DETECTING LATENT VOLATILITY CONTAGION
A PROJECTED SCORE ESTIMATOR FOR ROUGH VOLATILITY MODELS

J. VIDAL LLAURADÓ

Abstract. Roughness can be estimated from option surfaces or from realized-volatility paths, but directed volatility contagion is a different object. It is the source-screened component of the source-target covariance derivative after the target-only tangent space has been removed. This paper turns that component into a projected covariance-score estimator and applies it to a balanced Oxford-Man realized-volatility panel of eight global equity indices. Pricing visibility, path-space detectability, and observable-history screening remain distinct restrictions on the same local estimand.

The econometric experiment is a reduced local Gaussian block model. Local alternatives enter through covariance derivatives; target-only covariance scores are projected out; the remaining source-incremental derivative is estimated by a low-dimensional projected covariance-score GMM statistic. The score geometry yields local Gaussian efficiency constants, a rough-regime projected rank condition, pilot-adaptive transfer under logarithmic roughness accuracy, and a uniform minimax statement once the primitive Volterra reduction and projected-rank inputs hold uniformly. Synthetic experiments check the implementation against closed-form information and noncentrality constants.

On the Oxford-Man panel the estimated physical-measure roughnesses lie between about 0.04 and 0.09. For SPX the estimator gives $H_{\mathbb{P}} \approx 0.071$, while the external projected option-surface benchmark of Livieri, Mouti, Pallavicini, and Rosenbaum (2018) gives $H_{\mathrm{imp}} \approx 0.30$. The full-sample $\mathbb{P}$-side directed contagion map is dense: all 56 ordered pairs are source-visible at 5% FDR. The empirical objects are therefore intensity ranking and rolling stability, not sparse edge recovery. The Oxford-Man exercise remains a $\mathbb{P}$-side realized-volatility application; matched $\mathbb{P}/\mathbb{Q}$ classification requires a paired option-panel implementation of the same projected information geometry.

1. Introduction

Roughness is visible in more than one experiment. Option surfaces produce implied roughness estimates; high-frequency realized measures produce physical roughness estimates. Directed volatility contagion enters as the source-target covariance direction left after target-only covariance movements have been removed. That remaining direction is the estimand.

The empirical pressure comes from the $\mathbb{P}/\mathbb{Q}$ roughness discrepancy documented by [1] relative to the high-frequency benchmark of [2]. Realized and option-implied roughness estimates need not coincide. The econometric problem therefore separates roughness estimation from pricing visibility, path-space detectability, and source-screened dynamic observability. The Oxford-Man application below estimates the $\mathbb{P}$-side objects on realized-volatility indices and uses the SPX implied roughness only as an external benchmark. Because no matched option panel is present, the application does not assert pairwise $\mathbb{P}/\mathbb{Q}$ labels.

The estimand is the filtered source-target covariance direction. A raw cross-covariance would mix this direction with target-only covariance movements. Smoothing contagion can be invisible to leading option prices and still detectable in the volatility path law [3]. Rough directed contagion can also be lost when latent drivers are reduced to observable filtrations [4]. The source-screened covariance derivative is the residual direction left after that reduction. Projected covariance scores estimate this direction rather than a raw spillover coefficient.

Date: April 2026.

2020 *Mathematics Subject Classification.* 60G22, 62M15, 62F12, 91G80.

Key words and phrases. rough volatility, latent volatility contagion, projected covariance scores, source-screened inference, Gaussian Volterra models, high-frequency realized volatility, Oxford-Man realized library.

DOI: 10.2139/ssrn.6616118.

A scalar residual-product moment can be locally blind even when the reduced Gaussian experiment contains source-screened information. A matched Gaussian oracle can have nonzero projected information while the scalar moment remains nearly uninformative. The estimator therefore uses projected covariance scores. In the local Gaussian block experiment, the local alternative is a covariance derivative, the target-only tangent space is removed, and the retained score basis extracts the source-incremental component. Power loss is then measured by information geometry, not by the behavior of one residual-product statistic.

The distinction from adjacent literatures is at the level of the estimand. Univariate roughness work such as [2, 5] estimates a Hurst exponent from realized or option-implied volatility, often inside rough Heston or affine-Volterra classes [6, 7]. The present estimand is joint and directional. Contagion and connectedness papers such as [8, 9, 10, 11] summarize reduced-form spillovers or connectedness. Here the object is a nuisance-screened covariance derivative in a rough-volatility experiment. Orthogonal-score and semiparametric screening procedures such as [12, 13] remove nuisance bias; the projected covariance-score construction keeps that orthogonality while retaining explicit source-screened Fisher information and Wald noncentralities.

Recent econometric work also shows that roughness inference from realized measures is sensitive to proxy construction and to the interpretation of short-scale dependence [14]. The statistic studied here is therefore a screened directional component of the joint source-target covariance geometry, rather than an isolated univariate roughness estimate.

What the projected estimator targets. The scalar residual-product statistic is retained as a benchmark. It shows how observable screening converts latent source information into a feasible correction. The default empirical estimator is the projected covariance-score statistic, which keeps the same screening discipline and replaces the single moment by the local Gaussian covariance geometry. The full and target-only Gaussian covariance scores serve as diagnostics. They separate weak source-screened signal, target-only contamination, poor basis choice, finite-sample calibration, and the narrowness of the scalar moment class.

The construction starts from the full covariance oracle. It removes the target-only covariance oracle, keeps the source-orthogonal covariance component, and projects that component onto a low-dimensional covariance-score basis. The scalar residual-product statistic survives as a one-dimensional anchor direction inside this hierarchy. The projected information matrix supplies constants as well as rates. In the reduced LAN experiment it gives the noncentrality for each local direction, so the implementation can be checked by comparing observed mean Wald statistics with the theoretical noncentral chi-square mean.

The word latent is used in the dynamic source-screened sense. The estimator is built for the rough directed regime

$$0 < H_{XY} < H_Y < \frac{1}{2}.$$

It does not estimate the smoothing leading-price-invisible interval $H_Y < H_{XY} \le H_Y + \frac{1}{4}$, where the cross-channel lies below the leading short-maturity pricing exponent.

The assumptions sit higher in the regularity hierarchy than threshold asymptotics alone. Threshold and leading-price statements require only the first-order diagonal Volterra structure. The closed source-screening theorem uses C^2 near-diagonal control and endpoint-flat cutoffs. The projected covariance-score CLT inherits that C^2 screening structure and adds the sampling, proxy, and effective-information conditions of the feasible local experiment.

Main results and empirical scope. Four theorem blocks delimit the econometric claim. Theorem 9.2 gives the feasible projected-GMM limit and the closed-form Wald noncentrality in the reduced Gaussian experiment. Theorem 8.11 proves that, in the rough regime $H < 1/2$, the screened Volterra boundary-layer profiles retain full rank after target-only directions are removed. Theorem 8.7 gives the logarithmic condition under which a cross-fitted roughness pilot is oracle-equivalent,

$$|\widehat{H}_{XY} - H_{XY}|\,|\log h_n| \to 0.$$

Theorem 8.8 upgrades the local Gaussian efficiency statement to a uniform rough-class minimax result once the observed-to-Gaussian reduction and projected-rank inputs hold uniformly over the primitive Volterra class. The result is a pilot-adaptive projected score estimator with explicit local information constants and an explicit boundary for rough-class efficiency.

The empirical scope is part of the statement. The Oxford-Man application estimates the $\mathbb{P}$-side screened object and leaves $\mathbb{P}/\mathbb{Q}$ classification outside the realized-volatility evidence. It reports single-index roughness estimates, a full-sample directed contagion map, rolling observability diagnostics, and an external SPX $\mathbb{P}/\mathbb{Q}$ roughness comparison. The matched $\mathbb{P}/\mathbb{Q}$ material is an information interface for paired panels, not an empirical claim made from this realized-volatility application.

Theory and evidence. The mathematical development follows the estimator. Section 2 states the primitive Volterra observation model and the observed-data reduction feeding the econometric experiment. Section 3 constructs the local Gaussian block experiment and its covariance score. Section 4 removes target-only tangent directions and defines the source-orthogonal information bound. Sections 5 and 6 construct the projected GMM statistic and its feasible cross-fitted version. Section 7 proves the scalar projected LAN limit, and Sections 8 and 9 give the rough-regime identification, vector LAN theorem, Wald noncentrality, and $\mathbb{P}/\mathbb{Q}$ information interface. Section 10 records the noise-compatible and regularized observed-data extensions. Sections 11 and 12 show how the theory is implemented in synthetic experiments and in the Oxford-Man panel.

The assumptions enter at different levels. Local validity uses a projected-score mean expansion, HAC central limit theory, and feasible replacement. Source-screened efficiency uses Gaussian nuisance projection against the target-only tangent. Rough identification uses the boundary-layer profile Gram lower bound and dictionary approximation. Minimax efficiency uses uniform rough LAN, uniform rank, and feasible rough-local scaling, supplied either stratumwise or by a pilot roughness estimator. The $\mathbb{P}/\mathbb{Q}$ comparison uses consistent estimation of the projected information objects and separation from classification boundaries. The paper therefore proves the finite projected-score statements directly and states the primitive Volterra inputs needed for their uniform rough-class form.

2. Primitive Volterra and observed-data reduction

The estimator is defined on a reduced Gaussian block experiment. The primitive observation model is higher-frequency: log-volatilities are generated by a rough Gaussian Volterra system and observed through realized-volatility proxies. This section records the inputs inherited from [3, 4] and fixes the map from block proxies to the vector used by the projected score.

Assumption 2.1 (Primitive rough Volterra model)**.** Fix $T > 0$. The target and source log-volatilities are generated by a centered Gaussian Volterra system with kernels

$$K_i(t,s) = L_i(t,s)(t-s)^{H_i-\frac{1}{2}}, \qquad 0 \le s < t \le T, \qquad i \in \{Y, XY, X\}.$$

The rough exponents satisfy

$$0 < H_{XY} < H_Y < \frac{1}{2}, \qquad H_X \in (0, \tfrac{1}{2}),$$

and the amplitudes L_i are C^2 near the diagonal. Moreover

$$L_Y(t,t) \neq 0, \qquad L_{XY}(t,t) \neq 0, \qquad 0 < t < T,$$

and the source-side local covariance operator is nondegenerate on every nontrivial time interval.

The observed log-prices are

$$dP_t^Y = \sigma_t^Y \; dB_t^Y, \qquad dP_t^X = \sigma_t^X \; dB_t^X, \qquad \log \sigma_t^a = V_t^a, \quad a \in \{X, Y\}.$$

The baseline theory excludes additive microstructure noise. The noisy extension is handled by replacing the block realized-variance proxy below with a pre-averaged version. That extension sits outside the core projected-score construction. The measurement convention follows the realized-volatility literature in which high-frequency realized variance serves as a proxy for integrated variance under infill sampling [15, 16]. When market microstructure noise or jumps are non-negligible, the same block construction can be paired with realized kernels, pre-averaging, or bipower-type proxies [17, 18, 19].

Assumption 2.2 (Sampling, memory, and proxy scale)**.** Prices are observed on the regular grid

$$\{j\Delta_n : 0 \le j \le \lfloor T/\Delta_n \rfloor\}, \qquad \Delta_n \downarrow 0.$$

Let $h_n \downarrow 0$ be the block horizon and set

$$m_n := h_n/\Delta_n \in \mathbb{N}, \qquad m_n \to \infty.$$

For a fixed $\alpha \in (0,1)$, define the boundary-layer memory window

$$r_n := h_n^{\alpha}, \qquad J_n := \lfloor r_n/h_n \rfloor, \qquad J_n \to \infty, \qquad J_n h_n \to 0.$$

The effective number of evaluation blocks is $N_n \to \infty$. The disjoint asymptotic construction uses blocks separated by their memory windows, while the empirical implementation may evaluate on a denser overlapping grid and normalize by a HAC long-run covariance.

Definition 2.3 (Canonical realized-volatility block vector)**.** Let

$$I_{k,n} := [(k-1)h_n, kh_n), \qquad 1 \le k \le N_n + J_n + 1.$$

For $a \in \{X, Y\}$, define the block realized variance

$$\mathrm{RV}^a_{k,n} := \sum_{i:(i-1)\Delta_n,\ i\Delta_n \in I_{k,n}} \left(\Delta^n_i P^a\right)^2, \qquad \Delta^n_i P^a := P^a_{i\Delta_n} - P^a_{(i-1)\Delta_n},$$

and the block log-variance proxy

$$\widehat{V}^a_{k,n} := \frac{1}{2}\{\log \mathrm{RV}^a_{k,n} - \log h_n\}.$$

After discarding the first J_n blocks and the last future block, define

$$\widehat{Y}^{\mathrm{fut}}_{k,n} := h_n^{-H_Y}\big(\widehat{V}^Y_{k+J_n+1,n} - \widehat{V}^Y_{k+J_n,n}\big),$$
$$\widehat{Y}^{\mathrm{past}}_{k,j,n} := h_n^{-H_Y}\big(\widehat{V}^Y_{k+J_n+1-j,n} - \widehat{V}^Y_{k+J_n-j,n}\big), \qquad 1 \le j \le J_n,$$

and

$$\widehat{X}^{\mathrm{past}}_{k,j,n} := \widehat{V}^X_{k+J_n+1-j,n} - \widehat{V}^X_{k+J_n-j,n}, \qquad 1 \le j \le J_n.$$

The reduced block vector used in Section 3 is

$$Z_{k,n} = \big(\widehat{Y}^{\mathrm{fut}}_{k,n}, \widehat{\mathbf{Y}}^{\mathrm{past}}_{k,n}, \widehat{\mathbf{X}}^{\mathrm{past}}_{k,n}\big),$$

up to the optional replacement of the source and target lag vectors by lower-dimensional dictionaries.

Assumption 2.4 (Proxy dominance and local alternative scale)**.** The mesh and memory window are chosen so that realized-volatility proxy errors are negligible after the rough target normalization and after aggregation over the J_n-dimensional past window. A convenient sufficient set of conditions is

$$\sqrt{J_n}\, m_n^{-1/2} = o(h_n^{H_Y}), \qquad \sqrt{J_n}\, h_n^{H_Y} \to 0, \qquad \sqrt{J_n}\, h_n^{2H_X} \to 0, \qquad 2H_X > H_{XY}.$$

The local contagion alternatives are indexed by a bounded amplitude τ through

$$K^{(n,\tau)}_{XY}(t,s) = \theta_n \tau\, b(t,s)(t-s)^{H_{XY}-\frac{1}{2}} \mathbf{1}_{\{s<t\}}, \qquad \theta_n := \frac{1}{\sqrt{N_n}\, h_n^{H_{XY}}},$$

where b is a bounded C^2 profile with $b(t,t) \ne 0$. The sequence of laws is assumed contiguous around $\tau = 0$.

The first reduction identifies the experiment used below. After aggregation, feasible block proxies and latent Gaussian blocks have the same retained projected-score law.

Theorem 2.5 (Observed-to-Gaussian block reduction)**.** *Under Assumptions 2.1–2.4, the feasible block vectors from Definition 2.3 admit a reduced Gaussian block representation. Equivalently, there exist centered Gaussian latent block vectors*

$$Z_{k,n}^{\mathrm{lat}} = (Y_{k,n}^{\mathrm{lat}}, \mathbf{Y}_{k,n}^{\mathrm{lat}}, \mathbf{X}_{k,n}^{\mathrm{lat}})$$

such that

$$\frac{1}{N_n} \sum_{k=1}^{N_n} \mathbb{E}[\|Z_{k,n} - Z_{k,n}^{\mathrm{lat}}\|^2] \to 0.$$

Moreover, for every fixed finite collection of projected covariance-score directions whose whitening and target projections are stable under the null covariance, any aggregated score replacement satisfying Proposition B.4 can be transferred from the feasible blocks to the latent Gaussian blocks after $\sqrt{N_n}$-aggregation.

Proof. The block log-realized-variance expansion in Lemma B.1 replaces each feasible volatility proxy by its latent block-average counterpart with averaged L^2 error negligible under the rates in Assumption 2.4. The boundary-layer covariance asymptotics in Lemma B.2 identify the deterministic frozen Gaussian covariance approximations for the latent block vector on the shrinking memory window. Combining these two steps gives a centered Gaussian latent block vector $Z_{k,n}^{\mathrm{lat}}$ whose covariance matches the Volterra block covariance up to an operator-norm error that is negligible on the retained coordinates, and whose averaged distance to the feasible block vector satisfies the displayed L^2 bound.

The displayed block-level L^2 reduction is the primitive input used later for score-based statistics. The retained quadratic scores are continuous on the stabilized finite-dimensional whitened coordinate space, so the block replacement reduces to the aggregated mean, L^2, and dependence criteria in Proposition B.4. Once those criteria are verified for the retained score basis, the feasible and latent Gaussian score sums have the same $\sqrt{N_n}$-limit. ☐

Remark 2.6 (Primitive reduction input). Theorem 2.5 supplies the observed-data reduction inherited from the earlier Volterra proofs. It combines the block log-realized-variance expansion, boundary-layer covariance asymptotics, and proxy-noise replacement. The projected estimator starts after this reduction. The scalar residual-product score is replaced by projected covariance scores on the reduced block covariance geometry.

The discrete transfer reduction identifies the regression coefficient through which the Volterra source channel enters the covariance derivative. It also separates that coefficient from the target-only correction.

Theorem 2.7 (Discrete Volterra transfer representation)**.** *Under the primitive Volterra model, the observed-to-Gaussian reduction, and the local alternatives from Assumption 2.4, assume in addition that the null past covariance blocks*

$$C_{PP,n} := \mathrm{Cov}_0(P_{k,n}^{\mathrm{lat}})$$

are uniformly coercive on the growing discrete past coordinates,

$$\sup_n \|C_{PP,n}^{-1}\|_{\mathrm{op}} < \infty.$$

Then the latent target future block admits the local regression expansion

$$Y_{k,n}^{\mathrm{lat}} = \langle \beta_n^Y, \mathbf{Y}_{k,n}^{\mathrm{lat}} \rangle + \frac{\tau}{\sqrt{N_n}} \left\{ \langle \delta_n^Y, \mathbf{Y}_{k,n}^{\mathrm{lat}} \rangle + \langle \alpha_n^{XY}, \mathbf{X}_{k,n}^{\mathrm{lat}} \rangle \right\} + \zeta_{k,n} + r_{k,n}^{\mathrm{tr}},$$

where $\zeta_{k,n}$ is a centered Gaussian innovation orthogonal to $(\mathbf{Y}_{k,n}^{\mathrm{lat}}, \mathbf{X}_{k,n}^{\mathrm{lat}})$,

$$\mathrm{Var}(\zeta_{k,n}) \to \sigma_{\zeta,t}^2 \in (0, \infty), \qquad \frac{1}{N_n} \sum_{k=1}^{N_n} \mathbb{E}[(r_{k,n}^{\mathrm{tr}})^2] \to 0.$$

After boundary-layer rescaling, the coefficient profiles

$$u \mapsto J_n \alpha_{n,\lceil uJ_n \rceil}^{XY}, \qquad u \mapsto J_n \delta_{n,\lceil uJ_n \rceil}^{Y}, \qquad 0 < u \le 1,$$

converge locally in $L^2((0,1])$ to the frozen source-transfer profile and the frozen observable-screening correction profile from the continuous-time theory.

Proof. Let

$$P^{\mathrm{lat}}_{k,n} = (\mathbf{Y}^{\mathrm{lat}}_{k,n}, \mathbf{X}^{\mathrm{lat}}_{k,n})$$

denote the latent past vector. Since the reduced block vector is Gaussian, the regression decomposition is exact:

$$Y^{\mathrm{lat}}_{k,n} = b_n(\tau)^\top P^{\mathrm{lat}}_{k,n} + \zeta_{k,n}(\tau), \qquad b_n(\tau) = \mathrm{Cov}_\tau(Y^{\mathrm{lat}}_{k,n}, P^{\mathrm{lat}}_{k,n})\,\mathrm{Cov}_\tau(P^{\mathrm{lat}}_{k,n})^{-1}.$$

The Schur complement gives $\mathrm{Var}\{\zeta_{k,n}(0)\} \to \sigma^2_{\zeta,t} \in (0,\infty)$, and $\zeta_{k,n}(0)$ is orthogonal to $P^{\mathrm{lat}}_{k,n}$. In the expansion stated in the theorem, $\zeta_{k,n}$ is set equal to $\zeta_{k,n}(0)$. Uniform coercivity of $C_{PP,n}$ gives local Lipschitz control of the Schur complement and of the regression coefficient in the covariance blocks, so replacing $\zeta_{k,n}(\tau)$ by the null innovation contributes only to the averaged L^2 remainder. The local covariance expansion from the Volterra perturbation has the form

$$\Sigma_{n,\tau} = \Sigma_{n,0} + \frac{\tau}{\sqrt{N_n}}\dot{\Sigma}_n + o(N_n^{-1/2})$$

in the block covariance norm used by the regression map. Differentiating $C_{YP}C_{PP}^{-1}$ at $\tau = 0$ gives

$$b_n(\tau) = b_n(0) + \frac{\tau}{\sqrt{N_n}}\dot{b}_n + o(N_n^{-1/2}), \qquad \dot{b}_n = (\dot{C}_{YP} - b_n(0)\dot{C}_{PP})C_{PP}^{-1}.$$

The target coordinates of $b_n(0)$ are β^Y_n. Splitting $\dot{b}_n$ into its target and source coordinates gives $(\delta^Y_n, \alpha^{XY}_n)$. Since $\theta_n h_n^{H_{XY}} = N_n^{-1/2}$, this derivative enters at the LAN scale. The covariance remainder and the replacement of $\zeta_{k,n}(\tau)$ by the null innovation produce an averaged L^2 regression remainder

$$\frac{1}{N_n}\sum_{k=1}^{N_n} \mathbb{E}[(r^{\mathrm{tr}}_{k,n})^2] \to 0.$$

It remains to identify the limiting coefficient profiles. On the shrinking memory window $[t - r_n, t]$, Lemma B.2 turns the discrete covariance blocks and their local derivatives into Riemann-sum approximations of the frozen Volterra covariance operators. Write $\widetilde{C}_{PP,n}$ and $\widetilde{C}_{YP,n}$ for the corresponding frozen past covariance and future-past covariance blocks. Then

$$\|C_{PP,n} - \widetilde{C}_{PP,n}\|_{\mathrm{op}} \to 0, \qquad \|C_{YP,n} - \widetilde{C}_{YP,n}\|_{\ell^2} \to 0,$$

and the same holds for the covariance derivatives entering $\dot{b}_n$. The uniform coercivity assumption gives $\|C^{-1}_{PP,n}\|_{\mathrm{op}} = O(1)$. Hence, once $\|C_{PP,n} - \widetilde{C}_{PP,n}\|_{\mathrm{op}}$ is small enough, the frozen matrices are also uniformly invertible, because

$$\widetilde{C}_{PP,n} = C_{PP,n}\{I + C^{-1}_{PP,n}(\widetilde{C}_{PP,n} - C_{PP,n})\}$$

and the bracketed perturbation is invertible by a Neumann-series argument. The resolvent identity

$$C^{-1}_{PP,n} - \widetilde{C}^{-1}_{PP,n} = C^{-1}_{PP,n}(\widetilde{C}_{PP,n} - C_{PP,n})\widetilde{C}^{-1}_{PP,n}$$

then yields

$$\|C^{-1}_{PP,n} - \widetilde{C}^{-1}_{PP,n}\|_{\mathrm{op}} \to 0.$$

Combining this with the covariance-block approximation gives

$$\|b_n(0) - \widetilde{b}_n(0)\|_{\ell^2} \to 0, \qquad \|\dot{b}_n - \widetilde{\dot{b}}_n\|_{\ell^2} \to 0,$$

where $\widetilde{b}_n$ and $\widetilde{\dot{b}}_n$ are the frozen regression coefficients. The growing-window discrete regression coefficients are therefore controlled by the frozen boundary-layer regression problem. Away from the singular endpoint this gives ordinary operator convergence of the regression solutions. The endpoint contribution is controlled by the uniform L^2 envelope of the rough kernels $(t-s)^{H-1/2}$. The rescaled source and target derivative coefficients consequently converge locally in $L^2((0,1])$ to the frozen source-transfer and observable-screening correction profiles. ∎

Remark 2.8 (Where the covariance derivative comes from). The covariance derivative $\dot{\Sigma}_n$ used below is the covariance-level version of Theorem 2.7. The source profile α_n^{XY} generates future-source and target-source covariance directions, while δ_n^Y describes the target-only correction that must be projected away. The projected-score construction therefore reads the same discrete Volterra transfer map through covariance scores rather than through a single residual-product moment.

Assumption 2.9 (One-sided target-only observation channel)**.** Let $P_{n,\tau}^{Y,\mathrm{loc}}$ denote the localized reduced Gaussian target-only experiment and let $\mathbb{P}_{n,\tau}^{Y,\mathrm{only}}$ denote the feasible target-only proxy experiment generated by the observed target block proxies. For every compact local-amplitude set Θ_{loc}, there exists a Markov kernel $K_n^{Y,\mathrm{only}}$, not depending on τ, such that

$$\sup_{\tau\in\Theta_{\mathrm{loc}}} \|K_n^{Y,\mathrm{only}} P_{n,\tau}^{Y,\mathrm{loc}} - \mathbb{P}_{n,\tau}^{Y,\mathrm{only}}\|_{\mathrm{TV}} \to 0.$$

This is a one-sided transfer from the reduced Gaussian target-only experiment to the feasible target-only observation experiment. It is weaker than full two-sided asymptotic equivalence and is exactly what is needed to lift target-only impossibility statements to realized-volatility proxies.

3. Local Gaussian block experiment

3.1. **Observable block vector.**

The preceding reduction turns realized-volatility proxies into a reduced Gaussian block experiment. On each effective block define

$$Z_{k,n} = \begin{pmatrix} F_{k,n}^Y \\ P_{k,n}^Y \\ P_{k,n}^X \end{pmatrix} \in \mathbb{R}^{p_n}, \qquad p_n = 1 + J_n^Y + J_n^X.$$

Here $F_{k,n}^Y$ denotes the target future proxy, $P_{k,n}^Y \in \mathbb{R}^{J_n^Y}$ the target past proxy vector, and $P_{k,n}^X \in \mathbb{R}^{J_n^X}$ the source past proxy vector. In the simplest implementation $J_n^Y = J_n^X = J_n$, but the notation allows separate target and source dictionaries.

The coordinate split is

$$Z_{k,n} = (Z_{k,n}^{Y,\mathrm{only}}, Z_{k,n}^X), \qquad Z_{k,n}^{Y,\mathrm{only}} = (F_{k,n}^Y, P_{k,n}^Y).$$

The implementation uses this conservative target-only convention. Every covariance score generated by the target future and target past is removed before the source-screened statistic is formed. A less conservative theoretical variant could remove only the target past. The enlarged convention gives a sharper empirical separation.

