Supplemental Appendix to:

Multiple testing in rolling-window analysis via p-value combination

Xing Ling, Qian Wu*, Kun Chen

School of Statistics and Data Science, Southwestern University of Finance and Economics, China

This document provides proofs of the theoretical results in the main text. All notation and assumptions are as in the main text.

Appendix A1. Technical lemmas

Lemma 1 (Joint FCLT for window fluctuations)**.** *Under Assumptions 1–2, for any distinct windows $k \neq \ell$ that are not asymptotically fully overlapping,*

$$\begin{pmatrix} U_{k,m} \\ U_{\ell,m} \end{pmatrix} \Rightarrow N\left(0,\ \Sigma_{k\ell}\right) \quad \textit{with} \quad \Sigma_{k\ell} = \begin{pmatrix} \Gamma & \kappa_{k\ell}\Gamma \\ \kappa_{k\ell}\Gamma & \Gamma \end{pmatrix}, \tag{1}$$

where $\Sigma_{k\ell}$ is a $2d \times 2d$ block covariance matrix.

Proof. Recall $X_t = h(Z_t) - \theta_0 \in \mathbb{R}^d$, where $\theta_0 = \mathbb{E}[h(Z_0)]$ under the global null $\boldsymbol{H}_0$. Let $S_n := \sum_{t=1}^{n} X_t$ with the convention $S_0 = 0$ and define the partial-sum process

$$\mathbb{S}_m(t) := \frac{1}{\sqrt{m}} S_{\lfloor mt \rfloor}, \qquad t \geq 0,$$

where $\lfloor \cdot \rfloor$ denotes the floor function. Under Assumption 1, a functional central limit theorem

*Correspondence to: School of Statistics and Data Science, Southwestern University of Finance and Economics, Chengdu, China; E-mail address: wuqj@swufe.edu.cn (Q. Wu)

(FCLT) holds in $D([0,L],\mathbb{R}^d)$ for any fixed $L<\infty$,

$$\mathbb{S}_m(\cdot) \Rightarrow B(\cdot),$$

where $\Rightarrow$ denotes weak convergence and $B(\cdot)$ is a d-dimensional Brownian motion with $\mathrm{Cov}(B(t)) = t\,\Gamma$ (Herrndorf, 1985).

Each window fluctuation can be written as an increment of the partial-sum process. In the k-th window $W_k = \{k, k+1, \ldots, k+m-1\}$,

$$\widehat{\theta}_{k,m} = \frac{1}{m}\sum_{t\in W_k} h(Z_t) \quad \text{and} \quad U_{k,m} = \sqrt{m}\,(\widehat{\theta}_{k,m} - \theta_0).$$

Therefore,

$$U_{k,m} = \frac{1}{\sqrt{m}}\sum_{t=k}^{k+m-1} X_t = \frac{S_{k+m-1} - S_{k-1}}{\sqrt{m}},$$

and similarly $U_{\ell,m} = (S_{\ell+m-1} - S_{\ell-1})/\sqrt{m}$.

Let $s_{k\ell} := |\ell - k|$. By Assumption 2, $\delta_m = s_{k\ell}/m \to \delta_{k\ell} \in (0,\infty)$. Hence, the windows are not asymptotically fully overlapping. Since $\{X_t\}$ is strictly stationary under $\boldsymbol{H}_0$, the joint distribution of the two window increments depends on (k,ℓ) only through the separation $s_{k\ell}$. Therefore, we may shift indices and work with the representative pair of windows starting at 1 and $1+s_{k\ell}$, which yields

$$(U_{k,m}, U_{\ell,m}) \;\stackrel{d}{=}\; \Big(\frac{S_m - S_0}{\sqrt{m}},\; \frac{S_{m+s_{k\ell}} - S_{s_{k\ell}}}{\sqrt{m}}\Big),$$

where $\stackrel{d}{=}$ denotes equality in distribution. Choose $L > 1+\delta_{k\ell}$. Then $1+\delta_m \le L$ for all sufficiently large m. We have $\lfloor m\delta_m \rfloor = s_{k\ell}$ and $\lfloor m(1+\delta_m)\rfloor = m + s_{k\ell}$, so $(S_{m+s_{k\ell}} - S_{s_{k\ell}})/\sqrt{m} = \mathbb{S}_m(1+\delta_m) - \mathbb{S}_m(\delta_m)$. Thus,

$$(U_{k,m}, U_{\ell,m}) \stackrel{d}{=} \big(\mathbb{S}_m(1) - \mathbb{S}_m(0),\; \mathbb{S}_m(1+\delta_m) - \mathbb{S}_m(\delta_m)\big).$$

