ and $a > 1$.

For stable localization and confidence inversion near the population peak, we use a trimmed candidate set and assume the true change-point is away from the boundaries.

Assumption 2.3 (Trimmed search domain and interior change-point)**.** *Fix trimming parameters γ, ρ with*

$$0 < \gamma < \rho < \frac{1}{2}.$$

Define the trimmed candidate set

$$\mathcal{I}_n(\gamma) := \{k \in \{1, \dots, n-1\} : \ \gamma n \le k \le (1-\gamma)n\}.$$

Assume $\tau^\star \in \mathcal{I}_n(\rho)$.

2.3. Robust truncation, robust contrast, and population contrast

Define the scalar truncation map

$$\mathcal{T}_u(x) := \operatorname{sign}(x) \min\{|x|, u\}, \qquad u > 0,$$

with $\operatorname{sign}(0) = 0$, and apply it coordinatewise to vectors.

For each $k \in \mathcal{I}_n(\gamma)$, define truncated means

$$\bar{X}^{(u)}_{1:k} := \frac{1}{k}\sum_{t=1}^{k} \mathcal{T}_u(X_t), \qquad \bar{X}^{(u)}_{k+1:n} := \frac{1}{n-k}\sum_{t=k+1}^{n} \mathcal{T}_u(X_t),$$

and the robust empirical contrast

$$D^{(u)}(k) := \sqrt{\frac{k(n-k)}{n}}\left(\bar{X}^{(u)}_{1:k} - \bar{X}^{(u)}_{k+1:n}\right) \in \mathbb{R}^p. \tag{2.2}$$

Its population counterpart is

$$\Theta^{(u)}(k) := \sqrt{\frac{k(n-k)}{n}}\left(\mathbb{E}[\bar{X}^{(u)}_{1:k}] - \mathbb{E}[\bar{X}^{(u)}_{k+1:n}]\right). \tag{2.3}$$

The robust max-CUSUM statistic on the trimmed set is

$$T_u := \max_{k\in\mathcal{I}_n(\gamma)} \|D^{(u)}(k)\|_\infty. \tag{2.4}$$

Remark 2.4. *The truncation operator $\mathcal{T}_u$ limits the influence of extreme coordinates while preserving the direction of moderate observations. This coordinatewise robustification is especially useful under finite $(2+\delta)$ moments, where rare large excursions can distort the untruncated CUSUM contrast at finite sample sizes.*

Remark 2.5. *The population contrast $\Theta^{(u)}(k)$ isolates the deterministic signal generated by a mean break after truncation, while $D^{(u)}(k) - \Theta^{(u)}(k)$ represents the stochastic fluctuation around that signal. The normalization factor $\sqrt{k(n-k)/n}$ keeps the contrast on a comparable scale across split points in the trimmed set and centers the analysis on the interior region where localization is stable.*

2.4. Block self-normalization

Fix a block length $\ell \in \{1, \ldots, \lfloor n/2 \rfloor\}$ and let $m := \lfloor n/\ell \rfloor$. For each coordinate j, define block sums

$$B_{b,j} := \sum_{t=(b-1)\ell+1}^{b\ell} \mathcal{T}_u(X_{t,j}), \qquad b = 1, \ldots, m.$$

Let

$$\widetilde{B}_{b,j} := B_{b,j} - \frac{1}{m}\sum_{r=1}^{m} B_{r,j},$$

and define the scale proxy

$$\widehat{s_j^2} := \frac{1}{m\ell}\sum_{b=1}^{m} \widetilde{B}^2_{b,j}. \tag{2.5}$$

The self-normalized statistic on the trimmed set is

$$S_u := \max_{k\in\mathcal{I}_n(\gamma)} \max_{1\le j\le p} \frac{|D^{(u)}_j(k)|}{\widehat{s_j}}. \tag{2.6}$$

2.5. Estimator and confidence set

Define the trimmed argmax estimator

$$\widehat{\tau} \in \arg\max_{k \in \mathcal{I}_n(\gamma)} \|D^{(u)}(k)\|_\infty, \tag{2.7}$$

with the smallest maximizer selected in case of ties.

For an integer radius $r \geq 1$, define

$$\mathcal{K}(r) := \{k \in \mathcal{I}_n(\gamma) : |k - \widehat{\tau}| \leq r\}.$$

Given a threshold $\widehat{q}_\alpha(r) > 0$, define the localized confidence set

$$C_\alpha(r) := \left\{k \in \mathcal{K}(r) : \ \|D^{(u)}(\widehat{\tau})\|_\infty - \|D^{(u)}(k)\|_\infty \leq \widehat{q}_\alpha(r)\right\}. \tag{2.8}$$

3. Preliminaries

3.1. Block coupling for absolutely regular sequences

We use a standard coupling device for absolutely regular sequences. The following block-coupling statement follows from Berbee coupling; see [3, 18].

Lemma 3.1 (Block coupling)**.** *Let $(Y_t)_{t=1}^n$ be a real-valued process such that Y_t is measurable with respect to $\varepsilon_t, \varepsilon_{t-1}, \ldots$ through a measurable transformation preserving the β-mixing coefficients bound in Assumption 2.2. Partition $\{1, \ldots, n\}$ into m consecutive blocks of length ℓ (discarding a remainder of length $< \ell$), and define*

$$Z_b := \sum_{t=(b-1)\ell+1}^{b\ell} Y_t, \qquad b = 1, \ldots, m.$$

There exists a coupling $(Z_b, Z_b^\circ)_{b=1}^m$ on a common probability space such that:

(i) $(Z_b^\circ)_{b=1}^m$ are independent and $Z_b^\circ \overset{d}{=} Z_b$ for each b;
(ii) with $\mathcal{E} := \{\forall b,\ Z_b = Z_b^\circ\}$,

$$\mathbb{P}(\mathcal{E}^c) \leq (m-1)\beta(\ell).$$

Proof. Apply Berbee coupling sequentially to the block sigma-fields generated by the consecutive blocks. At each iteration, the coupling failure probability is bounded by $\beta(\ell)$ because adjacent block sigma-fields are separated by ℓ time indices. Summing the failure probabilities over the $m - 1$ couplings gives the bound. The construction is standard in the absolute-regularity literature; see [3, 18]. □

3.2. Bernstein inequality for bounded independent summands

Lemma 3.2 (Bernstein inequality)**.** *Let $W_1, \ldots, W_m$ be independent, centered random variables with $|W_b| \leq L$ almost surely and*

$$\sum_{b=1}^m \mathbb{E}[W_b^2] \leq V.$$

Then for every $t > 0$,

$$\mathbb{P}\left(\sum_{b=1}^{m} W_b \geq t\right) \leq \exp\left(-\frac{t^2}{2(V + Lt/3)}\right),$$

and the same bound holds for the lower tail

$$\mathbb{P}\left(\sum_{b=1}^{m} W_b \leq -t\right).$$

Proof. This is the standard Bernstein inequality for bounded independent summands; see [2, 19]. □

3.3. Truncation bias and tail remainder bounds

Lemma 3.3 (Truncation bias and tail remainders)**.** *Let Z be a real random variable with $\mathbb{E}[Z] = 0$ and $\mathbb{E}[|Z|^{2+\delta}] \leq M_{2+\delta}$ for some $\delta > 0$. Then for every $u > 0$,*

$$\left|\mathbb{E}[\mathcal{T}_u(Z)]\right| \leq \mathbb{E}[|Z|\mathbf{1}\{|Z| > u\}] \leq \frac{M_{2+\delta}}{u^{1+\delta}}, \tag{3.1}$$

$$\mathbb{E}\left[Z^2\mathbf{1}\{|Z| > u\}\right] \leq \frac{M_{2+\delta}}{u^{\delta}}, \tag{3.2}$$

$$\mathbb{E}\left[\mathcal{T}_u(Z)^2\right] \leq \mathbb{E}[Z^2] \leq M_{2+\delta}^{2/(2+\delta)}. \tag{3.3}$$

Proof. The pointwise inequalities

$$|Z|\,\mathbf{1}\{|Z| > u\} \leq |Z|^{2+\delta}u^{-(1+\delta)}, \qquad Z^2\mathbf{1}\{|Z| > u\} \leq |Z|^{2+\delta}u^{-\delta}$$

yield (3.1) and (3.2) after taking expectations. Since $|\mathcal{T}_u(Z)| \leq |Z|$, the first inequality in (3.3) holds. The second inequality in (3.3) follows from Lyapunov's inequality. □

4. Main results

The main results are now presented. We first derive a uniform deviation inequality for the trimmed robust contrast around its population version. We then obtain finite-sample guarantees for testing, detection power, localization, and confidence sets, including a diameter bound for the inverted localized confidence set.