Assumption 3.1 (Local covariance expansion)**.** Under the local alternatives indexed by a bounded amplitude τ, the reduced block vector is approximately centered Gaussian with covariance

$$\Sigma_n(\tau) = \Sigma_{0,n} + \frac{\tau}{\sqrt{N_n}}\dot{\Sigma}_n + o(N_n^{-1/2}),$$

uniformly for τ in compact sets. The null covariance $\Sigma_{0,n}$ is uniformly positive definite on the retained block coordinates, and the covariance derivative $\dot{\Sigma}_n$ is symmetric.

The covariance derivative $\dot{\Sigma}_n$ is the reduced object produced by the local contagion perturbation. It may move target autocovariances, source-target cross-covariances, and future-source covariances at the same local scale. The scalar corrected statistic uses only one component of this derivative.

3.2. **Gaussian covariance score.**

Let

$$W_n = \Sigma_{0,n}^{-1/2}, \qquad U_{k,n} = W_n Z_{k,n}, \qquad A_n = W_n \dot{\Sigma}_n W_n.$$

For any symmetric matrix $B \in \mathbb{R}^{p_n\times p_n}$, define the centered quadratic score

$$\ell_{k,n}(B) = \frac{1}{2}\left(U_{k,n}^\top B U_{k,n} - \mathrm{tr}(B)\right).$$

If the blocks were independent and exactly Gaussian, the local log likelihood in the full reduced block experiment would have the LAN form

$$\log \frac{d\mathbb{P}_{n,\tau}}{d\mathbb{P}_{n,0}} = \frac{\tau}{\sqrt{N_n}} \sum_{k=1}^{N_n} \ell_{k,n}(A_n) - \frac{\tau^2}{4} \|A_n\|_F^2 + o_{\mathbb{P}_{n,0}}(1),$$

so that the one-block Fisher information is

$$\mathcal{I}_n^{\text{full}} = \frac{1}{2} \|A_n\|_F^2.$$

With overlapping blocks the same expression remains the population score, but the variance of its sample sum is a long-run variance rather than the independent-block variance.

Definition 3.2 (Full covariance-score oracle)**.** The full reduced-Gaussian oracle statistic is the normalized score in the direction A_n:

$$T_n^{\text{full}} = (\mathcal{I}_n^{\text{full}})^{-1/2} \frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \ell_{k,n}(A_n)$$

in the independent-block normalization, with the corresponding HAC normalization on an overlapping grid. This statistic is an upper-bound diagnostic for the reduced Gaussian experiment; feasible screening uses the estimator defined below.

The score algebra fixes the Hilbert geometry for the covariance-score projections.

Lemma 3.3 (Gaussian covariance-score algebra)**.** *Let $U \sim N(0, I_p)$, and for symmetric matrices $B, C \in \mathbb{R}^{p\times p}$ define*

$$\ell(B) = \frac{1}{2}\{U^\top BU - \operatorname{tr}(B)\}.$$

Then

$$\mathbb{E}_0[\ell(B)] = 0, \qquad \operatorname{Cov}_0(\ell(B), \ell(C)) = \frac{1}{2} \operatorname{tr}(BC).$$

If the covariance of U is locally perturbed from I_p to $I_p + \tau A/\sqrt{N}$, then

$$\frac{\partial}{\partial\tau} \mathbb{E}_\tau\left[\sqrt{N}\, \ell(B)\right]\Big|_{\tau=0} = \frac{1}{2} \operatorname{tr}(BA).$$

Consequently the covariance-score Fisher information in direction A is $\frac{1}{2}\|A\|_F^2$, and the information captured by any linear score space is the squared Hilbert projection of A onto that space under the inner product $\langle B, C\rangle_{\text{score}} = \frac{1}{2} \operatorname{tr}(BC)$.

Proof. The first identity follows from $\mathbb{E}[U^\top BU] = \operatorname{tr}(B)$. The covariance identity is Isserlis' formula: $\operatorname{Cov}(U^\top BU, U^\top CU) = 2\operatorname{tr}(BC)$, multiplied by the two factors $1/2$. Under the local covariance perturbation, $\mathbb{E}_\tau[U^\top BU] = \operatorname{tr}(B) + \tau N^{-1/2} \operatorname{tr}(BA) + o(N^{-1/2})$, which gives the loading formula. The final statement is the finite-dimensional Hilbert-space projection identity for Gaussian LAN scores. □

4. Removing target-only information

4.1. **Target-only tangent space.** Let $S_{Y,n}$ select the target coordinates of $Z_{k,n}$. Depending on the result being proved, this target vector can be either $P_{k,n}^Y$ alone or the enlarged vector $(F_{k,n}^Y, P_{k,n}^Y)$. The conservative implementation uses the enlarged target vector, because it removes every score that could be constructed from the target series alone. Denote this conservative target coordinate set by $\mathcal{Y}_n$.

Definition 4.1 (Target-only covariance tangent)**.** The whitened target-only tangent space is

$$\mathcal{T}_{Y,n} = \left\{ W_n S_{Y,n}^\top C S_{Y,n} W_n : C = C^\top \right\}.$$

Let $\Pi_{Y,n}$ denote Frobenius projection onto $\mathcal{T}_{Y,n}$.

The target-only projection is the covariance-score analogue of residualizing the target future on the target past. It prevents the statistic from drawing power from information that the screening lower bound assigns to target-only procedures.

The lower-bound input identifies the component excluded from a source-screened statistic.

Theorem 4.2 (Target-only screening lower-bound input). *Let $\mathcal{E}_n^{Y,\mathrm{only}}$ denote the experiment generated by the feasible target-only block proxies, or by the corresponding localized reduced Gaussian target-only experiment. Under the primitive Volterra assumptions, the one-sided target-only observation channel, and the rougher source-screening regime $H_{XY} < H_Y < 1/2$, any sequence of target-only tests*

$$\phi_n = \phi_n(\mathcal{E}_n^{Y,\mathrm{only}})$$

is asymptotically powerless at the local scale

$$\theta_n = (\sqrt{N_n} h_n^{H_{XY}})^{-1}.$$

Equivalently, for every compact local-amplitude set Θ_{loc},

$$\limsup_{n\to\infty} \sup_{\tau\in\Theta_{\mathrm{loc}}} \left|\mathbb{E}_{n,\tau}[\phi_n] - \mathbb{E}_{n,0}[\phi_n]\right| = 0.$$

Consequently no target-only estimator or predictor can recover the rough contagion perturbation at the oracle local scale with uniformly nontrivial asymptotic accuracy.

Proof. In the localized reduced Gaussian target-only experiment, the log-likelihood ratio has a LAN expansion

$$\log \frac{dP_{n,\tau}^{Y,\mathrm{loc}}}{dP_{n,0}^{Y,\mathrm{loc}}} = \tau \Delta_n^{Y,\mathrm{only}} - \frac{1}{2}\tau^2 \mathcal{I}_n^{Y,\mathrm{only}} + o_{\mathbb{P}_{n,0}}(1),$$

where $\Delta_n^{Y,\mathrm{only}}$ is the target-only covariance score and $\mathrm{Var}_{n,0}(\Delta_n^{Y,\mathrm{only}}) = \mathcal{I}_n^{Y,\mathrm{only}}$. The observable-screening input from the rough Volterra analysis gives $\mathcal{I}_n^{Y,\mathrm{only}} \to 0$ at the source oracle scale. Hence $\Delta_n^{Y,\mathrm{only}} \xrightarrow{\mathbb{P}} 0$. In the Gaussian LAN shift, the likelihood ratio $L_{n,\tau}^{Y,\mathrm{loc}}$ also satisfies the sharper second-moment bound

$$\mathbb{E}_{n,0}\left[\left(L_{n,\tau}^{Y,\mathrm{loc}} - 1\right)^2\right] = \exp\{\tau^2 \mathcal{I}_n^{Y,\mathrm{only}}\} - 1 + o(1) \to 0,$$

uniformly for τ in compact sets. Therefore

$$\|P_{n,\tau}^{Y,\mathrm{loc}} - P_{n,0}^{Y,\mathrm{loc}}\|_{\mathrm{TV}} \to 0, \qquad \sup_{\tau\in\Theta_{\mathrm{loc}}} |\mathbb{E}_{n,\tau}\phi_n - \mathbb{E}_{n,0}\phi_n| \le \sup_{\tau\in\Theta_{\mathrm{loc}}} \|P_{n,\tau}^{Y,\mathrm{loc}} - P_{n,0}^{Y,\mathrm{loc}}\|_{\mathrm{TV}},$$

so no target-only test can separate $\tau \neq 0$ from $\tau = 0$ at this scale.

For feasible realized-volatility proxies, Assumption 2.9 transfers the reduced target-only experiment to the target-only proxy experiment with asymptotically vanishing one-sided deficiency. Given any feasible target-only test ϕ_n, define the randomized reduced-experiment test

$$\widetilde{\phi}_n(y) := \int \phi_n(x) K_n^{Y,\mathrm{only}}(y, dx).$$

The deficiency bound gives, uniformly over compact local-amplitude sets,

$$\left|\mathbb{E}_{\mathbb{P}_{n,\tau}^{Y,\mathrm{only}}}\phi_n - \mathbb{E}_{P_{n,\tau}^{Y,\mathrm{loc}}}\widetilde{\phi}_n\right| \to 0.$$

The reduced lower bound therefore transfers to the observed target-only experiment.

The estimator statement is the same separation argument applied to decision rules. If a target-only estimator were uniformly accurate at scale θ_n, then, for any fixed $\tau_1 \neq 0$, the rule that rejects when the estimated local coordinate is closer to τ_1 than to 0 would have asymptotic size zero and asymptotic power one. That rule is a target-only test at the same local scale, contradicting the total-variation collapse. $\square$

Remark 4.3 (Why the projection targets source information). Theorem 4.2 is not a simulation convention. After the observable-screening collapse of [4], the target-only likelihood has vanishing local Fisher information at the source oracle scale. The projected covariance-score estimator therefore removes the whole target-only covariance tangent space before measuring source information. The quotient is conservative. Any remaining information is source-incremental rather than a repackaged target-only signal.

Definition 4.4 (Screening-orthogonal covariance derivative)**.** The source-incremental covariance score direction is

$$A_n^{\perp} = A_n - \Pi_{Y,n} A_n.$$

The corresponding population information is

$$\mathcal{I}_n^{\mathrm{src}\perp Y} = \frac{1}{2} \|A_n^{\perp}\|_F^2.$$

Definition 4.5 (Target-only and source-orthogonal covariance oracles)**.** Whenever the corresponding information is positive, define

$$T_n^{Y,\mathrm{only}} = (\mathcal{I}_n^Y)^{-1/2} \frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \ell_{k,n}(\Pi_{Y,n} A_n), \qquad \mathcal{I}_n^Y = \frac{1}{2} \|\Pi_{Y,n} A_n\|_F^2,$$

and

$$T_n^{\mathrm{src}\perp Y} = (\mathcal{I}_n^{\mathrm{src}\perp Y})^{-1/2} \frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \ell_{k,n}(A_n^{\perp}).$$

The first statistic is the strongest covariance-score diagnostic available from the conservative target-only reduced experiment. The second is the infeasible oracle score for the source-incremental part of the local alternative. The feasible projected GMM estimator is judged against $T_n^{\mathrm{src}\perp Y}$, not against the full oracle.

The score Hilbert space separates full, target-only, and source-orthogonal information.

Proposition 4.6 (Orthogonal information decomposition)**.** *In the whitened Gaussian block experiment,*

$$\mathcal{I}_n^{\mathrm{full}} = \mathcal{I}_n^Y + \mathcal{I}_n^{\mathrm{src}\perp Y}.$$

Moreover $T_n^{Y,\mathrm{only}}$ is the most powerful one-dimensional covariance-score direction inside the target-only tangent space, while $T_n^{\mathrm{src}\perp Y}$ is the most powerful one-dimensional covariance-score direction among directions orthogonal to the target-only tangent space. In particular, if $A_n^{\perp} = 0$, no covariance-score test that respects target-only orthogonality has nonzero local information.

Proof. The projection $\Pi_{Y,n}$ is the Frobenius orthogonal projection onto $\mathcal{T}_{Y,n}$. Therefore

$$A_n = \Pi_{Y,n} A_n + (I - \Pi_{Y,n}) A_n$$

is an orthogonal decomposition in the score inner product $\langle B, C\rangle_{\mathrm{score}} = \frac{1}{2}\operatorname{tr}(BC)$. Taking squared norms gives

$$\frac{1}{2}\|A_n\|_F^2 = \frac{1}{2}\|\Pi_{Y,n} A_n\|_F^2 + \frac{1}{2}\|(I - \Pi_{Y,n}) A_n\|_F^2.$$

For any candidate direction B with score norm one, Lemma 3.3 gives local mean shift $\langle B, A_n\rangle_{\mathrm{score}}$. Cauchy–Schwarz is maximized by the normalized projection of A_n onto the specified subspace. Choosing the target-only subspace gives the first oracle, and choosing its orthogonal complement gives the source-orthogonal oracle. If $A_n^{\perp} = 0$, every target-orthogonal B has zero inner product with A_n, so the local information is zero. □

Target-only projection is the efficient-score residual in the Gaussian covariance experiment.

Theorem 4.7 (Nuisance-efficient source-screened covariance score)**.** *Consider the independent Gaussian block LAN experiment in which the source-contagion local coordinate has whitened covariance derivative A_n, while target-only perturbations are treated as nuisance directions in the tangent space $\mathcal{T}_{Y,n}$. Equivalently, local alternatives have whitened covariance derivative*

$$\tau A_n + C_n, \qquad C_n \in \mathcal{T}_{Y,n}.$$

Then the nuisance-efficient covariance score for τ is the score in the orthogonal residual direction

$$A_n^{\perp} = (I - \Pi_{Y,n}) A_n,$$

and the efficient source-screened information is

$$\mathcal{I}_{n,\mathrm{eff}}^{\tau|Y} = \left\|A_n^{\perp}\right\|_{\mathrm{score}}^2 = \frac{1}{2}\|A_n^{\perp}\|_F^2 = \mathcal{I}_n^{\mathrm{src}\perp Y}.$$

For an m-dimensional local coordinate with derivatives $A_{n,1},\dots,A_{n,m}$, the nuisance-efficient information matrix is

$$\mathcal{I}_{n,\mathrm{eff}}^{\vartheta|Y} = \left(\langle A_{n,i}^{\perp}, A_{n,j}^{\perp}\rangle_{\mathrm{score}}\right)_{i,j}, \qquad A_{n,j}^{\perp} = (I - \Pi_{Y,n})A_{n,j}.$$

Thus target-only projection is more than a conservative diagnostic convention. In the Gaussian covariance LAN experiment it is exactly the efficient-score projection that removes the target-only nuisance tangent.

Proof. The Gaussian covariance scores form a Hilbert space with inner product $\langle B, C\rangle_{\mathrm{score}} = \frac{1}{2}\operatorname{tr}(BC)$. In a LAN experiment with parameter direction A_n and nuisance tangent $\mathcal{T}_{Y,n}$, the efficient score is the component of the full score orthogonal to the nuisance tangent. Orthogonal projection in this Hilbert space therefore gives

$$A_n - \Pi_{Y,n}A_n = A_n^{\perp}.$$

The variance of the efficient score, equivalently the squared norm of the residual derivative, is $\|A_n^{\perp}\|_{\mathrm{score}}^2 = \frac{1}{2}\|A_n^{\perp}\|_F^2$, proving the scalar information formula. For a vector parameter, apply the same projection to each derivative column. The covariance matrix of the resulting efficient score vector is the Gram matrix of the residual derivatives, which gives the displayed efficient information matrix. □

Remark 4.8 (Reading the information decomposition). The information split

$$\mathcal{I}_n^{\mathrm{full}} = \frac{1}{2}\|\Pi_{Y,n}A_n\|_F^2 + \frac{1}{2}\|A_n^{\perp}\|_F^2$$

is the most direct way to read the numerical experiment. If the full information is large but the source-incremental component is small, the local alternative is mostly target-only at the tested scale. If the source-incremental component is large but the scalar residual-product moment has little information, then the scalar residual-product moment is not aligned with the screened local covariance direction.

5. Projected covariance-score GMM

5.1. Low-dimensional score basis.

The full matrix $A_n^{\perp}$ is too large and too model-dependent to be the default empirical estimator. The feasible statistic therefore starts from a fixed or slowly growing family of symmetric raw covariance directions

$$H_{1,n},\dots,H_{d_n,n}.$$

These are source-incremental directions: future-source cross moments, target-history/source cross moments, and a small number of boundary-layer perturbations. Whiten and target-orthogonalize them by setting

$$B_{r,n}^0 = W_n H_{r,n} W_n, \qquad B_{r,n} = B_{r,n}^0 - \Pi_{Y,n}B_{r,n}^0.$$

The score vector is

$$g_{k,n} = \left(\ell_{k,n}(B_{1,n}),\dots,\ell_{k,n}(B_{d_n,n})\right)^{\top}.$$

The basis contains the scalar residual-product estimator as a special case. If a_n is the target residual direction and q_n is the source instrument direction, the symmetrized cross matrix

$$H_n^{\mathrm{scalar}} = a_n q_n^{\top} + q_n a_n^{\top}$$

generates the residual-product score. The projected estimator expands this single direction into a small projected covariance dictionary. Once the score dictionary is fixed, the feasible statistic is a finite-dimensional optimally weighted GMM/Wald construction in the sense of [20]. The specific ingredient is the projected covariance-score geometry that determines the screened moments, nuisance removal, and local information constants.

Definition 5.1 (Residual-product scalar anchor)**.** Let a_n and q_n^X be whitened linear directions satisfying

$$R^Y_{k,n} = a_n^\top U_{k,n}, \qquad I^X_{k,n} = (q_n^X)^\top U_{k,n},$$

where $R^Y_{k,n}$ is the target future residual after target-only projection and $I^X_{k,n}$ is the first rough source-boundary loading. The scalar residual-product score is

$$S^{\mathrm{scalar}}_{k,n} = R^Y_{k,n} I^X_{k,n}.$$

Equivalently, up to the centering term in the Gaussian covariance score, it is the quadratic score generated by

$$B_n^{\mathrm{scalar}} = a_n (q_n^X)^\top + q_n^X a_n^\top.$$

Thus the scalar statistic remains as the one-dimensional anchor direction inside the projected covariance-score dictionary. Its empirical information $\mathcal{I}_n^{\mathrm{scalar}}$ is reported to quantify how much of the source-orthogonal covariance derivative is visible to the original residual-product moment.

5.2. **Projected basis information.** Let $\mathcal{B}_n = \mathrm{span}\{B_{1,n}, \ldots, B_{d_n,n}\}$ denote the retained source-incremental score space after target-only projection. In the reduced Gaussian experiment, the amount of source-orthogonal covariance information captured by this basis is

$$\mathcal{I}_n^{\mathcal{B}} = \Gamma_n^\top \Omega_n^{-1} \Gamma_n,$$

where Γ_n and Ω_n are defined below. If the matrices are Frobenius-orthonormal and the blocks are independent, this is exactly the squared norm of the projection of $A_n^\perp$ onto $\mathcal{B}_n$, up to the same factor $1/2$ used in the Gaussian Fisher information. Hence

$$0 \le \mathcal{I}_n^{\mathcal{B}} \le \mathcal{I}_n^{\mathrm{src}\perp Y}.$$

The difference $\mathcal{I}_n^{\mathrm{src}\perp Y} - \mathcal{I}_n^{\mathcal{B}}$ is therefore a basis-design loss, not a screening impossibility.

The scalar anchor has its own one-dimensional information

$$\mathcal{I}_n^{\mathrm{scalar}} = \frac{\Gamma^2_{n,\mathrm{scalar}}}{\Omega_{n,\mathrm{scalar}}}.$$

The diagnostic signature for the specification is

$$\mathcal{I}_n^{\mathrm{src}\perp Y} > 0, \qquad \mathcal{I}_n^{\mathcal{B}} \approx \mathcal{I}_n^{\mathrm{src}\perp Y}, \qquad \mathcal{I}_n^{\mathrm{scalar}} \ll \mathcal{I}_n^{\mathcal{B}}.$$

This pattern means that the source-screened signal exists, but it lies in covariance directions not captured by the scalar residual-product moment.

5.3. **Local loading and long-run covariance.** Define

$$\Gamma_n = \frac{\partial}{\partial \tau} \mathbb{E}_\tau \big[\sqrt{N_n}\, \bar{g}_n\big]\Big|_{\tau=0}, \qquad \bar{g}_n = \frac{1}{N_n} \sum_{k=1}^{N_n} g_{k,n}.$$

In the independent Gaussian idealization,

$$\Gamma_{r,n} = \frac{1}{2} \operatorname{tr}(B_{r,n} A_n).$$

In the feasible implementation, Γ_n can be obtained from the calibrated local Gaussian block model or by a plus/minus finite-difference pilot around $\tau = 0$.

Let Ω_n denote the long-run covariance of

$$\sqrt{N_n}\, \bar{g}_n.$$

For disjoint blocks this is the one-block covariance. For overlapping raw blocks it is estimated by a positive-definite HAC estimator $\widehat{\Omega}_n$ in the spirit of [21, 22]; the deterministic bandwidth rules used below follow the automatic-lag tradition of [23].

Definition 5.2 (Vector HAC and multiplier calibration)**.** For evaluation scores $g_{k,n} \in \mathbb{R}^{d_n}$, $1 \le k \le \widetilde{N}_n$, set

$$\bar{g}_n = \frac{1}{\widetilde{N}_n} \sum_{k=1}^{\widetilde{N}_n} g_{k,n}$$

and, for $0 \le \ell \le L_n^{\mathrm{hac}}$,

$$\widehat{C}_n(\ell) = \frac{1}{\widetilde{N}_n} \sum_{k=1}^{\widetilde{N}_n - \ell} (g_{k,n} - \bar{g}_n)(g_{k+\ell,n} - \bar{g}_n)^\top.$$

Given a compactly supported kernel K_{hac}, define

$$\widehat{\Omega}_n = \widehat{C}_n(0) + \sum_{\ell=1}^{L_n^{\mathrm{hac}}} K_{\mathrm{hac}}\left(\frac{\ell}{L_n^{\mathrm{hac}} + 1}\right) \{\widehat{C}_n(\ell) + \widehat{C}_n(\ell)^\top\}.$$

The iid Gaussian calibration uses $L_n^{\mathrm{hac}} = 0$. For overlapping Volterra blocks, L_n^{hac} is tied to the memory window, and the default kernel is Bartlett,

$$K_{\mathrm{hac}}(x) = (1 - |x|)_+.$$

Any dependent-multiplier calibration used with positive bandwidth must use multiplier weights whose long-run covariance matches the same kernel. A flat moving-average multiplier may serve as an implementation check; final positive-bandwidth calibration uses the matching kernel.

The multiplier proposition proves that resampling must use the same long-run covariance kernel as the statistic it calibrates.

Proposition 5.3 (Kernel-matched HAC multiplier validity)**.** *Assume $d_n \le d < \infty$, the centered projected score array satisfies the HAC CLT in Assumption 7.1, and the HAC bandwidth obeys the rate conditions under which $\widehat{\Omega}_n \xrightarrow{\mathbb{P}} \Omega$. Let $(\xi_{k,n})$ be a mean-zero stationary Gaussian multiplier sequence, independent of the data, with autocovariance*

$$\mathbb{E}^*[\xi_{k,n}\xi_{k+\ell,n}] = K_{\mathrm{hac}}\left(\frac{\ell}{L_n^{\mathrm{hac}} + 1}\right) \mathbf{1}_{\{|\ell| \le L_n^{\mathrm{hac}}\}} + o(1)$$

uniformly for $|\ell| \le L_n^{\mathrm{hac}}$, and negligible autocovariance outside the HAC window. Define the multiplier score

$$S_n^* = \frac{1}{\sqrt{\widetilde{N}_n}} \sum_{k=1}^{\widetilde{N}_n} \xi_{k,n}(g_{k,n} - \bar{g}_n).$$

Then, conditionally on the data,

$$\mathrm{Var}^*(S_n^*) = \widehat{\Omega}_n + o_{\mathbb{P}}(1), \qquad S_n^* \xrightarrow{d,\mathbb{P}} N(0, \Omega).$$

Consequently multiplier critical values for any fixed-dimensional studentized linear score or Wald statistic have the same first-order null limit as the corresponding HAC-normalized statistic.

Proof. The conditional covariance of S_n^* is

$$\frac{1}{\widetilde{N}_n} \sum_{k,j} \mathbb{E}^*(\xi_{k,n}\xi_{j,n})(g_{k,n} - \bar{g}_n)(g_{j,n} - \bar{g}_n)^\top.$$

By the assumed covariance matching, this equals the HAC sum in Definition 5.2, up to boundary terms and the displayed $o(1)$ multiplier covariance error. The boundary contribution is negligible under $L_n^{\mathrm{hac}}/\widetilde{N}_n \to 0$, or is absent when $L_n^{\mathrm{hac}} = 0$. Hence $\mathrm{Var}^*(S_n^*) = \widehat{\Omega}_n + o_{\mathbb{P}}(1)$, and HAC consistency gives convergence to Ω. Conditional on the data, S_n^* is Gaussian. Convergence of its conditional covariance therefore gives $S_n^* \xrightarrow{d,\mathbb{P}} N(0, \Omega)$. The linear-score and Wald conclusions follow by the continuous mapping theorem and Slutsky's theorem. □

Definition 5.4 (Projected covariance-score statistic)**.** Let $\widehat{\Gamma}_n$ and $\widehat{\Omega}_n$ be pilot estimates of Γ_n and Ω_n. Define

$$\widehat{\lambda}_n = \frac{\widehat{\Omega}_n^{-1}\widehat{\Gamma}_n}{\left(\widehat{\Gamma}_n^\top \widehat{\Omega}_n^{-1}\widehat{\Gamma}_n\right)^{1/2}},$$

and

$$T_n^{\mathrm{pcov}} = \widehat{\lambda}_n^\top \sqrt{N_n}\,\bar{g}_n.$$

The estimated projected information index is

$$\widehat{\Delta}_n^{\mathrm{pcov}} = \left(\widehat{\Gamma}_n^\top \widehat{\Omega}_n^{-1}\widehat{\Gamma}_n\right)^{1/2}.$$

The associated local limit target is

$$T_n^{\mathrm{pcov}} \xrightarrow{d} N(\tau\Delta^{\mathrm{pcov}}, 1), \qquad \Delta^{\mathrm{pcov}} = \lim_n \left(\Gamma_n^\top \Omega_n^{-1}\Gamma_n\right)^{1/2}.$$

The scalar residual-product statistic corresponds to the case $d_n = 1$ with the basis matrix H_n^{scalar}.

6. Feasible cross-fitted estimator

The projected covariance-score construction becomes an empirical estimator after the whitening, target projection, basis, loading, and evaluation sample are fixed. The implementation used below is regularized and cross-fitted.