We have $\mathbb{S}_m \Rightarrow B$ in $D([0,L],\mathbb{R}^d)$. Since δ_m is deterministic and $\delta_m \to \delta_{k\ell}$, we have $(\mathbb{S}_m, \delta_m) \Rightarrow (B, \delta_{k\ell})$ in $D([0,L],\mathbb{R}^d)\times\mathbb{R}$. Define the increment map $\Psi(x,\delta) := \big(x(1) - x(0),\; x(1+\delta) - x(\delta)\big)$ for $x \in D([0,L],\mathbb{R}^d)$. Ψ is almost surely continuous at $(B,\delta_{k\ell})$ because B has continuous sample paths. Therefore, by the continuous mapping theorem (Billingsley, 1999), we obtain

$$(U_{k,m}, U_{\ell,m}) \Rightarrow \big(B(1) - B(0),\ B(1+\delta_{k\ell}) - B(\delta_{k\ell})\big) =: (\Delta_k, \Delta_\ell),$$

where the limit (Δ_k, Δ_ℓ) is centered and jointly Gaussian.

We next compute the covariance of (Δ_k, Δ_ℓ). Since $\mathrm{Cov}(B(s), B(t)) = \min\{s,t\}\Gamma$, we have $\mathrm{Cov}\big(B(b) - B(a),\ B(d) - B(c)\big) = \big|[a,b] \cap [c,d]\big|\Gamma$. Therefore,

$$\mathrm{Var}(\Delta_k) = \mathrm{Var}(\Delta_\ell) = \Gamma \quad \text{and} \quad \mathrm{Cov}(\Delta_k, \Delta_\ell) = \big|[0,1] \cap [\delta_{k\ell}, 1+\delta_{k\ell}]\big|\Gamma = \kappa_{k\ell}\Gamma,$$

where $\kappa_{k\ell} = \max\{1 - \delta_{k\ell}, 0\}$. This completes the proof. □

Lemma 2 (Asymptotically tail-independent Gaussian norm exceedances)**.** *Let (Δ_1, Δ_2) be a centered jointly Gaussian vector in $\mathbb{R}^d \times \mathbb{R}^d$ with block covariance matrix*

$$\mathrm{Cov}\begin{pmatrix}\Delta_1 \\ \Delta_2\end{pmatrix} = \begin{pmatrix}\Gamma & \kappa\Gamma \\ \kappa\Gamma & \Gamma\end{pmatrix}, \qquad \kappa \in [0,1),$$

where Γ is positive semidefinite and satisfies $\Gamma_{jj} > 0$ for $j = 1, \ldots, d$. Let $\|\cdot\|$ denote the Euclidean norm. Then for any $\eta > 0$ satisfying $(1+\eta) < \sqrt{2/(1+\kappa)}$, as $x \to \infty$,

$$\mathbb{P}\big(\|\Delta_1\| \geq x,\ \|\Delta_2\| \geq x\big) = o\big(\mathbb{P}(\|\Delta_1\| \geq (1+\eta)x)\big).$$

As a consequence,

$$\mathbb{P}\big(\|\Delta_1\| \geq x,\ \|\Delta_2\| \geq x\big) = o\big(\mathbb{P}(\|\Delta_1\| \geq x)\big).$$