4.1. Tuning parameters and shorthand

Fix $\alpha \in (0, 1)$ and define

$$L_\alpha := \log\left(\frac{8np}{\alpha}\right). \tag{4.1}$$

Set the truncation level

$$u := \left(\frac{n}{L_\alpha}\right)^{1/(2+\delta)}. \tag{4.2}$$

By Lemma 3.3, this choice gives truncation bias of order

$$u^{-(1+\delta)} = \left(\frac{L_\alpha}{n}\right)^{(1+\delta)/(2+\delta)}.$$

Let ℓ be a block length such that $(m-1)\beta(\ell)$ is small, where $m = \lfloor n/\ell \rfloor$. Under geometric β-mixing, one may take $\ell \asymp \log(n/\alpha)$. Under polynomial β-mixing with exponent $a > 1$, one may take $\ell \asymp (n/\alpha)^{1/a}$ up to constant and logarithmic factors.

4.2. Uniform deviation for the trimmed robust contrast

Theorem 4.1 (Uniform deviation under β-mixing)**.** *Assume 2.2 and 2.3. Fix $u > 0$, a block length ℓ, and let $m = \lfloor n/\ell \rfloor$. For $t \in \{1, \dots, n\}$ and $j \in \{1, \dots, p\}$, define*

$$Y_{t,j} := \mathcal{T}_u(X_{t,j}) - \mathbb{E}[\mathcal{T}_u(X_{t,j})], \qquad \sigma^2_{u,\max} := \max_{1\le t\le n} \max_{1\le j\le p} \mathbb{E}[Y^2_{t,j}].$$

Then there exists an absolute constant $C_0 > 0$ such that for every $x > 0$,

$$\mathbb{P}\left(\max_{k\in \mathcal{I}_n(\gamma)} \|D^{(u)}(k) - \Theta^{(u)}(k)\|_\infty \ge x \right) \le 2np \, \exp\left(-\frac{x^2}{C_0\gamma^{-1}\sigma^2_{u,\max} + C_0 u\ell x/\sqrt{\gamma n}} \right) + (m-1)\beta(\ell). \tag{4.3}$$

Consequently, for a sufficiently large absolute constant $C_1 > 0$,

$$\eta_\alpha := C_1 \left(\frac{\sigma_{u,\max}}{\sqrt{\gamma}} \sqrt{L_\alpha} + \frac{u\ell}{\sqrt{\gamma n}} L_\alpha \right) \tag{4.4}$$

satisfies

$$\mathbb{P}\left(\max_{k\in \mathcal{I}_n(\gamma)} \|D^{(u)}(k) - \Theta^{(u)}(k)\|_\infty \ge \eta_\alpha \right) \le \alpha + (m-1)\beta(\ell). \tag{4.5}$$

Proof. Fix a coordinate j. For $r = 1, \dots, n$, define the partial sums

$$S_{r,j} := \sum_{t=1}^{r} Y_{t,j}, \qquad M_j := \max_{1\le r\le n} |S_{r,j}|.$$

Since $|Y_{t,j}| \le 2u$, we have $|S_{r,j} - S_{r-1,j}| \le 2u$.

For each $k \in \mathcal{I}_n(\gamma)$, write the empirical-population contrast difference in coordinate j as

$$D^{(u)}_j(k) - \Theta^{(u)}_j(k) = \sqrt{\frac{n-k}{n}} \frac{S_{k,j}}{\sqrt{k}} - \sqrt{\frac{k}{n}} \frac{S_{n,j} - S_{k,j}}{\sqrt{n-k}}. \tag{4.6}$$

The coefficients satisfy

$$\sqrt{\frac{n-k}{n}} \frac{1}{\sqrt{k}} \le \frac{1}{\sqrt{\gamma n}}, \qquad \sqrt{\frac{k}{n}} \frac{1}{\sqrt{n-k}} \le \frac{1}{\sqrt{\gamma n}},$$

because $k \in \mathcal{I}_n(\gamma)$. Also,

$$|S_{n,j} - S_{k,j}| \le |S_{n,j}| + |S_{k,j}| \le 2M_j.$$

Hence (4.6) yields

$$\max_{k\in \mathcal{I}_n(\gamma)} |D^{(u)}_j(k) - \Theta^{(u)}_j(k)| \le \frac{3M_j}{\sqrt{\gamma n}}. \tag{4.7}$$

Therefore,

$$\mathbb{P}\left(\max_{k\in I_n(\gamma)} |D_j^{(u)}(k) - \Theta_j^{(u)}(k)| \geq x\right) \leq \mathbb{P}\left(M_j \geq \frac{x\sqrt{\gamma n}}{3}\right). \tag{4.8}$$

It remains to control M_j. Partition $\{1,\dots,n\}$ into m consecutive blocks of length ℓ and define

$$Z_b := \sum_{t=(b-1)\ell+1}^{b\ell} Y_{t,j}, \qquad b = 1,\dots,m.$$

Then $|Z_b| \leq 2u\ell$. By Lemma 3.1, there exists a coupling $(Z_b, Z_b^\circ)_{b=1}^m$ with independent copies $Z_b^\circ \overset{d}{=} Z_b$ and event $\mathcal{E}$ such that

$$\mathbb{P}(\mathcal{E}^c) \leq (m-1)\beta(\ell).$$

For $r \leq m\ell$, write $r = q\ell + s$ with $q = \lfloor r/\ell \rfloor$ and $0 \leq s < \ell$. Then

$$S_{r,j} = \sum_{b=1}^{q} Z_b + \sum_{t=q\ell+1}^{q\ell+s} Y_{t,j},$$

so

$$|S_{r,j}| \leq \left|\sum_{b=1}^{q} Z_b\right| + 2us.$$

Taking the maximum over $r \leq m\ell$ gives

$$\max_{1\leq r\leq m\ell} |S_{r,j}| \leq \max_{0\leq q\leq m} \left|\sum_{b=1}^{q} Z_b\right| + 2u\ell.$$

Thus for every $z > 2u\ell$,

$$\begin{aligned}\mathbb{P}(M_j \geq z) &\leq \mathbb{P}\left(\max_{0\leq q\leq m} \left|\sum_{b=1}^{q} Z_b\right| \geq z - 2u\ell\right) \\ &\leq \mathbb{P}\left(\max_{0\leq q\leq m} \left|\sum_{b=1}^{q} Z_b^\circ\right| \geq z - 2u\ell\right) + \mathbb{P}(\mathcal{E}^c).\end{aligned} \tag{4.9}$$

For each fixed q, the variables $Z_1^\circ,\dots,Z_q^\circ$ are independent and centered. The variance proxy is bounded by

$$\begin{aligned}\sum_{b=1}^{q} \mathbb{E}[(Z_b^\circ)^2] = \sum_{b=1}^{q} \mathbb{E}[Z_b^2] &\leq \sum_{b=1}^{q} \ell \sum_{t=(b-1)\ell+1}^{b\ell} \mathbb{E}[Y_{t,j}^2] \\ &\leq \ell \cdot q\ell \cdot \sigma_{u,\max}^2 \leq n\,\sigma_{u,\max}^2.\end{aligned}$$

The inequality $\mathbb{E}[(\sum a_t)^2] \leq \ell \sum \mathbb{E}[a_t^2]$ was used in the first step. Applying Lemma 3.2 with $L = 2u\ell$ gives, for $s > 0$,

$$\mathbb{P}\left(\left|\sum_{b=1}^{q} Z_b^\circ\right| \geq s\right) \leq 2\exp\left(-\frac{s^2}{2(n\sigma_{u,\max}^2 + (2u\ell)s/3)}\right).$$

A union bound over $q = 1, \dots, m$ yields

$$\mathbb{P}\left(\max_{1\le q\le m} \left| \sum_{b=1}^{q} Z_b^{\circ} \right| \ge s \right) \le 2m \exp\left(-\frac{s^2}{2(n\sigma_{u,\max}^2 + (2u\ell)s/3)} \right) \le 2n \exp\left(-\frac{s^2}{C(n\sigma_{u,\max}^2 + u\ell s)} \right), \tag{4.10}$$

for an absolute constant $C > 0$.

