

Local-Time Limits and Calibration Limits of Spatial-Sign Delay Statistics

Sai Sharath Chandra Koppu
BML Munjal University

Luckshay Batra
Professor, BML Munjal University

Abstract

We study what happens when overlapping time-series windows are converted into directions by dividing each window by its length. The main example is the spatial-sign matrix and the entropy gap formed from two such matrices. For integrated Gaussian random walks with a fixed window length, we show that the resulting fluctuations are governed by Brownian local time, which records how often the limiting process visits the origin. We prove this result while accounting for the dependence created by overlapping windows and for the cancellation of the signed tails. We also identify a separate regime in which the initial level of the series is much larger than its random-walk fluctuations. In that case, the statistic shrinks at a different rate and its entropy gap collapses. For the spatial-sign matrices, the limiting entropy law depends on the correlation between the innovations. This has a direct calibration consequence: a lower-tail rule using a fixed positive cutoff at the square-root sample-size scale can reject with probability approaching one for some noncointegrated Gaussian pairs. Simulations show the slow finite-sample convergence and separate the effects of correlation and initial levels. A small empirical illustration shows how the reported screen changes with the reference construction. The paper characterizes this statistic and its calibration limits; it does not establish a general bootstrap cointegration test.

Keywords: Cointegration; Spatial Signs; Entropy; Local Time; Non-pivotal Inference; Near Unit Root; Local to Unity; Monte-Carlo critical values.

1 Introduction

Two economic series may each wander without bound while a linear combination of them remains stable. Cointegration formalizes this form of long-run equilibrium, the individual series are integrated of order one, $I(1)$, while some linear combination is stationary, $I(0)$ (Granger, 1981; Engle and Granger, 1987). Identifying such relationships matters in applications including purchasing-power parity, the term structure of interest rates, error-correction modelling, and pairs trading. It also requires careful calibration, because statistics constructed from integrated levels can display apparent similarity even when no stationary linear combination exists.

Normalizing overlapping delay windows of an integrated series produces bounded directional summaries. For most late observations, the entries within a window are nearly equal relative to

their common level, causing the spatial-sign outer product to approach a fixed rank-one matrix. Visits near the origin nevertheless continue to contribute to its fluctuations. Consequently, a small difference between two such summaries cannot by itself be interpreted as evidence of cointegration; its distribution must be calibrated under the relevant nonstationary model.

We study two asymptotic regimes and their implications for that calibration. With fixed initial levels, the joint Gaussian limits of the spatial-sign matrices and their entropy comparison retain Brownian local time and innovation-correlation dependence. When the initial levels dominate stochastic fluctuations, the compressed matrices follow a different normalization and the entropy gap collapses at the fixed-level scale. Together with the correlation obstruction, this separation shows why fixed-positive lower-tail calibration cannot be applied uniformly across the noncointegrated random-walk families considered here.

The underlying occupation, spatial-sign, and entropy tools are known, but their combination for overlapping delay windows requires additional verification. Conditional centering must account for the dependence created by overlapping marks; the signed coupling tail must cancel before integration; and the entropy expansion must hold uniformly over positive semidefinite matrices approaching a rank-one boundary. The large-level regime must also be separated from the fixed-level occupation regime. We verify these steps, derive the within-window coefficients, and obtain the resulting matrix and entropy limits.

The spatial-sign covariance matrix is not new to this paper. Its construction and eigenstructure come from Dürre et al. (2014) and Dürre et al. (2016), while its correlation variant is from Dürre et al. (2015), with an implementation in Dürre and Vogel (2025). Our contribution is studying its behavior under overlapping, nonstationary delay windows. Related nonlinear transformations were studied by Park and Phillips (1999) and Kartashov and Kulik (2007); we adapt these ideas to the angular, overlapping-window setting. The motivation is related to the concerns raised by Phillips (1986), but our calibration problem is different from spurious regression. Unlike topological methods such as Perea and Harer (2015), which study the full delay-embedded cloud, we focus on additive marginal statistics.

The paper makes five contributions.

1. **An explicit joint local-time characterization.** For fixed-dimensional delay windows of jointly Gaussian random walks with fixed initial levels, we derive the joint limit of the spatial-sign matrices,

$$\sqrt{T}(R_{X_i,T} - \Pi) \Rightarrow L_i(\Lambda_d - \kappa_d \Pi).$$

The result identifies both the common rank-one limit and the exact within-window coefficient governing its first-order fluctuations. Because the coordinate local times are coupled through the joint Brownian motion, the limit preserves innovation-correlation dependence rather than reducing to separate marginal limits.

2. **A rigorous reduction for overlapping directional windows.** The delay windows overlap, so their within-window marks cannot be treated as independent observations. We establish the required reduction by conditionally centering the marks with respect to the full innovation history and separating them into residue classes. We also verify the signed-tail cancellation of the matrix coupling term before applying Fubini. These arguments provide

the window-specific link between the spatial-sign statistic and the underlying occupation theorem.

3. **An exact entropy limit at a singular boundary.** The limiting spatial-sign matrix has rank one, where von Neumann entropy is not differentiable in the manner required by an interior Taylor expansion. We prove a uniform expansion over actual positive semidefinite trace-one matrices approaching this boundary. Combining that expansion with the joint matrix limit yields

$$\sqrt{T}Q_T \Rightarrow \kappa_d \Psi(L_1, L_2),$$

where the limiting functional is exact and retains the dependence between the two Brownian local times.

4. **A separation of fixed-level and large-level regimes.** The local-time limit applies when the initial levels remain fixed. When $|c_T|/(\sigma\sqrt{T}) \rightarrow \infty$, we instead show that the compressed spatial-sign matrix has a c_T^2 normalization and a deterministic limit, while the entropy comparison satisfies

$$\sqrt{T}Q_T \xrightarrow{p} 0.$$

Thus marginal scale invariance does not make the statistic invariant to additive levels, and the fixed-level occupation approximation cannot be extrapolated to dominating initial conditions.

5. **Two scoped obstructions to fixed-positive lower-tail calibration.** Within noncoin-integrated Gaussian random-walk families, the limiting entropy gap collapses as innovation correlation approaches one. It also collapses in the large-initial-level regime, even for independent trends. We therefore show that a deterministic positive lower-tail cutoff calibrated only by sample size and window dimension can have worst-case rejection probability tending to one over the stated null family. Numerical experiments and a frozen empirical illustration document these mechanisms and the sensitivity to reference construction. They are supporting evidence for the characterization, not a proposed bootstrap procedure or a claim of competitive testing power.

2 Setup and Geometry

Let $X_1, \dots, X_T$ be a time series containing T observations. Fix an integer $d \geq 2$, which denotes both the length of each overlapping window and the dimension of the resulting delay vectors. We treat d as fixed while T increases and assume $T \geq d$. The number of overlapping windows is therefore

$$N = T - d + 1.$$

For $n = 1, \dots, N$, define

$$w_{X,n} = (X_n, \dots, X_{n+d-1})', \quad R_{X,T} = \frac{1}{N} \sum_{n=1}^N \frac{w_{X,n} w_{X,n}'}{\|w_{X,n}\|^2}. \quad (2.1)$$

The vector $w_{X,n} \in \mathbb{R}^d$ is the n th chronological delay window. It collects the d consecutive observations beginning at time n . Because successive windows share $d - 1$ observations, the sequence of windows is overlapping.

For a nonzero window $w_{X,n}$, the normalized vector $w_{X,n}/\|w_{X,n}\|$ retains its direction while discarding its magnitude. Its outer product,

$$\frac{w_{X,n}w_{X,n}'}{\|w_{X,n}\|^2},$$

is the spatial-sign matrix associated with that window. The matrix $R_{X,T}$ is the average of these window-level matrices and hence summarizes the empirical distribution of the delay-window directions. It is a positive semidefinite $d \times d$ matrix with unit trace.

We assume that every window is nonzero whenever the statistic is evaluated. This condition holds almost surely under the Gaussian models considered below. With other types of data, an all-zero window has no defined direction and therefore requires an explicit convention. Our implementation reports the statistic as undefined when such a window occurs. Reversing the coordinate order of every window conjugates all the corresponding matrices by the same permutation matrix. The matrix comparisons considered below are consequently unchanged.

For a positive semidefinite matrix A , which need not have unit trace, define

$$\text{st}(A) = -\text{tr}(A \log A), \quad 0 \log 0 = 0.$$

The matrix logarithm is understood spectrally: if A has eigenvalues $\lambda_1, \dots, \lambda_d$, then

$$\text{st}(A) = -\sum_{j=1}^d \lambda_j \log \lambda_j.$$

When A has unit trace, we write $S(A)$ in place of $\text{st}(A)$. Thus S is the von Neumann entropy of a trace-one positive semidefinite matrix. All logarithms in the paper are natural logarithms.