6.1. **Regularized whitening and target projection.**

Let $\widehat{\Sigma}_{0,n}^{\mathcal{C}}$ be a null covariance estimate obtained on a calibration sample or from a fitted null Volterra model, and write

$$\widehat{\Sigma}_{0,n}^{\mathcal{C}} = Q_n \operatorname{diag}(\widehat{\lambda}_{1,n}, \dots, \widehat{\lambda}_{p_n,n}) Q_n^\top$$

for its spectral decomposition. For a deterministic ridge sequence $\rho_n > 0$, define

$$\widehat{\Sigma}_{0,n}^{\rho} = Q_n \operatorname{diag}(\widehat{\lambda}_{1,n} \vee \rho_n, \dots, \widehat{\lambda}_{p_n,n} \vee \rho_n) Q_n^\top, \qquad \widehat{W}_n = (\widehat{\Sigma}_{0,n}^{\rho})^{-1/2}.$$

The feasible target-only tangent space is

$$\widehat{\mathcal{T}}_{Y,n} = \left\{\widehat{W}_n S_{Y,n}^\top C S_{Y,n} \widehat{W}_n : C = C^\top\right\},$$

and $\widehat{\Pi}_{Y,n}$ is the Frobenius projection onto this space. The ridge sequence satisfies $\rho_n \downarrow 0$ slowly enough that the whitening map is stable on the retained coordinates. In finite samples ρ_n is also a diagnostic tuning parameter. If the projected information changes sharply over reasonable ridge values, the covariance calibration is not yet reliable.

6.2. **Canonical projected basis.**

The default basis is fixed before looking at the evaluation sample. It contains three blocks:

$$\mathcal{H}_n^{\mathrm{def}} = \mathcal{H}_n^{\mathrm{scalar}} \cup \mathcal{H}_n^{FX} \cup \mathcal{H}_n^{YX}.$$

The scalar anchor is

$$H_n^{\mathrm{scalar}} = a_n (q_{1,n}^X)^\top + q_{1,n}^X a_n^\top,$$

where a_n is the target-future residual direction under the null and $q_{1,n}^X$ is the leading source boundary-layer loading. The future-source block is

$$H_{r,n}^{FX} = e_F (q_{r,n}^X)^\top + q_{r,n}^X e_F^\top, \qquad 1 \le r \le d_X,$$

and the target-source block is

$$H_{s,r,n}^{YX} = q_{s,n}^Y (q_{r,n}^X)^\top + q_{r,n}^X (q_{s,n}^Y)^\top, \qquad 1 \le s \le d_Y, \quad 1 \le r \le d_X.$$

Here $q_{r,n}^X$ and $q_{s,n}^Y$ are deterministic low-dimensional boundary-layer loadings on the source and target lag grids. A practical default is $d_X, d_Y \in \{1, 2\}$, using the constant and front-loaded shapes. The baseline contagion model excludes the source-only covariance block because the local alternative leaves the marginal source law unchanged; the excluded block can be reported as a placebo in robustness exercises.

The estimator uses the projected basis matrices

$$\widehat{B}_{r,n} = \widehat{W}_n H_{r,n} \widehat{W}_n - \widehat{\Pi}_{Y,n}(\widehat{W}_n H_{r,n} \widehat{W}_n), \qquad H_{r,n} \in \mathcal{H}_n^{\mathrm{def}},$$

after dropping numerically null directions and orthonormalizing the retained matrices in the Frobenius inner product. The orthonormalization is not mathematically essential. It stabilizes the information decomposition and its numerical reading.

6.3. **Cross-fitted loading and evaluation score.** All nuisance quantities are estimated on a calibration sample $\mathcal{C}$, while the statistic is computed on a disjoint evaluation sample $\mathcal{E}$. The calibration sample provides

$$\widehat{\Sigma}_{0,n}^{\mathcal{C}}, \qquad \widehat{B}_{1,n}, \ldots, \widehat{B}_{\widehat{d},n}, \qquad \widehat{\Gamma}_n^{\mathcal{C}}.$$

The loading $\widehat{\Gamma}_n^{\mathcal{C}}$ may be analytic, using the fitted local Gaussian block model, or finite-difference based:

$$\widehat{\Gamma}_{r,n}^{\mathcal{C}} = \frac{\sqrt{N_n}\, \widehat{\mathbb{E}}_{+\delta_n}^{\mathcal{C}}[\widehat{g}_{r,k,n}] - \sqrt{N_n}\, \widehat{\mathbb{E}}_{-\delta_n}^{\mathcal{C}}[\widehat{g}_{r,k,n}]}{2\delta_n}.$$

The evaluation score is then

$$\widehat{g}_{r,k,n}^{\mathcal{E}} = \frac{1}{2}\left\{(\widehat{U}_{k,n}^{\mathcal{E}})^\top \widehat{B}_{r,n} \widehat{U}_{k,n}^{\mathcal{E}} - \operatorname{tr}(\widehat{B}_{r,n})\right\}, \qquad \widehat{U}_{k,n}^{\mathcal{E}} = \widehat{W}_n Z_{k,n}^{\mathcal{E}}.$$

Let $\widehat{\bar{g}}_n^{\mathcal{E}}$ be the evaluation-sample average and let $\widehat{\Omega}_n^{\mathcal{E}}$ be the HAC long-run covariance estimator computed from $(\widehat{g}_{k,n}^{\mathcal{E}})$. The feasible projected covariance-score statistic is

$$T_n^{\mathrm{pcov,def}} = (\widehat{\lambda}_n^{\mathrm{def}})^\top \sqrt{N_n}\widehat{\bar{g}}_n^{\mathcal{E}}, \qquad \widehat{\lambda}_n^{\mathrm{def}} = \frac{(\widehat{\Omega}_n^{\mathcal{E}})^{-1}\widehat{\Gamma}_n^{\mathcal{C}}}{[(\widehat{\Gamma}_n^{\mathcal{C}})^\top (\widehat{\Omega}_n^{\mathcal{E}})^{-1}\widehat{\Gamma}_n^{\mathcal{C}}]^{1/2}}.$$

The associated information index is

$$\widehat{\Delta}_n^{\mathrm{pcov,def}} = [(\widehat{\Gamma}_n^{\mathcal{C}})^\top (\widehat{\Omega}_n^{\mathcal{E}})^{-1}\widehat{\Gamma}_n^{\mathcal{C}}]^{1/2}.$$

Remark 6.1 (Why cross-fitting is part of the definition). The projected covariance basis can overfit if the same data estimate the covariance derivative, choose the basis, and evaluate the score. Cross-fitting fixes the score conditionally on the evaluation sample. The finite-dimensional GMM interpretation and the multiplier calibration then remain valid. For long time series, several calibration/evaluation folds can be averaged; the single-split version is used for the theory below.

Cross-fitted nuisance estimation leaves the local score law and the projected information index unchanged.

Proposition 6.2 (Cross-fitted feasibility reduction)**.** *Let $B_{1,n}, \ldots, B_{d,n}$, Γ_n, and Ω_n denote the oracle projected basis, local loading, and long-run covariance for the evaluation experiment, with d fixed and*

$$\Gamma_n^\top \Omega_n^{-1} \Gamma_n \to \Delta^2 > 0.$$

Suppose the cross-fitted nuisance estimates satisfy, on the retained basis,

$$\max_{1\le r\le d} \|\widehat{B}_{r,n} - B_{r,n}\|_F = o_{\mathbb{P}}(1), \qquad \widehat{\Gamma}_n^{\mathcal{C}} - \Gamma_n = o_{\mathbb{P}}(1), \qquad \widehat{\Omega}_n^{\mathcal{E}} - \Omega_n = o_{\mathbb{P}}(1),$$

and

$$\frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \{\widehat{g}_{k,n}^{\mathcal{E}} - g_{k,n}^{\mathcal{E}}\} = o_{\mathbb{P}}(1), \qquad \sqrt{N_n}\bar{g}_n^{\mathcal{E}} = O_{\mathbb{P}}(1),$$

under the local alternatives under consideration. Then

$$\widehat{\Delta}_n^{\mathrm{pcov,def}} - \Delta_n = o_{\mathbb{P}}(1), \qquad T_n^{\mathrm{pcov,def}} - T_n^{\mathrm{pcov}} = o_{\mathbb{P}}(1),$$

where

$$\Delta_n^2 = \Gamma_n^\top \Omega_n^{-1} \Gamma_n, \qquad T_n^{\mathrm{pcov}} = \lambda_n^\top \sqrt{N_n}\bar{g}_n^{\mathcal{E}}, \qquad \lambda_n = \frac{\Omega_n^{-1}\Gamma_n}{(\Gamma_n^\top \Omega_n^{-1}\Gamma_n)^{1/2}}.$$

Consequently the feasible cross-fitted statistic has the same local limit as the oracle projected statistic whenever the latter satisfies a LAN limit. The same conclusion holds for a fixed number of folds after averaging the foldwise statistics with deterministic weights.

Proof. The oracle statistic is defined on the retained nonsingular score basis, after numerically null directions have been discarded. On this basis the eigenvalues of Ω_n are bounded away from zero and $\Delta_n \to \Delta > 0$. Hence matrix inversion and the normalization map

$$(\Gamma, \Omega) \mapsto \frac{\Omega^{-1}\Gamma}{(\Gamma^\top \Omega^{-1}\Gamma)^{1/2}}$$

are continuous in a neighborhood of the oracle limit. The loading and covariance consistency therefore imply

$$\widehat{\Delta}_n^{\mathrm{pcov,def}} - \Delta_n = o_{\mathbb{P}}(1), \qquad \widehat{\lambda}_n^{\mathrm{def}} - \lambda_n = o_{\mathbb{P}}(1).$$

Decompose the statistic difference as

$$T_n^{\mathrm{pcov,def}} - T_n^{\mathrm{pcov}} = (\widehat{\lambda}_n^{\mathrm{def}} - \lambda_n)^\top \sqrt{N_n}\bar{g}_n^{\mathcal{E}} + (\widehat{\lambda}_n^{\mathrm{def}})^\top \frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \{\widehat{g}_{k,n}^{\mathcal{E}} - g_{k,n}^{\mathcal{E}}\}.$$

The first term is $o_{\mathbb{P}}(1)$ because the oracle score sum is tight. The second term is $o_{\mathbb{P}}(1)$ because $\widehat{\lambda}_n^{\mathrm{def}} = O_{\mathbb{P}}(1)$ and the cross-fitted score replacement error is $o_{\mathbb{P}}(1)$. Slutsky's theorem gives the asserted transfer of the local limit. For finitely many folds, the same argument applies conditionally on each calibration sample, and a finite weighted sum of $o_{\mathbb{P}}(1)$ remainders remains $o_{\mathbb{P}}(1)$. □

6.4. **Diagnostic report for the feasible estimator.** Every run of the feasible estimator reports the five-object hierarchy

$$T_n^{\mathrm{full}}, \qquad T_n^{Y,\mathrm{only}}, \qquad T_n^{\mathrm{src}\perp Y}, \qquad T_n^{\mathrm{scalar}}, \qquad T_n^{\mathrm{pcov,def}}.$$

These objects separate the existence of local information from the ability of the feasible estimator to recover it. The primary estimator is $T_n^{\mathrm{pcov,def}}$, and the primary diagnostic is the ratio

$$\frac{(\widehat{\Delta}_n^{\mathrm{pcov,def}})^2}{\widehat{\mathcal{I}}_n^{\mathrm{src}\perp Y}},$$

together with the scalar capture ratio

$$\frac{\widehat{\mathcal{I}}_n^{\mathrm{scalar}}}{\widehat{\mathcal{I}}_n^{\mathrm{src}\perp Y}}.$$

The projected estimator is informative only if the first ratio is stable and materially positive while the procedure maintains valid null calibration.

6.5. **Feasible algorithm.** The estimator is the finite-sample procedure below.

(1) Construct the block vector $Z_{k,n} = (F_{k,n}^Y, P_{k,n}^Y, P_{k,n}^X)$ on the chosen raw or overlapping evaluation grid.
(2) Estimate or import the null covariance $\widehat{\Sigma}_{0,n}$, and compute a stabilized inverse square root $\widehat{W}_n$.
(3) Build a small raw covariance dictionary $H_{1,n}, \dots, H_{d_n,n}$, including the scalar anchor and the target-source cross directions.
(4) Whiten and project each raw direction away from the target-only tangent space:

$$\widehat{B}_{r,n} = \widehat{W}_n H_{r,n} \widehat{W}_n - \widehat{\Pi}_{Y,n}(\widehat{W}_n H_{r,n} \widehat{W}_n).$$

(5) Evaluate the centered quadratic score vector

$$\widehat{g}_{k,n} = \left(\frac{1}{2} \{\widehat{U}_{k,n}^\top \widehat{B}_{r,n} \widehat{U}_{k,n} - \mathrm{tr}(\widehat{B}_{r,n})\} \right)_{r=1}^{d_n}.$$

(6) Estimate the local loading $\widehat{\Gamma}_n$ from the calibrated Gaussian block model or an independent plus/minus finite-difference pilot.

(7) Estimate the long-run covariance $\widehat{\Omega}_n$ by HAC, form $\widehat{\lambda}_n \propto \widehat{\Omega}_n^{-1}\widehat{\Gamma}_n$, and report both the asymptotic studentized statistic and a multiplier-calibrated finite-sample version.

The same score engine produces the full oracle, the target-only oracle, the source-orthogonal oracle, the scalar anchor, and the projected basis statistic by changing only the score direction or retained basis.

6.6. **Local coefficient estimation.** The projected score also estimates the local amplitude. Define

$$\widehat{\tau}_n^{\mathrm{pcov}} = \frac{\widehat{\Gamma}_n^\top \widehat{\Omega}_n^{-1}\sqrt{N_n}\,\bar{g}_n}{\widehat{\Gamma}_n^\top \widehat{\Omega}_n^{-1}\widehat{\Gamma}_n}.$$

Under the same local expansion,

$$\widehat{\tau}_n^{\mathrm{pcov}} \xrightarrow{d} N\left(\tau, (\Delta^{\mathrm{pcov}})^{-2}\right).$$

The projected covariance-score estimator recovers the local amplitude once the loading calibration is fixed. The coefficient in the original volatility model is the local-scale coefficient $\theta_n\widehat{\tau}_n^{\mathrm{pcov}}$, where θ_n is the shrinking alternative scale.

6.7. **Diagnostic information decomposition.** The numerical study reports information before power. The diagnostics separate model design, basis design, and implementation:

$$\mathcal{I}_n^{\mathrm{full}} = \frac{1}{2}\|A_n\|_F^2, \qquad \mathcal{I}_n^{Y} = \frac{1}{2}\|\Pi_{Y,n}A_n\|_F^2, \qquad \mathcal{I}_n^{\mathrm{src}\perp Y} = \frac{1}{2}\|A_n - \Pi_{Y,n}A_n\|_F^2.$$

For a chosen projected basis $B_{1,n}, \dots, B_{d_n,n}$, report

$$\mathcal{I}_n^{\mathcal{B}} = \Gamma_n^\top \Omega_n^{-1}\Gamma_n,$$

and for the scalar anchor report the one-dimensional analogue

$$\mathcal{I}_n^{\mathrm{scalar}} = \frac{\Gamma_{n,\mathrm{scalar}}^2}{\Omega_{n,\mathrm{scalar}}}.$$

The diagnostics separate the residual cases as follows.

(i) If $\mathcal{I}_n^{\mathrm{full}}$ is small, the local alternative is intrinsically weak at the tested scale.
(ii) If $\mathcal{I}_n^{\mathrm{full}}$ is large but $\mathcal{I}_n^{\mathrm{src}\perp Y}$ is small, the alternative is mostly target-only and does not support a strong source-screened estimator.
(iii) If $\mathcal{I}_n^{\mathrm{src}\perp Y}$ is large but $\mathcal{I}_n^{\mathcal{B}}$ is small, the selected basis misses the screened covariance direction.
(iv) If $\mathcal{I}_n^{\mathcal{B}}$ is large but empirical power is weak, the remaining problem is finite-sample calibration, HAC estimation, or proxy reduction rather than estimator design.
(v) If $\mathcal{I}_n^{\mathrm{scalar}}$ is small while $\mathcal{I}_n^{\mathcal{B}}$ is large, the scalar residual-product estimator is too narrow, while the projected covariance-score estimator still has identifying information.

6.8. **Default basis design.** A practical implementation keeps d_n small. The basis has three blocks.

(1) Future-source cross scores:

$$H_{r,n}^{FX} = e_F(q_{r,n}^X)^\top + q_{r,n}^X e_F^\top,$$

where e_F selects $F_{k,n}^Y$ and $q_{r,n}^X$ are source boundary-layer loadings.

(2) Target-history/source cross scores:

$$H_{r,s,n}^{YX} = q_{r,n}^Y(q_{s,n}^X)^\top + q_{s,n}^X(q_{r,n}^Y)^\top.$$

These directions are the main extension beyond the scalar residual statistic.

(3) Residualized scalar anchor:

$$H_n^{\mathrm{scalar}} = a_n q_{1,n}^{X\top} + q_{1,n}^X a_n^\top,$$

where a_n is the target-future residual direction from the scalar corrected score. Keeping this anchor embeds the scalar theorem statistic as a literal submodel of the implementation.

The default dictionary starts with the scalar anchor plus two or three low-frequency future-source and target-source directions. It does not include unrestricted source-source covariance scores unless the model implies that the local alternative also changes the marginal source law. In the baseline contagion design the source process is exogenous, so source-only scores are mainly useful as placebo diagnostics.

6.9. **Calibration components.** The feasible estimator has five calibration components.

(i) Estimate or import $\widehat{\Sigma}_{0,n}$ from the null Volterra calibration, a null Monte Carlo pilot, or a rolling local covariance estimator.
(ii) Build raw basis matrices $H_{r,n}$, whiten them by $\widehat{\Sigma}_{0,n}^{-1/2}$, and project them away from the estimated target-only tangent space.
(iii) Estimate $\widehat{\Gamma}_n$ by analytic covariance differentiation when the Gaussian block calibration is available, and otherwise by a plus/minus finite-difference pilot around $\tau = 0$.
(iv) Compute $g_{k,n}$ on the evaluation grid and estimate $\widehat{\Omega}_n$. In the iid Gaussian calibration experiment the correct default is bandwidth zero. For overlapping Volterra blocks, positive HAC bandwidths are tied to the memory window J_n and reported as dependence-robustness checks.
(v) Report both the asymptotic studentized score and a dependent-multiplier finite-sample calibration. In Monte Carlo work, empirical-null critical values are reported as diagnostics, not as the estimator definition.

For the rough-local scale itself, the implementation uses an ordered pilot strategy. Injected noisy pilots are used first to validate Theorem 8.7 directly. The baseline feasible pilot is an adaptive variogram-correlation estimator. A regime-adaptive upper-boundary extension is then activated only when the calibration path is long enough to support it; on shorter paths it collapses exactly to the adaptive pilot. This keeps the feasible construction aligned with the asymptotic logic. Extra upper-boundary flexibility is introduced only in the path-length regime where the synthetic calibration shows that it improves the projected information geometry.

The loading pilot is independent of the evaluation replications in simulation. In an application, the analogous separation is a split between the calibration window used to fit the local Gaussian block model and the evaluation window used to compute the statistic.

6.10. **Relation to the scalar statistic.** The scalar construction retains a separate role. It is the baseline object for proving the observable-screening mechanism and for connecting the discrete block experiment to the continuous-time boundary-layer theorem. The change concerns the default econometric implementation, not the scalar theorem.

The hierarchy is

(i) scalar residual-product statistic, the proof baseline;
(ii) screened bilinear GMM, the source-dictionary extension of the scalar score;
(iii) projected covariance-score GMM, the preferred finite-sample estimator;
(iv) full Gaussian covariance oracle, the diagnostic upper bound;
(v) target-only covariance oracle, the lower-bound comparator.

The numerical study compares these objects in that order. At a given design point, the projected covariance-score estimator is informative when it captures a material share of $\mathcal{I}_n^{\mathrm{src}\perp Y}$, controls null size after HAC/multiplier calibration, and improves on the scalar residual-product statistic without approaching the target-only oracle through excluded target-only directions.

7. Projected LAN theory

The limiting experiment separates the primitive rough-Volterra reduction from the finite-dimensional projected-GMM step. The assumptions collect the inputs supplied by the Gaussian block reduction, feasible proxy construction, and memory-window central limit theory. Conditional on these inputs, the projected estimator has a finite-dimensional LAN interpretation.

Assumption 7.1 (Projected score regularity)**.** Along the tested sequence of block designs, suppose:

(a) the basis dimension is bounded, $d_n \le d < \infty$, and the projected score matrices $B_{r,n}$ have uniformly bounded Frobenius norms;

(b) the long-run covariance matrix

$$\Omega_n = \mathrm{Var}_0\left(\frac{1}{\sqrt{N_n}}\sum_{k=1}^{N_n} g_{k,n}\right)$$

has eigenvalues bounded away from zero and infinity on the retained basis;

(c) under local alternatives, the projected score vector has mean expansion

$$\mathbb{E}_{n,\tau}\left[\frac{1}{\sqrt{N_n}}\sum_{k=1}^{N_n} g_{k,n}\right] = \tau\Gamma_n + o(1)$$

uniformly for τ in compact sets;

(d) the centered projected score vector satisfies a HAC-compatible central limit theorem,

$$\frac{1}{\sqrt{N_n}}\sum_{k=1}^{N_n}\{g_{k,n} - \mathbb{E}_{n,\tau} g_{k,n}\} \xrightarrow{d} N(0,\Omega)$$

with $\Omega_n \to \Omega$ and $\Gamma_n \to \Gamma$;

(e) the feasible proxy scores, pilot loading, and HAC estimator obey

$$\frac{1}{\sqrt{N_n}}\sum_{k=1}^{N_n}(\widehat{g}_{k,n} - g_{k,n}) \xrightarrow{\mathbb{P}} 0, \qquad \widehat{\Gamma}_n \xrightarrow{\mathbb{P}} \Gamma, \qquad \widehat{\Omega}_n \xrightarrow{\mathbb{P}} \Omega.$$

The next proposition identifies the primitive reductions needed for projected-score regularity.

Proposition 7.2 (Primitive verification of projected score regularity)**.** *Assume the observed-to-Gaussian reduction, the discrete Volterra transfer representation, the vector proxy replacement, and the projected-score HAC CLT from Appendix B. Suppose also that the estimated null covariance, regularized whitening map, target-only projection, retained basis, loading pilot, and HAC estimator are stable in the sense that*

$$\widehat{W}_n - W_n = o_{\mathbb{P}}(1), \qquad \widehat{\Pi}_{Y,n} - \Pi_{Y,n} = o_{\mathbb{P}}(1), \qquad \widehat{\Gamma}_n - \Gamma_n = o_{\mathbb{P}}(1), \qquad \widehat{\Omega}_n - \Omega_n = o_{\mathbb{P}}(1)$$

on the retained finite-dimensional score space. Then Assumption 7.1 holds.

Proof. The bounded basis dimension and bounded projected score norms are imposed by construction after dropping numerically null directions. The discrete Volterra transfer theorem gives the covariance derivative and hence the local mean expansion $\mathbb{E}_{n,\tau}[\sqrt{N_n}\bar{g}_n] = \tau\Gamma_n + o(1)$, with $\Gamma_{r,n} = \frac{1}{2}\operatorname{tr}(B_{r,n}A_n) + o(1)$ in the reduced Gaussian score normalization. The projected-score HAC CLT gives the centered Gaussian limit and identifies the long-run covariance Ω. Vector proxy replacement transfers the latent Gaussian score expansion to the feasible realized-volatility scores. Finally, the displayed stability conditions and Slutsky's theorem give the feasible score, loading, and covariance convergence in Assumption 7.1. □

The quasi-score statement separates Gaussian score optimality from the weaker GMM validity needed for the projected statistic.

Proposition 7.3 (Quasi-score validity beyond exact Gaussianity)**.** *Consider the same quadratic projected scores*

$$g_{r,k,n} = \frac{1}{2}\{U_{k,n}^\top B_{r,n} U_{k,n} - \operatorname{tr}(B_{r,n})\}, \qquad 1 \le r \le d_n,$$

but suppose the block law need not be exactly Gaussian. If the retained scores satisfy the mean expansion, HAC central limit theorem, and feasible replacement conditions in Assumption 7.1, then Theorem 7.6 and Corollary 7.7 remain valid with Γ equal to the true derivative of the projected-score mean and Ω equal to the true long-run covariance of the retained scores.

Proof. The proof of Theorem 7.6 uses only the convergence $\sqrt{N_n}\bar{g}_n \xrightarrow{d} N(\tau\Gamma, \Omega)$, the consistency of $\widehat{\Gamma}_n$ and $\widehat{\Omega}_n$, and the continuity of the GMM loading map $(\Gamma, \Omega) \mapsto \Omega^{-1}\Gamma/(\Gamma^\top\Omega^{-1}\Gamma)^{1/2}$. None of these steps requires the block vector itself to be Gaussian. Gaussianity is used earlier to identify the quadratic score as the exact covariance likelihood score and to compute the closed-form information constant $\Gamma_{r,n} = \frac{1}{2}\operatorname{tr}(B_{r,n}A_n)$. If the displayed regularity conditions are verified directly, the same Slutsky argument gives the studentized quasi-score and local-amplitude limits. □

Remark 7.4 (Outside the Gaussian reduced experiment). Outside exact Gaussianity, the statistic becomes a projected quadratic GMM or quasi-score estimator. The loading Γ must be estimated as the true derivative of the retained score mean, for example by finite differences or model calibration, and Ω is the true long-run covariance of the retained quadratic scores. The information $\Gamma^\top\Omega^{-1}\Gamma$ is then projected quasi-information, not the full semiparametric efficiency bound. Likewise, the Gaussian target-only tangent removes target-only covariance information; a full non-Gaussian target-only lower bound would require enlarging the tangent space to higher cumulant scores.

The retained projected span captures only part of the source-orthogonal information.

Proposition 7.5 (Projected information bound). *In the independent Gaussian block experiment, let $\mathcal{B}_n$ be the projected score space generated by $B_{1,n}, \ldots, B_{d_n,n}$. Then*

$$0 \le \Gamma_n^\top\Omega_n^{-1}\Gamma_n \le \mathcal{I}_n^{\mathrm{src}\perp Y}.$$

Equality holds if and only if the source-orthogonal covariance derivative $A_n^\perp$ belongs to the closed linear span of $\mathcal{B}_n$, up to directions with zero score variance.