Combine (4.9) and (4.10) with $z = x\sqrt{\gamma n}/3$ and $s = z - 2u\ell$. Enlarging constants absorbs the shift $2u\ell$ and extends the inequality to all $x > 0$:

$$\mathbb{P}\left(\max_{k\in\mathcal{I}_n(\gamma)} |D_j^{(u)}(k) - \Theta_j^{(u)}(k)| \ge x \right) \le 2n \exp\left(-\frac{x^2}{C_0\gamma^{-1}\sigma_{u,\max}^2 + C_0 u\ell x/\sqrt{\gamma n}} \right) + (m-1)\beta(\ell).$$

The coupling event depends on block sigma-fields of the underlying process and is common to all coordinates, so the coupling error term remains unchanged after a union bound over $j = 1, \dots, p$. Applying the union bound gives (4.3). The threshold (4.4) follows by solving the exponent at level L_α and choosing C_1 sufficiently large. □

Remark 4.2 (Operational calibration of the deviation threshold)**.** *The absolute constants C_0 and C_1 enter the analysis only through the envelope η_α in* (4.4)*. For implementation, the dependence of η_α on (n, p, γ, u, ℓ) provides the theoretical scaling, while the final numerical threshold may be calibrated empirically under a null design matched to the application. This preserves the robust contrast in* (2.4) *and transfers the unknown absolute constants into a model-specific critical value. Section 5 records this calibration step for the heavy-tailed dependent design used in the numerical study.*

4.3. Finite-sample level control for the self-normalized test

We add a block-scale nondegeneracy condition under H_0 for the denominator of the self-normalized statistic.

Assumption 4.3 (Block-scale nondegeneracy under H_0)**.** *Under H_0 (so $\mu^{(1)} = \mu^{(2)} =: \mu$), define for each coordinate j and block length ℓ*

$$v_{j,\ell}^{(u)} := \frac{1}{\ell} \operatorname{Var}\left(\sum_{t=1}^{\ell} \mathcal{T}_u(\mu_j + \varepsilon_{t,j}) \right).$$

There exists $\underline{v}_u > 0$ such that

$$\min_{1\le j\le p} v_{j,\ell}^{(u)} \ge \underline{v}_u^2.$$

Remark 4.4 (A sufficient regime for block-scale nondegeneracy)**.** *For a fixed coordinate j, set*

$$Y_{t,j}^{(u)} := \mathcal{T}_u(\mu_j + \varepsilon_{t,j}) - \mathbb{E}\left[\mathcal{T}_u(\mu_j + \varepsilon_{t,j})\right], \qquad \Gamma_j^{(u)} := \operatorname{Var}(Y_{1,j}^{(u)}) + 2\sum_{h=1}^{\infty} \operatorname{Cov}(Y_{1,j}^{(u)}, Y_{1+h,j}^{(u)}).$$

If there exist constants $\underline{\Gamma}_u > 0$ and $A_u < \infty$ such that, uniformly over j,

$$\Gamma_j^{(u)} \ge 2\underline{\Gamma}_u, \qquad \sum_{h=1}^{\infty} h \left| \operatorname{Cov}(Y_{1,j}^{(u)}, Y_{1+h,j}^{(u)}) \right| \le A_u,$$

then

$$v_{j,\ell}^{(u)} \geq \underline{\Gamma}_u \qquad \text{for every } \ell \geq \frac{4A_u}{\underline{\Gamma}_u}.$$

Indeed,

$$v_{j,\ell}^{(u)} = \Gamma_j^{(u)} - 2\sum_{h=\ell}^{\infty} \mathrm{Cov}(Y_{1,j}^{(u)}, Y_{1+h,j}^{(u)}) - \frac{2}{\ell}\sum_{h=1}^{\ell-1} h\ \mathrm{Cov}(Y_{1,j}^{(u)}, Y_{1+h,j}^{(u)}),$$

so

$$\left|v_{j,\ell}^{(u)} - \Gamma_j^{(u)}\right| \leq \frac{4}{\ell}\sum_{h=1}^{\infty} h\left|\mathrm{Cov}(Y_{1,j}^{(u)}, Y_{1+h,j}^{(u)})\right| \leq \frac{4A_u}{\ell}.$$

Hence $v_{j,\ell}^{(u)} \geq \Gamma_j^{(u)} - 4A_u/\ell \geq \underline{\Gamma}_u$ *for* $\ell \geq 4A_u/\underline{\Gamma}_u$. *In the AR(1) design of Section 5 with* $|\phi| < 1$ *and standardized Student-*t_5 *innovations, the truncated process* $Y_{t,j}^{(u)}$ *is bounded and inherits geometric decay of autocovariances. Once the truncated coordinate distribution is nondegenerate, this sufficient regime is satisfied for moderate block lengths.*

Theorem 4.5 (Finite-sample level for the self-normalized test)**.** *Assume 2.2, 2.3, and 4.3. Work under* $H_0 : \Delta = 0$. *Fix* $u > 0$ *and a block length* ℓ *with* $m = \lfloor n/\ell \rfloor$. *Suppose*

$$(m-1)\beta(\ell) \leq \frac{\alpha}{64p}, \qquad \frac{m\underline{v}_u^2}{\ell u^2} \geq C_\star \log\left(\frac{64p}{\alpha}\right), \tag{4.11}$$

for a sufficiently large absolute constant $C_\star > 0$. *Let*

$$\lambda_\alpha := \frac{4\eta_\alpha}{\underline{v}_u}, \tag{4.12}$$

where η_α *is defined in* (4.4) *with* $\sigma^2_{u,\max}$ *computed under* H_0. *Then the test that rejects* H_0 *when* $S_u \geq \lambda_\alpha$ *satisfies*

$$\mathbb{P}_{H_0}(S_u \geq \lambda_\alpha) \leq \alpha,$$

where S_u *is defined by 2.6.*

Proof. Fix a coordinate j. Under H_0, the process

$$Y_{t,j} := \mathcal{T}_u(X_{t,j}) - \mathbb{E}[\mathcal{T}_u(X_{t,j})]$$

is strictly stationary and centered. Define the block-centered sums

$$C_{b,j} := \sum_{t=(b-1)\ell+1}^{b\ell} Y_{t,j}, \qquad \bar{C}_j := \frac{1}{m}\sum_{b=1}^{m} C_{b,j}.$$

Since the centering by the block average removes the block mean, the denominator can be written as

$$\widehat{s}_j^2 = \frac{1}{m\ell}\sum_{b=1}^{m}(C_{b,j} - \bar{C}_j)^2 = \underbrace{\frac{1}{m\ell}\sum_{b=1}^{m} C_{b,j}^2}_{=:A_j} - \underbrace{\frac{\bar{C}_j^2}{\ell}}_{=:R_j}. \tag{4.13}$$

We control A_j from below and R_j from above.

Control of A_j. Let

$$\mu_{A,j} := \mathbb{E}[A_j] = \frac{1}{\ell}\mathbb{E}[C_{1,j}^2] = v_{j,\ell}^{(u)}.$$

Define

$$W_{b,j} := C_{b,j}^2 - \mathbb{E}[C_{b,j}^2].$$

Since $|Y_{t,j}| \le 2u$, we have $|C_{b,j}| \le 2\ell u$, hence

$$|W_{b,j}| \le |C_{b,j}|^2 + \mathbb{E}[C_{b,j}^2] \le 8\ell^2 u^2.$$

Also,

$$W_{b,j}^2 \le C_{b,j}^4 \le (2\ell u)^2 C_{b,j}^2 = 4\ell^2 u^2 C_{b,j}^2.$$

Therefore,

$$\sum_{b=1}^{m} \mathbb{E}[W_{b,j}^2] \le 4\ell^2 u^2 \sum_{b=1}^{m} \mathbb{E}[C_{b,j}^2] = 4m\ell^3 u^2\, \mu_{A,j}.$$

Apply Lemma 3.1 to the block sequence $(W_{b,j})_{b=1}^m$ and then Bernstein's inequality (Lemma 3.2) to the coupled independent version. For an absolute constant $c_2 > 0$,

$$\mathbb{P}\left(A_j \le \frac{1}{2}\mu_{A,j}\right) = \mathbb{P}\left(\sum_{b=1}^{m} W_{b,j} \le -\frac{1}{2}m\ell\mu_{A,j}\right) \le \exp\left(-c_2\frac{m\mu_{A,j}}{\ell u^2}\right) + (m-1)\beta(\ell).$$

Since $\mu_{A,j} \ge \underline{v}_u^2$, by Assumption 4.3,

$$\mathbb{P}\left(A_j \le \frac{1}{2}\underline{v}_u^2\right) \le \exp\left(-c_2\frac{m\underline{v}_u^2}{\ell u^2}\right) + (m-1)\beta(\ell). \tag{4.14}$$