Proof. Write $\langle \cdot, \cdot \rangle$ for the Euclidean inner product. Let $\mathbb{U} = \{u \in \mathbb{R}^d : \|u\| = 1\}$. Fix $\varepsilon \in (0,1)$ (to be chosen later). Let $\mathcal{N}_\varepsilon \subset \mathbb{U}$ be an ε-net of $\mathbb{U}$. A standard covering-number bound implies $|\mathcal{N}_\varepsilon| \leq (2/\varepsilon + 1)^d \leq (3/\varepsilon)^d$; see (Vershynin, 2018, Corollary 4.2.11). The net yields the following reduction: for any $z \in \mathbb{R}^d$ and any $x > 0$,

$$\{\|z\| \geq x\} \subseteq \bigcup_{u \in \mathcal{N}_\varepsilon} \{\langle u, z\rangle \geq (1-\varepsilon)x\}.$$

To see this, assume $z \neq 0$ and set $u^\star = z/\|z\| \in \mathbb{U}$. Choose $u \in \mathcal{N}_\varepsilon$ with $\|u - u^\star\| \leq \varepsilon$. Then

$$\langle u, z\rangle = \langle u^\star, z\rangle + \langle u - u^\star, z\rangle = \|z\| + \langle u - u^\star, z\rangle \geq \|z\| - \|u - u^\star\| \, \|z\| \geq (1-\varepsilon)\|z\|.$$

This verifies the reduction; the case $z = 0$ is trivial.

Fix $u, v \in \mathcal{N}_\varepsilon$. Set $A_u := \langle u, \Delta_1\rangle$ and $B_v := \langle v, \Delta_2\rangle$. Set $y := (1-\varepsilon)x$. Applying the reduction above to Δ_1 and Δ_2 and intersecting the resulting events yields

$$\{\|\Delta_1\| \geq x, \ \|\Delta_2\| \geq x\} \subseteq \bigcup_{u \in \mathcal{N}_\varepsilon} \bigcup_{v \in \mathcal{N}_\varepsilon} \{A_u \geq y, \ B_v \geq y\}.$$

By the union bound,

$$\mathbb{P}(\|\Delta_1\| \geq x, \ \|\Delta_2\| \geq x) \leq |\mathcal{N}_\varepsilon|^2 \max_{u,v \in \mathcal{N}_\varepsilon} \mathbb{P}\left(A_u \geq y, \ B_v \geq y\right). \tag{2}$$

We now bound $\mathbb{P}\left(A_u \geq y, B_v \geq y\right)$ uniformly over $u, v \in \mathcal{N}_\varepsilon$. Let $\sigma_u^2 := \mathrm{Var}(A_u) = u^\top \Gamma u$ and $\sigma_v^2 := \mathrm{Var}(B_v) = v^\top \Gamma v$. If $\sigma_u = 0$ then $A_u = 0$ almost surely, so $\mathbb{P}(A_u \geq y, B_v \geq y) = 0$ for $y > 0$; similarly for $\sigma_v = 0$. Hence we may assume $\sigma_u, \sigma_v > 0$ and standardize $X := A_u/\sigma_u$ and $Y := B_v/\sigma_v$. Since (Δ_1, Δ_2) is centered jointly Gaussian, (A_u, B_v) is centered bivariate normal; hence (X, Y) is standard bivariate normal with correlation ρ, where

$$\rho = \mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(A_u, B_v)}{\sigma_u \sigma_v} = \kappa \frac{u^\top \Gamma v}{\sqrt{u^\top \Gamma u}\sqrt{v^\top \Gamma v}}.$$

By the Cauchy–Schwarz inequality, $|\rho| \leq \kappa < 1$, and therefore $X + Y \sim N(0, 2 + 2\rho)$. Let Φ be the standard normal CDF, $\bar{\Phi} = 1 - \Phi$ its upper tail, and ϕ its density. Since $\{X \geq a, Y \geq b\} \subseteq \{X + Y \geq a + b\}$ for any $a, b \in \mathbb{R}$, we have

$$\mathbb{P}(X \geq a, \ Y \geq b) \leq \mathbb{P}(X + Y \geq a + b) \leq \bar{\Phi}\left(\frac{a+b}{\sqrt{2+2\rho}}\right).$$

Take $a = y/\sigma_u$ and $b = y/\sigma_v$. Let $\lambda_{\max} > 0$ be the largest eigenvalue of Γ and set $\sigma_{\max} := \sqrt{\lambda_{\max}}$.

