Adjacent-Window Fluctuations of Rolling-Covariance Spectral Projectors under a Static Gaussian Null

June 27, 2026

Abstract

We study how the rank-r spectral projector of a rolling-window sample covariance matrix changes between two consecutive, overlapping windows when the data are i.i.d. Gaussian. The two windows share all but one observation each, so the covariance difference has rank two; we quantify the resulting squared projector increment $\mathbb{E}\left\|\widehat{P}_t - \widehat{P}_{t-1}\right\|_F^2$ and show it scales as r/W^2, with an explicit constant. The self-contained core proved here consists of a conditional exchangeability identity for the two projectors given the shared observations, an exact conditional sign-randomization principle (which yields a finite-sample test for oriented contrasts), exact Gaussian formulas for the linearized projector variance both for a fixed population and conditionally around the random overlap covariance, and the algebra giving the r/W^2 rate. The perturbation, concentration, χ^2-tail, and Gaussian-chaos facts used along the way are standard and are cited at the point of use rather than restated; with those citations in place the r/W^2 rate (Theorem 3.10) holds outright. Only two items are specific to this work and remain to be written: a Lipschitz-continuity step for the conditional variance coefficient, needed for the sharp constant (Theorem 3.11), and the local-geometry details of a matching minimax lower bound.

1 Conventions on proof status

We use two labels. A *Proposition*, *Lemma*, *Corollary*, or *Theorem* is proved here in full from elementary conditioning, Gaussian moment identities, linear algebra, and standard results cited inline where they are used. A *Deferred step* is specific to this work and is presently only sketched; these are the items a final version still has to write out.

The results proved here are: the rank-two structure and conditional exchangeability (§3.1); exact conditional sign randomization and the resulting finite-sample test (§3.2); the exact fixed-population linearized variance (§3.3); the exact conditional linearized variance around the random overlap covariance, the overlap-separation lemma, and the order-r control of the variance coefficient (§3.4); and the r/W^2 rate together with the conditional constant-level expansion (§3.5). Standard perturbation, covariance-concentration, χ^2-tail, and Gaussian-chaos facts are invoked by citation where needed and are not restated as separate claims. After this accounting, the rate in Theorem 3.10 is unconditional; the sharp constant in Theorem 3.11 rests on a single continuity step (§4, Deferred step 4.1), and the matching lower bound on a second (Deferred step 4.2).

2 Setup and assumptions

For a window width W let $I_t = \{t-W+1, \dots, t\}$ and define the rolling sample covariance

$$\widehat{\Sigma}_t = \frac{1}{W} \sum_{s \in I_t} r_s r_s^\top.$$

For a symmetric matrix A whose r-th and $(r+1)$-st eigenvalues differ, let $P(A)$ be its top-r spectral projector; on the measure-zero set of eigenvalue ties we fix an arbitrary measurable tie-breaking rule. Write $\widehat{P}_t = P(\widehat{\Sigma}_t)$.

Assumption 2.1 (Static Gaussian null)**.** On $I_t \cup I_{t-1}$ the vectors r_s are independent and identically distributed $N(0, \Sigma)$.

Assumption 2.2 (Bulk-plus-spikes spectrum)**.** The rank r is fixed. With eigenvalues ordered as $\lambda_1(\Sigma) \ge \cdots \ge \lambda_n(\Sigma)$, there are constants $0 < m < M < \infty$ and $0 < c_\lambda \le C_\lambda < \infty$, independent of n and W, with

$$m \le \lambda_n(\Sigma) \le \cdots \le \lambda_{r+1}(\Sigma) \le M, \qquad c_\lambda n \le \lambda_r(\Sigma) \le \cdots \le \lambda_1(\Sigma) \le C_\lambda n.$$

Assumption 2.3 (Eigengap)**.** The gap $\delta = \lambda_r(\Sigma) - \lambda_{r+1}(\Sigma)$ satisfies $\delta \ge c_\delta n$ for a constant $c_\delta > 0$.