Control of R_j. We have

$$R_j > \frac{1}{4}\underline{v}_u^2 \quad\Longrightarrow\quad |\bar{C}_j| > \frac{1}{2}\sqrt{\ell}\,\underline{v}_u \quad\Longrightarrow\quad \left|\sum_{b=1}^{m} C_{b,j}\right| > \frac{m}{2}\sqrt{\ell}\,\underline{v}_u.$$

The variables $C_{b,j}$ satisfy $|C_{b,j}| \le 2\ell u$ and

$$\sum_{b=1}^{m} \mathbb{E}[C_{b,j}^2] = m\ell\, v_{j,\ell}^{(u)}.$$

A coupling-plus-Bernstein argument, identical in form to the one used for A_j, yields an absolute constant $c_3 > 0$ such that

$$\mathbb{P}\left(R_j > \frac{1}{4}\underline{v}_u^2\right) \le \exp\left(-c_3\frac{m\underline{v}_u^2}{\ell u^2}\right) + (m-1)\beta(\ell). \tag{4.15}$$

Lower tail of the denominator. From (4.13), if $A_j > \frac{1}{2}\underline{v}_u^2$ and $R_j \le \frac{1}{4}\underline{v}_u^2$, then $\widehat{s}_j^2 > \frac{1}{4}\underline{v}_u^2$. Thus

$$\mathbb{P}\left(\widehat{s}_j \le \frac{\underline{v}_u}{2}\right) \le \mathbb{P}\left(A_j \le \frac{1}{2}\underline{v}_u^2\right) + \mathbb{P}\left(R_j > \frac{1}{4}\underline{v}_u^2\right). \tag{4.16}$$

Using (4.14) and (4.15), and then (4.11), we obtain

$$\mathbb{P}\left(\widehat{s}_j \le \frac{\underline{v}_u}{2}\right) \le \frac{\alpha}{4p}. \tag{4.17}$$

Control of the numerator. Under H_0, $\Theta^{(u)}(k) \equiv 0$ for all k. By Theorem 4.1, after replacing the factor p in (4.3) by 1 for a fixed coordinate,

$$\mathbb{P}\left(\max_{k\in\mathcal{I}_n(\gamma)} |D_j^{(u)}(k)| \ge \eta_\alpha\right) \le \frac{\alpha}{4p},$$

for C_1 sufficiently large and under (4.11).

Combine the numerator and denominator. By the union bound,

$$\begin{aligned}\mathbb{P}\left(\max_{k\in\mathcal{I}_n(\gamma)} \frac{|D_j^{(u)}(k)|}{\widehat{s}_j} \ge \frac{4\eta_\alpha}{\underline{v}_u}\right) &\le \mathbb{P}\left(\max_{k\in\mathcal{I}_n(\gamma)} |D_j^{(u)}(k)| \ge \eta_\alpha\right) + \mathbb{P}\left(\widehat{s}_j \le \frac{\underline{v}_u}{2}\right)\\ &\le \frac{\alpha}{2p}.\end{aligned}$$

A union bound over $j = 1, \dots, p$ gives

$$\mathbb{P}_{H_0}(S_u \ge \lambda_\alpha) \le \alpha,$$

with $\lambda_\alpha = 4\eta_\alpha/\underline{v}_u$. □

4.4. A finite-sample power guarantee for robust max-CUSUM detection

The next theorem gives a finite-sample power guarantee for the thresholded robust max-CUSUM detector based on T_u.

Theorem 4.6 (Finite-sample power under a separated alternative)**.** *Assume 2.2 and 2.3. Fix $u > 0$, a block length ℓ, and let $m = \lfloor n/\ell \rfloor$. Let η_α be defined by* (4.4)*. Define the rejection rule*

$$\phi_\alpha = \mathbf{1}\{T_u \ge \eta_\alpha\}.$$

Let

$$\kappa_\infty^{(u)} := \|\Delta^{(u)}\|_\infty, \qquad \Delta^{(u)} := m_2^{(u)} - m_1^{(u)},$$

where $m_r^{(u)} := \mathbb{E}[\mathcal{T}_u(\mu^{(r)} + \varepsilon_1)]$, $r \in \{1, 2\}$. If

$$\sqrt{\gamma n}\,\kappa_\infty^{(u)} \ge 2\eta_\alpha, \tag{4.18}$$

then

$$\mathbb{P}_{H_1}(\phi_\alpha = 1) \ge 1 - \alpha - (m-1)\beta(\ell). \tag{4.19}$$

Proof. At the true change-point $k = \tau^\star$, the population contrast equals

$$\Theta^{(u)}(\tau^\star) = \sqrt{\frac{\tau^\star(n-\tau^\star)}{n}}\,\Delta^{(u)}.$$

Assumption 2.3 implies $\tau^\star(n-\tau^\star)/n \geq \rho(1-\rho)n \geq \gamma(1-\gamma)n \geq \gamma n/2$. In particular,

$$\|\Theta^{(u)}(\tau^\star)\|_\infty \geq \sqrt{\gamma n}\,\kappa_\infty^{(u)}.$$

Define the event

$$\mathcal{G}_\alpha := \left\{ \max_{k \in \mathcal{I}_n(\gamma)} \|D^{(u)}(k) - \Theta^{(u)}(k)\|_\infty \leq \eta_\alpha \right\}.$$

By Theorem 4.1, $\mathbb{P}(\mathcal{G}_\alpha) \geq 1 - \alpha - (m-1)\beta(\ell)$.

On $\mathcal{G}_\alpha$,

$$T_u \geq \|D^{(u)}(\tau^\star)\|_\infty \geq \|\Theta^{(u)}(\tau^\star)\|_\infty - \eta_\alpha \geq \sqrt{\gamma n}\,\kappa_\infty^{(u)} - \eta_\alpha.$$

Under (4.18), the right-hand side is at least η_α, so $T_u \geq \eta_\alpha$ on $\mathcal{G}_\alpha$. Hence

$$\mathbb{P}_{H_1}(\phi_\alpha = 1) \geq \mathbb{P}(\mathcal{G}_\alpha) \geq 1 - \alpha - (m-1)\beta(\ell).$$

□

4.5. Population margin on the trimmed set

We now derive a deterministic margin inequality for the population robust contrast. This margin is the key input for localization and confidence-set diameter control.

Define the robustified regime means

$$m_r^{(u)} := \mathbb{E}[\mathcal{T}_u(\mu^{(r)} + \varepsilon_1)] \in \mathbb{R}^p, \qquad r \in \{1, 2\},$$

and the robustified shift

$$\Delta^{(u)} := m_2^{(u)} - m_1^{(u)}.$$

Assumption 4.7 (Identifiability)**.** *There exists a coordinate $j^\star \in \{1, \ldots, p\}$ and a constant $\kappa > 0$ such that*

$$|\Delta_{j^\star}^{(u)}| \geq \kappa.$$

Lemma 4.8 (Deterministic population margin)**.** *Assume 2.3 and 4.7. Define*

$$c_\rho := \frac{1}{2}\sqrt{\frac{\rho}{1-\rho}}. \tag{4.20}$$

Then for every $k \in \mathcal{I}_n(\gamma)$,

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty \geq c_\rho\,\kappa\,\frac{|k-\tau^\star|}{\sqrt{n}}. \tag{4.21}$$

Proof. For the distinguished coordinate $j^\star$, we compute $\Theta_{j^\star}^{(u)}(k)$ exactly.

For $k \leq \tau^\star$,

$$\Theta_{j^\star}^{(u)}(k) = -(n-\tau^\star)\sqrt{\frac{k}{n(n-k)}}\,\Delta_{j^\star}^{(u)}. \tag{4.22}$$

For $k \geq \tau^\star$,

$$\Theta_{j^\star}^{(u)}(k) = -\tau^\star\sqrt{\frac{n-k}{nk}}\,\Delta_{j^\star}^{(u)}. \tag{4.23}$$

At $k = \tau^\star$,

$$|\Theta^{(u)}_{j^\star}(\tau^\star)| = \sqrt{\frac{\tau^\star(n-\tau^\star)}{n}}\,|\Delta^{(u)}_{j^\star}|. \tag{4.24}$$

We treat the two sides separately.