Proof. By Lemma 3.3, covariance scores form a finite-dimensional Hilbert space with inner product

$$\langle B, C\rangle_{\mathrm{score}} = \frac{1}{2}\operatorname{tr}(BC).$$

Work on the quotient of coefficient space obtained by removing directions with zero score variance. For

$$B(b) := \sum_{r=1}^{d_n} b_r B_{r,n},$$

the Gaussian score algebra gives

$$b^\top\Omega_n b = \|B(b)\|_{\mathrm{score}}^2, \qquad b^\top\Gamma_n = \langle B(b), A_n^\perp\rangle_{\mathrm{score}}.$$

Thus the retained information is the Rayleigh quotient

$$\Gamma_n^\top\Omega_n^{-1}\Gamma_n = \sup_{b\neq 0}\frac{\langle B(b), A_n^\perp\rangle_{\mathrm{score}}^2}{\|B(b)\|_{\mathrm{score}}^2}.$$

The supremum is the squared norm of the Riesz representative of the functional $B \mapsto \langle B, A_n^\perp\rangle_{\mathrm{score}}$ restricted to $\mathcal{B}_n$, hence

$$\Gamma_n^\top\Omega_n^{-1}\Gamma_n = \|\Pi_{\mathcal{B}_n}^{\mathrm{score}}A_n^\perp\|_{\mathrm{score}}^2.$$

Pythagoras gives

$$\|\Pi_{\mathcal{B}_n}^{\mathrm{score}}A_n^\perp\|_{\mathrm{score}}^2 \le \|A_n^\perp\|_{\mathrm{score}}^2 = \mathcal{I}_n^{\mathrm{src}\perp Y}.$$

Equality holds exactly when $(I - \Pi_{\mathcal{B}_n}^{\mathrm{score}})A_n^\perp$ has zero score norm, which is the stated closed-span condition modulo zero-variance score directions. □

The projected statistic has a scalar LAN limit and a one-sided local power curve.

Theorem 7.6 (Projected covariance-score local limit). *Assume Assumption 7.1, and suppose*

$$\Delta^2 := \Gamma^\top\Omega^{-1}\Gamma > 0.$$

Then the feasible projected covariance-score statistic satisfies

$$T_n^{\mathrm{pcov}} \xrightarrow{d} N(\tau\Delta, 1)$$

under the local alternatives indexed by bounded τ. Consequently the one-sided test that rejects for $T_n^{\mathrm{pcov}} > z_{1-\alpha}$ has asymptotic size α and local power

$$1 - \Phi(z_{1-\alpha} - \tau\Delta).$$

Proof. Let $S_n = \sqrt{N_n}\bar{g}_n$. Assumption 7.1 gives

$$S_n \xrightarrow{d} N(\tau\Gamma, \Omega).$$

The feasible score, loading, and HAC consistency assumptions give

$$\widehat{\Gamma}_n \xrightarrow{\mathbb{P}} \Gamma, \qquad \widehat{\Omega}_n \xrightarrow{\mathbb{P}} \Omega, \qquad \sqrt{N_n}(\widehat{\bar{g}}_n - \bar{g}_n) \xrightarrow{\mathbb{P}} 0.$$

Since $\Delta^2 = \Gamma^\top\Omega^{-1}\Gamma > 0$, the map

$$(\Gamma, \Omega) \mapsto \lambda(\Gamma, \Omega) = \frac{\Omega^{-1}\Gamma}{(\Gamma^\top\Omega^{-1}\Gamma)^{1/2}}$$

is continuous in a neighborhood of the limit. Hence Slutsky's theorem yields

$$T_n^{\mathrm{pcov}} = \widehat{\lambda}_n^\top \sqrt{N_n}\widehat{\bar{g}}_n \xrightarrow{d} \lambda^\top N(\tau\Gamma, \Omega), \qquad \lambda = \frac{\Omega^{-1}\Gamma}{\Delta}.$$

The limiting variance is

$$\lambda^\top\Omega\lambda = \Delta^{-2}\Gamma^\top\Omega^{-1}\Omega\Omega^{-1}\Gamma = 1,$$

and the limiting mean is

$$\tau\lambda^\top\Gamma = \tau\Delta^{-1}\Gamma^\top\Omega^{-1}\Gamma = \tau\Delta.$$

The stated size and local power formula follow from the normal limit. □

The same projected quotient gives the local-amplitude estimator.

Corollary 7.7 (Projected local-amplitude inference)**.** *Under the conditions of Theorem 7.6, define*

$$\widehat{\tau}_n^{\mathrm{pcov}} = \frac{\widehat{\Gamma}_n^\top\widehat{\Omega}_n^{-1}\sqrt{N_n}\,\widehat{\bar{g}}_n}{\widehat{\Gamma}_n^\top\widehat{\Omega}_n^{-1}\widehat{\Gamma}_n}.$$

Then

$$\widehat{\tau}_n^{\mathrm{pcov}} \xrightarrow{d} N(\tau, \Delta^{-2}).$$

Proof. The numerator equals

$$\widehat{\Gamma}_n^\top\widehat{\Omega}_n^{-1}\sqrt{N_n}\widehat{\bar{g}}_n = \Gamma^\top\Omega^{-1}(\tau\Gamma + \xi) + o_{\mathbb{P}}(1)$$

with $\xi \sim N(0, \Omega)$ in the limit. The denominator converges to $\Delta^2 = \Gamma^\top\Omega^{-1}\Gamma$. Therefore

$$\widehat{\tau}_n^{\mathrm{pcov}} = \tau + \Delta^{-2}\Gamma^\top\Omega^{-1}\xi + o_{\mathbb{P}}(1),$$

and the limiting variance of the noise term is

$$\Delta^{-4}\Gamma^\top\Omega^{-1}\Omega\Omega^{-1}\Gamma = \Delta^{-2}.$$

□

Remark 7.8 (Primitive rough-Volterra inputs)*.* Assumption 7.1 records the primitive inputs that connect the projected covariance-score estimator to the rough-Volterra results of [3, 4]: Gaussian block reduction for the enlarged covariance-score vector, stability of the target-only projection and whitening maps under feasible proxy estimation, and a memory-window HAC CLT for overlapping projected scores. Given these inputs, Theorem 7.6 is a finite-dimensional projected-GMM argument.

8. Efficiency scope and identification concepts

The efficiency claim is delimited. Local efficiency inside the projected Gaussian score experiment carries no global minimax claim over all rough Volterra alternatives.

8.1. **What efficiency means here.** Theorem 7.6 gives a LAN-style statement for a local amplitude τ in the reduced Gaussian block experiment. Conditional on the null covariance, the target projection, the basis, and the loading derivative, the statistic uses the efficient GMM loading $\Omega^{-1}\Gamma$. Theorem 4.7 identifies the source-orthogonal covariance derivative as the nuisance-efficient score after removing target-only directions, and Proposition 7.5 then identifies how much of that efficient Gaussian information the retained basis captures. If $A_n^{\perp}$ lies in the retained projected basis span, the estimator reaches the source-orthogonal Gaussian information bound for that reduced experiment.

The efficiency claim is local before it is global. A global minimax theorem over a rough Volterra class requires a loss function, a lower bound, and tuning rules for the block size, dictionary dimension, ridge, and HAC bandwidth. Such a theorem rests on three uniform statements:

(i) the high-frequency realized-volatility experiment is asymptotically equivalent to the reduced Gaussian block experiment at the proposed rough-local scale;
(ii) the source-orthogonal covariance derivative is approximated by the retained projected basis with error small relative to the stochastic scale;
(iii) no other estimator can estimate the same rough-local functional faster over the same class.

The finite-dimensional projected-LAN statement identifies the information constant, and the corresponding uniform minimax theorem applies once the rough Volterra reduction and projected-rank inputs are uniform over the chosen primitive class. The estimator is therefore locally efficient within the projected Gaussian score experiment, and conditionally minimax over any rough Volterra class on which those uniform primitive inputs have been verified.

The local efficiency statement is operational. It separates weak signal, target-only contamination, basis misspecification, calibration error, and failures of the scalar moment class. It also supplies the noncentrality constants used in the synthetic calibration.

8.2. **A uniform minimax theorem.** Here "global" means uniform over a primitive rough Volterra class. The parameter is still estimated at its local contiguity scale. Let $\mathcal{V} = \mathcal{V}(H_-, H_+, M, \lambda_*)$ be a class of bivariate Volterra volatility models satisfying the primitive regularity assumptions, with

$$0 < H_- \le H_{XY} \le H_+ < \frac{1}{2},$$

C^2 Volterra amplitudes bounded by M, diagonal source-channel amplitude bounded away from zero, admissible proxy and block tuning, and boundary-layer projected Gram eigenvalue at least $\lambda_* > 0$. For $P \in \mathcal{V}$, write

$$\theta_n(P) = (\sqrt{N_n} h_n^{H_{XY}(P)})^{-1}$$

and express the rough local coordinate as

$$\eta = \eta_0 + \theta_n(P)h, \qquad h \in K \subset \mathbb{R}^m,$$

where K is compact. The local-coordinate estimator $\widehat{h}_n$ is the primitive output of the projected score procedure. To convert it into a model-coordinate estimator one also needs a feasible local scale. A feasible scale gives two readings of the minimax theorem. It is stratumwise when H_{XY} is fixed or known and $\theta_n(P)$ is part of the local normalization. It is adaptive when a pilot scale estimate is available.

Assumption 8.1 (Feasible rough-local scale)**.** Either the roughness stratum is fixed so that the normalization $\theta_n(P)$ is known on the class under consideration, or there exists a pilot $\widehat{\theta}_n$, computed on an independent or cross-fitted calibration sample, such that

$$\sup_{P\in\mathcal{V}} \sup_{\|h\|\le R} \left| \frac{\widehat{\theta}_n}{\theta_n(P)} - 1 \right| \xrightarrow{\mathbb{P}} 0$$

for every fixed local ball. In the first case set $\widetilde{\theta}_n = \theta_n(P)$; in the second set $\widetilde{\theta}_n = \widehat{\theta}_n$. The projected estimator on the model scale is

$$\widehat{\eta}_n^{\mathrm{pcov}} = \eta_0 + \widetilde{\theta}_n \widehat{h}_n^{\mathrm{pcov}}.$$

Without this scale-feasibility input, the theorem gives a minimax statement for the local coordinate h, not for the model coordinate η.

The adaptive reading of Assumption 8.1 is not an oracle assumption. A pilot roughness estimator is enough if its error is small on the logarithmic scale at which the rough local normalization changes.

Assumption 8.2 (Pilot roughness and smooth projected geometry)**.** There exists an independent or cross-fitted pilot $\widehat{H}_{XY}$ such that, for every fixed local ball and every $\varepsilon > 0$,

$$\sup_{P\in\mathcal{V}} \sup_{\|h\|\le R} \mathbb{P}_{P_{n,h}} \Big\{ |\widehat{H}_{XY} - H_{XY}(P)|\,|\log h_n| > \varepsilon \Big\} \to 0.$$

For $H \in [H_-, H_+]$, let

$$\mathcal{G}_n(H) = \Big(\Pi^{\mathrm{score}}_{\mathcal{B},n}(H), \Gamma_n(H), \Omega_n(H), \mathcal{I}_{\mathcal{B},n}(H) \Big)$$

denote the retained score projection, loading matrix, score covariance, and projected information matrix computed with roughness H. Whenever $r_n|\log h_n| \to 0$,

$$\begin{aligned}\sup_{P\in\mathcal{V}} \sup_{\|h\|\le R} \sup_{|\Delta|\le r_n} \Big[&\|\Pi^{\mathrm{score}}_{\mathcal{B},n}(H_{XY}(P)+\Delta) - \Pi^{\mathrm{score}}_{\mathcal{B},n}(H_{XY}(P))\|_{\mathrm{op}} \\ &+ \|\Gamma_n(H_{XY}(P)+\Delta) - \Gamma_n(H_{XY}(P))\|_{\mathrm{op}} \\ &+ \|\Omega_n(H_{XY}(P)+\Delta) - \Omega_n(H_{XY}(P))\|_{\mathrm{op}} \\ &+ \|\mathcal{I}_{\mathcal{B},n}(H_{XY}(P)+\Delta) - \mathcal{I}_{\mathcal{B},n}(H_{XY}(P))\|_{\mathrm{op}} \Big] \to 0.\end{aligned}$$

The same continuity holds for the feasible cross-fitted versions of these objects, conditionally on the calibration folds. In addition, if $\widehat{\overline{g}}_n(H)$ denotes the evaluation-fold projected score mean computed with roughness input H, then

$$\sup_{P\in\mathcal{V}} \sup_{\|h\|\le R} \sup_{|\Delta|\le r_n} \sqrt{N_n}\, \|\widehat{\overline{g}}_n(H_{XY}(P)+\Delta) - \widehat{\overline{g}}_n(H_{XY}(P))\| \xrightarrow{\mathbb{P}} 0$$

whenever $r_n|\log h_n| \to 0$.

The rate proposition translates logarithmic roughness stability into ordinary pilot convergence.

Proposition 8.3 (Interpreting the pilot rate)**.** *Suppose there is a deterministic sequence $a_n \downarrow 0$ such that*

$$a_n|\log h_n| \to 0$$

and

$$\widehat{H}_{XY} - H_{XY}(P) = O_{\mathbb{P}_{P_{n,h}}}(a_n)$$

uniformly over $P \in \mathcal{V}$ and $\|h\| \le R$. Then the pilot-rate condition in Assumption 8.2 holds. In particular, if $h_n = n^{-\kappa}$ for some $\kappa > 0$, any uniformly consistent pilot with polynomial rate $a_n = n^{-\rho}$, $\rho > 0$, is more than sufficient; logarithmic rates $a_n = (\log n)^{-1-\delta}$ are also sufficient.

Proof. For every $\varepsilon > 0$,

$$\mathbb{P}_{P_{n,h}}\{|\widehat{H}_{XY} - H_{XY}(P)|\,|\log h_n| > \varepsilon\} = \mathbb{P}_{P_{n,h}}\{|\widehat{H}_{XY} - H_{XY}(P)| > \varepsilon/|\log h_n|\}.$$

Uniform $O_{\mathbb{P}}(a_n)$ and $a_n|\log h_n| \to 0$ imply that the right-hand side converges to zero uniformly on local balls. The examples follow because $(\log n)n^{-\rho} \to 0$ and $(\log n)(\log n)^{-1-\delta} \to 0$. □

The smoothness lemma controls the motion of boundary-layer profiles in the Hurst coordinate.

Lemma 8.4 (Smooth rough boundary-layer geometry)**.** *Let $0 < H_- < H_+ < 1/2$. For every integer $q \ge 0$,*

$$\sup_{H\in[H_-,H_+]} \int_0^1 u^{2H-1}|\log u|^q\, du \le \frac{q!}{(2H_-)^{q+1}} < \infty.$$

Consequently $H \mapsto u^{H-1/2}$ is continuously differentiable as an $L^2(0,1)$-valued map, with derivative $u^{H-1/2}\log u$, uniformly on $[H_-, H_+]$. The same statement holds for any finite boundary-layer *dictionary obtained from these profiles by multiplication by uniformly bounded C^1 weights and by adding finitely many logarithmic derivative profiles. If the whitening matrices, target-only projections, and retained-basis Gram matrices are uniformly well-conditioned, then the projected information functional*

$$H \mapsto \mathcal{I}_{\mathcal{B},n}(H) = \Gamma_n(H)^\top \Omega_n(H)^{-1}\Gamma_n(H)$$

is locally Lipschitz uniformly over $H \in [H_-, H_+]$. If the unnormalized Volterra block arrays contain the rough factor h_n^H, the Lipschitz constant is allowed to grow at most as $C(1 + |\log h_n|)$.

Proof. The integral identity follows by the change of variables $x = -\log u$:

$$\int_0^1 u^{2H-1} |\log u|^q \, du = \int_0^\infty e^{-2Hx} x^q \, dx = \frac{q!}{(2H)^{q+1}}.$$

Uniform boundedness on $[H_-, H_+]$ follows because H_- is positive. The derivative formula is the pointwise derivative of $u^{H-1/2}$ in H, and the preceding bound with $q = 2$ gives an L^2 envelope for the difference quotients on compact subintervals of $(0, 1/2)$. Dominated convergence proves C^1 continuity in $L^2(0, 1)$.

Multiplication by bounded C^1 weights and adding finitely many logarithmic profiles preserves the same L^2 differentiability because the integrals above remain finite for higher powers of $|\log u|$. The matrices $\Gamma_n(H)$ and $\Omega_n(H)$ are finite collections of inner products and covariance functionals of these profiles after bounded linear transformations. Hence they inherit the same local Lipschitz bound. On a uniformly nonsingular set,

$$\Omega_n(H + \Delta)^{-1} - \Omega_n(H)^{-1} = \Omega_n(H + \Delta)^{-1}\{\Omega_n(H) - \Omega_n(H + \Delta)\}\Omega_n(H)^{-1},$$

so inversion is Lipschitz as well. Combining this identity with the product formula for $\Gamma^\top\Omega^{-1}\Gamma$ gives the stated Lipschitz property of $\mathcal{I}_{\mathcal{B},n}(H)$. Finally, differentiating h_n^H contributes the factor $(\log h_n)h_n^H$, which is the only source of the displayed logarithmic growth. □

The GMM-stability lemma converts profile and covariance perturbations into estimator perturbations.

Lemma 8.5 (Stability of the projected GMM map)**.** *Let*

$$\Psi(S, \Gamma, \Omega) = (\Gamma^\top\Omega^{-1}\Gamma)^{-1}\Gamma^\top\Omega^{-1}S.$$

Fix constants $M < \infty$, $0 < \omega_- \le \omega_+ < \infty$, and $0 < i_- \le i_+ < \infty$. Then Ψ is Lipschitz on the set of triples (S, Γ, Ω) satisfying

$$\|S\| \le M, \qquad \|\Gamma\| \le M, \qquad \|\Omega\| \le \omega_+, \qquad \lambda_{\min}(\Omega) \ge \omega_-,$$

and

$$i_- \le \lambda_{\min}(\Gamma^\top\Omega^{-1}\Gamma) \le \lambda_{\max}(\Gamma^\top\Omega^{-1}\Gamma) \le i_+.$$

There exists $C = C(M, \omega_-, \omega_+, i_-, i_+)$ such that for any two such triples,

$$\|\Psi(S_1, \Gamma_1, \Omega_1) - \Psi(S_2, \Gamma_2, \Omega_2)\| \le C\big(\|S_1 - S_2\| + \|\Gamma_1 - \Gamma_2\| + \|\Omega_1 - \Omega_2\|\big).$$

Proof. Write

$$A(\Gamma, \Omega) = \Gamma^\top\Omega^{-1}\Gamma, \qquad K(\Gamma, \Omega) = A(\Gamma, \Omega)^{-1}\Gamma^\top\Omega^{-1}, \qquad \Psi(S, \Gamma, \Omega) = K(\Gamma, \Omega)S.$$

On the displayed set,

$$\|\Omega^{-1}\| \le \omega_-^{-1}, \qquad \|A(\Gamma, \Omega)^{-1}\| \le i_-^{-1}.$$

For two pairs (Γ_a, Ω_a), $a = 1, 2$,

$$\begin{aligned} A_1 - A_2 = {} & (\Gamma_1 - \Gamma_2)^\top\Omega_1^{-1}\Gamma_1 + \Gamma_2^\top\Omega_1^{-1}(\Gamma_1 - \Gamma_2) \\ & + \Gamma_2^\top(\Omega_1^{-1} - \Omega_2^{-1})\Gamma_2. \end{aligned}$$

The inverse-difference identity

$$\Omega_1^{-1} - \Omega_2^{-1} = \Omega_1^{-1}(\Omega_2 - \Omega_1)\Omega_2^{-1}$$

and the uniform norm bounds imply

$$\|A_1 - A_2\| \le C_1\big(\|\Gamma_1 - \Gamma_2\| + \|\Omega_1 - \Omega_2\|\big).$$

Applying the same inverse-difference identity to $A_1^{-1} - A_2^{-1}$ gives

$$\|A_1^{-1} - A_2^{-1}\| \le C_2\|A_1 - A_2\| \le C_3\big(\|\Gamma_1 - \Gamma_2\| + \|\Omega_1 - \Omega_2\|\big).$$

Therefore

$$\begin{aligned} K_1 - K_2 &= (A_1^{-1} - A_2^{-1})\Gamma_1^\top \Omega_1^{-1} + A_2^{-1}(\Gamma_1 - \Gamma_2)^\top \Omega_1^{-1} \\ &\quad + A_2^{-1}\Gamma_2^\top(\Omega_1^{-1} - \Omega_2^{-1}), \end{aligned}$$

so

$$\|K_1 - K_2\| \le C_4\big(\|\Gamma_1 - \Gamma_2\| + \|\Omega_1 - \Omega_2\|\big).$$

Finally,

$$\begin{aligned} \|\Psi_1 - \Psi_2\| &\le \|K_1\|\,\|S_1 - S_2\| + \|K_1 - K_2\|\,\|S_2\| \\ &\le C\big(\|S_1 - S_2\| + \|\Gamma_1 - \Gamma_2\| + \|\Omega_1 - \Omega_2\|\big), \end{aligned}$$

using the uniform bounds on $\|K_1\|$ and $\|S_2\|$. □

Assumption 8.6 (Uniform rough LAN and rank)**.** Uniformly over $P \in \mathcal{V}$ and $h \in K$, the primitive observed experiment is Le Cam equivalent, up to $o(1)$ deficiency distance, to the projected Gaussian shift

$$S_P = \mathcal{I}_{\mathcal{B}}(P)h + \xi_P, \qquad \xi_P \sim N(0, \mathcal{I}_{\mathcal{B}}(P)).$$

The information matrices are uniformly nonsingular and bounded:

$$0 < i_- \le \lambda_{\min}\{\mathcal{I}_{\mathcal{B}}(P)\} \le \lambda_{\max}\{\mathcal{I}_{\mathcal{B}}(P)\} \le i_+ < \infty, \qquad P \in \mathcal{V}.$$

The feasible projected-score objects $\widehat{\Gamma}_n$, $\widehat{\Omega}_n$, the whitening map, and the target-only projection satisfy the uniform analog of Assumption 9.1 over $P \in \mathcal{V}$ and $h \in K$. For the quadratic-risk assertion, the normalized local estimation errors of the projected estimator are uniformly square-integrable on local balls; equivalently, one may work with a version truncated to a radius $R_n \to \infty$ in local coordinates.

Roughness-pilot error is separated from projected local recovery in the adaptive statistic.

Theorem 8.7 (Pilot roughness transfer and adaptive projected minimax)**.** *Assume Assumptions 8.2 and 8.6. Define*

$$\widehat{\theta}_n = (\sqrt{N_n} h_n^{\widehat{H}_{XY}})^{-1}, \qquad \theta_n(P) = (\sqrt{N_n} h_n^{H_{XY}(P)})^{-1}.$$

Let $\widehat{h}_n^{\mathrm{pcov}}(H)$ be the projected local-coordinate estimator computed with roughness input H, and let the adaptive model-scale estimator be

$$\widehat{\eta}_n^{\mathrm{ad}} = \eta_0 + \widehat{\theta}_n \widehat{h}_n^{\mathrm{pcov}}(\widehat{H}_{XY}).$$

Then, on every fixed local ball,

$$\sup_{P\in\mathcal{V}} \sup_{\|h\|\le R} \left| \frac{\widehat{\theta}_n}{\theta_n(P)} - 1 \right| \xrightarrow{\mathbb{P}} 0$$

and

$$\sup_{P\in\mathcal{V}} \sup_{\|h\|\le R} \left\| \widehat{h}_n^{\mathrm{pcov}}(\widehat{H}_{XY}) - \widehat{h}_n^{\mathrm{pcov}}(H_{XY}(P)) \right\| \xrightarrow{\mathbb{P}} 0.$$

Consequently

$$\widehat{\eta}_n^{\mathrm{ad}} - \{\eta_0 + \theta_n(P)h\} = \theta_n(P)\{\widehat{h}_n^{\mathrm{pcov}}(H_{XY}(P)) - h\} + o_{\mathbb{P}}\{\theta_n(P)\}$$

uniformly over $P \in \mathcal{V}$ and $\|h\| \le R$. In particular, whenever the oracle projected estimator attains the upper bound in Theorem 8.8, the pilot-adaptive estimator $\widehat{\eta}_n^{\mathrm{ad}}$ attains the same model-scale rate and the same normalized quadratic-risk constant m.

Proof. For the scale, take logarithms:

$$\log \frac{\widehat{\theta}_n}{\theta_n(P)} = -\{\widehat{H}_{XY} - H_{XY}(P)\} \log h_n.$$

The pilot condition forces the right-hand side to be $o_{\mathbb{P}}(1)$ uniformly on every local ball. Continuity of the exponential gives $\widehat{\theta}_n/\theta_n(P) \xrightarrow{\mathbb{P}} 1$ uniformly.

Let

$$E_{n,R}(\varepsilon) = \left\{ |\widehat{H}_{XY} - H_{XY}(P)| \, |\log h_n| \leq \varepsilon \right\}.$$

The preceding assumption gives $\mathbb{P}_{P_{n,h}}\{E_{n,R}(\varepsilon_n)\} \to 1$ for some $\varepsilon_n \downarrow 0$, uniformly in P and h. On this event the roughness perturbation is $r_n = \varepsilon_n/|\log h_n|$, so the smooth-geometry condition gives

$$\|\mathcal{G}_n(\widehat{H}_{XY}) - \mathcal{G}_n(H_{XY}(P))\| = o(1)$$

in the combined projection/loading/covariance/information norm, and the same is true for the evaluation-fold score mean $\widehat{\bar{g}}_n(H)$. By Assumption 8.6, the scaled score means are uniformly tight on local balls and the feasible matrices are uniformly well-conditioned. Hence, for every $\delta > 0$, one can choose deterministic constants $M_\delta < \infty$, $0 < \omega_- \leq \omega_+ < \infty$, and $0 < i_- \leq i_+ < \infty$, independent of P and h, such that both the oracle and plug-in triples

$$(\sqrt{N_n}\widehat{\bar{g}}_n(H_{XY}(P)), \widehat{\Gamma}_n(H_{XY}(P)), \widehat{\Omega}_n(H_{XY}(P)))$$

and

$$(\sqrt{N_n}\widehat{\bar{g}}_n(\widehat{H}_{XY}), \widehat{\Gamma}_n(\widehat{H}_{XY}), \widehat{\Omega}_n(\widehat{H}_{XY}))$$

lie in the lemma's stability set with probability tending to at least $1 - \delta$, uniformly over $P \in \mathcal{V}$ and $\|h\| \leq R$. Writing

$$\Psi(S, \Gamma, \Omega) = (\Gamma^\top \Omega^{-1} \Gamma)^{-1} \Gamma^\top \Omega^{-1} S,$$

the estimator can be written as

$$\widehat{h}_n^{\mathrm{pcov}}(H) = \Psi(\sqrt{N_n}\widehat{\bar{g}}_n(H), \widehat{\Gamma}_n(H), \widehat{\Omega}_n(H)).$$

Lemma 8.5 therefore gives

$$\begin{aligned}
\|\widehat{h}_n^{\mathrm{pcov}}(\widehat{H}_{XY}) - \widehat{h}_n^{\mathrm{pcov}}(H_{XY}(P))\| \leq C\Big[&\sqrt{N_n}\|\widehat{\bar{g}}_n(\widehat{H}_{XY}) - \widehat{\bar{g}}_n(H_{XY}(P))\| \\
&+ \|\widehat{\Gamma}_n(\widehat{H}_{XY}) - \widehat{\Gamma}_n(H_{XY}(P))\| \\
&+ \|\widehat{\Omega}_n(\widehat{H}_{XY}) - \widehat{\Omega}_n(H_{XY}(P))\|\Big] = o_{\mathbb{P}}(1),
\end{aligned}$$

uniformly on local balls.