Assumption 2.4 (Asymptotic regime)**.** The rank r is fixed, n may grow, and $W = W_n \to \infty$. Wherever exponential sample-covariance concentration is invoked we assume $W \ge C \log n$ and $W \ge W_0$, with constants depending only on the spectral constants above and on r.

For a Hilbert-space-valued square-integrable random variable X and a sub-σ-algebra $\mathcal{G}$, write the conditional variance

$$V_F(X \mid \mathcal{G}) = \mathbb{E}\big[\|X - \mathbb{E}[X \mid \mathcal{G}]\|_F^2 \mid \mathcal{G}\big].$$

Lemma 2.5 (Conditional i.i.d. variance identity)**.** *If X_1, X_2 are conditionally i.i.d. and square-integrable given $\mathcal{G}$, then*

$$\mathbb{E}\big[\|X_1 - X_2\|_F^2 \mid \mathcal{G}\big] = 2\, V_F(X_1 \mid \mathcal{G}).$$

Proof. Let $\mu = \mathbb{E}[X_1 \mid \mathcal{G}] = \mathbb{E}[X_2 \mid \mathcal{G}]$. Conditional independence gives $\mathbb{E}[\langle X_1 - \mu, X_2 - \mu \rangle \mid \mathcal{G}] = 0$, so $\mathbb{E}[\|X_1 - X_2\|_F^2 \mid \mathcal{G}] = \mathbb{E}[\|X_1 - \mu\|_F^2 \mid \mathcal{G}] + \mathbb{E}[\|X_2 - \mu\|_F^2 \mid \mathcal{G}] = 2V_F(X_1 \mid \mathcal{G})$. □

3 Results proved here

3.1 Rank-two structure and conditional exchangeability

For adjacent windows,

$$\widehat{\Sigma}_t - \widehat{\Sigma}_{t-1} = \frac{1}{W}\big(r_t r_t^\top - r_{t-W} r_{t-W}^\top\big),$$

so the covariance difference has rank at most two: the windows differ only through the *entering* observation r_t and the *leaving* observation r_{t-W}. Their common part is the *overlap covariance*

$$C_t := \frac{1}{W} \sum_{s=t-W+1}^{t-1} r_s r_s^\top, \qquad \alpha_W := \frac{W-1}{W},$$

which satisfies $\mathbb{E} C_t = \alpha_W \Sigma$ and

$$\widehat{\Sigma}_t = C_t + \frac{1}{W} r_t r_t^\top, \qquad \widehat{\Sigma}_{t-1} = C_t + \frac{1}{W} r_{t-W} r_{t-W}^\top.$$

Let $\phi_{C_t}(v)$ denote the top-r projector of $C_t + W^{-1} v v^\top$, and set $\mathcal{G}_t = \sigma(r_{t-W+1}, \dots, r_{t-1})$.

Proposition 3.1 (Conditional exchangeability)**.** *Under Assumption 2.1, conditional on $\mathcal{G}_t$ the projectors*

$$\widehat{P}_t = \phi_{C_t}(r_t), \qquad \widehat{P}_{t-1} = \phi_{C_t}(r_{t-W})$$

are i.i.d. Consequently

$$\mathbb{E}[\widehat{P}_t - \widehat{P}_{t-1} \mid \mathcal{G}_t] = 0, \qquad \mathbb{E}\left\|\widehat{P}_t - \widehat{P}_{t-1}\right\|_F^2 = 2\,\mathbb{E}\big[V_F(\widehat{P}_t \mid \mathcal{G}_t)\big].$$

Proof. Given $\mathcal{G}_t$ the overlap covariance C_t is deterministic and the two boundary observations r_t, r_{t-W} are i.i.d. $N(0,\Sigma)$, independent of C_t. Hence $\widehat{P}_t$ and $\widehat{P}_{t-1}$ are conditionally i.i.d. The mean-zero statement is then immediate, and the second-moment statement is Lemma 2.5 followed by taking expectations. □