Then $\sigma_u, \sigma_v \le \sigma_{\max}$. Consequently,

$$\frac{a+b}{\sqrt{2+2\rho}} \ge \frac{2y/\sigma_{\max}}{\sqrt{2+2\kappa}} = (1-\varepsilon)\sqrt{\frac{2}{1+\kappa}}\,\frac{x}{\sigma_{\max}}.$$

Define $\beta := (1-\varepsilon)\sqrt{2/(1+\kappa)}$. Since $\bar{\Phi}$ is decreasing, the above inequality implies the uniform bound

$$\mathbb{P}(A_u \ge y,\ B_v \ge y) = \mathbb{P}(X \ge a,\ Y \ge b) \le \bar{\Phi}\left(\beta\,\frac{x}{\sigma_{\max}}\right).$$

Combining this inequality with the union bound in (2) yields

$$\mathbb{P}(\|\Delta_1\| \ge x,\ \|\Delta_2\| \ge x) \le \left(\frac{3}{\varepsilon}\right)^{2d} \bar{\Phi}\left(\beta\,\frac{x}{\sigma_{\max}}\right), \tag{3}$$

using $|\mathcal{N}_\varepsilon| \le (3/\varepsilon)^d$.

Next we lower bound the marginal tail. Let $w \in \mathbb{U}$ be a unit eigenvector corresponding to $\lambda_{\max}$. Then $\langle w, \Delta_1\rangle \sim N(0, \lambda_{\max})$. Because $\langle w, \Delta_1\rangle \le \|w\|\,\|\Delta_1\| = \|\Delta_1\|$, it follows that $\{\langle w, \Delta_1\rangle \ge s\} \subseteq \{\|\Delta_1\| \ge s\}$. Hence,

$$\mathbb{P}(\|\Delta_1\| \ge s) \ \ge\ \mathbb{P}\big(\langle w, \Delta_1\rangle \ge s\big) = \bar{\Phi}\left(\frac{s}{\sigma_{\max}}\right). \tag{4}$$

Fix $\eta > 0$ with $(1+\eta) < \sqrt{2/(1+\kappa)}$. Then we can choose $\varepsilon \in (0,1)$ small enough so that $\beta = (1-\varepsilon)\sqrt{2/(1+\kappa)} > 1+\eta$. Dividing (3) by (4) with $s = (1+\eta)x$ gives

$$\frac{\mathbb{P}(\|\Delta_1\| \ge x,\ \|\Delta_2\| \ge x)}{\mathbb{P}\big(\|\Delta_1\| \ge (1+\eta)x\big)} \le \left(\frac{3}{\varepsilon}\right)^{2d} \frac{\bar{\Phi}(\beta t)}{\bar{\Phi}\left((1+\eta)t\right)},$$

where we write $t = x/\sigma_{\max}$.

It remains to show that $\bar{\Phi}(\beta t)/\bar{\Phi}((1+\eta)t) \to 0$ as $t \to \infty$. Standard Mills' ratio bounds imply that, for any $r > 0$,

$$\frac{\phi(r)\,r}{1+r^2} \le \bar{\Phi}(r) \le \frac{\phi(r)}{r}.$$

Combining the upper bound for $\bar{\Phi}(\beta t)$ and the lower bound for $\bar{\Phi}((1+\eta)t)$, we obtain

$$\frac{\bar{\Phi}(\beta t)}{\bar{\Phi}((1+\eta)t)} \le \frac{\phi(\beta t)/(\beta t)}{\phi((1+\eta)t)\cdot(1+\eta)t/(1+(1+\eta)^2t^2)} = \frac{1+(1+\eta)^2t^2}{\beta(1+\eta)t^2}\exp\!\left(\frac{t^2}{2}\left((1+\eta)^2-\beta^2\right)\right).$$

Since $\beta > (1+\eta)$, the exponent is strictly negative and the right-hand side tends to 0 as $t \to \infty$.