Now decompose

$$\begin{aligned}
\widehat{\eta}_n^{\mathrm{ad}} - \eta_0 - \theta_n(P)h = {}& \theta_n(P)\{\widehat{h}_n^{\mathrm{pcov}}(H_{XY}(P)) - h\} \\
&+ \theta_n(P)\{\widehat{h}_n^{\mathrm{pcov}}(\widehat{H}_{XY}) - \widehat{h}_n^{\mathrm{pcov}}(H_{XY}(P))\} \\
&+ \{\widehat{\theta}_n - \theta_n(P)\}\widehat{h}_n^{\mathrm{pcov}}(\widehat{H}_{XY}).
\end{aligned}$$

The second term is $o_{\mathbb{P}}\{\theta_n(P)\}$ by the plug-in equivalence. The third term is also $o_{\mathbb{P}}\{\theta_n(P)\}$, because $\widehat{\theta}_n/\theta_n(P) - 1 = o_{\mathbb{P}}(1)$ and the local estimator is tight under the uniform LAN and rank assumptions. The displayed expansion follows. Substituting this expansion into the normalized quadratic risk shows that

$$\begin{aligned}
&\theta_n(P)^{-2}(\widehat{\eta}_n^{\mathrm{ad}} - \eta_0 - \theta_n(P)h)^\top \mathcal{I}_{\mathcal{B}}(P)(\widehat{\eta}_n^{\mathrm{ad}} - \eta_0 - \theta_n(P)h) \\
&\quad = (\widehat{h}_n^{\mathrm{pcov}}(H_{XY}(P)) - h)^\top \mathcal{I}_{\mathcal{B}}(P)(\widehat{h}_n^{\mathrm{pcov}}(H_{XY}(P)) - h) + o_{\mathbb{P}}(1)
\end{aligned}$$

uniformly on local balls. The square-integrability or truncation condition from Assumption 8.6 upgrades this $o_{\mathbb{P}}(1)$ statement to convergence of the corresponding risks. Hence any oracle upper bound, and in particular the sharp constant m from Theorem 8.8, transfers to the adaptive estimator. The lower bound is unchanged because the minimax infimum already ranges over all estimators, including procedures that use data-driven pilots. $\square$

The minimax theorem identifies the common information constant for the Gaussian lower-bound experiment and the feasible projected upper bound.

Theorem 8.8 (Uniform minimax theorem over rough Volterra classes)**.** *Assume Assumptions 8.1 and 8.6 on every local ball $K_R = \{h \in \mathbb{R}^m : \|h\| \le R\}$. Then $\theta_n(P)$ is the minimax rate for the rough model coordinate η over $\mathcal{V}$ at the tested local scale. The sharp lower bound is*

$$\lim_{R\to\infty} \liminf_{n\to\infty} \inf_{\widetilde{\eta}_n} \sup_{P\in\mathcal{V}} \sup_{\|h\|\le R} \theta_n(P)^{-2}\mathbb{E}_{P_{n,h}}\Big[(\widetilde{\eta}_n - \eta_0 - \theta_n(P)h)^\top \mathcal{I}_{\mathcal{B}}(P)(\widetilde{\eta}_n - \eta_0 - \theta_n(P)h)\Big] \ge m.$$

The projected covariance-score estimator attains this bound on every fixed local ball:

$$\lim_{n\to\infty} \sup_{P\in\mathcal{V}} \sup_{\|h\|\le R} \Big|\theta_n(P)^{-2}\mathbb{E}_{P_{n,h}}\Big[(\widehat{\eta}_n^{\mathrm{pcov}} - \eta_0 - \theta_n(P)h)^\top \mathcal{I}_{\mathcal{B}}(P)(\widehat{\eta}_n^{\mathrm{pcov}} - \eta_0 - \theta_n(P)h)\Big] - m\Big| = 0.$$

Equivalently, in local coordinates,

$$\widehat{h}_n^{\mathrm{pcov}} = h + \mathcal{I}_{\mathcal{B}}(P)^{-1/2}Z + o_{\mathbb{P}}(1), \qquad Z \sim N(0, I_m),$$

uniformly over the class, and no estimator can uniformly improve the normalized quadratic-risk constant below m in the projected rough local experiment.

Proof. For fixed P, the Gaussian limit experiment from Assumption 8.6 has observation

$$S_P = \mathcal{I}_{\mathcal{B}}(P)h + \xi_P, \qquad \xi_P \sim N(0, \mathcal{I}_{\mathcal{B}}(P)).$$

Equivalently,

$$Y_P := \mathcal{I}_{\mathcal{B}}(P)^{-1/2}S_P = \mathcal{I}_{\mathcal{B}}(P)^{1/2}h + Z, \qquad Z \sim N(0, I_m).$$

In that standardized shift experiment, the quadratic loss in local coordinates is simply $\|\widetilde{u} - \mathcal{I}_{\mathcal{B}}(P)^{1/2}h\|^2$, so the canonical equivariant estimator is $\widehat{u}_P = Y_P$, corresponding to

$$\widehat{h}_P = \mathcal{I}_{\mathcal{B}}(P)^{-1}S_P.$$

Its local error is

$$\widehat{h}_P - h = \mathcal{I}_{\mathcal{B}}(P)^{-1/2}Z, \qquad Z \sim N(0, I_m).$$

Therefore

$$\mathbb{E}[(\widehat{h}_P - h)^\top \mathcal{I}_{\mathcal{B}}(P)(\widehat{h}_P - h)] = \mathbb{E}\|Z\|^2 = m.$$

For the unrestricted Gaussian location experiment the exact minimax quadratic risk is m, and the same constant is approached on expanding balls $\|h\| \le R$; this is the classical finite-dimensional benchmark underlying the local asymptotic minimax theory [24, Chapter 8]. To transfer this lower bound back to the primitive experiment, fix $M < \infty$ and truncate the normalized quadratic loss at M. Uniform Le Cam equivalence then implies that the difference between the truncated minimax risks in the observed and Gaussian experiments is $o(1)$, uniformly over $P \in \mathcal{V}$ and $\|h\| \le R$. Sending first $n \to \infty$, then $R \to \infty$, and finally $M \to \infty$, the uniform square-integrability assumption removes the truncation and yields the displayed lower bound m.

For the upper bound, the uniform version of Theorem 9.2 gives

$$\widehat{h}_n^{\mathrm{pcov}} \xrightarrow{d} N(h, \mathcal{I}_{\mathcal{B}}(P)^{-1})$$

uniformly over the same index set. Equivalently,

$$\mathcal{I}_{\mathcal{B}}(P)^{1/2}\{\widehat{h}_n^{\mathrm{pcov}} - h\} \xrightarrow{d} N(0, I_m)$$

uniformly over $P \in \mathcal{V}$ and $\|h\| \le R$. The uniform square-integrability or truncation condition in Assumption 8.6 upgrades this distributional limit to convergence of the normalized local quadratic risk to $\mathbb{E}\|Z\|^2 = m$, where $Z \sim N(0, I_m)$. Assumption 8.1 gives

$$\widehat{\eta}_n^{\mathrm{pcov}} - \eta = \theta_n(P)(\widehat{h}_n^{\mathrm{pcov}} - h) + o_{\mathbb{P}}\{\theta_n(P)\}$$

uniformly on local balls. The scale remainder is negligible in the normalized quadratic risk by the uniform eigenvalue bounds and the square-integrability/truncation condition in Assumption 8.6. Thus the model-coordinate risk is $\theta_n(P)^2$ times the local-coordinate risk up to $o(1)$, uniformly over $P \in \mathcal{V}$ and $\|h\| \le R$, proving both the rate statement and the sharp constant. $\square$

Remark 8.9 (Primitive verification requirements). Theorem 8.8 is stated through primitive conditions. Once the observed high-frequency experiment is uniformly reduced to the projected Gaussian shift and rough projected rank is uniformly bounded away from zero, the projected estimator has no remaining efficiency gap inside that experiment. The primitive verification task is localized: prove the uniform Volterra boundary-layer reduction, prove the uniform rank theorem below, and choose tuning rules for which proxy error, dictionary approximation error, whitening regularization, and HAC error are all smaller than the local stochastic scale. If H_{XY} is unknown rather than fixed stratumwise, Theorem 8.7 gives the adaptation requirement. The pilot roughness estimate must satisfy $|\widehat{H}_{XY} - H_{XY}|\,|\log h_n| = o_{\mathbb{P}}(1)$, and the projected information geometry must make the plug-in score space first-order equivalent to the oracle score space.

8.3. **Gaussian reduction and quasi-score scope.** The covariance-score construction is exact in the Gaussian reduced experiment. In that experiment,

$$\ell(B) = \frac{1}{2}\{U^\top BU - \operatorname{tr}(B)\}$$

is the likelihood score for covariance perturbations, the score inner product is $\frac{1}{2}\operatorname{tr}(BC)$, and the target/source information decomposition is an exact orthogonal decomposition of Gaussian LAN information. On this reduced experiment, the projected estimator has closed-form information constants and a source-orthogonal oracle bound.

The feasible statistic also has a quasi-score interpretation. Proposition 7.3 shows that exact Gaussianity is not needed for the studentized GMM and Wald limits once the retained quadratic scores satisfy the mean expansion and HAC CLT in Assumption 7.1. Outside the Gaussian reduced experiment, the constants change their meaning. Γ and Ω are the true mean derivative and long-run covariance of the retained quadratic scores, and $\Gamma^\top\Omega^{-1}\Gamma$ is projected quasi-information rather than a likelihood information bound. The numerical constant-matching exercises check the exact Gaussian reduced experiment; a non-Gaussian robustness exercise would check the quasi-score extension.

8.4. **Multi-parameter projected information.** The construction extends to several local directions. Let $\vartheta \in \mathbb{R}^m$ collect local roughness and contagion parameters, for example a target roughness component and a source-to-target rough component. Define

$$A_{n,j}^\perp = (I - \Pi_{Y,n})W_n\dot{\Sigma}_{n,j}W_n, \qquad 1 \le j \le m,$$

be the target-orthogonal covariance derivatives. For projected basis scores $g_{r,k,n}$, define the loading matrix

$$\Gamma_{n,rj} = \frac{1}{2}\operatorname{tr}(B_{r,n}A_{n,j}^\perp), \qquad \Gamma_n \in \mathbb{R}^{d_n\times m}.$$

The projected information matrix is

$$\mathcal{I}_{\mathcal{B},n}(\vartheta) = \Gamma_n^\top\Omega_n^{-1}\Gamma_n.$$

The multi-parameter local model is identified by the projected estimator if

$$\operatorname{rank}(\Gamma_n) = m \quad \text{equivalently} \quad \mathcal{I}_{\mathcal{B},n}(\vartheta) \text{ is nonsingular.}$$

The projected information matrix is the joint-estimation object for target roughness and source-to-target roughness. A scalar statistic can be locally blind to one of these directions. The projected covariance score reports the rank, weak eigenvalues, and block directions carrying the identifying information.

8.5. **P- and Q-measure information objects.** The same projected score construction applies under physical and risk-neutral inputs. Under the physical measure $\mathbb{P}$, the null covariance and local derivative are estimated from realized-volatility proxies. Under the pricing measure $\mathbb{Q}$, the same objects come from an option-calibrated rough Volterra model or from an option-implied covariance surface. For each measure $\mu \in \{\mathbb{P}, \mathbb{Q}\}$ one obtains

$$\widehat{\mathcal{I}}_\mu^{\mathrm{src}\perp Y}, \qquad \widehat{\mathcal{I}}_\mu^{\mathcal{B}}, \qquad \widehat{\vartheta}_\mu^{\mathrm{pcov}}, \qquad \widehat{\mathcal{I}}_{\mathcal{B},\mu}.$$

These are the measure-specific inputs to the projected P/Q interface in Section 9.3. The Oxford-Man application below does not implement this comparison on market data. The common object is the orthogonalized information geometry; the controlled experiments check the estimator only where the constants are known. A matched $\mathbb{P}/\mathbb{Q}$ empirical application would compare two evaluations of the same source-orthogonal information geometry under different probability measures, not two unrelated estimators.

The corresponding mismatch functionals are already defined at the level of the estimator. With a small ridge $\varepsilon > 0$, set

$$\widehat{D}_I = \log \frac{\widehat{\mathcal{I}}_Q^{\mathrm{src}\perp Y} + \varepsilon}{\widehat{\mathcal{I}}_P^{\mathrm{src}\perp Y} + \varepsilon}, \qquad \widehat{D}_{\mathcal{B}} = \log \frac{\widehat{\mathcal{I}}_Q^{\mathcal{B}} + \varepsilon}{\widehat{\mathcal{I}}_P^{\mathcal{B}} + \varepsilon}.$$

These log-ratios measure whether source-orthogonal contagion is more visible under the pricing geometry or under the physical geometry. For local coordinates, define

$$\widehat{\delta}_{PQ} = \widehat{\vartheta}_Q^{\mathrm{pcov}} - \widehat{\vartheta}_P^{\mathrm{pcov}}.$$

If the $\mathbb{P}$- and $\mathbb{Q}$-side inputs are obtained from independent calibration samples or from asymptotically independent sources, the natural Hausman-type discrepancy statistic is

$$H_{PQ} = \widehat{\delta}_{PQ}^{\top} \left(\widehat{V}_P + \widehat{V}_Q \right)^{-1} \widehat{\delta}_{PQ}, \qquad \widehat{V}_\mu = \left(\widehat{\mathcal{I}}_{\mathcal{B},\mu} \right)^{-1}, \quad \mu \in \{P, Q\}.$$

Finally, when both sides are identified, directional agreement can be summarized by the projected cosine

$$\widehat{c}_{PQ} = \frac{(\widehat{\vartheta}_P^{\mathrm{pcov}})^{\top} \widehat{\mathcal{I}}_{\mathcal{B},0} \widehat{\vartheta}_Q^{\mathrm{pcov}}}{\left[(\widehat{\vartheta}_P^{\mathrm{pcov}})^{\top} \widehat{\mathcal{I}}_{\mathcal{B},0} \widehat{\vartheta}_P^{\mathrm{pcov}} \right]^{1/2} \left[(\widehat{\vartheta}_Q^{\mathrm{pcov}})^{\top} \widehat{\mathcal{I}}_{\mathcal{B},0} \widehat{\vartheta}_Q^{\mathrm{pcov}} \right]^{1/2}},$$

where $\widehat{\mathcal{I}}_{\mathcal{B},0}$ is a common reference information matrix, for example the pooled or null-calibrated projected information. Paired synthetic experiments can therefore validate P/Q mismatch estimation without claiming access to market P/Q data.

8.6. **Identification in the rough regime.** The rough regime is not automatic. For $H < 1/2$, the projected estimator is identified only if the rough-local covariance derivatives have nonzero source-orthogonal projection and satisfy the multi-parameter rank condition above. In primitive Volterra terms, the boundary-layer derivative generated by the source-to-target rough channel must not be absorbed by the target-only tangent space:

$$(I - \Pi_Y) A_{H_{XY}} \neq 0, \qquad \operatorname{rank}\Big(\langle B_r, (I - \Pi_Y) A_j \rangle_{\mathrm{score}} \Big)_{r,j} = m.$$

For $H < 1/2$, proving this requires a dedicated boundary-layer argument. The singular Volterra kernel generates distinctive near-diagonal covariance shapes, often including logarithmic profiles when differentiating with respect to H, and those shapes must survive observable screening and realized-volatility proxy error. The basis therefore includes constant, front-loaded, and roughness-sensitive boundary-layer shapes, including log-front-loaded profiles.

The rough-regime claim has three parts.

(i) the projected estimator is designed for the rough regime because it uses covariance geometry rather than pathwise smoothness or long-memory estimators built for $H > 1/2$;
(ii) identification for $H < 1/2$ is equivalent, in the projected LAN experiment, to nonsingularity of the projected information matrix;
(iii) the Volterra boundary-layer proof must verify the rank theorem Theorem 8.11 for the kernels and proxy scheme used here.

This formulation isolates the primitive rough analytic step. The finite-dimensional estimator, Wald limit, and information decomposition do not change in the rough case. The singular boundary-layer derivatives must survive target-only projection and be approximated by the retained dictionary. When that criterion is verified, the estimator is identified in the rough regime. When the criterion is omitted, the diagnostic hierarchy still reports whether the tested rough source-screened direction is present in the reduced Gaussian experiment and whether the chosen basis captures it.

Assumption 8.10 (Rough projected rank)**.** Let $\vartheta \in \mathbb{R}^m$ collect the rough local coordinates to be estimated, and let

$$A_{n,j}^{\perp} = (I - \Pi_{Y,n}) W_n \dot{\Sigma}_{n,j} W_n, \qquad 1 \le j \le m,$$

be the target-orthogonal covariance derivatives generated by the primitive Volterra model. There exists a fixed rough dictionary $B_{1,n}, \dots, B_{d,n}$, containing boundary-layer and log-boundary-layer shapes, such that the loading matrices

$$\Gamma_{n,rj} = \frac{1}{2} \operatorname{tr}(B_{r,n} A_{n,j}^{\perp})$$

converge to a matrix Γ with rank m. Equivalently, the projected information matrix

$$\mathcal{I}_{\mathcal{B}} = \Gamma^{\top} \Omega^{-1} \Gamma$$

is nonsingular.

The rank theorem identifies when boundary-layer Volterra profiles survive target-only projection.

Theorem 8.11 (Volterra boundary-layer projected rank)**.** *Let $\langle B, C \rangle_{\text{score}} = \frac{1}{2} \operatorname{tr}(BC)$. For each rough local coordinate ϑ_j, suppose the whitened target-orthogonal covariance derivative*

$$A_{n,j}^{\perp} = (I - \Pi_{Y,n}) W_n \dot{\Sigma}_{n,j} W_n$$

has a frozen Volterra boundary-layer profile

$$\psi_j^{\perp} = (I - \Pi_{\mathfrak{Y}}) \mathcal{W}_t \dot{\mathbb{G}}_{j,t} \mathcal{W}_t \in \mathfrak{H}_{\perp}, \qquad 1 \le j \le m.$$

Here $\mathfrak{H}_{\perp}$ is the Hilbert closure of target-orthogonal boundary-layer covariance profiles, $\mathcal{W}_t$ is the limiting whitening map, and $\Pi_{\mathfrak{Y}}$ is the limiting target-only projection. For an amplitude coordinate in the rough source channel, $\dot{\mathbb{G}}_{j,t}$ is generated by the frozen kernel

$$u \mapsto b_j(t,t) u^{H_{XY} - 1/2};$$

for a Hurst-coordinate derivative it is generated by the logarithmic boundary-layer kernel

$$u \mapsto b(t,t) u^{H_{XY} - 1/2} \log u.$$

Assume the following four conditions hold.

(i) Boundary-layer profile convergence is uniform over the local derivative sphere:

$$\sup_{\|c\|=1} \left| \left\| \sum_{j=1}^{m} c_j A_{n,j}^{\perp} \right\|_{\text{score}}^2 - \left\| \sum_{j=1}^{m} c_j \psi_j^{\perp} \right\|_{\mathfrak{H}_{\perp}}^2 \right| \to 0.$$

(ii) The limiting source-screened Gram matrix

$$G_{\perp} = \left(\langle \psi_i^{\perp}, \psi_j^{\perp} \rangle_{\mathfrak{H}_{\perp}} \right)_{i,j}$$

has $\lambda_{\min}(G_{\perp}) \ge \lambda_* > 0$.

(iii) The retained rough dictionary $\mathcal{B}_n = \operatorname{span}(B_{1,n}, \dots, B_{d,n})$, containing boundary-layer and log-boundary-layer shapes, approximates the whole derivative span:

$$\sup_{\|c\|=1} \left\| (I - \Pi_{\mathcal{B}_n}^{\text{score}}) \sum_{j=1}^{m} c_j A_{n,j}^{\perp} \right\|_{\text{score}} \to 0.$$

(iv) After replacing the retained dictionary by any score-orthonormal basis of the same span, the retained score covariance matrices satisfy

$$0 < \omega_- \le \lambda_{\min}(\Omega_n) \le \lambda_{\max}(\Omega_n) \le \omega_+ < \infty.$$

Then the rough projected-rank condition holds quantitatively:

$$\liminf_{n\to\infty} \lambda_{\min}\left(\Gamma_n^\top \Omega_n^{-1}\Gamma_n\right) \geq \frac{\lambda_*}{4\omega_+} > 0, \qquad \Gamma_{n,rj} = \frac{1}{2}\operatorname{tr}(B_{r,n}A_{n,j}^\perp).$$

In particular, the projected covariance-score experiment identifies the m rough local coordinates after target-only screening.

Proof. Let

$$A_{n,c}^\perp = \sum_{j=1}^m c_j A_{n,j}^\perp, \qquad \psi_c^\perp = \sum_{j=1}^m c_j \psi_j^\perp, \qquad \|c\| = 1.$$

By boundary-layer convergence and the Gram lower bound,

$$\inf_{\|c\|=1} \|A_{n,c}^\perp\|_{\text{score}}^2 \geq \frac{\lambda_*}{2}$$

for all large n. Let $P_n = \Pi_{\mathcal{B}_n}^{\text{score}}$. Since P_n is an orthogonal projection,

$$\|P_n A_{n,c}^\perp\|_{\text{score}}^2 = \|A_{n,c}^\perp\|_{\text{score}}^2 - \|(I-P_n)A_{n,c}^\perp\|_{\text{score}}^2 \geq \frac{\lambda_*}{4}$$

uniformly over $\|c\| = 1$, again for all large n. After score-orthonormalizing the retained basis, the coordinate vector of $P_n A_{n,c}^\perp$ is exactly $\Gamma_n c$. Hence

$$\inf_{\|c\|=1} \|\Gamma_n c\|^2 \geq \frac{\lambda_*}{4}.$$

Using $\lambda_{\max}(\Omega_n) \leq \omega_+$,

$$c^\top \Gamma_n^\top \Omega_n^{-1}\Gamma_n c \geq \omega_+^{-1}\|\Gamma_n c\|^2 \geq \frac{\lambda_*}{4\omega_+}$$

uniformly over the unit sphere. This proves the quantitative lower bound, hence full projected rank and Assumption 8.10. □

The corollary transfers pointwise profile rank to a uniform rough class.

Corollary 8.12 (Uniform rough projected rank)**.** *Suppose the four hypotheses of Theorem 8.11 hold uniformly over a rough Volterra class $\mathcal{V}$, with common constants $\lambda_* > 0$ and $\omega_+ < \infty$. Then the rank part of Assumption 8.6 holds uniformly:*

$$\inf_{P\in\mathcal{V}} \lambda_{\min}\{\mathcal{I}_{\mathcal{B}}(P)\} \geq \frac{\lambda_*}{4\omega_+} > 0.$$

Consequently, once the observed-to-Gaussian reduction and feasible covariance/projection stability are also uniform over $\mathcal{V}$, the projected estimator falls under the uniform minimax theorem.

Proof. Apply Theorem 8.11 with constants independent of P. The displayed lower bound is exactly the uniform version of its quantitative information lower bound. The final sentence is then the definition of Assumption 8.6. □

Projected rank gives a nondegenerate LAN experiment.

Theorem 8.13 (Rough-regime identification conditional on projected rank)**.** *Assume the primitive reduction inputs, Assumption 9.1, and Assumption 8.10. Then the projected covariance-score estimator identifies the m-dimensional rough local coordinate in the regime $H < 1/2$, and*

$$\widehat{\vartheta}_n^{\text{pcov}} \xrightarrow{d} N(\vartheta, \mathcal{I}_{\mathcal{B}}^{-1}).$$

The Wald statistic for $H_0 : \vartheta = 0$ has the local noncentral chi-square limit

$$W_n \xrightarrow{d} \chi_m^2(\vartheta^\top \mathcal{I}_{\mathcal{B}}\vartheta).$$

Proof. Assumption 8.10 gives nonsingularity of $\mathcal{I}_{\mathcal{B}} = \Gamma^\top\Omega^{-1}\Gamma$. The conclusion is then exactly the vector projected GMM limit in Theorem 9.2. The role of the rough Volterra analysis is not to change the finite-dimensional GMM proof; it verifies that, for $H < 1/2$, the projected covariance derivatives created by the boundary-layer kernels survive target-only projection and span the local coordinates. □

Remark 8.14 (Connection to minimax efficiency). Theorem 8.13 is the pointwise identification and local-efficiency statement in the projected LAN experiment. Theorem 8.8 gives the uniform minimax upgrade. If the boundary-layer rank lower bound and the observed-to-Gaussian reduction hold uniformly over a rough Volterra class, the same projected estimator attains the sharp local quadratic-risk constant over that class. The asymptotic statement is complete at the projected-score level. Applying it to a specified primitive class requires verification of the uniform Volterra reduction hypotheses.

9. Vector LAN and P/Q information interface

The scalar projected score gives the simplest LAN statement. The vector extension separates the remaining econometric objects. Theorem 4.7 identifies the source-screened nuisance removal, the rough-rank condition gives identification in the rough regime, and the $\mathbb{P}/\mathbb{Q}$ comparison procedure gives the matched-measure interface for option-panel applications.

9.1. Vector projected LAN.

Assumption 9.1 (Vector projected LAN)**.** Let $\vartheta \in \mathbb{R}^m$ be the local coordinate vector and let $g_{k,n} \in \mathbb{R}^{d_n}$ be the retained projected score vector, with $d_n \leq d < \infty$. Along the tested sequence of local experiments,

$$\sqrt{N_n}\,\bar{g}_n = \Gamma_n \vartheta + \xi_n + o_{\mathbb{P}_{n,\vartheta}}(1), \qquad \xi_n \xrightarrow{d} N(0,\Omega),$$

uniformly for ϑ in compact sets, with $\Gamma_n \to \Gamma$ and $\Omega_n \to \Omega$ positive definite. The feasible objects satisfy

$$\widehat{\Gamma}_n \xrightarrow{\mathbb{P}} \Gamma, \qquad \widehat{\Omega}_n \xrightarrow{\mathbb{P}} \Omega, \qquad \sqrt{N_n}(\widehat{\bar{g}}_n - \bar{g}_n) \xrightarrow{\mathbb{P}} 0.$$

The vector theorem identifies joint local recovery and the Wald noncentrality determined by the projected information matrix.