3.2 Sign randomization and a finite-sample test

Theorem 3.2 (Exact conditional sign randomization)**.** *Under Assumption 2.1, let $\mathcal{A}$ be any measurable Hilbert-space-valued functional of two rank-r projectors with $\mathcal{A}(Q_1, Q_2) = -\mathcal{A}(Q_2, Q_1)$. Then, conditional on $\mathcal{G}_t$,*

$$\mathcal{A}(\widehat{P}_t, \widehat{P}_{t-1}) \overset{d}{=} -\mathcal{A}(\widehat{P}_t, \widehat{P}_{t-1}).$$

This applies to oriented statistics such as $\widehat{P}_t - \widehat{P}_{t-1}$ and its linear contrasts. It says nothing about the nonnegative magnitude $\left\|\widehat{P}_t - \widehat{P}_{t-1}\right\|_F$, which is an even functional.

Proof. Swapping the two conditionally i.i.d. boundary observations r_t and r_{t-W} preserves the conditional law and exchanges the two projectors; antisymmetry of $\mathcal{A}$ then negates the statistic. □

Corollary 3.3 (Finite-sample sign test for oriented contrasts)**.** *Under Assumption 2.1, fix a deterministic symmetric matrix A or a $\mathcal{G}_t$-measurable symmetric matrix A_t, and set $T_t = \langle A_t, \widehat{P}_t - \widehat{P}_{t-1}\rangle_F$. Then $T_t \overset{d}{=} -T_t$ conditional on $\mathcal{G}_t$. Hence if $\mathbb{P}(T_t = 0 \mid \mathcal{G}_t) = 0$ the conditional signs of T_t are exactly balanced under the null; in the presence of an atom at zero, treating zeros as non-rejections gives a conservative test.*

Proof. Apply Theorem 3.2 to $\mathcal{A}(Q_1, Q_2) = \langle A_t, Q_1 - Q_2\rangle_F$, conditioning on $\mathcal{G}_t$ when A_t is random. □

3.3 Exact linearized variance: fixed population

Definition 3.4 (Linearized-projector variance coefficient)**.** With $\Sigma = \sum_{k=1}^n \lambda_k u_k u_k^\top$, $\lambda_1 \geq \cdots \geq \lambda_n$ and $\lambda_r > \lambda_{r+1}$, set

$$\mathfrak{d}_{\mathrm{sub}}(\Sigma, r) = 2 \sum_{i \leq r < j} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2}.$$

Proposition 3.5 (Exact fixed-population linearized variance)**.** *Let $P = P(\Sigma)$ and let*

$$\mathcal{L}_\Sigma(E) = \sum_{i \leq r < j} \frac{u_j u_j^\top E\, u_i u_i^\top + u_i u_i^\top E\, u_j u_j^\top}{\lambda_i - \lambda_j}$$

be the first-order projector perturbation. Then

$$\|\mathcal{L}_\Sigma(E)\|_F^2 = 2 \sum_{i \leq r < j} \frac{\left|u_j^\top E u_i\right|^2}{(\lambda_i - \lambda_j)^2}.$$

If $\widehat{\Sigma} = W^{-1}\sum_{s=1}^{W} r_s r_s^\top$ *with i.i.d.* $r_s \sim N(0,\Sigma)$*, then*

$$\mathbb{E}\left\|\mathcal{L}_\Sigma(\widehat{\Sigma}-\Sigma)\right\|_F^2 = \frac{\mathfrak{d}_{\mathrm{sub}}(\Sigma,r)}{W}.$$

If $H = W^{-1}(r_+ r_+^\top - r_- r_-^\top)$ *with* $r_+, r_- \overset{\text{iid}}{\sim} N(0,\Sigma)$*, then*

$$\mathbb{E}\,\|\mathcal{L}_\Sigma(H)\|_F^2 = \frac{2\,\mathfrak{d}_{\mathrm{sub}}(\Sigma,r)}{W^2}.$$