Now let $Y_1, \dots, Y_T$ be a second series, and construct $R_{Y,T}$ from its overlapping windows in the same way as $R_{X,T}$. We compare the two spatial-sign matrices using

$$Q_T = S\left(\frac{R_{X,T} + R_{Y,T}}{2}\right) - \frac{S(R_{X,T}) + S(R_{Y,T})}{2}. \quad (2.2)$$

The statistic Q_T is an entropy divergence between the empirical angular geometries of the two series. It is zero when $R_{X,T} = R_{Y,T}$, although the converse need not be asserted for the arguments below. More generally,

$$0 \leq Q_T \leq \log 2.$$

The lower bound follows from the concavity of entropy. For the upper bound, mixing two matrices with equal weights can increase entropy over their average component entropy by at most the entropy of the two-point mixing distribution, which is $\log 2$. Equivalently, the mixture dominates each component divided by two, and operator monotonicity of the logarithm bounds the associated relative entropies by $\log 2$. The result for singular matrices follows by regularization and continuity.

To describe the geometry of $R_{X,T}$, define

$$u = d^{-1/2}(1, \dots, 1)', \quad \Pi = uu', \quad P = I - \Pi,$$

where I is the $d \times d$ identity matrix. The unit vector u represents the common-level direction: a window lies in this direction when all its coordinates are equal. The matrix Π is the orthogonal

projector onto the one-dimensional space spanned by u , whereas P projects onto its orthogonal complement $u^\perp$. Consequently, P removes the common-level component of a window and retains only its within-window deviations.

Define the scalar defect $D_{X,T}$ and the compressed matrix $H_{X,T}$ by

$$D_{X,T} = 1 - u'R_{X,T}u = \text{tr}(PR_{X,T}P), \quad H_{X,T} = PR_{X,T}P. \quad (2.3)$$

The matrix $H_{X,T}$ is the part of $R_{X,T}$ supported on $u^\perp$. It therefore records the directional variation that remains after the common-level direction has been removed. Its trace is the scalar defect:

$$D_{X,T} = \text{tr}(H_{X,T}).$$

Thus $D_{X,T}$ measures the total mass of $R_{X,T}$ lying away from the common-level direction. In particular, as $R_{X,T}$ approaches the rank-one matrix Π , both $H_{X,T}$ and $D_{X,T}$ approach zero. When referring to the eigenvalues of $H_{X,T}$, we identify the compression with its restriction to the $(d-1)$ -dimensional space $u^\perp$.

The same decomposition can be described at the level of an individual window. For any $w \in \mathbb{R}^d$, write

$$w = su + v, \quad s = u'w, \quad v = Pw.$$

Here su is the common-level component of the window and $v \in u^\perp$ contains its within-window deviations. Since u is a unit vector and $u'v = 0$,

$$\|w\|^2 = s^2 + \|v\|^2.$$

For a nonzero window w , define its individual directional defect by

$$\delta(w) = \frac{\|v\|^2}{s^2 + \|v\|^2}. \quad (2.4)$$

This quantity is the fraction of the squared window magnitude lying outside the common-level direction. It equals zero when all coordinates of the window are equal and increases as within-window variation becomes large relative to the window level. The construction therefore depends on within-window fluctuations relative to the prevailing level, rather than simply on a stationary covariance matrix of the increments.

The normalization in (2.1) also determines which transformations leave the construction unchanged. For every nonzero scalar a ,

$$R_{aX,T} = R_{X,T},$$

because multiplying every observation by a changes the magnitude of each window but not its normalized outer product. Hence the spatial-sign matrix is exactly invariant to rescaling.

A constant shift behaves differently. Replacing X_t by $X_t + c$ changes the direction of its delay windows and generally changes both the numerator and denominator in (2.1). The construction is therefore not translation invariant. Global demeaning can impose translation invariance in a finite-sample calculation, but it also changes the stochastic object being analyzed. This distinction is material for integrated series because their sample mean is of order $\sqrt{T}$. Demeaning such a

series cannot be treated as a negligible degrees-of-freedom adjustment when deriving the raw-level asymptotic limit.

We finally introduce a functional that will be used to express the entropy limit on the compressed space. For positive semidefinite matrices H and K , define

$$F(H, K) = \text{st} \left(\frac{H + K}{2} \right) - \frac{\text{st}(H) + \text{st}(K)}{2}. \quad (2.5)$$

Unlike $R_{X,T}$ and $R_{Y,T}$, the compressed matrices H and K need not have unit trace, which is why the functional is written using st rather than S . The trace-dependent terms generated by entropy scaling cancel in this difference. Consequently,

$$F(aH, aK) = aF(H, K), \quad a \geq 0.$$

For nonnegative scalars x and y , define

$$\Psi(x, y) = \frac{x \log x + y \log y}{2} - \frac{x + y}{2} \log \left(\frac{x + y}{2} \right), \quad (2.6)$$

with the convention $0 \log 0 = 0$. If Λ is positive semidefinite, then $x\Lambda$ and $y\Lambda$ share the same matrix shape and differ only in scale. Direct substitution into (2.5) yields the exact identity

$$F(x\Lambda, y\Lambda) = \text{tr}(\Lambda) \Psi(x, y). \quad (2.7)$$

This identity reduces a matrix entropy comparison to a scalar comparison whenever the limiting compressed matrices have a common matrix factor and differ only through their occupation factors.

3 The Gaussian Matrix and Entropy Limits

Assumption 1. For $i = 1, \dots, m$, $X_{i,k} = c_i + \sum_{j=1}^k \varepsilon_{i,j}$, where m, d are fixed, the c_i are fixed real constants, and the innovation vectors are iid centered Gaussian. Each marginal variance σ_i^2 is positive. Their correlation matrix is C , which may be singular.

Let $B = (B_1, \dots, B_m)$ be Brownian motion with covariance C . Its coordinate local times have the normalization

$$\int_0^t f(B_i(s)) ds = \int_{\mathbb{R}} f(a) L_i(t, a) da, \quad L_i = L_i(1, 0).$$

Use independent standard normal variables $e_1, \dots, e_{d-1}$ to define

$$\Gamma = (0, e_1, e_1 + e_2, \dots, e_1 + \dots + e_{d-1})', \quad v = P\Gamma,$$

and the deterministic coefficients

$$\Lambda_d = \frac{\pi}{\sqrt{d}} \mathbb{E} \frac{vv'}{\|v\|}, \quad \kappa_d = \text{tr}(\Lambda_d) = \frac{\pi}{\sqrt{d}} \mathbb{E} \|v\|. \quad (3.1)$$

The matrix Λ_d is zero on u and positive definite on $u^\perp$. Finiteness follows from the first moment of $\|v\|$. Positivity follows because every nonzero linear functional on the projected Gaussian vector has positive variance.

Theorem 1 (Joint matrix limit). *Under Assumption 1, jointly for $i = 1, \dots, m$,*

$$\sqrt{T}(R_{X_i,T} - \Pi) \xrightarrow{d} L_i(\Lambda_d - \kappa_d \Pi). \quad (3.2)$$

In particular, $\sqrt{T}D_{X_i,T} \xrightarrow{d} \kappa_d L_i$ and $\sqrt{T}PR_{X_i,T}u \xrightarrow{p} 0$.

The common coefficient across the marginal variances arises because uniformly scaling the innovations proportionally rescales the within-window variation. This invariance is restricted to the Gaussian family and does not extend to the innovation distribution itself. For instance, in dimension two, the integrated kernel mass normalized by the innovation standard deviation is $\pi \mathbb{E}|\varepsilon|/(2\sigma)$, which differs between Gaussian and Laplace innovations.

The entropy step uses the following deterministic result. The norm in its statement is the operator norm, and the constant may depend on fixed dimension.

Lemma 1 (Expansion at a rank-one boundary). *For density matrices R, S let $\eta = \max(\|R - \Pi\|_{\text{op}}, \|S - \Pi\|_{\text{op}})$. As $\eta \downarrow 0$,*

$$S((R + S)/2) - \{S(R) + S(S)\}/2 = F(PRP, PSP) + O_d(\eta^2 \log(1/\eta)), \quad (3.3)$$

uniformly over actual density matrices in this neighborhood.

Uniformity over the actual matrices avoids a hidden positivity issue when considering perturbations from a rank-one matrix. A norm bound on the perturbation direction does not guarantee a common positive step size for every ray leaving the rank-one matrix. Similarly, differentiability of entropy at a positive definite matrix does not justify applying the same expansion at a boundary point with a zero eigenvalue.

Theorem 2 (Entropy limit). *For any two coordinates under Assumption 1,*

$$\sqrt{T}Q_T \xrightarrow{d} \kappa_d \Psi(L_1, L_2). \quad (3.4)$$

Proof. Theorem 1 gives $\eta_T = O_p(T^{-1/2})$. The remainder in Lemma 1, multiplied by $\sqrt{T}$, is $o_p(1)$. Continuity and positive homogeneity of F give the limit $F(L_1\Lambda_d, L_2\Lambda_d)$, which equals the stated expression by (2.7). $\square$

For independent random walks, the local times are independent, with each following the distribution of $|N(0, 1)|$. Consequently, they are not asymptotically equal relative to their magnitudes. Therefore, replacing Ψ with a quadratic expansion around the diagonal changes the resulting limiting distribution. At perfect positive or negative correlation, the local times coincide, causing the first-order entropy gap to vanish. We do not claim any finer convergence rate at these boundary points.