Theorem 9.2 (Vector projected GMM and Wald limit)**.** *Assume Assumption 9.1 and suppose*

$$\mathcal{I}_{\mathcal{B}} := \Gamma^\top \Omega^{-1} \Gamma$$

is nonsingular. Define

$$\widehat{\vartheta}_n^{\mathrm{pcov}} = (\widehat{\Gamma}_n^\top \widehat{\Omega}_n^{-1} \widehat{\Gamma}_n)^{-1} \widehat{\Gamma}_n^\top \widehat{\Omega}_n^{-1} \sqrt{N_n}\,\widehat{\bar{g}}_n.$$

Then

$$\widehat{\vartheta}_n^{\mathrm{pcov}} \xrightarrow{d} N(\vartheta, \mathcal{I}_{\mathcal{B}}^{-1}).$$

Moreover, for the Wald statistic

$$W_n = (\widehat{\vartheta}_n^{\mathrm{pcov}})^\top \widehat{\mathcal{I}}_{\mathcal{B},n} \widehat{\vartheta}_n^{\mathrm{pcov}}, \qquad \widehat{\mathcal{I}}_{\mathcal{B},n} = \widehat{\Gamma}_n^\top \widehat{\Omega}_n^{-1} \widehat{\Gamma}_n,$$

we have, under the local coordinate ϑ,

$$W_n \xrightarrow{d} \chi^2_m(\lambda(\vartheta)), \qquad \lambda(\vartheta) = \vartheta^\top \mathcal{I}_{\mathcal{B}} \vartheta.$$

In particular, under the null $W_n \xrightarrow{d} \chi^2_m$, while along a one-dimensional local path $\vartheta = \tau w$,

$$\lambda(\tau) = \tau^2 w^\top \mathcal{I}_{\mathcal{B}} w.$$

Proof. Write $\widehat{S}_n = \sqrt{N_n}\widehat{\bar{g}}_n$ and $S_n = \sqrt{N_n}\bar{g}_n$. By Assumption 9.1,

$$S_n = \Gamma_n \vartheta + \xi_n + o_{\mathbb{P}}(1), \qquad \xi_n \xrightarrow{d} N(0,\Omega).$$

The feasible replacement condition gives $\widehat{S}_n - S_n \xrightarrow{\mathbb{P}} 0$, so also

$$\widehat{S}_n = \Gamma_n \vartheta + \xi_n + o_{\mathbb{P}}(1) = \Gamma \vartheta + \xi_n + o_{\mathbb{P}}(1).$$

Now write

$$\Psi(S,\Gamma,\Omega) = (\Gamma^\top \Omega^{-1} \Gamma)^{-1} \Gamma^\top \Omega^{-1} S.$$

Because $\mathcal{I}_{\mathcal{B}} = \Gamma^\top \Omega^{-1} \Gamma$ is nonsingular and $\widehat{\Gamma}_n \xrightarrow{\mathbb{P}} \Gamma$, $\widehat{\Omega}_n \xrightarrow{\mathbb{P}} \Omega$, the score vectors are tight and the matrices are uniformly well-conditioned on events of arbitrarily high probability. Therefore, for every $\delta > 0$, one can choose deterministic constants $M_\delta, \omega_\pm, i_\pm$ such that the triples $(\widehat{S}_n, \widehat{\Gamma}_n, \widehat{\Omega}_n)$

and $(\Gamma\vartheta + \xi_n, \Gamma, \Omega)$ lie in the stability region of Lemma 8.5 with probability at least $1 - \delta - o(1)$. Applying the lemma on that event, and then letting $\delta \downarrow 0$, gives

$$\widehat{\vartheta}_n^{\mathrm{pcov}} = \Psi(\widehat{S}_n, \widehat{\Gamma}_n, \widehat{\Omega}_n) = \Psi(\Gamma\vartheta + \xi_n, \Gamma, \Omega) + o_{\mathbb{P}}(1) = \vartheta + \mathcal{I}_{\mathcal{B}}^{-1}\Gamma^\top\Omega^{-1}\xi_n + o_{\mathbb{P}}(1).$$

The limiting noise is centered Gaussian with covariance

$$\mathcal{I}_{\mathcal{B}}^{-1}\Gamma^\top\Omega^{-1}\Omega\Omega^{-1}\Gamma\mathcal{I}_{\mathcal{B}}^{-1} = \mathcal{I}_{\mathcal{B}}^{-1}.$$

This proves the asymptotic normality. For the Wald statistic, multiply the limit by $\mathcal{I}_{\mathcal{B}}^{1/2}$:

$$\mathcal{I}_{\mathcal{B}}^{1/2}\widehat{\vartheta}_n^{\mathrm{pcov}} \xrightarrow{d} N(\mathcal{I}_{\mathcal{B}}^{1/2}\vartheta, I_m).$$

The squared Euclidean norm of this Gaussian vector is $\chi^2_m(\vartheta^\top\mathcal{I}_{\mathcal{B}}\vartheta)$. The one-dimensional path formula follows by substituting $\vartheta = \tau v$ into $\lambda(\vartheta) = \vartheta^\top\mathcal{I}_{\mathcal{B}}\vartheta$. □

The diagnostic corollary separates observed Wald ratios from size-adjusted local power.

Corollary 9.3 (Wald-ratio and size-adjusted power diagnostics)**.** *Assume the conditions of Theorem 9.2. Fix a local coordinate ϑ and suppose the Wald statistics are uniformly integrable under this local sequence. Then*

$$\mathbb{E}_\vartheta W_n \to m + \lambda(\vartheta), \qquad \lambda(\vartheta) = \vartheta^\top\mathcal{I}_{\mathcal{B}}\vartheta.$$

Consequently, for an independent Monte Carlo experiment with $R_n \to \infty$ replications at the same local coordinate,

$$\frac{R_n^{-1}\sum_{r=1}^{R_n} W_{n,r}}{m + \lambda(\vartheta)} \xrightarrow{\mathbb{P}} 1$$

whenever $m + \lambda(\vartheta) > 0$. If $\widehat{q}^{\,0}_{n,\alpha}$ is an empirical null critical value satisfying

$$\widehat{q}^{\,0}_{n,\alpha} \xrightarrow{\mathbb{P}} q_{m,1-\alpha},$$

where $q_{m,1-\alpha}$ is the $1-\alpha$ quantile of χ^2_m, then the size-adjusted rejection probability under ϑ converges to

$$\mathbb{P}\left\{\chi^2_m(\lambda(\vartheta)) > q_{m,1-\alpha}\right\}.$$

Along a one-dimensional path $\vartheta = \tau w$, this limit is the noncentral chi-square power curve with noncentrality $\tau^2 w^\top\mathcal{I}_{\mathcal{B}} w$.

Proof. Theorem 9.2 gives

$$W_n \xrightarrow{d} \chi^2_m(\lambda(\vartheta)).$$

Uniform integrability upgrades distributional convergence to convergence of first moments, and the mean of a noncentral chi-square with m degrees of freedom and noncentrality λ is $m+\lambda$. The Monte Carlo ratio follows from the law of large numbers, conditionally on the design, together with the preceding moment convergence. For the size-adjusted power statement, Slutsky's theorem gives $W_n - \widehat{q}^{\,0}_{n,\alpha} \xrightarrow{d} \chi^2_m(\lambda(\vartheta)) - q_{m,1-\alpha}$. The noncentral chi-square distribution has no atom at $q_{m,1-\alpha}$, so the rejection probabilities converge to

$$\mathbb{P}\left\{\chi^2_m(\lambda(\vartheta)) > q_{m,1-\alpha}\right\}.$$

This is the displayed size-adjusted power limit. □

The pilot-stability corollary proves that the Wald diagnostics do not change under logarithmically accurate roughness replacement.

Corollary 9.4 (Pilot-stable Wald diagnostics)**.** *Assume the conditions of Theorem 8.7 and Corollary 9.3. Write $H = H_{XY}(P)$, and let*

$$W_n(H) = \widehat{\vartheta}_n(H)^\top\widehat{\mathcal{I}}_{\mathcal{B},n}(H)\widehat{\vartheta}_n(H)$$

denote the projected Wald statistic computed with roughness input H. If the oracle and pilot-input Wald statistics are uniformly integrable, then

$$W_n(\widehat{H}_{XY}) - W_n(H) \xrightarrow{\mathbb{P}} 0, \qquad \widehat{\lambda}_n(\widehat{H}_{XY}) - \widehat{\lambda}_n(H) \xrightarrow{\mathbb{P}} 0,$$

where

$$\widehat{\lambda}_n(H) = \vartheta^\top \widehat{\mathcal{I}}_{\mathcal{B},n}(H)\vartheta$$

for a fixed local coordinate ϑ. Consequently the plug-in Wald-ratio diagnostic, the empirical-null size-adjusted power curve, and the noncentrality estimate have the same first-order limits as their oracle-H counterparts.

Proof. Theorem 8.7 gives

$$\widehat{\vartheta}_n(\widehat{H}_{XY}) - \widehat{\vartheta}_n(H) = o_{\mathbb{P}}(1), \qquad \widehat{\mathcal{I}}_{\mathcal{B},n}(\widehat{H}_{XY}) - \widehat{\mathcal{I}}_{\mathcal{B},n}(H) = o_{\mathbb{P}}(1)$$

on the uniformly nonsingular event. The oracle estimator is tight under the vector LAN theorem, so the quadratic map $(v, I) \mapsto v^\top I v$ is continuous on the random compact sets used in the argument with probability tending to one. This proves $W_n(\widehat{H}_{XY}) - W_n(H) = o_{\mathbb{P}}(1)$. The same information convergence gives $\widehat{\lambda}_n(\widehat{H}_{XY}) - \widehat{\lambda}_n(H) = o_{\mathbb{P}}(1)$. Corollary 9.3 then transfers the Wald-ratio and size-adjusted power limits by Slutsky's theorem; uniform integrability gives the corresponding mean statements. □

The efficiency statement is conditional on the projected score space. On that space, the estimator is the efficient GMM estimator for the local coordinates. The eigenvalues of $\mathcal{I}_{\mathcal{B}}$ are part of the empirical output because weak eigenvalues mark local directions with limited projected visibility. The Wald-ratio diagnostic in Section 11 is the empirical version of Corollary 9.3.

9.2. **Primitive rough-rank input.** The remaining primitive input for the rough regime is the boundary-layer rank condition in Assumption 8.10 and Theorem 8.11. For the rough Volterra class studied here, and in particular for $H < 1/2$, the projected covariance derivatives generated by the rough parameters must have full rank after target-only projection:

$$\operatorname{rank}\left(\frac{1}{2}\operatorname{tr}(B_{r,n}A_{n,j}^{\perp})\right)_{1\le r\le d_n,\,1\le j\le m} = m$$

for a fixed low-dimensional rough dictionary, uniformly over a compact subset of admissible rough parameters. The natural rough dictionary includes constant, front-loaded, and log-front-loaded boundary-layer shapes. This input upgrades the estimator from a projected Gaussian procedure to a primitive rough-volatility estimator. The finite-dimensional GMM proof is already complete, and the rough analysis supplies the survival and rank of the boundary-layer covariance derivatives after target-only projection and feasible proxy error.

9.3. **Projected P/Q information interface.** The matched-measure interface compares two projected information geometries. When physical and risk-neutral inputs are available, the same projected score construction is applied under the physical measure $\mathbb{P}$ and under the pricing measure $\mathbb{Q}$. For each asset pair, compute

$$\widehat{\mathcal{I}}_P^{\mathrm{src}\perp Y}, \quad \widehat{\mathcal{I}}_Q^{\mathrm{src}\perp Y}, \quad \widehat{\mathcal{I}}_{\mathcal{B},P}, \quad \widehat{\mathcal{I}}_{\mathcal{B},Q}, \quad \widehat{D}_I, \quad \widehat{D}_{\mathcal{B}}, \quad H_{PQ}, \quad \widehat{c}_{PQ}, \quad \widehat{\vartheta}_P^{\mathrm{pcov}}, \quad \widehat{\vartheta}_Q^{\mathrm{pcov}}.$$

The resulting classification rule is

(i) dynamically observable: both $\mathbb{P}$- and $\mathbb{Q}$-projected source information are large, so latent contagion is both priced and path-detectable; if $\widehat{c}_{PQ}$ is also close to one, the same projected direction is present under both measures;
(ii) pricing-visible only: $\mathbb{Q}$-projected information is large while $\mathbb{P}$-projected information is small, equivalently $\widehat{D}_I$ and $\widehat{D}_{\mathcal{B}}$ are strongly positive, so the channel is priced but weak in realized paths;
(iii) path-detectable only: $\mathbb{P}$-projected information is large while $\mathbb{Q}$-projected information is small, equivalently $\widehat{D}_I$ and $\widehat{D}_{\mathcal{B}}$ are strongly negative, so the channel is visible in realized dynamics but not option-implied;
(iv) direction-mismatched: both sides contain source-orthogonal projected information, but $\widehat{c}_{PQ}$ is small or H_{PQ} is large, so the measures disagree about the local direction;
(v) target-only or confounded: full information is large but source-orthogonal projected information is small, so the apparent signal is not source-screened;
(vi) unidentified: both $\mathbb{P}$- and $\mathbb{Q}$-projected source information are small at the tested scale.

These labels compare physical and risk-neutral observability using the same projected information objects. They are an output specification for matched panels. The Oxford-Man application below uses only the $\mathbb{P}$-side part of the interface, because its option information is limited to an external SPX roughness benchmark.

The matched-measure label map is stable away from its classification boundaries.

Theorem 9.5 (Consistency of separated P/Q regime labels)**.** *Fix thresholds for active projected information, positive and negative log-information mismatch, directional alignment, and Hausman discrepancy. Let $L(\mathcal{Z})$ be the population label assigned by the classification rule above to the population vector*

$$\mathcal{Z} = \left(\mathcal{I}_P^{\mathrm{src}\perp Y}, \mathcal{I}_Q^{\mathrm{src}\perp Y}, \mathcal{I}_{\mathcal{B},P}, \mathcal{I}_{\mathcal{B},Q}, D_I, D_{\mathcal{B}}, H_{PQ}, c_{PQ}, \vartheta_P, \vartheta_Q\right),$$

and let $\widehat{L}_n = L(\widehat{\mathcal{Z}}_n)$ be the plug-in label computed from the estimated quantities. Suppose

$$\widehat{\mathcal{Z}}_n \xrightarrow{\mathbb{P}} \mathcal{Z}$$

componentwise, with the convention that c_{PQ} is evaluated only on regimes where both population projected coordinates are nonzero. If $\mathcal{Z}$ is separated from every classification boundary by a margin $\delta > 0$, then

$$\mathbb{P}\{\widehat{L}_n = L(\mathcal{Z})\} \to 1.$$

The synthetic $\mathbb{P}/\mathbb{Q}$ design is therefore a consistency check for the operational labels. When the population regime is separated, label recovery follows from consistent estimation of the projected information, local coordinates, and mismatch functionals.

Proof. The classification rule is a finite collection of strict inequalities involving continuous functions of $\mathcal{Z}$ on each identified regime. Separation by δ means that all population inequalities defining the true label remain true after perturbations of size smaller than δ. Componentwise convergence in probability implies $\|\widehat{\mathcal{Z}}_n - \mathcal{Z}\| < \delta$ with probability tending to one, after restricting to the identified coordinates for c_{PQ}. On that event every inequality in the rule has the same sign for $\widehat{\mathcal{Z}}_n$ as for $\mathcal{Z}$, so the plug-in label equals the population label. □

The adaptive-label corollary transfers the same regime classification through the roughness pilot.

Corollary 9.6 (Pilot-adaptive P/Q label consistency)**.** *Suppose the $\mathbb{P}$- and $\mathbb{Q}$-side projected estimators each satisfy the conditions of Theorem 8.7, with cross-fitted pilots $\widehat{H}_{XY}^{\mathbb{P}}$ and $\widehat{H}_{XY}^{\mathbb{Q}}$. Let $\widehat{\mathcal{Z}}_n^{\mathrm{or}}$ be the oracle-roughness vector entering Theorem 9.5, and let $\widehat{\mathcal{Z}}_n^{\mathrm{ad}}$ be the same vector computed with the two pilots. If*

$$\widehat{\mathcal{Z}}_n^{\mathrm{or}} \xrightarrow{\mathbb{P}} \mathcal{Z}, \qquad \mathcal{Z} \text{ is separated from all classification boundaries},$$

then

$$\mathbb{P}\{L(\widehat{\mathcal{Z}}_n^{\mathrm{ad}}) = L(\mathcal{Z})\} \to 1.$$

Thus the $\mathbb{P}/\mathbb{Q}$ classification rule remains first-order valid when the roughness inputs are estimated, provided the two pilot errors are negligible at the logarithmic rough-local scale.

Proof. Apply Theorem 8.7 separately on the $\mathbb{P}$ and $\mathbb{Q}$ sides. The estimated local coordinates, projected information matrices, noncentrality functionals, and directional alignment terms are continuous functions of the side-specific projected geometries on the separated, nonsingular regime. Therefore

$$\widehat{\mathcal{Z}}_n^{\mathrm{ad}} - \widehat{\mathcal{Z}}_n^{\mathrm{or}} = o_{\mathbb{P}}(1)$$

componentwise. Since $\widehat{\mathcal{Z}}_n^{\mathrm{or}} \xrightarrow{\mathbb{P}} \mathcal{Z}$, the adaptive vector also converges to $\mathcal{Z}$. The conclusion is then exactly Theorem 9.5 applied to $\widehat{\mathcal{Z}}_n^{\mathrm{ad}}$. □

Scope of the efficiency claim. The baseline claim is local asymptotic efficiency in the finite-dimensional projected LAN submodel, validated at the level of constants in the Gaussian experiment. The Volterra, non-Gaussian, and paired $\mathbb{P}/\mathbb{Q}$ panels extend the same projected information geometry. They do not replace the separate global minimax theorem above. The supplementary diagnostics are collected in Appendix A, leaving the asymptotic results and empirical interpretation in the main development.

10. Observed-data robustness and regularization

The core estimator is stated for the clean high-frequency return experiment and the full target projection. Two transfers make the construction compatible with high-frequency data applications: pre-averaging for noisy prices and sieve compression for the high-dimensional target projection. Both transfers preserve the projected covariance-score geometry. They change the construction of $Z_{k,n}$, $\widehat{W}_n$, and $\widehat{\Pi}_{Y,n}$, not the LAN or GMM argument.

Assumption 10.1 (Noisy observation model)**.** In the noisy observation model, the econometrician observes

$$\widetilde{P}^a_{i\Delta_n} = P^a_{i\Delta_n} + \epsilon^a_{i,n}, \qquad a \in \{X, Y\},$$

where the noise sequence is centered, independent of the latent Volterra system, and has finite moments of order $4+\delta$ for some $\delta > 0$. In the baseline noisy theorem the noise is iid across observation times, with asset-specific variances.

Definition 10.2 (Pre-averaged block proxies)**.** Let $g_{\mathrm{pa}}(u) = u \wedge (1-u)$ on $[0,1]$, and let $k_n \to \infty$ satisfy

$$k_n\Delta_n \to 0, \qquad k_n^{-1} + m_n^{-1/2} = o(h_n^{H_{XY}}).$$

Let $\varpi_n \downarrow 0$ satisfy $\varpi_n/h_n \to 0$. Define the positive-floor pre-averaged quadratic variation by $\mathrm{PRV}^{a,+}_{k,n} := \mathrm{PRV}^a_{k,n} \vee \varpi_n$. Define the pre-averaged increments

$$\overline{\Delta}^n_{i,k_n}\widetilde{P}^a = \sum_{\ell=1}^{k_n-1} g_{\mathrm{pa}}\left(\frac{\ell}{k_n}\right)(\widetilde{P}^a_{(i+\ell)\Delta_n} - \widetilde{P}^a_{(i+\ell-1)\Delta_n}),$$

and let $\mathrm{PRV}^a_{k,n}$ be the standard bias-corrected pre-averaged quadratic variation on $I_{k,n}$. The pre-averaged log-volatility proxy is

$$\widehat{V}^{a,\mathrm{pa}}_{k,n} = \frac{1}{2}\{\log \mathrm{PRV}^{a,+}_{k,n} - \log h_n\}.$$

The pre-averaged block vector $Z^{\mathrm{pa}}_{k,n}$ is obtained from Definition 2.3 by replacing $\widehat{V}^a_{k,n}$ with $\widehat{V}^{a,\mathrm{pa}}_{k,n}$.

Pre-averaging changes the proxy construction. The projected-score quotient is unchanged.

Theorem 10.3 (Pre-averaged projected-score reduction)**.** *Assume the primitive Volterra conditions, the proxy-dominance conditions, and Assumption 10.1. Under the pre-averaging scale from Definition 10.2, with the usual additional mesh condition ensuring*

$$\sqrt{J_n}\, k_n^{-1} = o(h_n^{H_Y}),$$

the pre-averaged block vectors satisfy the same reduced Gaussian approximation as Theorem 2.5:

$$\frac{1}{N_n}\sum_{k=1}^{N_n} \mathbb{E}[\|Z^{\mathrm{pa}}_{k,n} - Z^{\mathrm{lat}}_{k,n}\|^2] \to 0.$$

Consequently, for any fixed projected covariance-score basis satisfying the stability conditions of Assumption 7.1 and the score-replacement criterion in Proposition B.4,

$$\frac{1}{\sqrt{N_n}}\sum_{k=1}^{N_n}\{g^{\mathrm{pa}}_{k,n} - g^{\mathrm{lat}}_{k,n}\} \xrightarrow{\mathbb{P}} 0.$$

The pre-averaged projected estimator therefore has the same first-order null and local-alternative limits as the clean-price projected estimator.

Proof. Write
$$e^{a,\mathrm{pa}}_{k,n} = \widehat{V}^{a,\mathrm{pa}}_{k,n} - \overline{V}^{a}_{k,n}.$$
The pre-averaging expansion under Assumption 10.1 and the scale in Definition 10.2 gives the same proxy-dominance bounds needed in Lemma B.1, with the raw realized-variance error replaced by the pre-averaged error. The independence and finite $4+\delta$ moment condition put the pre-averaging noise remainder in the same averaged L^2 class as the clean discretization error. On each block, $\mathrm{PRV}^a_{k,n} = h_n e^{2\overline{V}^a_{k,n}} + o_{\mathbb{P}}(h_n)$. Since $\varpi_n = o(h_n)$, the event on which the positive floor binds has probability tending to zero after averaging over blocks. The floor therefore does not change the first-order proxy expansion. Thus
$$\sqrt{J_n}\, h_n^{-H_Y} \left(\frac{1}{N_n} \sum_k \mathbb{E}[(e^{Y,\mathrm{pa}}_{k,n})^2] \right)^{1/2} \to 0, \qquad \sqrt{J_n} \left(\frac{1}{N_n} \sum_k \mathbb{E}[(e^{X,\mathrm{pa}}_{k,n})^2] \right)^{1/2} \to 0.$$
Bias correction removes the leading microstructure-noise term, while $\sqrt{J_n} k_n^{-1} = o(h_n^{H_Y})$ forces the residual pre-averaging discretization error to be negligible after the rough target normalization. Consequently
$$\frac{1}{N_n} \sum_{k=1}^{N_n} \mathbb{E}\left[\| Z^{\mathrm{pa}}_{k,n} - Z^{\mathrm{lat}}_{k,n} \|^2 \right] \to 0.$$
The fixed projected-score matrices have uniformly bounded operator norm after the assumed whitening and target-projection stability. Applying the quadratic-form replacement argument from Proposition B.4 to $Z^{\mathrm{pa}}_{k,n}$ and $Z^{\mathrm{lat}}_{k,n}$ gives the aggregated score equivalence. The GMM estimator and Wald limits then follow from Theorem 9.2 by Slutsky's theorem. $\square$

Definition 10.4 (Sieve-compressed target projection)**.** Let $(\varphi_\ell)_{\ell \geq 1}$ be the cosine basis of $L^2([0,1])$,
$$\varphi_1(u) = 1, \qquad \varphi_{\ell+1}(u) = \sqrt{2}\cos(\pi \ell u).$$
For $0 < \rho < 1$, set $L_n = \lfloor J_n^\rho \rfloor$ and define the compressed target-past coordinates
$$Z^Y_{k,\ell,n} = \frac{1}{\sqrt{J_n}} \sum_{j=1}^{J_n} \varphi_\ell\!\left(\frac{j}{J_n}\right) \widehat{Y}^{\mathrm{past}}_{k,j,n}, \qquad 1 \leq \ell \leq L_n.$$
The sieve target tangent space is the covariance tangent generated by
$$(F^Y_{k,n}, Z^Y_{k,1,n}, \ldots, Z^Y_{k,L_n,n})$$
instead of the full J_n-dimensional target past. The corresponding projection is denoted $\Pi^{L_n}_{Y,n}$.

The sieve theorem characterizes when compressed target projection leaves the retained score law unchanged.

Theorem 10.5 (Sieve-stable projected estimator)**.** *Suppose the target-only regression profile admits a boundary-layer transfer density with Sobolev smoothness $s > 0$, and choose $\rho \in (0,1)$ so that*
$$J_n^{-2s\rho} + \frac{J_n^\rho}{N_n} \to 0.$$
Assume also that the compressed target Gram matrices remain uniformly well conditioned. Let $g^{L_n}_{k,n}$ be the projected covariance-score vector obtained by replacing $\Pi_{Y,n}$ with $\Pi^{L_n}_{Y,n}$. Then
$$\frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \{ g^{L_n}_{k,n} - g_{k,n} \} \xrightarrow{\mathbb{P}} 0.$$
Consequently the sieve-compressed projected estimator has the same first-order local limit, projected information matrix, and Wald noncentrality as the full target-projection estimator.

Proof. Let $\Pi_{Y,n}$ and $\Pi_{Y,n}^{L_n}$ be the full and compressed target-only score projections. On the retained finite-dimensional score span, the difference between these projections is governed by two terms. The deterministic term is the cosine truncation error for the boundary-layer target regression profile. Sobolev smoothness s gives

$$\|(I - \Pi_{L_n})f\|_{L^2}^2 = O(L_n^{-2s}) = O(J_n^{-2s\rho})$$

for each target-transfer profile f entering the retained score directions. The stochastic term is the error in the L_n-dimensional compressed Gram system. The uniform conditioning assumption and finite-dimensional covariance concentration under the retained moment envelope give operator error $O_{\mathbb{P}}(L_n/N_n)^{1/2}$. Hence, on the retained score span,

$$\|\Pi_{Y,n}^{L_n} - \Pi_{Y,n}\|_{\text{score}} = O(J_n^{-s\rho}) + O_{\mathbb{P}}\left(\sqrt{\frac{J_n^{\rho}}{N_n}}\right) = o_{\mathbb{P}}(1).$$

Therefore the projected score matrices generated by the full and compressed target projections satisfy

$$\max_{1 \le r \le d} \|B_{r,n}^{L_n} - B_{r,n}\|_{\text{score}} \xrightarrow{\mathbb{P}} 0.$$

For Gaussian quadratic scores, the one-block variance of $\ell(B_{r,n}^{L_n}) - \ell(B_{r,n})$ is proportional to $\|B_{r,n}^{L_n} - B_{r,n}\|_{\text{score}}^2$, and the same bound applies to the mean derivative under local alternatives. The dependence envelope used for the retained projected scores applies to the difference array because only the projection map changes on a fixed score span. The variance calculation in Proposition B.4 therefore yields, since d is fixed,

$$\frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \{g_{k,n}^{L_n} - g_{k,n}\} \xrightarrow{\mathbb{P}} 0.$$

The compressed and full projected GMM problems have the same Γ, Ω, and local score limit up to $o_{\mathbb{P}}(1)$, so the estimator and Wald conclusions follow from Theorem 9.2. □

Remark 10.6 (Role of the extensions). The noisy and sieve extensions are robustness transfers, not new estimators. Once the projected covariance-score estimator is checked in the baseline reduced Gaussian experiment, the same first-order object survives the two high-frequency operations needed here: pre-averaging noisy prices and regularizing the high-dimensional target-only projection.