Proof. The Frobenius identity follows from the orthogonality of the matrix units $u_j u_i^\top$ in the reduced-resolvent expression. For Gaussian data and $i \neq j$,

$$\mathbb{E}\left|u_j^\top(\widehat{\Sigma}-\Sigma)u_i\right|^2 = \frac{\lambda_i\lambda_j}{W}, \qquad \mathbb{E}\left|u_j^\top H u_i\right|^2 = \frac{2\lambda_i\lambda_j}{W^2},$$

since $u_i^\top r_s$ and $u_j^\top r_s$ are independent mean-zero Gaussians with variances λ_i, λ_j. Summing over $i \le r < j$ gives both claims. □

Corollary 3.6 (Bulk-plus-spikes scale)**.** *Under Assumptions 2.2–2.3,* $\mathfrak{d}_{\mathrm{sub}}(\Sigma,r) = \Theta(r)$*. Hence the single-window linearized variance scales as* r/W *and the adjacent-window linearized variance as* r/W^2*.*

Proof. For $i \le r < j$ one has $\lambda_i \asymp n$, $\lambda_j \asymp 1$, and $\lambda_i - \lambda_j \asymp n$, so each summand is $\asymp 1/n$. There are $r(n-r)$ summands, giving $\Theta(r)$ for fixed r. □

3.4 Exact conditional linearized variance: random overlap

Write $C_t = \sum_{k=1}^n \widetilde{\lambda}_k \widetilde{u}_k \widetilde{u}_k^\top$ with $\widetilde{\lambda}_1 \ge \cdots \ge \widetilde{\lambda}_n$. On the event $\widetilde{\lambda}_r > \widetilde{\lambda}_{r+1}$ define $\mathcal{L}_{C_t}$ as in Proposition 3.5 but with the eigendata of C_t, and set

$$a_i = \widetilde{u}_i^\top \Sigma \widetilde{u}_i, \qquad b_j = \widetilde{u}_j^\top \Sigma \widetilde{u}_j, \qquad c_{ij} = \widetilde{u}_i^\top \Sigma \widetilde{u}_j.$$

Proposition 3.7 (Exact conditional linearized variance)**.** *Under Assumption 2.1, on* $\widetilde{\lambda}_r > \widetilde{\lambda}_{r+1}$*, with* $H_t = W^{-1}(r_t r_t^\top - r_{t-W} r_{t-W}^\top)$*,*

$$\mathbb{E}\big[\|\mathcal{L}_{C_t}(H_t)\|_F^2 \mid \mathcal{G}_t\big] = \frac{4}{W^2}\sum_{i\le r<j} \frac{a_i b_j + c_{ij}^2}{(\widetilde{\lambda}_i - \widetilde{\lambda}_j)^2}.$$

Proof. For one boundary term $\Delta = xx^\top/W$ with $x \sim N(0,\Sigma)$,

$$\|\mathcal{L}_{C_t}(\Delta)\|_F^2 = \frac{2}{W^2}\sum_{i\le r<j} \frac{\left|\widetilde{u}_j^\top x\right|^2 \left|\widetilde{u}_i^\top x\right|^2}{(\widetilde{\lambda}_i - \widetilde{\lambda}_j)^2}.$$

With $g_i = \widetilde{u}_i^\top x$, Isserlis' formula gives $\mathbb{E}[g_i^2 g_j^2 \mid \mathcal{G}_t] = a_i b_j + 2c_{ij}^2$, while $\|\mathbb{E}[\mathcal{L}_{C_t}(\Delta) \mid \mathcal{G}_t]\|_F^2 = \frac{2}{W^2}\sum_{i\le r<j} c_{ij}^2/(\widetilde{\lambda}_i - \widetilde{\lambda}_j)^2$. Subtracting gives the conditional variance of one term; the two terms are conditionally i.i.d., so the second moment of their difference is twice that, which is the displayed expression. □

Define the conditional variance coefficient

$$Q(C_t,\Sigma) = 4\sum_{i\le r<j}\frac{a_ib_j + c_{ij}^2}{(\widetilde{\lambda}_i-\widetilde{\lambda}_j)^2}, \qquad \text{so} \qquad \mathbb{E}\big[\|\mathcal{L}_{C_t}(H_t)\|_F^2 \mid \mathcal{G}_t\big] = \frac{Q(C_t,\Sigma)}{W^2},$$

and the concentration event

$$\mathcal{E}_t = \big\{\|C_t - \alpha_W\Sigma\|_{\mathrm{op}} \le \eta\,\alpha_W n\big\},$$

where $\eta > 0$ is small relative to c_λ and c_δ.