The full matrix limit also explains what would happen with a Frobenius comparison. With $A_d = \Lambda_d - \kappa_d \Pi$,

$$\sqrt{T}\|R_{X,T} - R_{Y,T}\|_F \xrightarrow{d} \|A_d\|_F |L_1 - L_2|. \quad (3.5)$$

Both comparisons therefore retain local-time dependence. Their limit functionals are different: the entropy gap also depends on the total local-time mass. Neither removes the correlation nuisance through boundedness or marginal scale invariance.

4 Levels and the boundary of the null

The local-time characterization concerns fixed starting levels. We now give a different result when the initial level dominates stochastic fluctuations. It applies beyond Gaussian innovations.

Theorem 3 (Large-level regime). *Let $X_{k,T} = c_T + \sum_{j=1}^k \varepsilon_j$ with centered iid innovations of finite positive variance σ^2 , and suppose $|c_T|/\sqrt{T} \rightarrow \infty$. Let Γ be its within-window partial-sum vector and $v = P\Gamma$. On events where the statistic is defined,*

$$c_T^2 P R_{X,T} P \xrightarrow{p} \frac{\mathbb{E} v v'}{d}, \quad c_T^2 D_{X,T} \xrightarrow{p} \frac{\sigma^2(d^2 - 1)}{6d}. \quad (4.1)$$

The probability that all windows are nonzero tends to one. For two series satisfying these assumptions with their respective levels, $\sqrt{T}Q_T \xrightarrow{p} 0$, regardless of their cross-series dependence.

Proof. The maximal inequality for centered independent sums gives

$$\frac{1}{|c_T|} \max_{k \leq T} \left| \sum_{j \leq k} \varepsilon_j \right| \xrightarrow{p} 0.$$

Consequently $\|w_n\|^2/(dc_T^2) \xrightarrow{p} 1$ uniformly in n , and all windows are nonzero with probability tending to one. The stationary finite-dependent sequence $v_n v'_n$ has integrable entries. Its sample average converges in probability to $\mathbb{E} v v'$, by splitting into finitely many independent residue classes. The uniform denominator error is bounded by its supremum times the sample mean of $\|v_n\|^2$. This proves the matrix limit. Direct summation of the covariance of the partial-sum vector gives $\mathbb{E}\|v\|^2 = \sigma^2(d^2 - 1)/6$ and hence the trace limit.

For any trace-one R with defect D , pinching into $\text{span}(u)$ and $u^\perp$ increases entropy. Indeed, for $U = \Pi - P$, the pinched matrix is $(R + URU')/2$, so this follows from concavity and unitary invariance. The entropy of its compressed block is maximized at equal eigenvalues. Thus

$$S(R) \leq h_2(D) + D \log(d - 1), \quad (4.2)$$

where $h_2(D) = -D \log D - (1 - D) \log(1 - D)$. This also follows at $D = 0$ by continuity; when $d = 2$ the last term is zero. The mixture of the two sample matrices has defect $\bar{D} = (D_{X,T} + D_{Y,T})/2 = O_p(b_T)$, where $b_T = c_{X,T}^{-2} + c_{Y,T}^{-2} = o(T^{-1})$. Since $0 \leq Q_T \leq S((R_X + R_Y)/2)$, (4.2) bounds it by $O_p(b_T \log(1/b_T))$. This bound is $o_p(T^{-1/2})$. $\square$

The normalization governing the distinction is $c_T/(\sigma\sqrt{T})$, not the unscaled starting level. A fixed nonzero level is in the first regime eventually. Intermediate levels proportional to $\sqrt{T}$ lead to shifted local-time questions. We do not use an intermediate-regime theorem or an estimated-centering theorem in the claims of this paper.

The compressed limit in (4.1) does not automatically yield a c_T^2 entropy limit. The coupling contains the sample mean of v_n divided by c_T , which can have a different order from the compression. Bounding the full distance to Π by the scalar defect would miss this distinction. The entropy bound above supplies exactly the collapse needed for inference, without asserting an unproved finer rate.

4.1 Correlation dependence of the entropy law

Let B, W be independent standard Brownian motions and $B_\rho = \rho B + \sqrt{1 - \rho^2} W$. Denote their local times at zero at time one by L, L_ρ . For $|\rho| < 1$, the occupation-density calculation gives

$$\mathbb{E}(LL_\rho) = \int_0^1 \int_0^1 \frac{ds dt}{2\pi \sqrt{st - \rho^2 \min(s, t)^2}} = \frac{2 \arcsin \rho}{\pi \rho}, \quad (4.3)$$

with value $2/\pi$ at zero. To justify the calculation, replace each local time by a Gaussian mollification. The determinant is at least $(1 - \rho^2)st$, so dominated convergence applies. On the triangle $s \leq t$, substitute $s = tr$ and integrate; symmetry supplies the other triangle. This is a specialization of occupation-moment methods, not a new moment formula (Fitzsimmons and Pitman, 1999).

The marginal local-time law is half-normal, by Lévy's identity and the reflection principle (Mörters and Peres, 2010, Chapter 6). Hence $\mathbb{E}L^2 = 1$ and

$$\mathbb{E}(L - L_\rho)^2 = 2 - \frac{4 \arcsin \rho}{\pi \rho} \longrightarrow 0 \quad \text{as } \rho \uparrow 1. \quad (4.4)$$

Continuity of Ψ and tightness imply $\kappa_d \Psi(L, L_\rho) \xrightarrow{p} 0$. At $\rho = 0$, independent continuous local times are unequal almost surely and strict convexity gives a strictly positive gap. This establishes variation of the entropy law itself; a changing cross moment on its own would not establish that claim. Changing the sign of the second Brownian coordinate leaves its zero local time unchanged, so the gap law is even in ρ .

5 Consequences for Inference

Theorem 4 (Fixed-cutoff obstruction). *Fix $d \geq 2$ and $q > 0$. Let $\mathcal{P}$ be the Gaussian random-walk family in Assumption 1 with two coordinates, fixed initial levels, unit marginal variances and $|\rho| < 1$. Then*

$$\lim_{T \rightarrow \infty} \sup_{P_\rho \in \mathcal{P}} P_\rho(\sqrt{T}Q_T < q) = 1. \quad (5.1)$$

The result also holds for deterministic cutoffs converging to $q > 0$. Every member of $\mathcal{P}$ is noncointegrated.

Proof. Positive definiteness of the innovation covariance for $|\rho| < 1$ gives positive innovation variance for every nonzero linear combination. Its partial sum is therefore nonstationary. For any $\epsilon > 0$, (4.4) permits a fixed $\rho_\epsilon < 1$ such that the limiting gap is below q with probability greater than $1 - \epsilon$. The open-set part of the Portmanteau theorem and Theorem 2 give

$$\liminf_T P_{\rho_\epsilon}(\sqrt{T}Q_T < q) > 1 - \epsilon.$$

The supremum is at least this probability and at most one. Let $\epsilon \downarrow 0$. For a convergent deterministic cutoff, first use $q/2$. $\square$

The theorem does not assert that every test fails, that endpoint quantiles are stochastically ordered, or that there is size failure on every compact interval $|\rho| \leq \bar{\rho} < 1$. Correlation-dependent or estimated calibration is outside its conclusion. Its target is the practice of treating a positive

lower-tail cutoff indexed only by sample size and window dimension as valid across the unrestricted noncointegration null.

Theorem 3 gives a second obstruction for that same rule. Two independent Gaussian walks with levels satisfying its assumptions remain noncointegrated, while $\sqrt{T}Q_T \xrightarrow{p} 0$. Rejection then tends to one for a positive cutoff from the fixed-level regime. No correlation misspecification is needed for this mechanism.

6 Proof of the matrix characterization

We provide the essential arguments here. Fix a coordinate with variance σ^2 , anchor a window at X_n , and write

$$w_n = X_n \mathbf{1} + \Gamma_n, \quad m_n = d^{-1} \mathbf{1}' \Gamma_n, \quad v_n = P \Gamma_n, \quad V_n = \|v_n\|^2.$$

The mark Γ_n uses innovations $n+1, \dots, n+d-1$ and is independent of the full vector innovation history through n . For a generic mark write m, v, V without subscripts. The compressed block, coupling and scalar defect at anchor x are

$$\begin{aligned} H(x, \Gamma) &= \frac{vv'}{d(x+m)^2 + V}, & g(x, \Gamma) &= \frac{\sqrt{d}(x+m)v}{d(x+m)^2 + V}, \\ h(x, \Gamma) &= \frac{V}{d(x+m)^2 + V}. \end{aligned} \tag{6.1}$$

A bar denotes expectation over the mark.

Lemma 2 (Averaged kernels). *The entries of $\bar{H}, \bar{g}$ and $\bar{h}$ are bounded and belong to $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Their masses are*

$$\int \bar{H} = \sigma \Lambda_d, \quad \int \bar{h} = \sigma \kappa_d, \quad \int \bar{g} = 0. \tag{6.2}$$

Proof. All unaveraged entries are bounded. On $A_x = \{\|\Gamma\| \leq \sqrt{d}|x|/2\}$ the denominator is at least $dx^2/4$. On its complement, Markov's inequality gives probability at most $4\mathbb{E}\|\Gamma\|^2/(dx^2)$. Hence $\bar{h}(x) \leq \min(1, C_d \sigma^2/x^2)$, and every absolute entry of $\bar{H}$ is bounded by $\bar{h}$.