Case 1: $k \le \tau^\star$. Using (4.22) and (4.24),

$$\begin{aligned}|\Theta^{(u)}_{j^\star}(\tau^\star)| - |\Theta^{(u)}_{j^\star}(k)| &= |\Delta^{(u)}_{j^\star}|\left[\sqrt{\frac{\tau^\star(n-\tau^\star)}{n}} - (n-\tau^\star)\sqrt{\frac{k}{n(n-k)}}\right] \\ &= \frac{n-\tau^\star}{\sqrt{n}}\,|\Delta^{(u)}_{j^\star}|\left[\sqrt{\frac{\tau^\star}{n-\tau^\star}} - \sqrt{\frac{k}{n-k}}\right].\end{aligned}$$

Set

$$a := \frac{\tau^\star}{n-\tau^\star}, \qquad b := \frac{k}{n-k}.$$

Then $a \ge b$ and

$$\sqrt{a} - \sqrt{b} = \frac{a-b}{\sqrt{a}+\sqrt{b}} = \frac{n(\tau^\star - k)}{(n-\tau^\star)(n-k)(\sqrt{a}+\sqrt{b})}.$$

Hence

$$|\Theta^{(u)}_{j^\star}(\tau^\star)| - |\Theta^{(u)}_{j^\star}(k)| = |\Delta^{(u)}_{j^\star}| \cdot \frac{\sqrt{n}(\tau^\star - k)}{(n-k)(\sqrt{a}+\sqrt{b})}.$$

Since $k \le \tau^\star \le (1-\rho)n$, we have $n-k \le n$ and

$$\sqrt{a} + \sqrt{b} \le 2\sqrt{a} \le 2\sqrt{\frac{1-\rho}{\rho}}.$$

Therefore

$$|\Theta^{(u)}_{j^\star}(\tau^\star)| - |\Theta^{(u)}_{j^\star}(k)| \ge \frac{1}{2}\sqrt{\frac{\rho}{1-\rho}}\,|\Delta^{(u)}_{j^\star}|\frac{\tau^\star - k}{\sqrt{n}}.$$

Case 2: $k \ge \tau^\star$. Using (4.23) and (4.24),

$$\begin{aligned}|\Theta^{(u)}_{j^\star}(\tau^\star)| - |\Theta^{(u)}_{j^\star}(k)| &= |\Delta^{(u)}_{j^\star}|\left[\sqrt{\frac{\tau^\star(n-\tau^\star)}{n}} - \tau^\star\sqrt{\frac{n-k}{nk}}\right] \\ &= \frac{\tau^\star}{\sqrt{n}}\,|\Delta^{(u)}_{j^\star}|\left[\sqrt{\frac{n-\tau^\star}{\tau^\star}} - \sqrt{\frac{n-k}{k}}\right].\end{aligned}$$

Set

$$a' := \frac{n-\tau^\star}{\tau^\star}, \qquad b' := \frac{n-k}{k}.$$

Then $a' \ge b'$ and

$$\sqrt{a'} - \sqrt{b'} = \frac{a'-b'}{\sqrt{a'}+\sqrt{b'}} = \frac{n(k-\tau^\star)}{\tau^\star k\,(\sqrt{a'}+\sqrt{b'})}.$$

Thus

$$|\Theta_{j^\star}^{(u)}(\tau^\star)| - |\Theta_{j^\star}^{(u)}(k)| = |\Delta_{j^\star}^{(u)}| \cdot \frac{\sqrt{n}(k-\tau^\star)}{k(\sqrt{a'} + \sqrt{b'})}.$$

Since $k \le n$ and $\tau^\star \in [\rho n, (1-\rho)n]$,

$$\sqrt{a'} + \sqrt{b'} \le 2\sqrt{a'} \le 2\sqrt{\frac{1-\rho}{\rho}}.$$

Therefore

$$|\Theta_{j^\star}^{(u)}(\tau^\star)| - |\Theta_{j^\star}^{(u)}(k)| \ge \frac{1}{2}\sqrt{\frac{\rho}{1-\rho}}\,|\Delta_{j^\star}^{(u)}|\frac{k-\tau^\star}{\sqrt{n}}.$$

In both cases,

$$|\Theta_{j^\star}^{(u)}(\tau^\star)| - |\Theta_{j^\star}^{(u)}(k)| \ge c_\rho |\Delta_{j^\star}^{(u)}|\frac{|k-\tau^\star|}{\sqrt{n}}.$$

Assumption 4.7 gives $|\Delta_{j^\star}^{(u)}| \ge \kappa$. Since

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty \ge |\Theta_{j^\star}^{(u)}(\tau^\star)| - |\Theta_{j^\star}^{(u)}(k)|,$$

the result follows. □

Proposition 4.9 (Concrete sparse break and sharp margin scaling)**.** *Assume that $\mu^{(1)} = 0$ and that exactly one coordinate $j^\star$ changes, with*

$$\mu_j^{(2)} = \begin{cases} \vartheta, & j = j^\star, \\ 0, & j \ne j^\star, \end{cases} \qquad 0 < \vartheta \le u.$$

Assume also that

$$\mathbb{P}\Big(|\varepsilon_{1,j^\star}| \le u - \vartheta\Big) > 0.$$

Then

$$0 < \Delta_{j^\star}^{(u)} = \mathbb{E}\Big[\mathcal{T}_u(\vartheta + \varepsilon_{1,j^\star}) - \mathcal{T}_u(\varepsilon_{1,j^\star})\Big] \le \vartheta.$$

Moreover, for every $k \in \mathcal{I}_n(\gamma)$,

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty = \Delta_{j^\star}^{(u)} \times \begin{cases} \sqrt{\dfrac{\tau^\star(n-\tau^\star)}{n}} - (n-\tau^\star)\sqrt{\dfrac{k}{n(n-k)}}, & k \le \tau^\star, \\ \sqrt{\dfrac{\tau^\star(n-\tau^\star)}{n}} - \tau^\star\sqrt{\dfrac{n-k}{nk}}, & k \ge \tau^\star. \end{cases} \tag{4.25}$$

If $\lambda^\star := \tau^\star/n \in (\gamma, 1-\gamma)$ is fixed and $|k-\tau^\star| = o(n)$, then

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty = \frac{\Delta_{j^\star}^{(u)}}{2\sqrt{\lambda^\star(1-\lambda^\star)}}\frac{|k-\tau^\star|}{\sqrt{n}} + o\left(\frac{|k-\tau^\star|}{\sqrt{n}}\right). \tag{4.26}$$

Thus the population margin is linear in $|k-\tau^\star|/\sqrt{n}$ to first order, which shows that the scaling in Lemma 4.8 is sharp.

Proof. Because $\mathcal{T}_u$ is monotone increasing and 1-Lipschitz,

$$0 \le \mathcal{T}_u(\vartheta + x) - \mathcal{T}_u(x) \le \vartheta \qquad \text{for every } x \in \mathbb{R}.$$

On the event $\{|x| \le u - \vartheta\}$ one has $\mathcal{T}_u(\vartheta + x) - \mathcal{T}_u(x) = \vartheta$, so the positivity assumption yields $0 < \Delta^{(u)}_{j^\star} \le \vartheta$ after taking expectations.

All coordinates other than $j^\star$ have zero shift, hence their population contrasts vanish. Therefore the ℓ_∞ population contrast is attained at $j^\star$, and (4.25) follows directly from (4.22)–(4.24). For the local expansion, write $k = n\lambda_k$ with $\lambda_k \to \lambda^\star$. When $k \le \tau^\star$,

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty = \Delta^{(u)}_{j^\star} \sqrt{n}\,(1 - \lambda^\star)\left[\sqrt{\frac{\lambda^\star}{1 - \lambda^\star}} - \sqrt{\frac{\lambda_k}{1 - \lambda_k}}\right].$$

A first-order Taylor expansion of $f(\lambda) := \sqrt{\lambda/(1-\lambda)}$ at $\lambda^\star$ gives

$$f(\lambda^\star) - f(\lambda_k) = \frac{\lambda^\star - \lambda_k}{2\sqrt{\lambda^\star}(1 - \lambda^\star)^{3/2}} + o(|\lambda^\star - \lambda_k|),$$

which yields

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty = \frac{\Delta^{(u)}_{j^\star}}{2\sqrt{\lambda^\star(1 - \lambda^\star)}} \frac{\tau^\star - k}{\sqrt{n}} + o\left(\frac{|\tau^\star - k|}{\sqrt{n}}\right).$$

The case $k \ge \tau^\star$ is identical with $g(\lambda) := \sqrt{(1-\lambda)/\lambda}$, since

$$g'(\lambda) = -\frac{1}{2\lambda^{3/2}\sqrt{1 - \lambda}}.$$

This yields the same first-order coefficient and proves (4.26). □

4.6. Localization of the robust argmax estimator

Theorem 4.10 (Nonasymptotic localization)**.** *Assume 2.2, 2.3, and 4.7. Fix $u > 0$ and a block length ℓ with $m = \lfloor n/\ell \rfloor$. Let η_α be defined by* (4.4)*, and assume*

$$(m - 1)\beta(\ell) \le \alpha.$$

Then, with probability at least $1 - \alpha$,

$$|\widehat{\tau} - \tau^\star| \le \left\lceil \frac{2\sqrt{n}\,\eta_\alpha}{c_\rho\,\kappa} \right\rceil, \tag{4.27}$$

where c_ρ is defined in (4.20)*.*

Proof. Define the deviation event

$$\mathcal{G}_\alpha := \left\{ \max_{k \in \mathcal{I}_n(\gamma)} \|D^{(u)}(k) - \Theta^{(u)}(k)\|_\infty \le \eta_\alpha \right\}.$$

By Theorem 4.1, $\mathbb{P}(\mathcal{G}_\alpha) \ge 1 - \alpha$.