This proves

$$\mathbb{P}(\|\Delta_1\| \geq x,\ \|\Delta_2\| \geq x) = o\big(\mathbb{P}(\|\Delta_1\| \geq (1+\eta)x)\big).$$

In particular, since $\mathbb{P}(\|\Delta_1\| \geq (1+\eta)x) \leq \mathbb{P}(\|\Delta_1\| \geq x)$,

$$\frac{\mathbb{P}(\|\Delta_1\| \geq x,\ \|\Delta_2\| \geq x)}{\mathbb{P}(\|\Delta_1\| \geq x)} \leq \frac{\mathbb{P}(\|\Delta_1\| \geq x,\ \|\Delta_2\| \geq x)}{\mathbb{P}(\|\Delta_1\| \geq (1+\eta)x)} \to 0.$$

This completes the proof.

□

Lemma 3 (Comparability of tail radii)**.** *Under Assumption 3, define*

$$\delta(r) := \sup_{\|u\|\geq r} \big|\tau(u,\xi_0) - g(\|u\|)\big|, \qquad r \geq M.$$

Then $\delta(r) \to 0$ as $r \to \infty$. Define the envelope functions on $[M,\infty)$ by

$$\underline{g}(r) := g(r) - \delta(r), \qquad \overline{g}(r) := g(r) + \delta(r),$$

and the generalized inverse radii

$$r_+(x) := \inf\{r \geq M : \overline{g}(r) \geq x\}, \qquad r_-(x) := \inf\{r \geq M : \underline{g}(r) \geq x\}.$$

Assume the following inverse-shift regularity for g: for every fixed $a > 0$,

$$\frac{g^{-1}(y+a)}{g^{-1}(y)} \to 1 \quad \text{as} \quad y \to \infty,$$

where $g^{-1}(y) := \inf\{r \geq M : g(r) \geq y\}$. Now fix $\varepsilon > 0$. Then for any $\eta > 0$ there exists y_0 such that for all $y \geq y_0$,

$$r_-(y+\varepsilon) \leq (1+\eta)\, r_+(y-\varepsilon).$$

Proof. Fix $\varepsilon > 0$ and $\eta > 0$, and let $y \to \infty$. Since $\delta(r) \to 0$, choose $R \geq M$ such that $\delta(r) \leq \varepsilon/4$ for all $r \geq R$. Since g is nondecreasing and tends to infinity, $g^{-1}(y) \to \infty$ as $y \to \infty$.

We first derive a lower bound on $r_+(y-\varepsilon)$. Let $r_0 := g^{-1}(y - 5\varepsilon/4)$. We claim that $\overline{g}(r) < y - \varepsilon$

for every $r < r_0$ and all sufficiently large y. Fix $r < r_0$. Then $g(r) < y - 5\varepsilon/4$. Consider two cases. If $r \geq R$, then $\delta(r) \leq \varepsilon/4$, and hence

$$\overline{g}(r) = g(r) + \delta(r) \; < \; \left(y - \frac{5\varepsilon}{4}\right) + \frac{\varepsilon}{4} = y - \varepsilon.$$

If $r < R$, then $\overline{g}(r) \leq \sup_{s\in[M,R]} \overline{g}(s) < \infty$, so $\overline{g}(r) < y - \varepsilon$ for all sufficiently large y. Thus, in both cases, no $r < r_0$ can satisfy $\overline{g}(r) \geq y - \varepsilon$, so the infimum defining $r_+(y - \varepsilon)$ is at least r_0. Hence $r_+(y - \varepsilon) \geq r_0$.

Next we derive an upper bound on $r_-(y + \varepsilon)$. Let $r_1 := g^{-1}(y + 5\varepsilon/4)$, so $g(r_1) \geq y + 5\varepsilon/4$. Since $r_1 \to \infty$, we have $r_1 \geq R$ for all large y, and thus $\delta(r_1) \leq \varepsilon/4$. It follows that

$$\underline{g}(r_1) = g(r_1) - \delta(r_1) \geq \left(y + \frac{5\varepsilon}{4}\right) - \frac{\varepsilon}{4} = y + \varepsilon.$$

Hence the infimum defining $r_-(y + \varepsilon)$ is at most r_1, that is, $r_-(y + \varepsilon) \leq r_1$.