The coupling needs a cancellation. The smooth map $w \mapsto (Pw)(u'w)/\|w\|^2$ has first-order expansion $v/(\sqrt{d}x)$ at $w = x\mathbf{1}$. Along the segment to $x\mathbf{1} + \Gamma$ on A_x , its second derivatives are bounded by C_d/x^2 . Its remainder is consequently bounded by $C_d\|\Gamma\|^2/x^2$. On A_x^c , use $\|g\| \leq 1/2$ and

$$|x|^{-1} \mathbb{E}[\|\Gamma\| \mathbf{1}_{A_x^c}] \leq C_d x^{-2} \mathbb{E}\|\Gamma\|^2.$$

Since $\mathbb{E}v = 0$, subtracting the leading term before taking expectations proves $\|\bar{g}(x)\| \leq C_d \sigma^2/x^2$. Boundedness near zero establishes both integrability requirements.

Integrating the first two kernels by completing the square gives

$$\int_{\mathbb{R}} \frac{dx}{d(x+m)^2 + V} = \frac{\pi}{\sqrt{dV}}.$$

Tonelli applies to the defect; absolute integrability of the matrix entries follows from their domination by it. Scaling Γ by σ gives the first two identities in (6.2).

For the last identity, let $a = \sqrt{V/d}$ and $q_a(x) = x/(x^2 + a^2)$. Central symmetry of Γ implies $\mathbb{E}[vq_a(x)] = 0$. Thus

$$\bar{g}(x) = \mathbb{E} \left[\frac{v}{\sqrt{d}} \{q_a(x+m) - q_a(x)\} \right].$$

For $a > 0$, $\int |q'_a| = 2/a$. The integral over x of the norm inside this expectation is at most $2|m|$, with finite mean. Fubini now applies to the difference. Its signed integral is zero, as follows from the logarithmic primitive of q_a . This proves the coupling mass without interchanging the non-absolutely-integrable original kernel. $\square$

Lemma 3 (Overlapping marks). *For an entry f of H, g , or $-h$, set $M_n = f(X_n, \Gamma_n) - \bar{f}(X_n)$. Then*

$$\mathbb{E} \left| T^{-1/2} \sum_{n=1}^N M_n \right|^2 = O_d(T^{-1/2}). \quad (6.3)$$

Proof. Let $\mathcal{F}_n$ be the full vector history through n . We have $\mathbb{E}(M_n \mid \mathcal{F}_n) = 0$, and M_n is measurable at $n+d-1$. Within each residue class modulo $d-1$, the variables are orthogonal in L^2 . Cauchy–Schwarz across the classes gives

$$\mathbb{E} \left| \sum M_n \right|^2 \leq (d-1) \sum \mathbb{E} M_n^2.$$

Each entry satisfies $|f|^2 \leq h$: in particular $\|g\|^2 = h(1-h)$ and $\|H\|_F^2 = h^2$. Conditional variance is bounded by the conditional second moment, so $\mathbb{E} M_n^2 \leq \mathbb{E} h(X_n, \Gamma_n)$. The Gaussian anchor density is bounded by $(2\pi n \sigma^2)^{-1/2}$, uniformly in its mean. Lemma 2 therefore gives $\mathbb{E} h(X_n, \Gamma_n) \leq Cn^{-1/2}$. Summation and division by T yield (6.3). $\square$

The occupation theorem we use states, in the present Gaussian specialization, that for a fixed bounded $f \in L^1 \cap L^2$,

$$T^{-1/2} \sum_{n=1}^T f(c + \sigma S_n^*) \xrightarrow{d} \sigma^{-1} \left(\int f \right) L_B(1, 0), \quad (6.4)$$

where S_n^* is a standard Gaussian walk and c is fixed. The common-space conclusion of Theorem 2.1 in Wang and Phillips (2006), under its Assumptions 2.1–2.3 and 2.2*, provides the required joint version. This is the discussion-paper numbering; Wang and Phillips (2009) is the published article.

Here is the assumption mapping. We construct a vector Brownian motion with covariance matrix C and represent the normalized Gaussian walks on the grid n/T using this common probability space. By continuity of Brownian paths, the resulting processes converge uniformly. Conditional on the complete vector past, a standardized increment of a future coordinate is independent and follows a standard normal distribution. Its density is therefore bounded and uniformly continuous. For a gap of length $l-k$, the relevant scale is $\sqrt{(l-k)/T}$. The corresponding inverse-scale sum, after division by T , remains bounded, while the contribution from gaps of size at most aT is bounded by $2\sqrt{a} + o(1)$. Outside these small gaps, the scale is bounded away from zero. These conditions provide the required controls on the conditional densities and their associated scales. The theorem is applied with kernel scale $\sigma\sqrt{T}$, which diverges as $T \rightarrow \infty$, while its ratio to T converges to zero.

We then apply the common-space result to each of the finitely many coordinate kernels. A union bound preserves joint convergence to the same vector Brownian motion, without requiring independence between coordinates. If a kernel has zero mass, it can be expressed as the difference of two kernels belonging to $L^1 \cap L^2$, each with nonzero mass, by adding a positive integrable kernel. Applying the result to both kernels on the same probability space and taking their difference establishes the zero-mass case. Alternatively, a fixed initial level can be incorporated into a fixed translated kernel, since translation does not change its total mass.

Combining (6.4), Lemmas 2 and 3, and the identity of the top block minus one with $-h$, proves Theorem 1. Using N instead of T and omitting finitely many endpoints changes the normalized sums by $o_p(1)$, since all window entries and kernels are bounded. This completes the stochastic argument.

6.1 Extension to additive angular features

The marked reduction extends beyond matrix entries. Write $e = \mathbf{1}$ and let $\tau : \mathbb{R}^d \setminus \{0\} \rightarrow \mathbb{R}^q$ be bounded and measurable, with $\tau(aw) = \tau(w)$ for every nonzero real a . Suppose τ is C^2 near e ; global smoothness is unnecessary. Put $\tau_0 = \tau(e)$, $A = D\tau(e)$ and $V_{i,T} = N^{-1} \sum_{n=1}^N \tau(w_{X_i,n})$. Homogeneity gives $Ae = 0$.

Theorem 5 (Additive angular features). *Under Assumption 1 and the preceding map conditions, jointly over finitely many maps and coordinates,*

$$\sqrt{T}(V_{i,T} - \tau_0) \xrightarrow{d} K_\tau L_i, \quad K_\tau = \int_{\mathbb{R}} \mathbb{E}[\tau(ye + \Gamma^*) - \tau_0] dy,$$

where Γ^* is the standard Gaussian within-window mark. The coefficient is finite and independent of marginal innovation scale, but may be zero. The centered marked remainder has L^2 norm $O(T^{-1/4})$ after normalization by $T^{-1/2}$.

Proof. For a mark with variance scale σ , define $f(x, \Gamma) = \tau(xe + \Gamma) - \tau_0$, $h_\sigma(x) = \mathbb{E}f(x, \Gamma)$ and $b_\sigma(x) = \mathbb{E}\|f(x, \Gamma)\|^2$. On $\{\|\Gamma\| \leq r|x|\}$, for sufficiently small fixed $r > 0$, Taylor expansion and signed scale invariance give

$$f(x, \Gamma) = A\Gamma/x + O(\|\Gamma\|^2/x^2).$$

On the complement, boundedness of f and

$$\Pr(\|\Gamma\| > r|x|) \leq \frac{\mathbb{E}\|\Gamma\|^2}{r^2 x^2}, \quad \frac{\mathbb{E}[\|\Gamma\| \mathbf{1}_{\{\|\Gamma\| > r|x|\}}]}{|x|} \leq \frac{\mathbb{E}\|\Gamma\|^2}{rx^2}$$

control the discarded terms. Since $\mathbb{E}\Gamma = 0$, subtracting $A\Gamma/x$ before averaging proves $\|h_\sigma(x)\| = O(x^{-2})$; the squared Taylor bound gives $b_\sigma(x) = O(x^{-2})$. Both functions are bounded near zero. Thus $h_\sigma \in L^1 \cap L^2$ entrywise and $b_\sigma \in L^1$.

For $M_n = f(X_n, \Gamma_n) - h_\sigma(X_n)$, the residue-class argument of Lemma 3 applies with b_σ in place of $\bar{h}$. The Gaussian density bound gives $\sum \mathbb{E}\|M_n\|^2 = O(\sqrt{T})$ and hence the stated L^2 rate. Equation (6.4), on the same joint coupling, gives the limit with coefficient $\sigma^{-1} \int h_\sigma$. Changing variables $x = \sigma y$ identifies K_τ . The bounded endpoint terms are negligible. $\square$

This is a sufficient class, not a minimal one. The reduction only needs the displayed kernel integrability and squared-mark envelope, together with finite look-ahead centering and the occupation input. Smoothness verifies these conditions rather than assuming the desired reduction. The marked L^2 rate is not a rate of convergence to local time. When $K_\tau = 0$, the conclusion is $o_p(T^{-1/2})$, with no universal finer rate.