Fix an outcome in $\mathcal{G}_\alpha$. For every $k \in \mathcal{I}_n(\gamma)$,

$$\|D^{(u)}(k)\|_\infty \le \|\Theta^{(u)}(k)\|_\infty + \eta_\alpha, \qquad \|D^{(u)}(\tau^\star)\|_\infty \ge \|\Theta^{(u)}(\tau^\star)\|_\infty - \eta_\alpha. \tag{4.28}$$

Subtracting the two inequalities gives

$$\|D^{(u)}(\tau^\star)\|_\infty - \|D^{(u)}(k)\|_\infty \ge \|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty - 2\eta_\alpha. \tag{4.29}$$

By Lemma 4.8,

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty \ge c_\rho \kappa \frac{|k-\tau^\star|}{\sqrt{n}}.$$

Hence

$$\|D^{(u)}(\tau^\star)\|_\infty - \|D^{(u)}(k)\|_\infty \ge c_\rho \kappa \frac{|k-\tau^\star|}{\sqrt{n}} - 2\eta_\alpha. \tag{4.30}$$

If $|k-\tau^\star| > 2\sqrt{n}\,\eta_\alpha/(c_\rho\kappa)$, then the right-hand side of (4.30) is strictly positive, so

$$\|D^{(u)}(\tau^\star)\|_\infty > \|D^{(u)}(k)\|_\infty.$$

Therefore every maximizer of $k \mapsto \|D^{(u)}(k)\|_\infty$ over $\mathcal{I}_n(\gamma)$ lies inside the stated radius. In particular,

$$\lceil \widehat{\tau} - \tau^\star \rceil \le \left\lceil \frac{2\sqrt{n}\,\eta_\alpha}{c_\rho \kappa} \right\rceil.$$

The event $\mathcal{G}_\alpha$ has a probability of at least $1-\alpha$, which completes the proof. □

4.7. A finite-sample confidence set by inversion

Theorem 4.11 (Finite-sample confidence set)**.** *Assume 2.2, 2.3, and 4.7. Fix $u > 0$ and a block length ℓ with $m = \lfloor n/\ell \rfloor$. Let η_α be defined by* (4.4)*, and assume $(m-1)\beta(\ell) \le \alpha$. Choose*

$$r \ge \left\lceil \frac{2\sqrt{n}\,\eta_\alpha}{c_\rho \kappa} \right\rceil, \qquad \widehat{q}_\alpha(r) := 4\eta_\alpha. \tag{4.31}$$

Then the confidence set $C_\alpha(r)$ in (2.8) *satisfies*

$$\mathbb{P}(\tau^\star \in C_\alpha(r)) \ge 1-\alpha.$$

Proof. On the event $\mathcal{G}_\alpha$ from the proof of Theorem 4.10, Theorem 4.10 implies $\lceil \widehat{\tau} - \tau^\star \rceil \le r$. Hence $\tau^\star \in \mathcal{K}(r)$ on $\mathcal{G}_\alpha$.

It remains to verify the contrast inequality in the definition of $C_\alpha(r)$ at $k = \tau^\star$. Since $\widehat{\tau}$ maximizes $\|D^{(u)}(k)\|_\infty$ over $\mathcal{I}_n(\gamma)$,

$$\|D^{(u)}(\widehat{\tau})\|_\infty \ge \|D^{(u)}(\tau^\star)\|_\infty.$$

On $\mathcal{G}_\alpha$,

$$\|D^{(u)}(\widehat{\tau})\|_\infty \le \|\Theta^{(u)}(\widehat{\tau})\|_\infty + \eta_\alpha, \qquad \|D^{(u)}(\tau^\star)\|_\infty \ge \|\Theta^{(u)}(\tau^\star)\|_\infty - \eta_\alpha.$$

Combining these inequalities gives

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(\widehat{\tau})\|_\infty \le 2\eta_\alpha.$$

Now compute

$$\begin{aligned}\|D^{(u)}(\widehat{\tau})\|_\infty - \|D^{(u)}(\tau^\star)\|_\infty &\le (\|\Theta^{(u)}(\widehat{\tau})\|_\infty + \eta_\alpha) - (\|\Theta^{(u)}(\tau^\star)\|_\infty - \eta_\alpha)\\ &= \|\Theta^{(u)}(\widehat{\tau})\|_\infty - \|\Theta^{(u)}(\tau^\star)\|_\infty + 2\eta_\alpha\\ &\le 4\eta_\alpha = \widehat{q}_\alpha(r).\end{aligned}$$

Thus $\tau^\star \in C_\alpha(r)$ on $\mathcal{G}_\alpha$. Since $\mathbb{P}(\mathcal{G}_\alpha) \ge 1-\alpha$, the result follows. □

4.8. A finite-sample diameter bound for the confidence set

The next theorem adds an explicit bound on the radius and cardinality of the confidence set. This result sharpens the qualitative coverage statement by quantifying the size of the inferred uncertainty region.

Theorem 4.12 (Confidence-set diameter bound)**.** *Assume the hypotheses of Theorem 4.11. On the event*

$$\mathcal{G}_\alpha = \left\{ \max_{k \in \mathcal{I}_n(\gamma)} \|D^{(u)}(k) - \Theta^{(u)}(k)\|_\infty \le \eta_\alpha \right\},$$

the confidence set $C_\alpha(r)$ with $\widehat{q}_\alpha(r) = 4\eta_\alpha$ obeys

$$C_\alpha(r) \subseteq \left\{ k \in \mathcal{I}_n(\gamma) : \ |k - \tau^\star| \le \left\lceil \frac{8\sqrt{n}\,\eta_\alpha}{c_\rho \kappa} \right\rceil \right\}. \tag{4.32}$$

Consequently, with probability at least $1-\alpha$,

$$|C_\alpha(r)| \le 2\left\lceil \frac{8\sqrt{n}\,\eta_\alpha}{c_\rho \kappa} \right\rceil + 1. \tag{4.33}$$

Proof. Fix an outcome in $\mathcal{G}_\alpha$, and let $k \in C_\alpha(r)$. By definition of $C_\alpha(r)$,

$$\|D^{(u)}(\widehat{\tau})\|_\infty - \|D^{(u)}(k)\|_\infty \le 4\eta_\alpha. \tag{4.34}$$

As in the proof of Theorem 4.11, the maximizing property of $\widehat{\tau}$ and the event $\mathcal{G}_\alpha$ imply

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(\widehat{\tau})\|_\infty \le 2\eta_\alpha. \tag{4.35}$$

We now bound the population contrast gap at k:

$$\begin{aligned}\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty &= (\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(\widehat{\tau})\|_\infty) + (\|\Theta^{(u)}(\widehat{\tau})\|_\infty - \|\Theta^{(u)}(k)\|_\infty)\\ &\le 2\eta_\alpha + (\|D^{(u)}(\widehat{\tau})\|_\infty + \eta_\alpha) - (\|D^{(u)}(k)\|_\infty - \eta_\alpha)\\ &= 2\eta_\alpha + (\|D^{(u)}(\widehat{\tau})\|_\infty - \|D^{(u)}(k)\|_\infty) + 2\eta_\alpha\\ &\le 2\eta_\alpha + 4\eta_\alpha + 2\eta_\alpha = 8\eta_\alpha,\end{aligned}$$

where (4.34) was used in the final line.