Lemma 3.8 (Overlap-covariance separation). *Under Assumptions 2.1–2.3, for $\eta > 0$ small there are $c, C > 0$ with $\mathbb{P}(\mathcal{E}_t^c) \le Ce^{-cW}$ whenever $W \ge C\log n$ and $W \ge W_0$.*

Proof. Write $C_t = \alpha_W S_t$ with $S_t = (W-1)^{-1}\sum_{s=t-W+1}^{t-1} r_s r_s^\top$, the sample covariance of $W-1$ i.i.d. $N(0,\Sigma)$ vectors. Then $\mathcal{E}_t^c = \{\|S_t-\Sigma\|_{\mathrm{op}} > \eta n\}$, and the bound is the exponential operator-norm tail of the Gaussian sample-covariance concentration inequality (Koltchinskii and Lounici [3]; Vershynin [5], Ch. 4–6) applied with $N = W-1$ and effective rank $\mathbf{r}(\Sigma) = O(1)$ under Assumption 2.2. □

Lemma 3.9 (Order-r control of the variance coefficient). *Under Assumptions 2.1–2.3, on $\mathcal{E}_t$, and with $W \ge W_0$ so that α_W is bounded away from zero,*

$$cr \le Q(C_t,\Sigma) \le Cr.$$

Consequently, by Lemma 3.8, $\mathbb{E}\big[\|\mathcal{L}_{C_t}(H_t)\|_F^2\,\mathbf{1}_{\mathcal{E}_t}\big] \asymp r/W^2$.

Proof. On $\mathcal{E}_t$, Weyl's inequality gives $\widetilde{\lambda}_r - \widetilde{\lambda}_{r+1} \ge \alpha_W\delta - 2\eta\alpha_W n \ge \frac{1}{2}\alpha_W\delta \asymp n$, so $\big|\widetilde{\lambda}_i - \widetilde{\lambda}_j\big| \gtrsim n$ for $i \le r < j$; together with $0 \le \widetilde{\lambda}_j \le \widetilde{\lambda}_i \lesssim n$ the denominators are of order n^2.

By Cauchy–Schwarz for the form $\langle\cdot,\cdot\rangle_\Sigma$, $c_{ij}^2 \le a_ib_j$, so $\sum_{i\le r<j}(a_ib_j + c_{ij}^2) \le 2(\sum_{i\le r} a_i)(\sum_{j>r} b_j)$. Since Σ has r spikes of size $O(n)$ and bulk trace $O(n)$, $\sum_{i\le r} a_i \lesssim rn$ and $\sum_{j>r} b_j \lesssim n$, giving the upper bound.

For the lower bound, on $\mathcal{E}_t$ one has $\widetilde{\lambda}_i \gtrsim n$ for $i \le r$, and since $\widetilde{u}_i$ is an eigenvector of C_t, $\alpha_W a_i = \widetilde{u}_i^\top(\alpha_W\Sigma)\widetilde{u}_i \ge \widetilde{\lambda}_i - \|C_t - \alpha_W\Sigma\|_{\mathrm{op}}$, so $a_i \gtrsim n$ for $i \le r$ after fixing η small. By Ky Fan, $\sum_{j>r} b_j = \mathrm{tr}((I - P(C_t))\Sigma) \ge \sum_{k=r+1}^n \lambda_k(\Sigma) \gtrsim n$. Dropping $c_{ij}^2 \ge 0$ and bounding denominators by Cn^2 yields the lower bound. The expectation statement combines this with Lemma 3.8. □

3.5 Rate and constant

Theorem 3.10 (Adjacent-window rate). *Assume Assumptions 2.1–2.3 with r fixed. Then, for $W \ge C\log n$ and $W \ge W_0$,*

$$c\,\frac{r}{W^2} \le \mathbb{E}\left\|\widehat{P}_t - \widehat{P}_{t-1}\right\|_F^2 \le C\,\frac{r}{W^2} + Cre^{-cW}.$$