11. Synthetic calibration evidence

The validation record has two parts. Section 11 checks the projected estimator in a controlled Gaussian experiment; Section 12 applies the feasible estimator to the Oxford-Man realized-volatility panel under the $\mathbb{P}$-side scope limits stated below. In the baseline synthetic design every ingredient of the projected LAN construction is known: Σ_0, the covariance derivative, the target-only tangent space, the projected basis, and the local information constants. The synthetic experiment is therefore an implementation check at the level of the reduced experiment. A failure in the projected score, whitening, target projection, variance normalization, loading computation, or multiplier calibration would appear first in the null-size and Wald-ratio diagnostics.

The synthetic exercises check the components least identifiable in real data. The null experiment must be correctly sized under the covariance normalization appropriate to the design, the observed Wald mean must match the theoretical noncentral chi-square mean $m + \lambda(\tau)$, and size-adjusted power must track the noncentral chi-square detection boundary. The scalar anchor is reported on the same panels to distinguish failures of the moment class from failures of the local experiment. Pilot roughness experiments check the transfer theorem with injected noisy pilots and then compare feasible pilot constructions on matched synthetic cells. The test is whether replacing the oracle rough-local scale by $\widehat{H}_{XY}$ leaves projected information and Wald ratios unchanged at first order. Paired synthetic P/Q inputs are used only to check the mismatch functionals, not to claim real-data P/Q evidence.

Adaptive pilot choice. The retained synthetic tables and figures address the estimator's local Gaussian geometry. They report null size, Wald-ratio calibration, and detection-boundary power. The pilot-roughness transfer diagnostics, the matched-cell feasible-pilot comparison, and the synthetic P/Q recovery design are reported in Appendix A. Those appendix diagnostics support the empirical default used below. Adaptive variogram-correlation is the default feasible baseline. The length-aware regime-adaptive extension is a robustness enhancement for long upper-boundary calibration paths, not a uniformly preferable replacement.

11.1. **Null size before power.** The synthetic panels are iid Gaussian. Therefore the baseline covariance normalization uses bandwidth zero; positive HAC bandwidths are reserved for dependent or overlapping-block designs. The null experiment separates calibration failure from unnecessary finite-sample HAC noise. The empirical sizes in Table 1 are for the vector projected Wald test at nominal level 5%.

Table 1. Null size for the vector projected Wald statistic.

N	HAC	chi-square	multiplier	interpretation
256	0	0.054	0.055	calibrated iid baseline
256	2	0.078	0.069	finite-sample HAC noise
512	2	0.060	0.057	HAC noise shrinks

Table 1 identifies the source of the size distortion. With bandwidth zero, both the asymptotic chi-square and dependent-multiplier critical values are within Monte Carlo error of nominal size. At $N = 256$, HAC lags make the test mildly liberal even though the data are iid. Increasing N to 512 moves the HAC-based sizes back toward nominal. The distortion therefore comes from finite-sample noise in estimating lag covariances that are zero in the baseline design, rather than from the projected score, the target-only projection, or the multiplier scheme.

11.2. **Wald-ratio calibration.** For the source-component vector estimator let

$$W_n = \widehat{\vartheta}_n^\top \widehat{\mathcal{I}}_{\mathcal{B},n} \widehat{\vartheta}_n$$

be the projected vector Wald statistic. In the local Gaussian experiment, if the true local coordinate is $\vartheta(\tau)$, then the theoretical noncentrality is

$$\lambda(\tau) = \vartheta(\tau)^\top \mathcal{I}_{\mathcal{B}} \vartheta(\tau),$$

and the limiting prediction is

$$W_n \approx \chi^2_m(\lambda(\tau)), \qquad \mathbb{E}W_n \approx m + \lambda(\tau).$$

The end-to-end diagnostic is therefore

$$\mathrm{Ratio}(\tau) = \frac{\widehat{\mathbb{E}}_\tau W_n}{m + \lambda(\tau)}.$$

This ratio checks the covariance derivative, whitening, target projection, projected information matrix, variance normalization, and Monte Carlo implementation simultaneously.

Table 2. Wald-ratio calibration for the 4-fold iid baseline.

τ	effective τ	$\lambda(\tau)$	Wald ratio	size-adjusted rejection
-8	-8	3.0167	0.9968	0.306
-4	-4	0.7542	1.0604	0.128
-2	-2	0.1885	1.0815	0.074
0	0	0	1.0301	0.048
2	2	0.1885	0.9876	0.052
4	4	0.7542	1.0875	0.102
8	8	3.0167	0.9915	0.302

The largest-amplitude cells are $\tau = \pm 8$, where the local noncentrality is about 3.02. The observed mean Wald statistic agrees with the theoretical mean $m + \lambda(\tau)$ to within one percent. The constant-level check is severe. The score construction, whitening, projection, and variance normalization must all be aligned for this ratio to be close to one. At smaller amplitudes the noncentrality is close to zero, so Monte Carlo noise in the null component has a larger relative effect on the ratio. The power comparison below is therefore read through size-adjusted rejection rather than raw rejection alone.

11.3. **Detection boundary and scalar comparison.** The unit projected noncentrality in the default source direction is

$$w^\top \mathcal{I}_{\mathcal{B}} w = 0.0471.$$

Consequently $\tau = 2$ produces only

$$\lambda(2) = 4 \times 0.0471 \simeq 0.1885,$$

which is too close to the null to generate much power. By contrast, $\tau = 8$ gives

$$\lambda(8) = 64 \times 0.0471 \simeq 3.0167,$$

for which the noncentral χ^2_2 prediction gives power around 0.30 at five percent size. The simulation matches this prediction.

TABLE 3. Size-adjusted power and scalar-baseline comparison.

τ	$\lambda(\tau)$	vector Wald	scalar anchor	one-sided projected score
-8	3.0167	0.306	0.034	0.002
-2	0.1885	0.074	0.020	0.024
0	0	0.048	0.048	0.048
2	0.1885	0.052	0.020	0.072
8	3.0167	0.302	0.016	0.272

The vector Wald test is sign-invariant and therefore detects both positive and negative source directions. The scalar anchor and the one-sided projected score are direction-specific comparators. The scalar anchor remains essentially flat because its unit information is only

$$\mathcal{I}^{\text{scalar}} \simeq 1.8 \times 10^{-6},$$

so even $|\tau| = 8$ leaves it close to the null. The one-sided projected score detects the positive aggregate direction but, by construction, does not reject negative alternatives.

Low power at moderate amplitudes measures the local information content of the reduced experiment. At $\tau = 2$, the noncentrality is only 0.19 and theoretical power is essentially nominal. At $|\tau| = 8$, the noncentrality is about 3.02, the Wald mean matches the noncentral chi-square prediction, and the size-adjusted rejection rate is about 0.30. The projected estimator follows the LAN geometry.

Supplementary synthetic design. The synthetic $\mathbb{P}/\mathbb{Q}$ recovery design and the matched-cell feasible-pilot comparison are moved to Appendix A. They remain part of the scope and implementation record. The paired $\mathbb{P}/\mathbb{Q}$ design checks whether the classification functionals recover known synthetic regimes, and the pilot table records why adaptive variogram-correlation remains the default feasible baseline while the length-aware regime-adaptive extension serves as a robustness check for longer upper-boundary paths. The main calibration section stays focused on the three synthetic claims used for the final estimator: null calibration, constant-level Wald-ratio calibration, and detection-boundary power.

12. Real-data application: global realized-volatility panel

The Oxford-Man application applies the feasible projected estimator to market data. The design is asymmetric across measures. The application estimates the $\mathbb{P}$-side projected geometry from realized-volatility proxies; the $\mathbb{Q}$-side enters only through the external SPX roughness benchmark reported by [1]. The application is therefore a real-data $\mathbb{P}$-side exercise; full pairwise $\mathbb{P}/\mathbb{Q}$ classification lies outside the design.

12.1. **Data, proxy choice, and balanced panel construction.** The main panel keeps eight global equity indices: `.SPX`, `.IXIC`, `.FTSE`, `.GDAXI`, `.FCHI`, `.STOXX50E`, `.N225`, and `.HSI`. It then forms their balanced intersection over the Oxford-Man realized-volatility library [25]. The common sample runs from 2000-02-02 to 2018-06-27 and contains $T = 4081$ daily observations, hence $8 \times 7 = 56$ ordered source-target pairs. The baseline volatility proxy is the logged `rk_parzen` series, representing the realized-kernel family [17]. Any nonpositive realized-volatility proxy is treated as missing before logging, and the logged `rv5_ss` and `bv` series are reserved for robustness checks spanning the realized-variance and bipower-variation proxy families [15, 19]. The feasible estimator uses four-fold cross-fitting, the adaptive variogram-correlation pilot as the baseline roughness input, the length-aware regime-adaptive pilot as a robustness rerun, and a Bartlett HAC bandwidth

$$b_T = \left\lfloor 4(T/100)^{2/9} \right\rfloor$$

with asymptotic chi-square critical values. Benjamini-Hochberg FDR control at 5% [26] is applied only to the 56 full-sample ordered-pair tests. Because the full-sample test statistics are far beyond the rejection thresholds, even the dependence-corrected procedure of [27] would not change the conclusion that the balanced $\mathbb{P}$-side network is dense. The balanced design is restrictive by construction. It sacrifices a larger universe of indices in exchange for a common sample, fixed rolling windows, and directly comparable roughness and projected-score diagnostics across all ordered pairs. Panel coverage details are reported in Appendix A.

12.2. **P-side roughness estimates and the SPX physical-implied comparison.** Table 4 reports the feasible $\mathbb{P}$-side roughness estimates. Across the eight indices the estimated physical-measure roughnesses lie between roughly 0.040 and 0.090, and the long-window subsampling intervals preserve the same rough ordering across markets. The SPX row gives the single-index physical-implied comparison available in this data design. The Oxford-Man projected estimator gives $\widehat{H}_{\mathbb{P}} = 0.071$, while the external option-implied benchmark from [1] is $H_{\text{imp}} = 0.300$, for a gap of about -0.229. The quantity H_{imp} enters here as a projected maturity-surface calibration exponent. A primitive Volterra kernel exponent under an equivalent change of measure would require a matched model and panel. The discrepancy motivates the interface, but the comparison here is single-index; pairwise option-side contagion estimation lies outside this data design.

Table 4. Feasible $\mathbb{P}$-side roughness estimates on the balanced Oxford-Man panel.

Index	$\widehat{H}^P$	95% interval
SPX	0.071	[0.037, 0.108]
IXIC	0.090	[0.031, 0.132]
FTSE	0.083	[0.070, 0.098]
GDAXI	0.082	[0.067, 0.125]
FCHI	0.053	[0.038, 0.084]
STOXX50E	0.070	[0.062, 0.097]
N225	0.064	[0.042, 0.083]
HSI	0.040	[0.020, 0.051]

Note. 95% overlapping-window subsampling interval, window length 1512 trading days, step 21, 123 windows.

For SPX, the external option-implied benchmark reported by [1] gives $\widehat{H}_{\mathbb{P}} = 0.071$, $H_{\mathbb{Q}}^{\text{ext}} = 0.300$, and $\widehat{H}_{\mathbb{P}} - H_{\mathbb{Q}}^{\text{ext}} = -0.229$. The robustness pilot agrees with the baseline to the third decimal place throughout, so the main table reports uncertainty bands rather than duplicating a nearly identical robustness column. That stability is consistent with the pilot-transfer theory and with the matched synthetic pilot comparisons reported in Appendix A.

12.3. **Full-sample directed contagion map.** On the full balanced sample the $\mathbb{P}$-side directed network is dense. All 56 ordered pairs are source-visible at 5% FDR under the baseline pilot; none are labeled target-only/confounded or unidentified. The regime-adaptive robustness rerun changes no pair label, and the two alternative realized-volatility proxies, `rv5_ss` and `bv`, also retain all 56 source-visible pairs with complete overlap against the baseline edge set; the appendix reports those supporting tables in full. The empirical content is therefore relative intensity ranking inside a market where every ordered pair is source-visible over the full 2000–2018 sample, rather than sparse edge selection. Dense volatility networks are familiar in spillover and connectedness studies [8, 9, 10, 11]. The difference here is the estimand. The projected score does not summarize forecast-error variance shares or raw comovement; it isolates the source-screened component that remains after removing the entire target-only covariance tangent. It addresses the same contamination concern formalized by the contagion-versus-interdependence critique of [28], but at the level of the local covariance derivative rather than at the level of unconditional correlation shifts.

Table 5. Largest directed source-screened effects on the balanced Oxford-Man sample.

Source	Target	Wald	projected info.
STOXX50E	FCHI	900.1	1.683
SPX	IXIC	804.9	2.363
FCHI	STOXX50E	792.4	2.093
STOXX50E	GDAXI	769.9	1.905
GDAXI	STOXX50E	751.9	2.264
GDAXI	FCHI	749.1	1.967
FCHI	GDAXI	683.8	2.083
IXIC	SPX	607.3	1.809
FTSE	STOXX50E	473.0	1.641
STOXX50E	FTSE	433.7	1.807

The largest full-sample Wald directions are concentrated in the Euro-area core and the US pair. Even the weakest visible pair, HSI to GDAXI, has Wald 53.1, far above the calibration threshold. All ten directions in Table 5 have $q < 10^{-16}$ and keep the same source-visible label under the pilot and proxy reruns. The table is an intensity ranking rather than a second significance screen. Figure 1 is a directed intensity map, not a sparse selection display.

12.4. **Rolling observability and regime stability.** The rolling analysis uses 504-trading-day windows stepped every 21 trading days, which yields 171 windows for each pre-registered case-study pair. The rolling results are reported as continuous Wald paths rather than rolling multiple-testing labels. Because the full-sample map is dense, the empirical object is the timing of strengthened or weakened projected visibility, rather than first edge appearance.

Figure 2 shows the four fixed case studies. The two core transatlantic links from SPX to STOXX50E and from STOXX50E to SPX both peak in the window from 2009-01-21 to 2011-04-13, with Wald statistics 125.5 and 134.6 respectively. That build-up is gradual rather than point-mass, but its crest falls in the neighborhood of the May 6, 2010 flash-crash episode. The continental European pair from GDAXI to FCHI peaks later, over 2012-02-22 to 2014-05-27, at Wald 172.2. By contrast, the link from SPX to HSI is much weaker, peaking only at 30.2 in the 2015-05-08 to 2017-08-04 window and dropping below 1 in a near-null early-2000s window. The rolling design therefore records pronounced time variation even though the full-sample graph is uniformly visible.

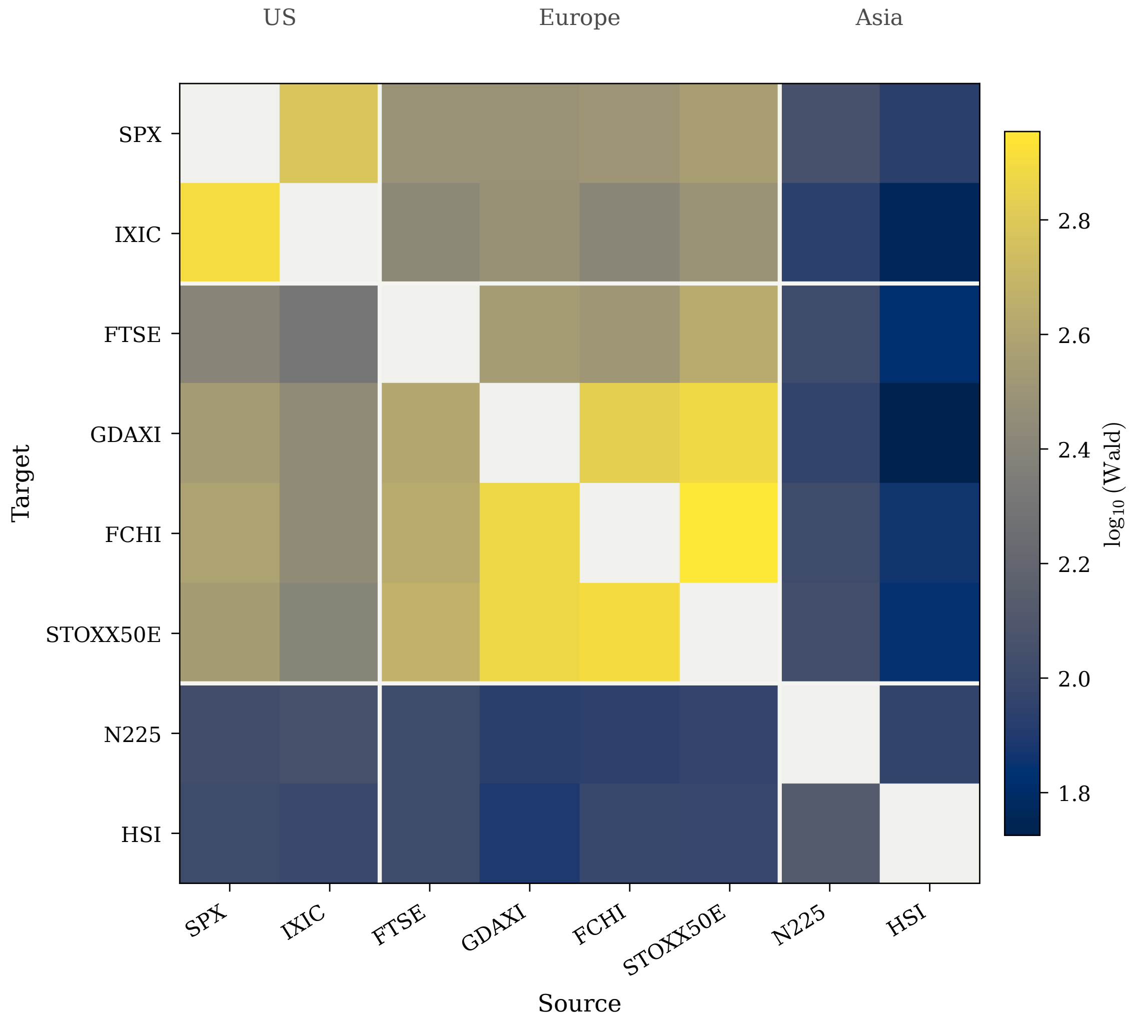

FIGURE 1. Directed source-screened intensity on the balanced Oxford-Man panel. Darker cells denote larger $\log_{10}$(Wald) values on the full 2000–2018 sample; the diagonal is suppressed.

Source. Oxford-Man Institute realized-volatility library; authors' projected-score calculations.

TABLE 6. Rolling summaries for the four pre-registered case-study pairs.

Pair	full-sample Wald	median rolling	peak Wald	peak window
SPX to STOXX50E	348.8	56.5	125.5	2009-01 to 2011-04
SPX to HSI	103.4	13.9	30.2	2015-05 to 2017-08
GDAXI to FCHI	749.1	103.9	172.2	2012-02 to 2014-05
STOXX50E to SPX	364.8	56.5	134.6	2009-01 to 2011-04

12.5. **Scope of the P/Q interface in the Oxford-Man application.** The Oxford-Man application is a $\mathbb{P}$-side market exercise. It estimates realized roughness and source-screened directed contagion under the physical measure, and it compares SPX only to an externally cited projected option-surface roughness benchmark. It does *not* estimate a matched option-panel covariance geometry and therefore does not produce pairwise pricing-visible, path-detectable-only, or dynamically observable labels under both measures. Those labels belong to the matched-measure interface developed above, not to the realized-volatility panel used here.

The scope boundary matters. Because the full-sample $\mathbb{P}$-side network is dense, realized-volatility data alone identify relative directed intensity and temporal stability. A claim that

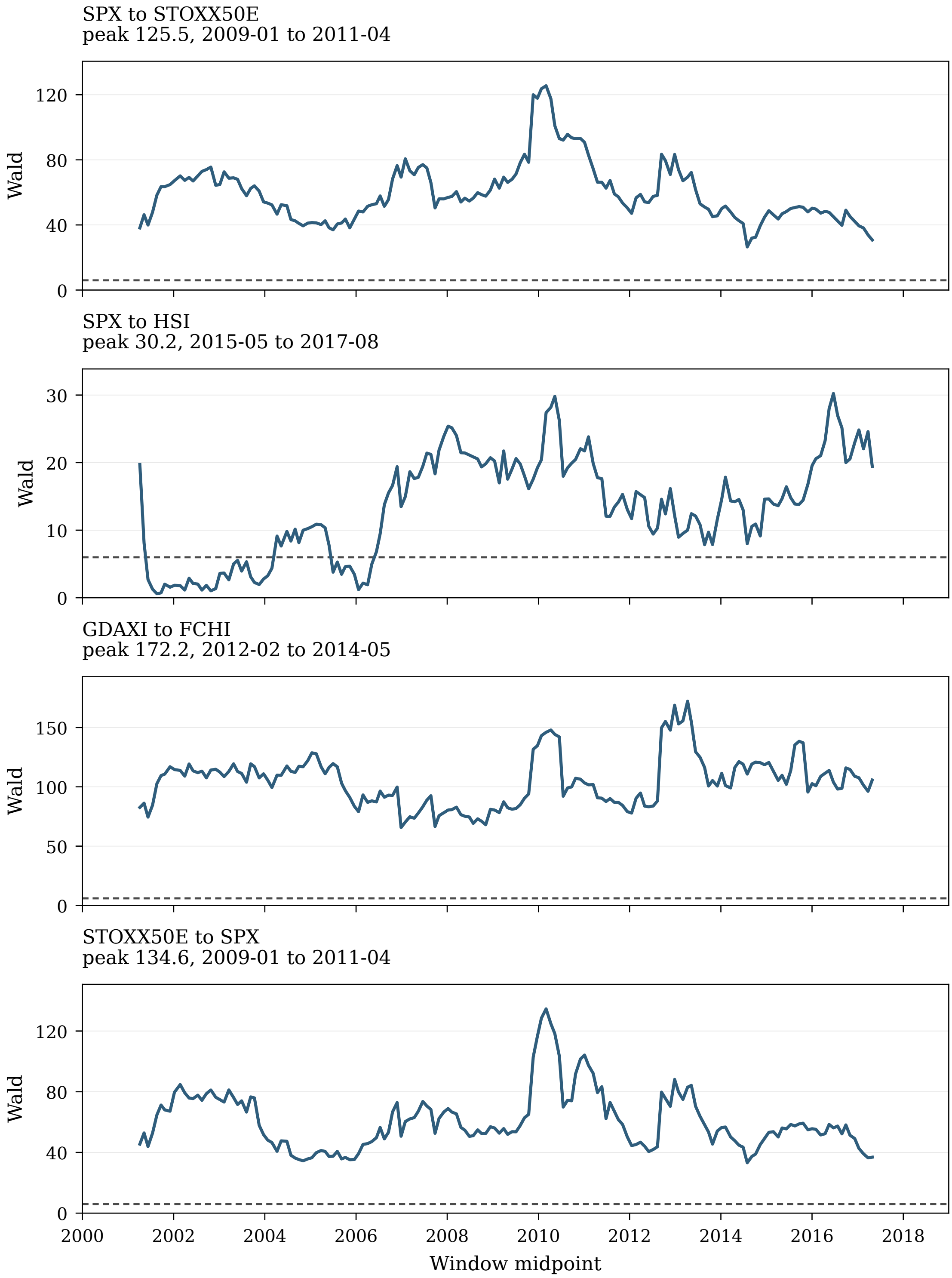

FIGURE 2. Rolling source-screened Wald statistics for the four pre-registered case-study pairs. Each panel reports one directed pair, and the dashed line marks the asymptotic 5% chi-square threshold.

Source. Oxford-Man Institute realized-volatility library; authors' projected-score calculations.

a pair is visible under $\mathbb{Q}$ but not under $\mathbb{P}$, or vice versa, requires a matched option-panel implementation of the same projected geometry. This section is a $\mathbb{P}$-side empirical benchmark. It shows that the projected estimator is implementable on a balanced global panel and that the SPX roughness discrepancy survives as a measured $\mathbb{P}$-versus-external-implied gap.

13. Conclusion

The threshold analysis separates pricing-visible smoothing contagion from contagion that is only path-detectable [3]. The dynamic observability analysis identifies the source-screened object that survives the reduction to observable filtrations [4]. Here that object becomes an estimator. The scalar residual-product statistic remains the analytical benchmark, but the empirical estimator is the projected covariance-score GMM statistic.

The governing object is the projected information matrix

$$\mathcal{I}_{\mathcal{B}} = \Gamma^{\top}\Omega^{-1}\Gamma.$$

It determines local efficiency in the reduced Gaussian experiment, supplies the Wald noncentrality constants, and measures how much source-screened information remains after the target-only tangent has been removed. Detection is therefore a statement about a covariance direction in a projected LAN experiment, not about a raw spillover coefficient.

The asymptotic contribution has four parts. The vector projected-GMM theorem gives the feasible LAN and Wald limit. The boundary-layer projected-rank theorem gives rough-regime identification for $H < 1/2$. The pilot roughness transfer theorem proves oracle equivalence under logarithmic roughness accuracy. The uniform minimax theorem upgrades the local Gaussian statement to uniform rough-class optimality once the primitive reduction hypotheses are verified.

The estimator changes the empirical reading of contagion. Weak rejection may be proximity to the detection boundary rather than estimator failure. Dense $\mathbb{P}$-side panels should be summarized by screened intensity and temporal stability rather than forced into sparse-edge language. A $\mathbb{P}/\mathbb{Q}$ discrepancy is a mismatch between projected information geometries, not an informal comparison of unrelated Hurst estimates.

The evidence stays inside those boundaries. The synthetic experiments check the estimator at the level of constants: null calibration, Wald-ratio calibration, and detection-boundary power. The Oxford-Man application shows that the same geometry is implementable on market data. Physical roughness remains stably rough, the SPX gap against the external implied benchmark is large, and the global contagion map is dense but not temporally uniform. Matched pairwise $\mathbb{P}/\mathbb{Q}$ classification requires option-panel data. Its econometric object is already specified: a projected information vector with separated labels for pricing-visible, path-detectable, dynamically observable, target-confounded, and unidentified regimes.