An exact coefficient can also be written as an absolutely integrable compensated line integral:

$$K_\tau = \sigma^{-1} \mathbb{E} \int_{\mathbb{R}} \left\{ f(x, \Gamma) - \frac{x A \Gamma}{1 + x^2} \right\} dx.$$

For $|x| \geq 1$ the expected absolute integrand is bounded by $Cx^{-2} + C|x|^{-3}$ using the proof above; on compact sets it is integrable. Fubini is therefore valid. The uncompensated line integral need not be absolutely integrable for individual marks.

The class contains the spatial-sign outer product and all fixed even angular moments $\tau(w) = (w/\|w\|)^{\otimes 2k}$. These are finite-dimensional additive features. A smooth function of their averages has the corresponding delta-method limit, possibly zero; a singular function requires a separate expansion, as entropy does here. This does not extend the theorem to arbitrary statistics of a delay point cloud.

Proposition 1 (Derivative-dependent large-level rates). *For a Gaussian walk with $|c_T|/\sqrt{T} \rightarrow \infty$, under the same map conditions, let $a = (0, 1, \dots, d-1)'$. Then*

$$c_T \sqrt{T} (V_T - \tau_0) \xrightarrow{d} N_q(0, \sigma^2(Aa)(Aa)').$$

If $A = 0$, the stronger conclusion is

$$c_T^2 (V_T - \tau_0) \xrightarrow{p} \frac{1}{2} \mathbb{E} D^2 \tau(e) [\Gamma, \Gamma].$$

The latter conclusion needs only centered iid finite-variance innovations, with an arbitrary bounded value assigned to $\tau(0)$; the probability of encountering a zero window tends to zero.

Thus a common level boundary does not imply a common large-level rate. In dimension two, $\tau(w) = (w_2^2 - w_1^2)/\|w\|^2$ has $Aa = 1$, so its c_T^2 normalization is not tight. The compressed spatial-sign map has $A = 0$; the full matrix does not. Even $A \neq 0$ need not imply nondegeneracy, because Aa can vanish through overlapping-filter cancellation.

Scale invariance alone is insufficient even at fixed levels. The bounded even map $1_{\{Pw \neq 0\}}$ equals one almost surely on Gaussian windows but is zero at e . The continuous map $\|Pw\|/\|w\|$, whose angular restriction is Lipschitz, has averaged tail proportional to $1/|x|$, which is not integrable. Finally, invariance only under positive scalars allows different limits at e and $-e$, and order-one sign occupation. These examples rule out a theorem asserting the same collapse and rate for every scale-invariant delay statistic.

6.2 Proof of the boundary expansion

In the decomposition $\text{span}(u) \oplus u^\perp$, write

$$R = \begin{pmatrix} a & g' \\ g & H \end{pmatrix}, \quad a = 1 - \text{tr}(H).$$

If $\|R - \Pi\|_{\text{op}} \leq \eta < 1/4$, then $a \geq 1 - \eta$, $0 \leq H \leq \eta I$, and $\|g\| \leq \eta$. There is one eigenvalue of R above $1 - \eta$ and all remaining eigenvalues lie in $[0, \eta]$. For $x \leq \eta$, block elimination by a congruence shows that the number of eigenvalues of R below x is the number of negative eigenvalues of

$$H - xI - \frac{gg'}{a - x}.$$

The subtracted matrix lies between zero and $2\|g\|^2 I$. Comparing ordered eigenvalue counts, and then using continuity at endpoints, bounds each small eigenvalue of R between the corresponding eigenvalue of H minus $2\eta^2$ and that eigenvalue itself. Nonnegativity of the actual eigenvalues remains available even if the lower bound is negative. Trace one gives the largest eigenvalue as $1 - \text{tr}(H) + O_d(\eta^2)$.

For $b(t) = -t \log t$, $b(0) = 0$, integration of $|1 + \log t|$ gives

$$|b(x) - b(y)| \leq h \log(1/h) + O(h), \quad 0 \leq x, y \leq 1, \quad |x - y| \leq h \leq 1/2.$$

Apply this bound to the small eigenvalues with $h = 2\eta^2$, and use smoothness of b near one for the top eigenvalue. Then

$$S(R) = \text{tr}(H) + \text{st}(H) + O_d(\eta^2 \log(1/\eta)).$$

The same bound applies to S and $(R + S)/2$. The three trace terms cancel, proving Lemma 1. This argument includes singular compressions and rotating pure states and requires no gap between the small eigenvalues.

6.3 Proof of derivative-dependent level rates

We prove the main paper's proposition entitled "Derivative-dependent large-level rates". Write $X_{n,T} = c_T + S_n$, $A = D\tau(e)$ and $B = D^2\tau(e)$. The maximal inequality gives $\max_{n \leq T} |S_n|/|c_T| \xrightarrow{p} 0$. The within-window marks are differences of partial sums, so $\max_n \|\Gamma_n\|/|c_T| \xrightarrow{p} 0$ as well. Consequently the Taylor expansion around e holds uniformly with probability tending to one, and all windows are nonzero on this event.

If $A = 0$, uniform second-order Taylor expansion gives

$$\tau(X_n e + \Gamma_n) - \tau_0 = \frac{B[\Gamma_n, \Gamma_n]}{2X_n^2} + r_n \frac{\|\Gamma_n\|^2}{X_n^2}, \quad \max_n \|r_n\| \xrightarrow{p} 0.$$

Uniformly replacing c_T^2/X_n^2 by one, and applying the finite-dependent law of large numbers to $B[\Gamma_n, \Gamma_n]$ and $\|\Gamma_n\|^2$, proves the deterministic limit. This argument needs only the finite second moment and no upper bound on the growth of $|c_T|$.

For general A , Taylor expansion instead yields

$$V_T - \tau_0 = N^{-1} \sum_{n=1}^N A \Gamma_n / X_n + O_p(c_T^{-2}).$$

The remainder vanishes after multiplication by $|c_T| \sqrt{T}$. Replacement of $1/X_n$ by $1/c_T$ in the linear sum requires cancellation, not an absolute-value bound. Fix $\eta \in (0, 1/2)$ and set

$$b_n = (X_n^{-1} - c_T^{-1}) 1_{\{\max_{j \leq n} |S_j| \leq \eta |c_T|\}}.$$

This coefficient is measurable with respect to the full innovation history through n and satisfies $|b_n| \leq 2\eta/|c_T|$. Since Γ_n uses subsequent innovations, $b_n A \Gamma_n$ is conditionally centered with look-ahead $d - 1$. Residue-class orthogonality therefore gives

$$\mathbb{E} \left\| \frac{c_T}{\sqrt{T}} \sum_{n=1}^N b_n A \Gamma_n \right\|^2 \leq C\eta^2.$$

For each fixed η the stopping event has probability tending to zero. Chebyshev's inequality followed by $\eta \downarrow 0$ proves the replacement in probability. The coordinate of $\sum \Gamma_n$ corresponding to lag j equals $j \sum_{n=1}^N \varepsilon_n$ plus finitely many endpoint innovations. Their normalized contribution vanishes. The iid central limit theorem gives the covariance $\sigma^2 aa'$, and multiplication by A proves the Gaussian limit. Using $N = T - d + 1$ changes neither argument.

For the explicit counterexample $\tau(w) = (w_2^2 - w_1^2)/\|w\|^2$, differentiation at $(1, 1)'$ gives $A = (-1, 1)$ and $Aa = 1$. Thus $c_T \sqrt{T} V_T$ has a nondegenerate continuous Gaussian limit. Since $|c_T|/\sqrt{T} \rightarrow \infty$, $c_T^2 V_T$ is not tight. This is a full-map counterexample, not a claim about the finer entropy rate.

6.4 Source and proof dependency map

The spatial-sign representation is an established construction of Dürre et al. (2014), which treats known and estimated location in a fixed-distribution sampling setting. The overlapping nonstationary windows in the main paper are not covered by importing that asymptotic normality result. The elliptical eigenvalue identification in Dürre et al. (2016) concerns a population matrix with squared radius in the denominator. The occupation coefficient here contains $\mathbb{E}(vv'/\|v\|)$, so that identification result is not transferred to it.

For the occupation step, the precise accessible source is Theorem 2.1 and Assumptions 2.1–2.3, including the common-space condition 2.2, in the author discussion paper (Wang and Phillips, 2006). The final publication is Wang and Phillips (2009). The main text verifies the Gaussian density and scale conditions and supplies the finite-mark reduction. It identifies the theorem numbering as belonging to the discussion paper rather than silently assuming that numbering was unchanged in publication.

The entropy background is Watrous (2018, Chapter 5). Propositions 5.19–5.20 give isometric invariance and direct-sum identities; equation (5.98) gives entropy scaling on unnormalized PSD matrices; Proposition 5.22 and Theorem 5.23 give nonnegativity of relative entropy for density matrices and entropy concavity. Theorem 5.26 states the Fannes-Audenaert bound with half the trace norm. The book uses base-two logarithms; the present paper uses natural logarithms. These standard properties are not claimed as original. The rank-one remainder is proved by spectral comparison in the main paper.