Proof. Let $\gamma_t = \widetilde{\lambda}_r - \widetilde{\lambda}_{r+1}$ and $\Delta_t^+ = W^{-1}r_tr_t^\top$, $\Delta_t^- = W^{-1}r_{t-W}r_{t-W}^\top$, $H_t = \Delta_t^+ - \Delta_t^-$. Throughout we use the second-order perturbation expansion of a top-r spectral projector with a gap-controlled quadratic remainder: for symmetric A with gap $\gamma = \mu_r - \mu_{r+1} > 0$ and $\|E\|_{\mathrm{op}} \le c_0\gamma$,

$$\|P(A+E) - P(A) - \mathcal{L}_A(E)\|_F \le C_r\,\|E\|_{\mathrm{op}}^2/\gamma^2,$$

with $c_0 \in (0, 1/4]$ and $C_r < \infty$ (Kato [1], Ch. II; Stewart and Sun [2], Ch. V).

Fix $\rho \in (0, c_0]$ and set $\mathcal{H}_t = \mathcal{E}_t \cap \{\|\Delta_t^+\|_{\mathrm{op}} \vee \|\Delta_t^-\|_{\mathrm{op}} \leq \rho\gamma_t\}$. Lemma 3.8 controls $\mathcal{E}_t^c$, and the boundary norms are controlled by a weighted-χ^2 tail: since $\|\Delta_t^\pm\|_{\mathrm{op}} \leq \|r_\pm\|_2^2 / W$ and $\mathbb{E}\,\|r_\pm\|_2^2 = \operatorname{tr}\Sigma \asymp n$, one has $\mathbb{P}\{\|r_\pm\|_2^2 > \kappa n W\} \leq C_\kappa e^{-c_\kappa W}$ for each fixed $\kappa > 0$ (Laurent and Massart [6], Lemma 1; equivalently the Hanson–Wright inequality, Rudelson and Vershynin [7]). Hence $\mathbb{P}(\mathcal{H}_t^c) \leq Ce^{-cW}$.

On $\mathcal{H}_t$, the projector expansion applied at C_t to each boundary term gives $\widehat{P}_t - \widehat{P}_{t-1} = L_t + R_t^\Delta$ with $L_t = \mathcal{L}_{C_t}(H_t)$, $R_t^\Delta = R_t^+ - R_t^-$, and $\|R_t^\pm\|_F \leq C_r \|\Delta_t^\pm\|_{\mathrm{op}}^2 / \gamma_t^2$. Since $\gamma_t \asymp n$ on $\mathcal{H}_t$, the Gaussian eighth-moment bound yields $\mathbb{E}[\|R_t^\Delta\|_F^2 \mathbf{1}_{\mathcal{H}_t}] \leq C_r W^{-4}$, while Lemma 3.9 gives $\mathbb{E}[\|L_t\|_F^2 \mathbf{1}_{\mathcal{E}_t}] \asymp r/W^2$. Conditionally on $\mathcal{G}_t$, L_t is a degree-two Gaussian chaos, whose L^4 and L^2 norms are equivalent by hypercontractivity (Janson [8], Ch. 5–6); with the second moment $\asymp r/W^2$ this gives $\mathbb{E}[\|L_t\|_F^4 \mid \mathcal{G}_t] \leq Cr^2/W^4$ on $\mathcal{E}_t$, so the loss on $\mathcal{E}_t \cap \mathcal{H}_t^c$ is exponentially small and $\mathbb{E}[\|L_t\|_F^2 \mathbf{1}_{\mathcal{H}_t}] \asymp r/W^2$.