Lemma 4.8 gives the lower bound

$$\|\Theta^{(u)}(\tau^\star)\|_\infty - \|\Theta^{(u)}(k)\|_\infty \ge c_\rho \kappa \frac{|k - \tau^\star|}{\sqrt{n}}.$$

Combining the upper and lower bounds yields

$$c_\rho \kappa \frac{|k - \tau^\star|}{\sqrt{n}} \le 8\eta_\alpha,$$

which proves (4.32). The cardinality bound (4.33) follows immediately. □

5. Numerical study and implementation guidance

This section examines the finite-sample behavior of the proposed statistics in a controlled heavy-tailed dependent design and records practical implementation guidance. All simulations are reproducible from the Python script delivered with the manuscript.

5.1. Simulation design

We generate a p-dimensional AR(1) process with a sparse mean break,

$$X_{t,j} - \mu_{t,j} = \phi(X_{t-1,j} - \mu_{t-1,j}) + \xi_{t,j}, \qquad \phi = 0.35,$$

where the innovations $\xi_{t,j}$ are independent across j and follow a standardized Student-t_5 law. The baseline design uses $n = 400$, $p = 40$, $\tau^\star = 200$, and a sparse shift acting on the first four coordinates:

$$\mu_{t,j} = \begin{cases} 0, & t \le \tau^\star, \\ \vartheta, & t > \tau^\star,\ j \le 4, \\ 0, & t > \tau^\star,\ j > 4. \end{cases}$$

The truncation level is fixed at the theoretical scale

$$u = \left(\frac{n}{\log(8np/0.05)}\right)^{1/3} \approx 3.00.$$

For the self-normalized detector S_u, each pair (γ, ℓ) is calibrated by the empirical 0.95 null quantile from 250 Monte Carlo replications. Power is then estimated from 250 replications under the alternative. This calibration keeps the same robust statistic as in the theory and isolates the practical effect of (γ, ℓ) on finite-sample performance.

5.2. Empirical size and power

Figure 1 and Table 1 show that the empirical size remains tightly anchored near the nominal 0.05 level across the full tuning grid. The same experiments reveal a clear dependence of power on the block length ℓ. For $\gamma = 0.10$, the rejection probability decreases from 0.812 at $\ell = 5$ to 0.024 at $\ell = 40$. For $\ell = 20$, the trim level has a milder effect, with power ranging from 0.232 at $\gamma = 0.05$ to 0.436 at $\gamma = 0.20$. In this design, short to moderate blocks and an interior trim level in the range 0.10–0.20 deliver the strongest performance.

The robust max-CUSUM detector without self-normalization is also stable in this design. Using the empirical 0.95 null quantile for T_u, the rejection probability under the alternative lies between 0.804 and 0.904 across $\gamma \in \{0.05, 0.10, 0.15, 0.20\}$. This pattern indicates that the main numerical sensitivity enters through block-scale calibration rather than through the robust contrast itself.

Table 1. Empirical size and power of the self-normalized detector at nominal level 0.05 under the heavy-tailed AR(1) design described in Section 5. The left block fixes $\gamma = 0.10$ and varies ℓ, while the right block fixes $\ell = 20$ and varies γ.

$\gamma = 0.10$				$\ell = 20$			
ℓ	threshold	size	power	γ	threshold	size	power
5	4.147	0.052	0.812	0.05	3.818	0.052	0.232
10	3.865	0.052	0.748	0.10	3.705	0.052	0.384
20	3.705	0.052	0.384	0.15	3.668	0.052	0.432
30	3.947	0.052	0.068	0.20	3.657	0.052	0.436
40	4.126	0.052	0.024				

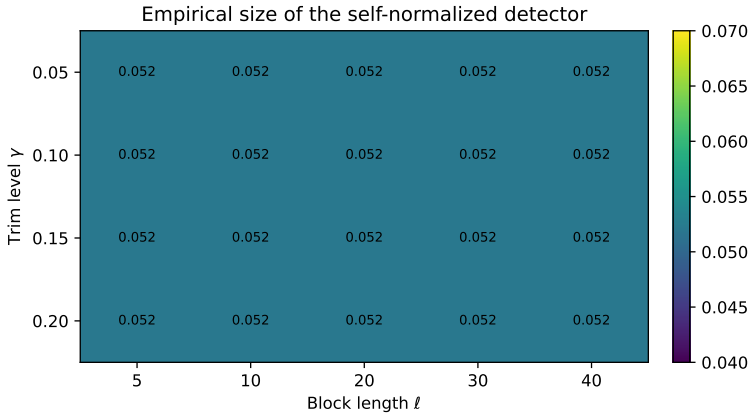

Figure 1. Empirical size of the self-normalized detector over the tuning grid (γ, ℓ).

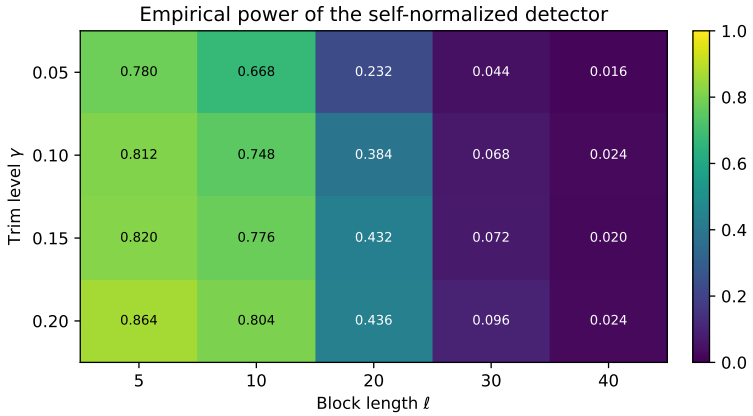

Figure 2. Empirical power of the self-normalized detector over the tuning grid (γ, ℓ) for a sparse mean shift with $\vartheta = 0.50$.

5.3. Representative contrast trajectories and localization

Figure 3 displays a representative sample path with one large outlier injected at time $t = 75$. The raw max-CUSUM curve attains its argmax at $k = 64$, while the truncated robust curve peaks at the true

location $k = 200$. This example illustrates how coordinatewise truncation suppresses spurious peaks created by rare large observations and preserves the structural break signal.

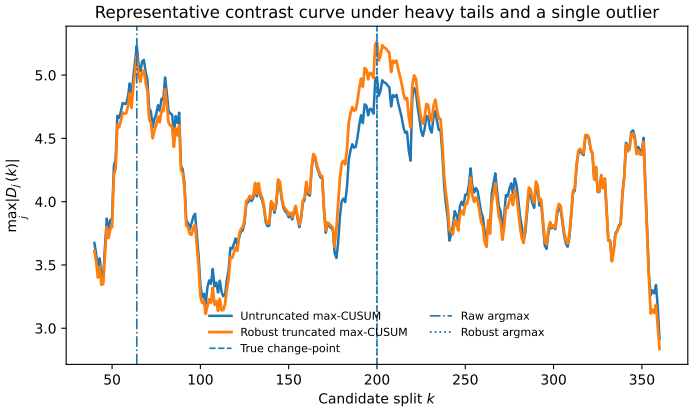

Figure 3. Representative raw and robust max-CUSUM curves under heavy tails with a single large contamination.

Table 2 and Figure 4 quantify the localization accuracy of the robust argmax estimator. As the shift strength increases from $\vartheta = 0.25$ to $\vartheta = 0.70$, the mean absolute localization error decreases from 74.75 to 14.07, and the median error decreases from 72.00 to 8.00. The monotone contraction of the error distribution is consistent with the population margin mechanism proved in Lemma 4.8 and Theorem 4.10.

Table 2. Localization error of the robust argmax estimator for $\gamma = 0.10$ as the sparse mean shift strength ϑ varies.

ϑ	mean $\lvert\widehat{\tau} - \tau^\star\rvert$	median $\lvert\widehat{\tau} - \tau^\star\rvert$	0.90-quantile	0.95-quantile
0.25	74.75	72.00	151.20	156.55
0.35	51.79	31.00	133.00	149.00
0.50	25.05	12.50	71.00	103.50
0.70	14.07	8.00	36.10	55.55

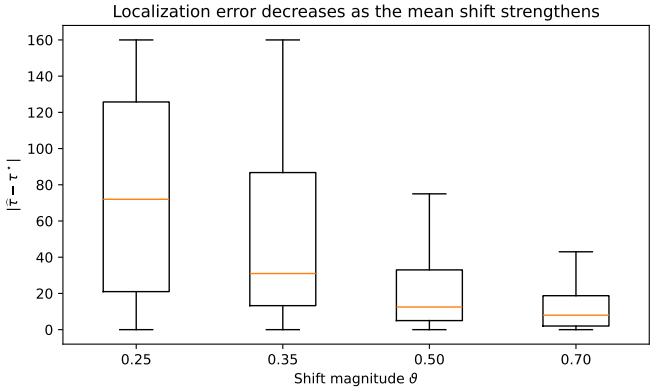

Figure 4. Localization error of the robust argmax estimator as the sparse mean shift strengthens.