Combining $r_+(y - \varepsilon) \geq r_0$ and $r_-(y + \varepsilon) \leq r_1$, we obtain that for all sufficiently large y,

$$\frac{r_-(y+\varepsilon)}{r_+(y-\varepsilon)} \leq \frac{g^{-1}(y + 5\varepsilon/4)}{g^{-1}(y - 5\varepsilon/4)} = \frac{g^{-1}(y + 5\varepsilon/4)}{g^{-1}(y)} \cdot \frac{g^{-1}(y)}{g^{-1}(y - 5\varepsilon/4)}.$$

By the inverse-shift regularity, the right-hand side converges to 1, so there exists y_0 such that for all $y \geq y_0$ the ratio is at most $1 + \eta$, i.e.,

$$\frac{r_-(y+\varepsilon)}{r_+(y-\varepsilon)} \leq 1 + \eta.$$

This completes the proof. □

Lemma 4 (Tail inclusions for p-value events)**.** *Under Assumptions 3–4, write $T_{k,m} = \tau(U_{k,m}, \xi_0) + R_{k,m}$ with $R_{k,m} = o_p(1)$ under $\boldsymbol{H}_0$. Define $c_m(u) := \inf\{t \in \mathbb{R} : \pi_m(t) \leq u\}$. Fix $\varepsilon > 0$. Let $\delta, \overline{g}, \underline{g}, r_+$, and r_- be defined as in Lemma 3. Choose u small enough so that, for all sufficiently large m, $c_m(u) \pm \varepsilon$ fall in the region where Assumption 3 applies; equivalently, $r_+(c_m(u) - \varepsilon)$ and $r_-(c_m(u) + \varepsilon)$ exceed M. The following statements hold:*

- *Upper inclusion.*

$$\{p_{k,m} \le u\} \subseteq \big\{\|U_{k,m}\| \ge r_+\big(c_m(u) - \varepsilon\big)\big\} \;\cup\; \{|R_{k,m}| > \varepsilon\};$$

  consequently, for any ℓ*,*

$$\mathbb{P}(p_{k,m} \le u,\; p_{\ell,m} \le u) \le \mathbb{P}\Big(\|U_{k,m}\| \ge r_+\big(c_m(u) - \varepsilon\big),\; \|U_{\ell,m}\| \ge r_+\big(c_m(u) - \varepsilon\big)\Big) + o(1). \quad (5)$$

- *Lower inclusion.*

$$\big\{\|U_{k,m}\| \ge r_-(c_m(u) + \varepsilon)\big\} \subseteq \{p_{k,m} \le u\} \;\cup\; \{|R_{k,m}| > \varepsilon\};$$

  consequently,

$$\mathbb{P}(p_{k,m} \le u) \ge \mathbb{P}\Big(\|U_{k,m}\| \ge r_-\big(c_m(u) + \varepsilon\big)\Big) - o(1). \quad (6)$$

Here $o(1) \to 0$ *as* $m \to \infty$*.*

Proof. Since π_m is strictly decreasing on the relevant upper tail, the event $\{p_{k,m} \le u\}$ is equivalent to $\{T_{k,m} \ge c_m(u)\}$. We first prove the upper inclusion and the joint probability bound. Write

$$\{T_{k,m} \ge c_m(u)\} \subseteq \big(\{T_{k,m} \ge c_m(u)\} \cap \{|R_{k,m}| \le \varepsilon\}\big) \;\cup\; \{|R_{k,m}| > \varepsilon\}.$$

On the intersection event of the above decomposition, we have $\tau(U_{k,m}, \xi_0) = T_{k,m} - R_{k,m} \ge c_m(u) - \varepsilon$. By the definition of δ and $\overline{g}$, we have $\tau(v, \xi_0) \le \overline{g}(\|v\|)$ for all v with $\|v\| \ge M$. Applying this with $v = U_{k,m}$ gives $\overline{g}(\|U_{k,m}\|) \ge c_m(u) - \varepsilon$. This implies $\|U_{k,m}\| \ge r_+(c_m(u) - \varepsilon)$ by the definition of r_+. Therefore

$$\big(\{T_{k,m} \ge c_m(u)\} \cap \{|R_{k,m}| \le \varepsilon\}\big) \subseteq \{\|U_{k,m}\| \ge r_+(c_m(u) - \varepsilon)\},$$

which yields the upper inclusion. The same argument applies to window ℓ. Intersecting the two inclusions yields

$$\{p_{k,m} \le u,\ p_{\ell,m} \le u\} \subseteq \big\{\|U_{k,m}\| \ge r_+\big(c_m(u)-\varepsilon\big),\ \|U_{\ell,m}\| \ge r_+\big(c_m(u)-\varepsilon\big)\big\}$$
$$\cup \{|R_{k,m}| > \varepsilon\} \cup \{|R_{\ell,m}| > \varepsilon\}.$$