For the upper bound use $\|L_t + R_t^\Delta\|_F^2 \leq 2\|L_t\|_F^2 + 2\|R_t^\Delta\|_F^2$ on $\mathcal{H}_t$ and $\left\|\widehat{P}_t - \widehat{P}_{t-1}\right\|_F^2 \leq 2r$ off $\mathcal{H}_t$. For the lower bound, $\|L_t + R_t^\Delta\|_F^2 \geq \|L_t\|_F^2 - 2\left|\langle L_t, R_t^\Delta\rangle\right|$ and Cauchy–Schwarz give a cross term of order $\sqrt{r}\,W^{-3} = o(r/W^2)$. □

Theorem 3.11 (Constant-level expansion; conditional on Deferred step 4.1)**.** *Assume Assumptions 2.1–2.3 with r fixed, $W \to \infty$, $W \geq C \log n$. Granting Deferred step 4.1,*

$$\mathbb{E}\left\|\widehat{P}_t - \widehat{P}_{t-1}\right\|_F^2 = \frac{2}{\alpha_W^2}\frac{\mathfrak{d}_{\mathrm{sub}}(\Sigma, r)}{W^2} + o(W^{-2}) = \frac{2\mathfrak{d}_{\mathrm{sub}}(\Sigma, r)}{W^2} + o(W^{-2}),$$

the last equality because $\alpha_W \to 1$ and $\mathfrak{d}_{\mathrm{sub}}(\Sigma, r) = O(1)$ for fixed r.

Proof. Use the event $\mathcal{H}_t$ and decomposition $\widehat{P}_t - \widehat{P}_{t-1} = L_t + R_t^\Delta$ of Theorem 3.10, with $\mathbb{E}[\|R_t^\Delta\|_F^2 \mathbf{1}_{\mathcal{H}_t}] \leq CW^{-4}$ and $\mathbb{P}(\mathcal{H}_t^c) \leq Ce^{-cW}$. By Proposition 3.7, $\mathbb{E}[\|L_t\|_F^2 \mid \mathcal{G}_t] = Q(C_t, \Sigma)/W^2$ on the gap event. Deferred step 4.1 with Lemma 3.8 gives $\mathbb{E}\,\|L_t\|_F^2 = W^{-2}\big(2\alpha_W^{-2}\mathfrak{d}_{\mathrm{sub}}(\Sigma, r) + o(1)\big) + O(e^{-cW})$; the restriction loss to $\mathcal{H}_t$ is exponentially small by the conditional fourth-moment bound (hypercontractivity; Janson [8], Ch. 5–6). Expanding $\|L_t + R_t^\Delta\|_F^2$, the remainder square is $O(W^{-4})$ and the cross term $O(W^{-3})$, both $o(W^{-2})$. □

Remark 3.12 (Consistency check on the constant)**.** Replacing the random overlap covariance by its mean $C_t = \alpha_W \Sigma$ gives $\widetilde{u}_i = u_i$, $a_i = \lambda_i$, $b_j = \lambda_j$, $c_{ij} = 0$, hence $Q(\alpha_W\Sigma, \Sigma) = 4\sum_{i \leq r < j} \lambda_i\lambda_j / (\alpha_W^2(\lambda_i - \lambda_j)^2) = 2\alpha_W^{-2}\mathfrak{d}_{\mathrm{sub}}(\Sigma, r)$, matching Theorem 3.11. Deferred step 4.1 is precisely the statement that this replacement is asymptotically valid.

4 Deferred steps (specific to this work)

These two statements are not off-the-shelf theorems: their ingredients are standard, but the assembled quantitative claims are particular to this problem and still have to be written out.

Deferred step 4.1 (Continuity of the variance coefficient)**.** *Under Assumptions 2.1–2.3 with r fixed, $W \to \infty$, $W \geq C \log n$,*

$$\mathbb{E}\left[|[|\Big]Q(C_t, \Sigma) - \frac{2}{\alpha_W^2}\mathfrak{d}_{\mathrm{sub}}(\Sigma, r)\,\mathbf{1}_{\mathcal{E}_t}\right] \leq C\,\frac{\mathbb{E}\,\|C_t - \alpha_W\Sigma\|_{\mathrm{op}}}{n} = O(W^{-1/2}) = o(1).$$