5.4. Computational comparison with a WBS-style extension

To assess the computational implications of extending the robust contrast to multiple change-points, we compared one full scan of the proposed single-change statistic with a WBS-style extension based on 250 random intervals. Figure 5 reports the observed runtimes. Over the range $n = 200$ to $n = 1000$, the average runtime of the single full scan increases from about 0.15 ms to 0.77 ms, while the WBS-style extension increases from about 12.8 ms to 31.3 ms. The gap is expected because the multiple-change extension repeatedly reevaluates local contrasts over many intervals. In practice, the present single-change procedure is computationally lightweight, and a multiple-change deployment should balance interval count against desired detection granularity.

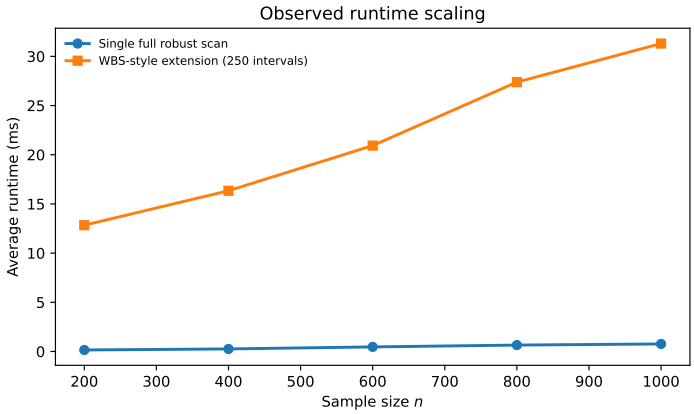

Figure 5. Observed runtime scaling for one full robust scan and a WBS-style extension using 250 random intervals.

5.5. Practical guidance and numerical limitations

The numerical evidence in this section suggests three implementation lessons. First, moderate trimming improves stability near the boundaries and sharpens localization. Second, the block length should reflect the short-range dependence scale; overly large blocks dilute power in finite samples. Third, the robust contrast itself remains strong across the tested trim levels, while calibration through the block denominator deserves the most tuning effort.

The current study covers one sparse AR(1) dependence pattern, one heavy-tail family, empirical null calibration, and a single structural break. Cross-sectional dependence, multiple change-points, data-driven block selection, and confidence-set calibration through dependent resampling are natural next steps for a broader empirical map of the method.

6. Conclusions and future directions

This paper develops a finite-sample framework for robust high-dimensional change-point inference under heavy tails and temporal dependence. The simulation study complements the theory by showing calibrated finite-sample size, a clear tuning response in power, improved robustness of the truncated contrast under contamination, and fast runtime for the single-change scan. The method combines coordinatewise truncation, a trimmed max-CUSUM contrast, and block self-normalization. The analysis

provides explicit probability bounds under β-mixing, finite-sample level control for the self-normalized test under a block-scale nondegeneracy condition, a finite-sample power guarantee for the robust max-CUSUM detector, a localization theorem for the robust argmax estimator, and a confidence-set construction with both coverage and diameter guarantees.

The trimming parameter plays a structural role in the analysis. It yields uniform control of normalized partial sums over the candidate domain and produces a clean deterministic margin for the population contrast. This structure is valuable in heavy-tailed dependent settings because the same trimming condition stabilizes both concentration and localization. The resulting bounds retain explicit dependence on (n, p, γ), the truncation level, the block length, and the mixing profile, which is useful for implementation and calibration.

Several directions are natural next steps. One direction is a data-driven tuning scheme for (u, ℓ, γ) that preserves finite-sample guarantees while improving practical sensitivity. A second direction is a sharper self-normalized calibration strategy using block multiplier methods or dependent Gaussian approximations, which may reduce conservativeness in the current thresholds. A third direction is extension to structured cross-sectional dependence, where sparse or low-rank covariance structure can improve detection and localization rates.

Multiple change-points form another important extension. The robust contrast and the finite-sample deviation bounds in this paper fit naturally with segmentation algorithms such as seeded or wild binary segmentation, and the present arguments suggest a path toward heavy-tail and dependence-robust guarantees in that setting. Additional extensions include changes in covariance, quantiles, and low-dimensional projections, where truncation and self-normalization remain effective ingredients. These directions will strengthen the robust nonasymptotic toolkit for modern dependent high-dimensional data.

Author contributions

All the authors contributed equally. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2026).

Conflict of interest

The authors declare that there are no conflict of interests.

References

1. J. Bai, P. Perron, Estimating and testing linear models with multiple structural changes, *Econometrica*, **66** (1998), 47–78. https://doi.org/10.2307/2998540
2. S. Boucheron, G. Lugosi, P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*, Oxford: Oxford University Press, 2013.
3. R. C. Bradley, Basic properties of strong mixing conditions. A survey and some open questions, *Probab. Surveys*, **2** (2005), 107–144. http://doi.org/10.1214/154957805100000104
4. B. E. Brodsky, B. S. Darkhovsky, *Nonparametric Methods in Change-Point Problems*, Dordrecht: Kluwer Academic, 1993.
5. O. Catoni, Challenging the empirical mean and empirical variance: A deviation study, *Ann. Institut Henri Poincaré, Probab. Statist.*, **48** (2012), 1148–1185. https://doi.org/10.1214/11-AIHP454
6. S. X. Chen, Y. L. Qin, A two-sample test for high-dimensional data with applications to gene-set testing, *Ann. Statist.*, **38** (2010), 808–835. https://doi.org/10.1214/09-AOS716
7. H. Cho, P. Fryzlewicz, Multiple-change-point detection for high-dimensional time series via sparsified binary segmentation, *J. Royal Statist. Soc.: Ser. B*, **77** (2015), 475–507. https://doi.org/10.1111/rssb.12079
8. M. Csörgő, L. Horváth, *Limit Theorems in Change-Point Analysis*, Hoboken: Wiley, 1997.
9. H. Dehling, R. Fried, M. Wendler, A robust method for shift detection in time series, *Biometrika*, **107** (2020), 647–660. https://doi.org/10.1093/biomet/asaa004
10. P. Fryzlewicz, Wild binary segmentation for multiple change-point detection, *Ann. Statist.*, **42** (2014), 2243–2281. https://doi.org/10.1214/14-AOS1245
11. F. Jiang, R. Wang, X. Shao, Robust inference for change points in high dimension, *J. Multivar. Anal.*, **193** (2023), 105114.
12. M. Jirak, Uniform change point tests in high dimension, *Ann. Statist.*, **43** (2015), 2451–2483.
13. D. V. Hinkley, Inference about the change-point in a sequence of random variables, *Biometrika*, **57** (1970), 1–17. https://doi.org/10.1093/biomet/57.1.1
14. P. J. Huber, Robust estimation of a location parameter, In: *Breakthroughs in Statistics: Methodology and Distribution*, New York: Springer, 1992. 492–518. https://doi.org/10.1007/978-1-4612-4380-9_35
15. S. Kovács, P. Bühlmann, H. Li, A. Munk, Seeded binary segmentation: A general methodology for fast and optimal change point detection, *Biometrika*, **110** (2023), 249–256. https://doi.org/10.1093/biomet/asac052
16. F. Merlevède, M. Peligrad, E. Rio, Bernstein inequality and moderate deviations under strong mixing conditions, *High Dimensional Probab.*, **5** (2009), 273–292. https://doi.org/10.1214/09-IMSCOLL518
17. E. S. Page, Continuous inspection schemes, *Biometrika*, **41** (1954), 100–115. https://doi.org/10.2307/2333009

18. D. Paulin, Concentration inequalities for Markov chains by Marton couplings and spectral methods, *Electron. J. Probab.*, **20** (2015), 1–32. http://doi.org/10.1214/EJP.v20-4039

19. R. Vershynin, *High-Dimensional Probability*, Cambridge: Cambridge University Press, 2018.

20. R. Wang, C. Zhu, S. Volgushev, X. Shao, Inference for change points in high-dimensional data via self-normalization, *Ann. Statist.*, **50** (2022), 781–806.

AIMS Press