The final claim is therefore bounded. The paper gives a screened directional contagion estimand, its projected local information geometry, and a feasible $\mathbb{P}$-side estimator. The matched-measure interface maps the same estimator to risk-neutral pricing data, while the empirical application remains faithful to the realized-volatility panel actually observed.

Appendix A. Supplementary calibration and data details

This appendix collects supplementary diagnostics and supporting tables. They document the feasible construction and robustness checks without interrupting the main text.

A.1. **Matched-cell feasible-pilot comparison.** The roughness pilot comparison is run on matched synthetic cells so that all pilot types are evaluated on the same Monte Carlo draws. Pilot improvements are overstated when different methods are compared on different random realizations. Table 7 fixes the baseline empirical choice. Adaptive variogram-correlation is the default because it is stable across short and medium paths, while the length-aware regime-adaptive extension serves as a targeted robustness enhancement for longer upper-boundary cells where the correlation channel becomes more informative.

The trade-off is explicit. The length-aware regime-adaptive construction is best on the estimation-oriented metrics, but the simpler adaptive baseline is slightly tighter on null Wald-ratio calibration. The reported implementation therefore uses the adaptive pilot as the default feasible implementation and treats the length-aware regime-adaptive pilot as a robustness rerun rather than a replacement.

TABLE 7. Matched-cell feasible-pilot comparison in the projected-score experiment. Smaller is better in every column.

Pilot	mean rel. info.	mean theta-gap	mean null Wald dist.	mean abs. log err.
Variogram	0.005885	0.041504	0.003780	0.133098
Adaptive var.-corr.	0.005593	0.038625	0.003878	0.124716
Length-aware regime-ad.	0.005529	0.038152	0.004270	0.122312

A.2. **Synthetic P/Q recovery design.** The P/Q mismatch functionals from Section 8 can be validated without market data by constructing paired reduced experiments $(\mathbb{P},\mathbb{Q})$ with known covariance derivatives (A_P, A_Q). The synthetic design is a controlled check that, when supplied with two measure-specific covariance inputs, the projected estimator recovers the intended regime. The paired synthetic design reports

$$\widehat{\mathcal{I}}_P^{\mathrm{src}\perp Y},\quad \widehat{\mathcal{I}}_Q^{\mathrm{src}\perp Y},\quad \widehat{\mathcal{I}}_P^{\mathcal{B}},\quad \widehat{\mathcal{I}}_Q^{\mathcal{B}},\quad \widehat{D}_I,\quad \widehat{D}_{\mathcal{B}},\quad H_{PQ},\quad \widehat{c}_{PQ}.$$

The minimal regime grid is shown in Table 8. Each row is generated by choosing synthetic derivatives with the displayed source-orthogonal information pattern and then checking whether the mismatch functionals recover the correct label. The formal target is Theorem 9.5: labels are expected to be stable only when the synthetic population regime is separated from the classification boundaries.

TABLE 8. Synthetic P/Q mismatch design.

Regime	Synthetic information pattern	Mismatch signature	Expected label
P/Q equality	$A_P^{\perp} = A_Q^{\perp} \neq 0$	$D_I, D_{\mathcal{B}} \approx 0,\ c_{PQ} \approx 1$	aligned
Pricing-visible	$A_Q^{\perp} \neq 0,\ A_P^{\perp} \approx 0$	$D_I, D_{\mathcal{B}} \gg 0$	Q-only
Path-detectable	$A_P^{\perp} \neq 0,\ A_Q^{\perp} \approx 0$	$D_I, D_{\mathcal{B}} \ll 0$	P-only
Dynamically observable	$A_P^{\perp}, A_Q^{\perp} \neq 0$, aligned	$c_{PQ} \approx 1$	both
Direction mismatch	$A_P^{\perp}, A_Q^{\perp} \neq 0$, nonaligned	c_{PQ} small, H_{PQ} large	mismatch
Target-confounded	$\lVert A_\mu \rVert > 0,\ \lVert A_\mu^{\perp} \rVert \approx 0$	source information small	confounded

A.3. **Oxford-Man panel coverage and robustness checks.** Table 9 records the balanced-panel coverage used in the main application. Table 10 records the edge-count comparisons underlying the statement that the full-sample $\mathbb{P}$-side network is dense and stable across the primary pilot and proxy reruns.

TABLE 9. Balanced-panel coverage for the eight-index Oxford-Man application.

Symbol	market	raw dates	balanced dates
`.SPX`	S&P 500	4620	4081
`.IXIC`	Nasdaq Composite	4616	4081
`.FTSE`	FTSE 100	4639	4081
`.GDAXI`	DAX	4670	4081
`.FCHI`	CAC 40	4691	4081
`.STOXX50E`	Euro Stoxx 50	4691	4081
`.N225`	Nikkei 225	4501	4081
`.HSI`	Hang Seng	4510	4081

A.4. **Supplementary calibration outputs.** The supplementary exercises report five procedure families on matched synthetic panels: the full covariance oracle, the target-only covariance oracle, the source-orthogonal covariance oracle, the scalar-anchor statistic, and the feasible projected-basis estimator. Across these procedures the reported diagnostics are organized around

TABLE 10. Full-sample visibility checks on the Oxford-Man panel.

Run	visible pairs	baseline overlap	label changes
Baseline (`rk_parzen`, adaptive)	56	56	0
Pilot rerun (`rk_parzen`, regime-adaptive)	56	56	0
Proxy rerun (`rv5_ss`, adaptive)	56	56	–
Proxy rerun (`bv`, adaptive)	56	56	–

the same decision tree used throughout the analysis: information decomposition, projected-rank and conditioning, null calibration, Wald-ratio calibration, feasible-pilot stability, and, when supplied, paired synthetic P/Q interface functionals. Each experiment records:

(i) information decomposition: $\mathcal{I}^{\mathrm{full}}$, $\mathcal{I}^{Y}$, $\mathcal{I}^{\mathrm{src}\perp Y}$, $\mathcal{I}^{\mathcal{B}}$, $\mathcal{I}^{\mathrm{scalar}}$, and capture ratios;
(ii) projected identification: $\mathcal{I}_{\mathcal{B},n}$, its eigenvalues, condition number, and numerical rank;
(iii) calibration: null size under chi-square, multiplier, and empirical-null critical values, reported separately for bandwidth zero and positive HAC bandwidths;
(iv) constant calibration: theoretical noncentrality, expected Wald mean, observed Wald mean, and Wald ratio across a local-amplitude grid;
(v) robustness extensions: theorem-matched Volterra blocks for $H < 1/2$, pilot roughness checks reporting $|\widehat{H} - H||\log h_n|$, $\widehat{\theta}_n/\theta_n$, plug-in/oracle information error, and Wald-ratio stability, non-Gaussian quasi-score panels with finite-difference loadings, and paired synthetic P/Q interface recovery with separated-label checks.

These diagnostics are the supporting record for the synthetic section; the appendix tables above retain only the pieces needed to document pilot choice and empirical robustness checks.

APPENDIX B. PRIMITIVE REDUCTION PROOFS

This appendix records the auxiliary reduction results used in the article. They are stated in the notation of the projected covariance-score estimator so that the econometric argument remains self-contained.

The first primitive lemma identifies the feasible realized-volatility proxy error at the block level.

Lemma B.1 (Block log-realized-variance expansion)**.** *Let*

$$\overline{V}^a_{k,n} = \frac{1}{h_n}\int_{I_{k,n}} V^a_s\,ds, \qquad a \in \{X, Y\}.$$

Under Assumptions 2.1–2.4, the block log-variance proxy admits the decomposition

$$\widehat{V}^a_{k,n} = \overline{V}^a_{k,n} + e^a_{k,n},$$

with

$$\frac{1}{N_n}\sum_k \mathbb{E}[(e^a_{k,n})^2] = O(m_n^{-1}) + O(h_n^{4H_a}) + O(m_n^{-2}).$$

Consequently the normalized target future and target past proxy errors are

$$O\big(h_n^{-H_Y} m_n^{-1/2} + h_n^{H_Y}\big) = o(1)$$

in averaged L^2, while the source past proxy errors are

$$O(m_n^{-1/2} + h_n^{2H_X}) = o(1)$$

coordinatewise, and remain negligible after summing over the J_n-dimensional memory window.

Proof. Conditional on the volatility path over $I_{k,n}$, the return increments are centered Gaussians with variances $\Delta_n e^{2V^a}$, up to the discretization error. The realized variance is therefore a weighted chi-square perturbation of the integrated variance. Taylor expansion of the logarithm gives a martingale fluctuation of order $m_n^{-1/2}$ and a quadratic remainder of order m_n^{-1}. The replacement of integrated variance by $\exp(2\overline{V}^a_{k,n})h_n$ costs $O(h_n^{2H_a})$ in first order, hence $O(h_n^{4H_a})$ in squared error, by the L^2-Hölder regularity of the Volterra field. The target normalization multiplies adjacent block differences by $h_n^{-H_Y}$; the source vector is left on its unnormalized proxy scale. The rate conditions in Assumption 2.4 make the resulting averaged vector errors vanish. $\square$

The covariance lemma identifies the frozen boundary-layer operator that enters every retained score.

Lemma B.2 (Boundary-layer covariance asymptotics)**.** *For each local evaluation time t, the latent Gaussian block covariance admits deterministic frozen boundary-layer approximations*

$$\mathbb{G}_t^{YY,(n)}, \quad \mathbb{G}_t^{XX,(n)}, \quad \mathbb{G}_t^{YX,(n)}, \quad \mathfrak{g}_t^{Y,(n)}, \quad \mathfrak{g}_t^{XY,(n)}$$

such that, along the memory window $J_n h_n \to 0$,

$$\|\Sigma_n^{YY} - \mathbb{G}_t^{YY,(n)}\|_{\mathrm{op}} \to 0, \qquad \|\Sigma_n^{XX} - \mathbb{G}_t^{XX,(n)}\|_{\mathrm{op}} \to 0,$$
$$\|h_n^{-H_{XY}}\Sigma_n^{YX} - \mathbb{G}_t^{YX,(n)}\|_{\mathrm{op}} \to 0, \qquad \|h_n^{-H_{XY}}\gamma_n^{X} - \mathfrak{g}_t^{XY,(n)}\|_{\ell^2} \to 0.$$

The target-target and source-source blocks remain order one under their canonical normalizations, whereas the source-target blocks enter at rough scale $h_n^{H_{XY}}$.

Proof. Write each covariance entry as a double integral of Volterra kernels over two adjacent blocks. Because $j, \ell \le J_n$ and $J_n h_n \to 0$, all times used by the block vector lie in a shrinking boundary layer near t. Replacing the smooth amplitudes $L_i(u,v)$ by their diagonal values and rescaling time by h_n gives deterministic frozen kernels on the discrete lag grid. The singular powers $(u-v)^{H_i-1/2}$ determine the normalizing factors. The C^2 amplitude remainder and the discarded off-layer terms vanish in operator norm by Schur bounds on the resulting kernel matrices. $\square$

The primitive rank proposition proves that the retained Volterra profiles have full projected rank under nonsingularity of the profile Gram matrix.

Proposition B.3 (Primitive verification of boundary-layer projected rank)**.** *Let $\mathcal{V}$ be a compact rough Volterra class with $H_{XY} \in [H_-, H_+] \subset (0, 1/2)$. Suppose Lemma B.2 holds uniformly over $\mathcal{V}$, and let $\iota_n : \mathfrak{H}_\perp \to \mathbb{R}^{p_n \times p_n}_{\mathrm{sym}}$ denote the boundary-layer discretization map from limiting covariance profiles to the whitened block covariance coordinates. Assume ι_n is asymptotically isometric on the finite-dimensional span of the derivative profiles and that the feasible whitening and target-only projection are stable:*

$$\sup_{\|c\|=1} \left\| \sum_{j=1}^{m} c_j A_{n,j}^{\perp} - \iota_n\left(\sum_{j=1}^{m} c_j \psi_j^{\perp}\right) \right\|_{\mathrm{score}} \to 0$$

uniformly over $\mathcal{V}$. For amplitude coordinates and Hurst coordinates the limiting profiles are generated respectively by

$$u^{H_{XY}-1/2} \quad \textit{and} \quad u^{H_{XY}-1/2}\log u,$$

multiplied by the frozen diagonal Volterra amplitudes and then transformed by $(I - \Pi_{\mathfrak{Y}})\mathcal{W}_t(\cdot)\mathcal{W}_t$.

If the source-screened profile Gram is uniformly nondegenerate,

$$\inf_{P \in \mathcal{V}} \lambda_{\min}\left(\left(\langle \psi_i^{\perp}, \psi_j^{\perp} \rangle_{\mathfrak{H}_\perp} \right)_{i,j} \right) \ge \lambda_* > 0,$$

and the retained dictionary approximates the derivative span uniformly,

$$\sup_{P \in \mathcal{V}} \sup_{\|c\|=1} \left\| (I - \Pi^{\mathrm{score}}_{\mathcal{B}_n}) \sum_{j=1}^{m} c_j A_{n,j}^{\perp} \right\|_{\mathrm{score}} \to 0,$$

then the hypotheses of Theorem 8.11 hold uniformly over $\mathcal{V}$, apart from the retained score covariance eigenvalue bound, which is supplied by the projected-score CLT and HAC consistency result.

Proof. The uniform approximation display and the asymptotic isometry of ι_n imply, uniformly over $P \in \mathcal{V}$ and $\|c\| = 1$,

$$\left| \left\| \sum_{j=1}^{m} c_j A_{n,j}^{\perp} \right\|_{\mathrm{score}}^2 - \left\| \sum_{j=1}^{m} c_j \psi_j^{\perp} \right\|_{\mathfrak{H}_\perp}^2 \right| \to 0.$$

This is the profile-convergence hypothesis in Theorem 8.11. The displayed Gram lower bound is exactly the source-screened nondegeneracy hypothesis, and the dictionary display is exactly the span approximation hypothesis. Finally, differentiating the frozen rough kernel $u^{H_{XY}-1/2}$ with respect to H_{XY} gives the logarithmic profile $u^{H_{XY}-1/2} \log u$, while amplitude differentiation only changes the diagonal multiplier. Thus amplitude and Hurst coordinates generate the profile family stated in the main theorem. □

The proxy-transfer proposition separates block-proxy error from covariance-score aggregation.

Proposition B.4 (Vector proxy replacement and projected-score transfer)**.** *Let $Z_{k,n}^{\mathrm{lat}}$ be any centered Gaussian block vector, defined on the same probability space as $Z_{k,n}$, such that*

$$\frac{1}{N_n} \sum_{k=1}^{N_n} \mathbb{E}[\|Z_{k,n} - Z_{k,n}^{\mathrm{lat}}\|^2] \to 0.$$

Fix a projected covariance-score basis and write

$$d_{k,n} := g_{k,n} - g_{k,n}^{\mathrm{lat}}.$$

If the null covariance estimate, whitening map, target projection, and retained basis are stable in operator norm, and if in addition

$$\frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \mathbb{E}[d_{k,n}] \to 0, \qquad \varepsilon_n^2 := \frac{1}{N_n} \sum_{k=1}^{N_n} \mathbb{E}[\|d_{k,n}\|^2] \to 0,$$

and either

(i) the proof is run on a disjoint block grid, so the score differences are independent across blocks, or

(ii) there is a covariance envelope $\widetilde{\rho}_n(\ell)$ such that

$$\sup_n \sum_{\ell \ge 1} \widetilde{\rho}_n(\ell) < \infty$$

and

$$\sup_k \|\mathrm{Cov}(d_{k,n}, d_{k+\ell,n})\| \le \widetilde{\rho}_n(\ell)\, \varepsilon_n^2, \qquad \ell \ge 1,$$

then

$$\frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \{g_{k,n} - g_{k,n}^{\mathrm{lat}}\} \xrightarrow{\mathbb{P}} 0.$$

Proof. The deterministic component is the assumed aggregated mean remainder:

$$\frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \mathbb{E}[d_{k,n}] \to 0.$$

This condition is needed at the $\sqrt{N_n}$ scale; blockwise L^2 convergence alone controls variance, but it does not control an aligned deterministic bias. If the proof is run on a disjoint grid, then

$$\mathrm{Var}\left(\frac{1}{\sqrt{N_n}} \sum_{k=1}^{N_n} \{d_{k,n} - \mathbb{E} d_{k,n}\} \right) = \frac{1}{N_n} \sum_{k=1}^{N_n} \mathrm{Var}(d_{k,n}) \le \varepsilon_n^2 \to 0.$$

Under the covariance-envelope alternative,

$$\begin{aligned}\operatorname{Var}\left(\frac{1}{\sqrt{N_n}}\sum_{k=1}^{N_n}\{d_{k,n}-\mathbb{E}d_{k,n}\}\right) &\le \frac{1}{N_n}\sum_{k=1}^{N_n}\mathbb{E}[\|d_{k,n}\|^2] \\ &\quad + \frac{2}{N_n}\sum_{\ell=1}^{N_n-1}\sum_{k=1}^{N_n-\ell}\|\operatorname{Cov}(d_{k,n},d_{k+\ell,n})\| \\ &\le \varepsilon_n^2\left(1+2\sum_{\ell\ge 1}\widetilde{\rho}_n(\ell)\right)\to 0.\end{aligned}$$

Chebyshev's inequality therefore yields

$$\frac{1}{\sqrt{N_n}}\sum_{k=1}^{N_n}\{g_{k,n}-g_{k,n}^{\mathrm{lat}}\}\xrightarrow{\mathbb{P}} 0.$$

□

The HAC lemma characterizes the long-run covariance limit for the retained projected-score array.

Lemma B.5 (Projected-score HAC CLT and consistency)**.** *Let $g_{k,n}\in\mathbb{R}^d$ be the retained projected score vector, with d fixed. Suppose that, after discarding $O(1)$ boundary blocks, each row of the triangular array is stationary in the block index k, and that the centered array has uniformly bounded $4+\varepsilon$ moments and a dependence envelope $\rho_n(\ell)$ satisfying*

$$\sup_k\|\operatorname{Cov}(g_{k,n},g_{k+\ell,n})\|\le\rho_n(\ell),\qquad \sup_n\sum_{\ell\ge 1}\rho_n(\ell)<\infty,\qquad \sum_{\ell>L}\sup_n\rho_n(\ell)\to 0,$$

with the usual triangular-array Lindeberg condition. Then

$$\frac{1}{\sqrt{N_n}}\sum_{k=1}^{N_n}\{g_{k,n}-\mathbb{E}g_{k,n}\}\xrightarrow{d} N(0,\Omega),$$

where the rowwise long-run covariance matrices

$$\Omega_n:=\operatorname{Cov}(g_{1,n},g_{1,n})+\sum_{\ell\ge 1}\{\operatorname{Cov}(g_{1,n},g_{1+\ell,n})+\operatorname{Cov}(g_{1+\ell,n},g_{1,n})\}$$

converge to Ω. If $L_n^{\mathrm{hac}}\to\infty$ and $L_n^{\mathrm{hac}}/N_n\to 0$, the HAC estimator from Definition 5.2 is consistent for Ω. In the iid Gaussian calibration case, the bandwidth-zero estimator is the correct specialization.

Proof. Fix a vector $a\in\mathbb{R}^d$. The scalar array $a^\top\{g_{k,n}-\mathbb{E}g_{k,n}\}$ has uniformly bounded $4+\varepsilon$ moments, a summable covariance envelope, and the stated Lindeberg condition. The rowwise triangular-array CLT for weakly dependent stationary rows gives

$$\frac{1}{\sqrt{N_n}}\sum_{k=1}^{N_n}a^\top\{g_{k,n}-\mathbb{E}g_{k,n}\}\xrightarrow{d} N(0,a^\top\Omega a).$$

Cramér–Wold gives the vector CLT. In the Gaussian quadratic-score case the required moment bound follows from hypercontractivity of Gaussian chaoses of order two.

For HAC consistency, decompose the estimator error into the finite-lag sample covariance error, the kernel truncation bias, and the long-lag covariance tail. The finite-lag term vanishes because $L_n^{\mathrm{hac}}/N_n\to 0$ and the rowwise moments are uniformly bounded. The kernel bias vanishes because $K_{\mathrm{hac}}(\ell/(L_n^{\mathrm{hac}}+1))\to 1$ for each fixed lag ℓ. The tail term is controlled by $\sum_{\ell>L}\sup_n\rho_n(\ell)\to 0$. When the panel is iid, all lag covariances are zero and bandwidth zero avoids estimating nonexistent dependence. □

Remark B.6 (Role of the scalar derivations). The scalar residual-product derivations enter only through reusable ingredients: proxy reduction, boundary-layer covariance, the target-only lower bound, scalar-anchor embedding, HAC normalization, pre-averaging, and sieve stability. The scalar residual-product testing theory is no longer the main econometric object. It is the benchmark submodel against which the projected covariance-score GMM estimator is compared.

References

[1] G. Livieri, S. Mouti, A. Pallavicini, and M. Rosenbaum, *Rough volatility: Evidence from option prices*, IISE Transactions **50** (2018), no. 9, 767–776, doi:10.1080/24725854.2018.1444297.

[2] J. Gatheral, T. Jaisson, and M. Rosenbaum, *Volatility is rough*, Quantitative Finance **18** (2018), no. 6, 933–949, doi:10.1080/14697688.2017.1393551.

[3] J. Vidal Llauradó (2026). *Latent Volatility Contagion in Rough Volatility Models: Spectral Thresholds and Detectability.* SSRN. https://doi.org/10.2139/ssrn.6520718.

[4] J. Vidal Llauradó (2026). *Dynamic Observability of Latent Contagion: Sequential Revelation and Observable Screening.* SSRN. https://doi.org/10.2139/ssrn.6539199.

[5] C. Bayer, P. Friz, and J. Gatheral, *Pricing under rough volatility*, Quantitative Finance **16** (2016), no. 6, 887–904, doi:10.1080/14697688.2015.1099717.

[6] O. El Euch and M. Rosenbaum, *The characteristic function of rough Heston models*, Mathematical Finance **29** (2019), no. 1, 3–38, doi:10.1111/mafi.12173.

[7] E. Abi Jaber, M. Larsson, and S. Pulido, *Affine Volterra processes*, Annals of Applied Probability **29** (2019), no. 5, 3155–3200, doi:10.1214/19-AAP1477.

[8] F. X. Diebold and K. Yılmaz, *Measuring financial asset return and volatility spillovers, with application to global equity markets*, The Economic Journal **119** (2009), no. 534, 158–171, doi:10.1111/j.1468-0297.2008.02208.x.

[9] F. X. Diebold and K. Yılmaz, *Better to give than to receive: Predictive directional measurement of volatility spillovers*, International Journal of Forecasting **28** (2012), no. 1, 57–66, doi:10.1016/j.ijforecast.2011.02.006.

[10] F. X. Diebold and K. Yılmaz, *On the network topology of variance decompositions: Measuring the connectedness of financial firms*, Journal of Econometrics **182** (2014), no. 1, 119–134, doi:10.1016/j.jeconom.2014.04.012.

[11] J. Baruník and T. Křehlík, *Measuring the frequency dynamics of financial connectedness and systemic risk*, Journal of Financial Econometrics **16** (2018), no. 2, 271–296, doi:10.1093/jjfinec/nby001.

[12] W. K. Newey, *The asymptotic variance of semiparametric estimators*, Econometrica **62** (1994), no. 6, 1349–1382, doi:10.2307/2951752.

[13] V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins, *Double/debiased machine learning for treatment and structural parameters*, The Econometrics Journal **21** (2018), no. 1, C1–C68, doi:10.1111/ectj.12097.

[14] R. Cont and P. Das, *Rough volatility: Fact or artefact?*, Sankhyā B **86** (2024), no. 1, 191–223, doi:10.1007/s13571-024-00322-2.

[15] O. E. Barndorff-Nielsen and N. Shephard, *Econometric analysis of realized volatility and its use in estimating stochastic volatility models*, Journal of the Royal Statistical Society: Series B **64** (2002), no. 2, 253–280, doi:10.1111/1467-9868.00336.

[16] T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys, *Modeling and forecasting realized volatility*, Econometrica **71** (2003), no. 2, 579–625, doi:10.1111/1468-0262.00418.

[17] O. E. Barndorff-Nielsen, P. R. Hansen, A. Lunde, and N. Shephard, *Designing realized kernels to measure the ex post variation of equity prices in the presence of noise*, Econometrica **76** (2008), no. 6, 1481–1536, doi:10.3982/ECTA6495.

[18] J. Jacod, Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter, *Microstructure noise in the continuous case: The pre-averaging approach*, Stochastic Processes and their Applications **119** (2009), no. 7, 2249–2276, doi:10.1016/j.spa.2008.11.004.

[19] O. E. Barndorff-Nielsen and N. Shephard, *Power and bipower variation with stochastic volatility and jumps*, Journal of Financial Econometrics **2** (2004), no. 1, 1–37, doi:10.1093/jjfinec/nbh001.

[20] L. P. Hansen, *Large sample properties of generalized method of moments estimators*, Econometrica **50** (1982), no. 4, 1029–1054, doi:10.2307/1912775.

[21] W. K. Newey and K. D. West, *A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix*, Econometrica **55** (1987), no. 3, 703–708, doi:10.2307/1913610.

[22] D. W. K. Andrews, *Heteroskedasticity and autocorrelation consistent covariance matrix estimation*, Econometrica **59** (1991), no. 3, 817–858, doi:10.2307/2938229.

[23] W. K. Newey and K. D. West, *Automatic lag selection in covariance matrix estimation*, The Review of Economic Studies **61** (1994), no. 4, 631–653, doi:10.2307/2297912.

[24] A. W. van der Vaart, *Asymptotic Statistics*, Cambridge University Press, 1998, doi:10.1017/CBO9780511802256.

[25] G. Heber, A. Lunde, N. Shephard, and K. Sheppard, *Oxford-Man Institute's realized library*, version 0.1, Oxford-Man Institute, University of Oxford, 2009, https://realized.oxford-man.ox.ac.uk/.

[26] Y. Benjamini and Y. Hochberg, *Controlling the false discovery rate: A practical and powerful approach to multiple testing*, Journal of the Royal Statistical Society: Series B **57** (1995), no. 1, 289–300, doi:10.1111/j.2517-6161.1995.tb02031.x.

[27] Y. Benjamini and D. Yekutieli, *The control of the false discovery rate in multiple testing under dependency*, The Annals of Statistics **29** (2001), no. 4, 1165–1188, doi:10.1214/aos/1013699998.

[28] K. J. Forbes and R. Rigobon, *No contagion, only interdependence: Measuring stock market comovements*, The Journal of Finance **57** (2002), no. 5, 2223–2261, doi:10.1111/0022-1082.00494.

Email address: `vidal.llaurado@gmail.com`