The local-time mixed moment can be obtained from additive-functional methods (Fitzsimmons and Pitman, 1999), but the main text provides the Gaussian density integral and its domination explicitly. Deterministic terminal time can be represented by killing the space-time process at time one; it should not be confused with independent exponential killing. The half-normal marginal follows from Lévy's local-time identity and reflection (Mörters and Peres, 2010, Chapter 6).

The dependency chain is: averaged kernel masses and overlapping-mark bound; the joint Gaussian occupation input; the full matrix limit; the deterministic rank-one entropy expansion;

the entropy limit; and the correlation-boundary obstruction. The large-level theorem instead uses a maximal inequality, a finite-dependent law of large numbers, and an entropy bound after pinching. It does not require the Gaussian occupation theorem.

7 Numerical and empirical evidence

7.1 Characterization simulation design

The base experiment has 16 cells: $T = 256, 1024, 4096, 16384$, $d = 2, 10$, and innovation correlation zero or 0.9. Each cell uses 1,000 independent pairs. Standard normal innovation pairs are transformed linearly to attain the stated correlation, and then cumulated from zero. The same realizations receive an additive level $T^{3/4}$ for the large-level comparison. There is no estimated drift in these simulations.

The seed root is 20260913, with NumPy `SeedSequence` keys consisting of sample size, dimension, integer-coded correlation, and fixed chunk identity. A chunk has 50 replications and does not change with worker count. The $d = 10$ coefficient uses an independent stream of one million window marks. The scripts report its Monte Carlo standard error separately from the errors of the time-series simulation means.

The large- T extension was selected after inspecting the base grid. It adds $T = 65,536$ and 262,144 at $d = 10$ and zero correlation, again with 1,000 pairs. It is stored separately in `results/characterization_large_T.csv`. This extension is evidence about convergence at larger samples, not an independent confirmation selected before observing the original discrepancy. Both experiments are retained in the package.

To calculate the limiting mean under independence, independent local times have the law of the absolute values of two independent standard normals. In polar coordinates their angle is uniform on $[0, \pi/2]$ and their independent radius has mean $\sqrt{\pi/2}$. Homogeneity gives

$$\mathbb{E}\Psi(L_1, L_2) = \sqrt{\frac{\pi}{2}} \frac{2}{\pi} \int_0^{\pi/2} \Psi(\cos \theta, \sin \theta) d\theta.$$

Numerical quadrature gives 0.1115792734. The comparison in the main figure is to this limiting-law moment; it does not claim that weak convergence establishes moment convergence or that central-moment agreement establishes lower-tail accuracy.

7.2 Complete correlation and location grid

For each sample size the reference draws and evaluation draws use distinct streams with seed root 20260911. Reference pools have 5,000 draws; evaluation samples have 3,000. The independent reference has correlation zero, while the oracle reference uses the evaluation correlation. Both references start at zero. Evaluation paths start at either zero or $10\sqrt{T}$. The dimension is ten. “Mean” in the table denotes global demeaning of each full series, also applied separately to every reference draw.

Entries are the fractions with plus-one lower-tail reference probability no greater than 0.05. The conditional standard error is $\sqrt{\hat{p}(1 - \hat{p})/3000}$. This does not include variability of the shared reference pool. A reported rate of one or zero has a plug-in standard error of zero but does not mean that the population probability is known exactly. No failure was discarded.

Table 1: Complete Gaussian reference grid. Level denotes $c/\sqrt{T}$.

T	ρ	Level	Adjust.	Indep.	Oracle
50	0	0	mean	0.049	0.049
50	0	0	raw	0.051	0.051
50	0	10	mean	0.049	0.049
50	0	10	raw	1.000	1.000
50	0.9	0	mean	0.487	0.049
50	0.9	0	raw	0.406	0.043
50	0.9	10	mean	0.487	0.049
50	0.9	10	raw	1.000	1.000
50	0.99	0	mean	0.963	0.041
50	0.99	0	raw	0.857	0.048
50	0.99	10	mean	0.963	0.041
50	0.99	10	raw	1.000	1.000
300	0	0	mean	0.056	0.056
300	0	0	raw	0.050	0.050
300	0	10	mean	0.056	0.056
300	0	10	raw	1.000	1.000
300	0.9	0	mean	0.285	0.061
300	0.9	0	raw	0.266	0.050
300	0.9	10	mean	0.285	0.061
300	0.9	10	raw	1.000	1.000
300	0.99	0	mean	0.802	0.048
300	0.99	0	raw	0.694	0.050
300	0.99	10	mean	0.802	0.048
300	0.99	10	raw	1.000	1.000
1256	0	0	mean	0.048	0.048
1256	0	0	raw	0.045	0.045
1256	0	10	mean	0.048	0.048
1256	0	10	raw	1.000	1.000
1256	0.9	0	mean	0.172	0.051
1256	0.9	0	raw	0.193	0.050
1256	0.9	10	mean	0.172	0.051
1256	0.9	10	raw	1.000	1.000
1256	0.99	0	mean	0.550	0.052
1256	0.99	0	raw	0.530	0.051
1256	0.99	10	mean	0.550	0.052
1256	0.99	10	raw	1.000	1.000
4096	0	0	mean	0.043	0.043
4096	0	0	raw	0.043	0.043
4096	0	10	mean	0.043	0.043

T	ρ	Level	Adjust.	Indep.	Oracle
4096	0	10	raw	1.000	1.000
4096	0.9	0	mean	0.122	0.051
4096	0.9	0	raw	0.158	0.053
4096	0.9	10	mean	0.122	0.051
4096	0.9	10	raw	1.000	1.000
4096	0.99	0	mean	0.398	0.053
4096	0.99	0	raw	0.415	0.055
4096	0.99	10	mean	0.398	0.053
4096	0.99	10	raw	1.000	1.000

7.3 Focused classical comparison

The seed root is 20260914. Separate reference streams use 5,000 Gaussian paths at correlations zero and 0.99. Every statistic in those two rules is globally demeaned. Evaluation streams are indexed by design and fixed 50-replication chunks. The null has Gaussian innovations with the stated correlation. Under the alternative, $Y_t = X_t + Z_t$, with independent stationary $Z_t = \phi Z_{t-1} + 0.3e_t$. The initial variance of Z is $0.09/(1 - \phi^2)$.

Engle–Granger uses a constant, AIC lag selection, and maximum lag ten, with Y regressed on X . Johansen uses the rank-zero trace statistic, `det_order=0` and one lagged difference. Phillips–Ouliaris uses the `arch` implementation with a constant and its default kernel and bandwidth.

The comparison uses nominal classical critical values and probabilities, not empirical size adjustment. Johansen’s excess null rejection remains visible. The table must not be interpreted as a size-matched ranking. Nor is the fixed-correlation entropy endpoint a proved uniformly valid reference. The main conclusion is the sensitivity of apparent performance to the nuisance class used for calibration.

Table 2: Complete focused comparison with conditional Monte Carlo errors.

Design	Param.	Rule	Flags	Rate	MCSE
Null	0.0	Q_0	58	0.058	0.0074
Null	0.0	$Q_{.99}$	0	0.000	0.0000
Null	0.0	EG	65	0.065	0.0078
Null	0.0	Joh.	137	0.137	0.0109
Null	0.0	PO	67	0.067	0.0079
Null	0.9	Q_0	288	0.288	0.0143
Null	0.9	$Q_{.99}$	0	0.000	0.0000
Null	0.9	EG	52	0.052	0.0070
Null	0.9	Joh.	116	0.116	0.0101
Null	0.9	PO	53	0.053	0.0071
Null	0.99	Q_0	815	0.815	0.0123
Null	0.99	$Q_{.99}$	56	0.056	0.0073

Design	Param.	Rule	Flags	Rate	MCSE
Null	0.99	EG	66	0.066	0.0079
Null	0.99	Joh.	116	0.116	0.0101
Null	0.99	PO	66	0.066	0.0079
Alt.	0.7	Q_0	997	0.997	0.0017
Alt.	0.7	$Q_{.99}$	47	0.047	0.0067
Alt.	0.7	EG	999	0.999	0.0010
Alt.	0.7	Joh.	1000	1.000	0.0000
Alt.	0.7	PO	1000	1.000	0.0000
Alt.	0.97	Q_0	753	0.753	0.0136
Alt.	0.97	$Q_{.99}$	7	0.007	0.0026
Alt.	0.97	EG	201	0.201	0.0127
Alt.	0.97	Joh.	247	0.247	0.0136
Alt.	0.97	PO	214	0.214	0.0130
Alt.	0.99	Q_0	553	0.553	0.0157
Alt.	0.99	$Q_{.99}$	5	0.005	0.0022
Alt.	0.99	EG	83	0.083	0.0087
Alt.	0.99	Joh.	148	0.148	0.0112
Alt.	0.99	PO	93	0.093	0.0092

There are three distinct kinds of reference calculation. Simulation under a fully specified known Gaussian model can produce a finite-sample Monte Carlo test using the rank of the observation among exchangeable draws. Estimating a nuisance parameter and inserting it into a reference law requires additional justification. Resampling from an estimated serial model requires a conditional approximation for the marked occupation statistic, not merely an ordinary functional central limit theorem. Smoothing residuals and restoring covariance rank do not themselves prove that approximation.