Taking probabilities and using $\mathbb{P}(|R_{k,m}| > \varepsilon) + \mathbb{P}(|R_{\ell,m}| > \varepsilon) = o(1)$ yields the joint probability bound (5).

We next prove the lower inclusion and the marginal probability bound. On the event $\|U_{k,m}\| \ge r_-(c_m(u)+\varepsilon)$, the definition of r_- gives $\underline{g}(\|U_{k,m}\|) \ge c_m(u)+\varepsilon$. Since $\tau(v,\xi_0) \ge \underline{g}(\|v\|)$ for all $\|v\| \ge M$, it follows that $\tau(U_{k,m},\xi_0) \ge c_m(u)+\varepsilon$. Consequently, on $\{|R_{k,m}| \le \varepsilon\}$, we have $T_{k,m} = \tau(U_{k,m},\xi_0) + R_{k,m} \ge c_m(u)$. Thus,

$$\big\{\|U_{k,m}\| \ge r_-\big(c_m(u)+\varepsilon\big)\big\} \subseteq \{T_{k,m} \ge c_m(u)\} \ \cup\ \{|R_{k,m}| > \varepsilon\},$$

which proves the lower inclusion. Taking probabilities and using $\mathbb{P}(|R_{k,m}| > \varepsilon) = o(1)$ yields the marginal bound (6), completing the proof. □

Appendix A2. Proof of Proposition 1

Proof. We fix two distinct windows $k \neq \ell$ that are not asymptotically fully overlapping. Lemma 4 converts the lower-tail events of $(p_{k,m}, p_{\ell,m})$ into norm exceedances of the window fluctuations $(U_{k,m}, U_{\ell,m})$, while Lemma 3 aligns the two radii appearing in the resulting upper and lower bounds. Passing to the limit $m \to \infty$, Lemma 1 provides a joint Gaussian limit, and Lemma 2 yields tail independence for the corresponding Gaussian norm exceedances. Finally, letting $u \downarrow 0$ completes the argument.

We first relate p-values to window fluctuations. Fix $\varepsilon > 0$ and u small enough so that Lemma 4 applies. Define $x_+(u,m) := r_+(c_m(u)-\varepsilon)$ and $x_-(u,m) := r_-(c_m(u)+\varepsilon)$. By Lemma 4,

$$\mathbb{P}(p_{k,m} \le u,\ p_{\ell,m} \le u) \le \mathbb{P}\big(\|U_{k,m}\| \ge x_+(u,m),\ \|U_{\ell,m}\| \ge x_+(u,m)\big) + o(1) \tag{7}$$

and

$$\mathbb{P}(p_{k,m} \le u) \ge \mathbb{P}\big(\|U_{k,m}\| \ge x_-(u,m)\big) - o(1), \tag{8}$$

where $o(1) \to 0$ as $m \to \infty$.

We next transfer the window-fluctuation exceedances to the Gaussian limit. By Lemma 1, we have $(U_{k,m}, U_{\ell,m})$ converges weakly to (Δ_k, Δ_ℓ), where the limit is jointly Gaussian with block correlation parameter $\kappa_{k\ell} \in [0,1)$. By Assumption 4, π_m converges uniformly to π on the relevant upper tail, so $c_m(u) \to c(u) := \inf\{t : \pi(t) \le u\}$ as $m \to \infty$. For u small enough, $c(u)$ lies in the tail region. Hence $x_+(u,m) \to x_+(u) := r_+\big(c(u) - \varepsilon\big)$ and $x_-(u,m) \to x_-(u) := r_-\big(c(u) + \varepsilon\big)$. For any fixed $a < x_+(u) < b$ we have $a < x_+(u,m) < b$ for all sufficiently large m. Therefore the joint event is eventually sandwiched between the corresponding fixed-threshold joint events at levels a and b. The same sandwiching argument applies to the marginal event because $x_-(u,m) \to x_-(u)$. Applying the Portmanteau theorem to these fixed-threshold events and using that the Gaussian law assigns probability zero to the relevant boundaries yields