The load-bearing step. *The claim is that the basis-free map* $A \mapsto Q(A,\Sigma) = \mathbb{E}\left\|\mathcal{L}_A(xx^\top - yy^\top)\right\|_F^2$, *with* $x, y \overset{\text{iid}}{\sim} N(0,\Sigma)$, *is Lipschitz at* $A_0 = \alpha_W \Sigma$ *with modulus* C/n *on an operator-norm ball of radius* $O(n)$*. The intended proof combines the Riesz-projector representation of* $\mathcal{L}_A$ *and the resolvent identity (the perturbation machinery of Kato [1], Ch. II) with fixed-rank bulk-plus-spikes moment bounds (Koltchinskii and Lounici [3]). This is the one step that should be written in full before submission.*

Deferred step 4.2 (Minimax lower bound for the projector increment)**.** *Fix* r *with* $n \geq 2r+1$*. There is a Gaussian adjacent-window subclass meeting the bulk-plus-spikes and eigengap conditions uniformly such that, with* $D_t = P(\bar{\Sigma}_t) - P(\bar{\Sigma}_{t-1})$,

$$\inf_{\widehat{D}} \sup \mathbb{E}\left\|\widehat{D} - D_t\right\|_F^2 \geq c\,\frac{r}{W^2}.$$

Status. *The intended proof uses the local chart* $U_\Theta = \binom{I_r}{\Theta}(I_r + \Theta^\top\Theta)^{-1/2}$, $P_\Theta = U_\Theta U_\Theta^\top$, $\Sigma_\Theta = I_n + nP_\Theta$, *with the overlap covariance set to* Σ_0 *and the two boundary covariances to* $\Sigma_{\pm\Theta}$*. The standard machinery is Assouad's lemma together with the Gaussian KL formula (Tsybakov [9], §2.7); the work specific to this paper is the local projector-geometry equivalence, the first-order expansion of* $D_t(\Theta)$*, the exact KL between adjacent hypercube vertices, and the Assouad reduction.*

5 Combined statement

Theorem 5.1 (Adjacent-window projector increment)**.** *Assume Assumptions 2.1–2.4. The rate*

$$\mathbb{E}\left\|\widehat{P}_t - \widehat{P}_{t-1}\right\|_F^2 \asymp \frac{r}{W^2}$$

holds unconditionally by Theorem 3.10. Granting the single continuity step (Deferred step 4.1), the sharp constant is

$$\mathbb{E}\left\|\widehat{P}_t - \widehat{P}_{t-1}\right\|_F^2 = \frac{2\mathfrak{d}_{\text{sub}}(\Sigma,r)}{W^2} + o(W^{-2}), \qquad \mathfrak{d}_{\text{sub}}(\Sigma,r) = \Theta(r).$$

Remark 5.2. Under the static null the conditional mean of the increment is zero (Proposition 3.1), so the r/W^2 scale is the baseline sampling-fluctuation level of consecutive projectors, against which a genuine change in the leading subspace must be detected. With the standard results cited inline, the rate is settled; the remaining gaps are the constant (Deferred step 4.1) and the matching lower bound (Deferred step 4.2). Distributional and concentration properties of the spectral projectors of a single sample covariance, the closest prior object, are studied by Koltchinskii and Lounici [4].

References

[1] T. Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics. Springer, Berlin, 1995. Reprint of the 1980 second edition.

[2] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, Boston, 1990.

[3] V. Koltchinskii and K. Lounici. Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli*, 23(1):110–133, 2017.

[4] V. Koltchinskii and K. Lounici. Normal approximation and concentration of spectral projectors of sample covariance. *Ann. Statist.*, 45(1):121–157, 2017.

[5] R. Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2018.

[6] B. Laurent and P. Massart. Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.*, 28(5):1302–1338, 2000.

[7] M. Rudelson and R. Vershynin. Hanson–Wright inequality and sub-gaussian concentration. *Electron. Commun. Probab.*, 18(82):1–9, 2013.

[8] S. Janson. *Gaussian Hilbert Spaces*. Cambridge Tracts in Mathematics 129. Cambridge University Press, Cambridge, 1997.

[9] A. B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, New York, 2009.