The plus-one formula $(1 + \#\{Q_b \leq Q_{\text{obs}}\})/(B + 1)$ avoids a zero reported probability and gives a conservative rank test in an exchangeable continuous setup, up to its attainable grid. It does not repair a misspecified reference law. Similarly, a confidence interval for a simulated reference probability measures Monte Carlo precision conditional on the fitted law; it is not an interval for the truth of cointegration.

The frozen price dates are 1 June 2021 through 29 June 2026. The equity columns are CVX, GDX, GLD, IEF, KO, PEP, SLV, TLT and XOM. Their common snapshot has 1,275 rows and no missing price entries. The showcase snapshot has 1,277 rows; the control snapshot has 1,275. Full column lists, file hashes and missingness are in `results/data_panel_metadata.json` and the input manifest. Prices are transformed by the natural logarithm before calculations.

The equity list contains exchange-traded products as well as individual equities; the inherited panel name is a file identifier, not a claim of homogeneous asset type. The showcase pairs are GLD/IAU, SPY/IVV, GLD/GC=F, SLV/SI=F, EWA/EWC and XLE/VDE. The ten controls are UNG/TLT, UNG/KO, UNG/GLD, UNG/IEF, EWJ/GLD, DBA/IEF, DBA/TLT, USO/TIP, EWJ/TLT and USO/KO. Selection was inherited from the research project. No

trading-performance claim is made.

The raw independent reference contains 2,000 standard Gaussian random walks, with root seed 20260826. Each fitted reference uses 999 draws with pair seed $20260826 + 101i$, where i is the stored pair ordering. The observed-origin per-series sieve is kept as a diagnostic implementation. The paired initialization audit uses 499 draws per pair, seeds $20260910 + i$, and the same simulated innovations at zero and observed origins.

For the joint fitted reference, a VAR is estimated on bivariate first differences with AIC order selection over zero through ten. Centered residuals are resampled and smoothed. The fitted long-run covariance is computed from the VAR transfer matrix at one. A target covariance retains its diagonal and replaces the correlation by the estimated value clipped in magnitude at 0.99, or by the same-sign endpoint 0.99. Residuals are linearly recolored to implement the target covariance under the fitted dynamics. The recursion uses a burn-in of 100 increments, retains the observed mean increment, and integrates from the observed initial values. Global demeaning is applied to the observed and simulated level series before computing the comparison.

These algorithmic steps define a reproducible reference law; they do not prove its validity. Invertibility of the transfer matrix at one is not a sufficient VAR stability condition. Recoloring can alter marginal higher moments. Retaining an estimated drift can matter at first order for an integrated path. Smoothing does not establish a uniform local limit theorem. The code’s historical name “least_favourable” is retained for traceability, but no stochastic ordering result supports that label here.

The 999-draw results reproduce the archived counts and appear in the next table. All raw independent-reference probabilities flag the 36 equity pairs. Joint and endpoint values in the table use adjusted statistics and are conditional reference probabilities. Classical columns use the same frozen pair samples.

Table 3: All 36 equity pairs: reexecuted reference probabilities.

Pair	Raw Q	Joint	Endpoint	EG	PO
CVX/GDX	7.18e-06	0.900	0.998	0.822	0.861
CVX/GLD	1.35e-06	0.976	1.000	0.869	0.882
CVX/IEF	4.25e-06	0.620	0.780	0.793	0.763
CVX/KO	5.97e-07	0.885	0.999	0.697	0.589
CVX/PEP	1.15e-06	0.270	0.344	0.133	0.148
CVX/SLV	8.33e-06	0.877	0.999	0.784	0.838
CVX/TLT	1.95e-06	0.719	0.981	0.230	0.242
CVX/XOM	2.70e-07	0.876	0.960	0.436	0.294
GDX/GLD	1.39e-05	0.528	0.555	0.117	0.173
GDX/IEF	1.98e-05	0.491	0.523	0.532	0.521
GDX/KO	1.14e-05	0.064	0.239	0.074	0.098
GDX/PEP	1.30e-05	0.803	0.955	0.069	0.080
GDX/SLV	6.06e-08	0.249	0.731	0.069	0.052
GDX/TLT	1.46e-05	0.400	0.284	0.613	0.585
GDX/XOM	5.30e-06	0.478	0.713	0.234	0.236

Pair	Raw Q	Joint	Endpoint	EG	PO
GLD/IEF	1.13e-06	0.649	0.610	0.563	0.544
GLD/KO	1.72e-07	0.419	0.928	0.117	0.055
GLD/PEP	2.45e-07	0.895	0.934	0.069	0.083
GLD/SLV	1.54e-05	0.606	0.648	0.202	0.212
GLD/TLT	6.73e-07	0.580	0.442	0.652	0.622
GLD/XOM	2.50e-06	0.736	0.996	0.305	0.324
IEF/KO	1.90e-06	0.418	0.340	0.945	0.970
IEF/PEP	1.16e-06	0.476	0.730	0.040	0.028
IEF/SLV	2.15e-05	0.339	0.377	0.932	0.958
IEF/TLT	8.82e-07	0.276	0.308	0.985	0.998
IEF/XOM	6.26e-06	0.011	0.023	0.940	0.917
KO/PEP	2.08e-07	0.781	0.785	0.115	0.136
KO/SLV	1.28e-05	0.102	0.311	0.095	0.109
KO/TLT	7.36e-07	0.308	0.209	0.628	0.657
KO/XOM	1.48e-06	0.436	0.962	0.298	0.278
PEP/SLV	1.45e-05	0.708	0.928	0.950	0.975
PEP/TLT	1.65e-07	0.542	0.579	0.666	0.676
PEP/XOM	2.43e-06	0.429	0.472	0.838	0.875
SLV/TLT	1.62e-05	0.265	0.174	0.614	0.592
SLV/XOM	6.27e-06	0.333	0.521	0.218	0.223
TLT/XOM	3.56e-06	0.136	0.307	0.313	0.263

Selected pairs were reevaluated at 10,000 draws per reference, with fixed batches of 500 and independent seeds. The results and exact binomial intervals are in `results/empirical_mc_precision.csv`. They measure conditional Monte Carlo precision and should not be read as confidence intervals for cointegration. The full-panel count remains the 999-draw result; mixing selected higher-precision values into that count would define a different reporting rule.

7.4 Rates and finite-sample approximation

The characterization experiment uses $d \in \{2, 10\}$, innovation correlations 0 and 0.9, and $T \in \{256, 1024, 4096, 16384\}$, with 1,000 independent pairs per cell. The same paths are also evaluated after adding $c_T = T^{3/4}$ to each series. After a visible discrepancy from the limiting mean at $d = 10$, a separately recorded extension uses $T = 65,536$ and $262,144$ for independent walks. No sample was discarded for a nonfinite result.

At $d = 2$, $\kappa_2 = \sqrt{\pi/2}$. For $d = 10$, one million independent Gaussian marks give $\hat{\kappa}_{10} = 3.718057$, with Monte Carlo standard error 0.001568. Integrating (2.6) against two independent half-normal laws gives $\mathbb{E}\Psi = 0.1115793$. Thus the mean of the $d = 10$ limiting entropy law is approximately 0.414858. We compare to a moment of the limit law; weak convergence alone is not a proof of convergence of unbounded moments.

At $T = 16,384$, the mean of $\sqrt{T}Q_T$ is 0.357434 (standard error 0.014666) for $d = 10$ and independent walks. It is 0.389430 (0.014089) at 65,536 and 0.408649 (0.015469) at 262,144. Figure 1