$$\mathbb{P}\big(\|U_{k,m}\| \ge x_+(u,m),\ \|U_{\ell,m}\| \ge x_+(u,m)\big) \to \mathbb{P}\big(\|\Delta_k\| \ge x_+(u),\ \|\Delta_\ell\| \ge x_+(u)\big),$$

and

$$\mathbb{P}\big(\|U_{k,m}\| \ge x_-(u,m)\big) \to \mathbb{P}\big(\|\Delta_k\| \ge x_-(u)\big).$$

Taking $\limsup_{m\to\infty}$ in (7)–(8) and using the convergence results above yields

$$\limsup_{m\to\infty} \frac{\mathbb{P}(p_{k,m} \le u,\ p_{\ell,m} \le u)}{\mathbb{P}(p_{k,m} \le u)} \ \le\ \frac{\mathbb{P}\big(\|\Delta_k\| \ge x_+(u),\ \|\Delta_\ell\| \ge x_+(u)\big)}{\mathbb{P}\big(\|\Delta_k\| \ge x_-(u)\big)}. \tag{9}$$

By Lemma 3, for any $\eta > 0$, there exists y_0 such that for all $y \ge y_0$, $r_-(y+\varepsilon) \le (1+\eta)\, r_+(y-\varepsilon)$. Since π is strictly decreasing on the upper tail, $c(u) \to \infty$ as $u \downarrow 0$. Hence for sufficiently small u, $c(u) \ge y_0$ and Lemma 3 applied with $y = c(u)$ gives

$$x_-(u) = r_-\big(c(u) + \varepsilon\big) \le (1+\eta)\, r_+\big(c(u) - \varepsilon\big) = (1+\eta)\, x_+(u).$$

Since the tail probability $P(\|\Delta_k\| \geq t)$ is nonincreasing in t, this yields

$$\mathbb{P}\big(\|\Delta_k\| \geq x_-(u)\big) \geq \mathbb{P}\big(\|\Delta_k\| \geq (1+\eta)x_+(u)\big).$$

Combining this with (9) yields, for u small,

$$\limsup_{m\to\infty} \frac{\mathbb{P}(p_{k,m} \leq u,\ p_{\ell,m} \leq u)}{\mathbb{P}(p_{k,m} \leq u)} \ \leq \ \frac{\mathbb{P}\big(\|\Delta_k\| \geq x_+(u),\ \|\Delta_\ell\| \geq x_+(u)\big)}{\mathbb{P}\big(\|\Delta_k\| \geq (1+\eta)x_+(u)\big)}.$$

Finally, choose $\eta > 0$ small enough so that $(1+\eta) < \sqrt{2/(1+\kappa_{k\ell})}$. Since $x_+(u) \to \infty$ as $u \downarrow 0$, Lemma 2 applies, and thus the right-hand side tends to 0 as $u \downarrow 0$. Therefore,

$$\lim_{u\downarrow 0}\ \limsup_{m\to\infty}\ \frac{\mathbb{P}(p_{k,m} \leq u,\ p_{\ell,m} \leq u)}{\mathbb{P}(p_{k,m} \leq u)} = 0.$$

This proves the proposition.

□

References

Billingsley, P., 1999. Convergence of probability measures. John Wiley & Sons. doi:https://doi.org/10.1002/9780470316962.

Herrndorf, N., 1985. A functional central limit theorem for strongly mixing sequences of random variables. Z. Wahrsch. Verw. Gebiete 69, 541–550. doi:https://doi.org/10.1007/BF00532665.

Vershynin, R., 2018. High-dimensional probability: An introduction with applications in data science. volume 47. Camb. university press. doi:https://doi.org/10.1017/9781108231596.