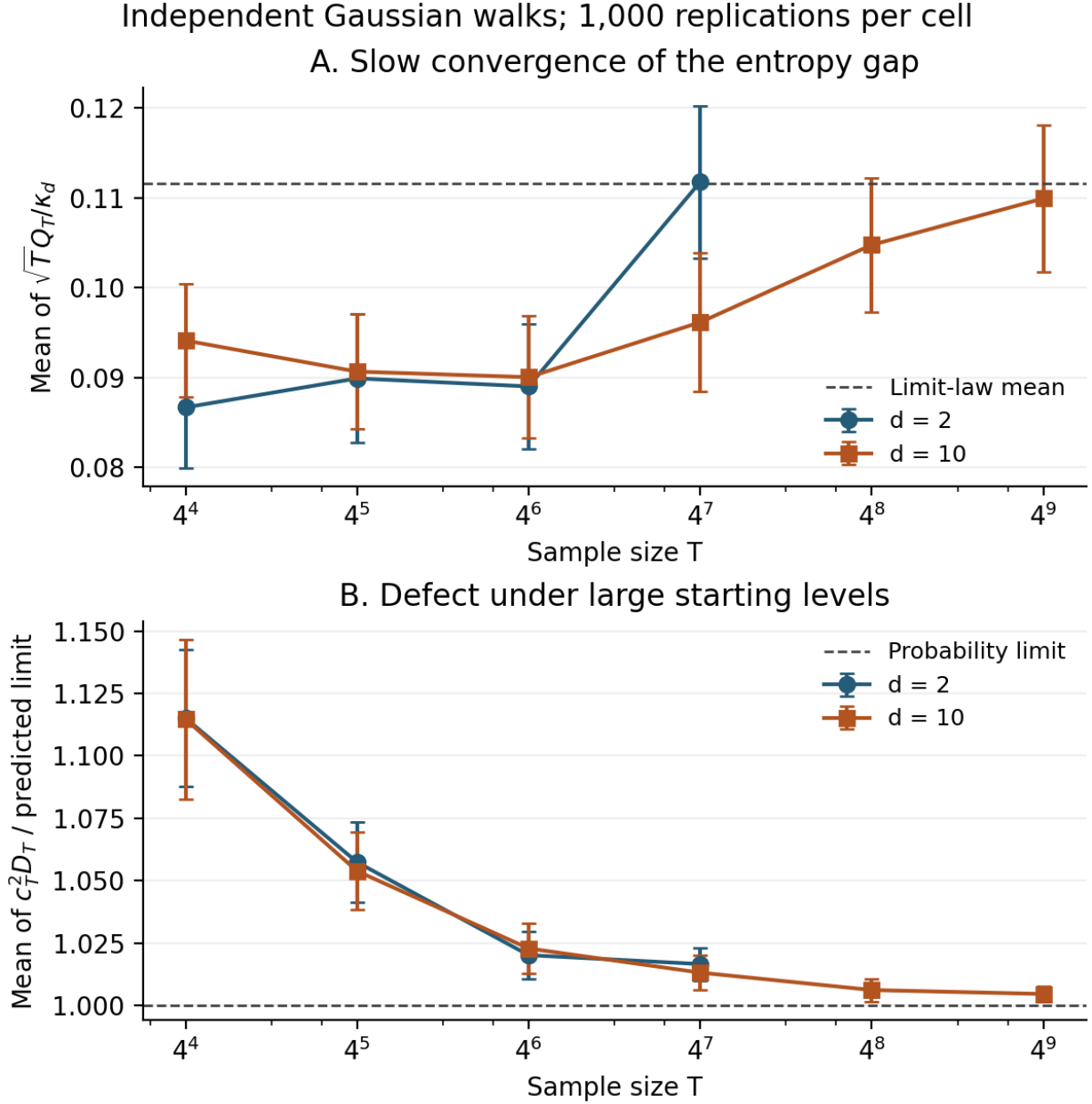


Figure 1: Characterization checks for independent Gaussian walks. Top: the scaled entropy gap. Bottom: the large-level defect divided by its predicted coefficient, with $c_T = T^{3/4}$. Bars are 1.96 Monte Carlo standard errors of simulation means. The estimated $d = 10$ coefficient's additional uncertainty is not included in the bars. The large- T extension was run after inspecting the original grid.

makes the finite-sample discrepancy visible rather than presenting only the largest sample. The first-order characterization should not be read as evidence that its lower quantiles are accurate at financial-panel sample sizes.

For the shifted paths the predicted defect coefficient is $(d^2 - 1)/(6d)$, equal to 1.65 at $d = 10$. The largest-sample estimate is 1.657495 (standard error 0.002715). The mean of $\sqrt{T}Q_T$ falls to approximately 5.27×10^{-9} . These results agree with the direction of Theorem 3, while its proof is independent of the numerical exercise.

7.5 Calibration and empirical reference sensitivity

At $T = 4096$ and innovation correlation 0.99, the raw statistic rejects 41.5% of independent evaluation draws against the independence reference, compared with 5.53% against the known-correlation reference. Demeaning leaves rates of 39.77% and 5.27%. The focused classical comparison uses $T = 300$, 1000 evaluation pairs per cell and separate 5000-draw entropy references. It reports raw rejection frequencies, including null distortions, rather than a size-matched power ranking.

The frozen nine-asset panel has 1275 observations and 36 pairs. The raw independent reference flags all 36; the adjusted joint fitted reference flags IEF/XOM alone. Engle-Granger (Engle and Granger, 1987), Johansen (Johansen, 1988) and Phillips-Ouliaris (Phillips and Ouliaris, 1990) each flag IEF/PEP alone. Holding simulated increments fixed, changing the inherited sieve’s origin changes its count from 36 to one. These are conditional reference outcomes, not cointegration labels. Selected 10000-draw precision checks remain separate from the 999-draw full panel; shared assets also make pair outcomes dependent.

8 Discussion

For fixed-dimensional integrated windows, bounded spatial normalization leads to a singular common limit. The fluctuations about that limit preserve occupation near the origin. The entropy comparison has an exact first-order local-time functional, with a coefficient determined by within-window innovations. This characterization explains both the scale invariance and the remaining dependence on the joint law.

The two calibration obstructions are complementary. Correlation can collapse the gap within a noncointegrated Gaussian family even when levels are fixed. Large initial levels can collapse it for independent trends. Global centering addresses additive levels in the computation, but changes the occupation object and leaves the dependence question intact. Neither boundedness nor normalization alone supplies a pivotal test.

The assumptions set the limits of the results. Dimension is fixed; increments are Gaussian and iid for the stochastic characterization; initial levels are deterministic; and the large-level result concerns compression and entropy collapse, not a universal finer entropy expansion. The paper does not establish serially dependent occupation limits, a conditional sieve-bootstrap theorem, estimated-centering asymptotics, or a stationary-error power theorem.

A useful extension would verify conditional marked occupation approximations for a specified dependent process class, including its centering and initialization. That extension must track both the local increment law and the long-run trend covariance. The present characterization identifies

the terms such a theory must approximate and the restrictions a valid calibration must confront.

Data and code. The accompanying replication package contains the frozen input snapshots, isolated computation scripts, numerical results, parameter and seed records, and a manifest of file hashes. The data statement identifies the inherited price source and the limitations on provenance and redistribution. A full rerun does not require downloading current market data. All empirical probabilities are interpreted under the reference models stated above.

References

- Alexander Dürre and Daniel Vogel. *sscor: Robust Correlation Estimation and Testing Based on Spatial Signs*, 2025. URL <https://cran.r-project.org/package=sscor>. R package version 0.2.1.
- Alexander Dürre, Daniel Vogel, and David E. Tyler. The spatial sign covariance matrix with unknown location. *Journal of Multivariate Analysis*, 130:107–117, 2014. doi: 10.1016/j.jmva.2014.05.004. URL <https://arxiv.org/abs/1307.5706>.
- Alexander Dürre, Daniel Vogel, and Roland Fried. Spatial sign correlation. *Journal of Multivariate Analysis*, 135:89–105, 2015. doi: 10.1016/j.jmva.2014.12.002.
- Alexander Dürre, David E. Tyler, and Daniel Vogel. On the eigenvalues of the spatial sign covariance matrix in more than two dimensions. *Statistics & Probability Letters*, 111:80–85, 2016. doi: 10.1016/j.spl.2016.01.009. URL <https://arxiv.org/abs/1512.02863>.
- Robert F. Engle and Clive W. J. Granger. Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2):251–276, 1987. doi: 10.2307/1913236.
- P. J. Fitzsimmons and Jim Pitman. Kac’s moment formula and the Feynman–Kac formula for additive functionals of a Markov process. *Stochastic Processes and their Applications*, 79:117–134, 1999. doi: 10.1016/S0304-4149(98)00081-7.
- Clive W. J. Granger. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16(1):121–130, 1981. doi: 10.1016/0304-4076(81)90079-8.
- Søren Johansen. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2–3):231–254, 1988. doi: 10.1016/0165-1889(88)90041-3.
- Yuri N. Kartashov and Alexey M. Kulik. Invariance principle for additive functionals of Markov chains, 2007. URL <https://arxiv.org/abs/0704.0508>. arXiv:0704.0508, version 1.
- Peter Mörters and Yuval Peres. *Brownian Motion*. Cambridge University Press, 2010. doi: 10.1017/CBO9780511750489.
- Joon Y. Park and Peter C. B. Phillips. Asymptotics for nonlinear transformations of integrated time series. *Econometric Theory*, 15:269–298, 1999. doi: 10.1017/S0266466699153015.

- Jose A. Perea and John Harer. Sliding windows and persistence: An application of topological methods to signal analysis. *Foundations of Computational Mathematics*, 15:799–838, 2015. doi: 10.1007/s10208-014-9206-z.
- Peter C. B. Phillips. Understanding spurious regressions in econometrics. *Journal of Econometrics*, 33:311–340, 1986. doi: 10.1016/0304-4076(86)90001-1.
- Peter C. B. Phillips and Sam Ouliaris. Asymptotic properties of residual based tests for cointegration. *Econometrica*, 58(1):165–193, 1990. doi: 10.2307/2938339.
- Qiyang Wang and Peter C. B. Phillips. Asymptotic theory for local time density estimation and nonparametric cointegrating regression. Technical Report 1594, Cowles Foundation, Yale University, 2006. URL <https://cowles.yale.edu/sites/default/files/2022-08/d1594.pdf>.
- Qiyang Wang and Peter C. B. Phillips. Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory*, 25(3):710–738, 2009. doi: 10.1017/S0266466608090269.
- John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. doi: 10.1017/9781316848142.