Self-Normalized Inference for Relevance-Based Functional Granger Causality

Jiajing Sun Abderrahim Taamouti Zhuo Lin

February 28, 2026

Affiliations

Jiajing Sun

MOE Social Science Laboratory of Digital Economic Forecasts and Policy Simulation, and
School of Economics and Management, University of Chinese Academy of Sciences,
Zhongguancun Nanyitiao, Haidian District, Beijing 100190, China.
Email: jiajing.sun@gmail.com.

Abderrahim Taamouti

Management School, University of Liverpool, L69 7ZH, UK.
Email: Abderrahim.Taamouti@liverpool.ac.uk.

Zhuo Lin

Academy of Mathematics and Systems Science, Chinese Academy of Sciences,
No. 55 Zhongguancun East Road, Haidian District, Beijing 100190, China; and
School of Economics and Management, University of Chinese Academy of Sciences,
Zhongguancun Nanyitiao, Haidian District, Beijing 100190, China.
Email: linzhuo2021@amss.ac.cn.

Correspondence to: Zhuo Lin, Academy of Mathematics and Systems Science, Chinese Academy of Sciences,
No. 55 Zhongguancun East Road, Haidian District, Beijing 100190, China. Telephone: +86 1660 1393 178.
Email: linzhuo2021@amss.ac.cn.

Abstract

We propose relevance-based tests of functional Granger causality for functional time series, assessing whether causal effects are economically meaningful rather than only statistically significant. Causal lag operators are estimated in Hilbert space via ridge regularization, preserving functional structure without truncation. Inference relies on self-normalization, yielding tuning-free tests robust to serial dependence and heavy-tailed innovations. Simulations demonstrate accurate boundary size control and strong power, particularly with the adjusted-range self-normalizer. In cryptocurrency markets, implied-volatility surfaces predict realized volatility at moderate relevance thresholds but not at higher thresholds, and no economically meaningful reverse or flow-based causal effects are detected.

Keywords: Functional time series; Functional VARX; Granger causality; Relevant hypotheses; Self-normalization.

JEL: C32, C58, G13.

Introduction

Many modern measurements in economics and finance are naturally recorded as curves or surfaces observed repeatedly over time: yield curves and term structures (Lengwiler and Lenz, 2010; Bardsley et al., 2017), intraday volatility profiles (Müller et al., 2011), implied-volatility smiles/surfaces, load curves, and other high-resolution objects. Functional data analysis (FDA) treats each observation as a random element in a function space and therefore preserves the shape information that is often lost in low-dimensional summaries; see, for example, Ramsay and Silverman (2005) and Ferraty and Vieu (2006). When such curves arrive sequentially, they form a functional time series (FTS), for which dependence, prediction, and inference are most naturally formulated in a separable Hilbert space $\mathbb{H}$ (Bosq, 2000; Horváth and Kokoszka, 2012; Aue et al., 2015). Working in $\mathbb{H}$ makes the statistical objects of interest—models, hypotheses, and effect sizes—invariant to the observation grid; any discretization or basis expansion is used only to compute inner products and apply operators, not as a modeling assumption.

This raises a natural directional-predictability question: does an entire curve observed today improve forecasts of another curve tomorrow, beyond what the target's own past already provides? In option markets, for example, one may ask whether the implied-volatility smile contains economically meaningful information about next-day realized volatility, and whether any predictive gain is large enough to matter for risk management rather than merely statistically detectable. Since Granger (1969), Wiener–Granger causality has become a workhorse tool for studying predictive linkages, policy transmission, and spillovers across markets and macroeconomic aggregates; see Lütkepohl (2005) for the VAR/VECM perspective and Shojaie and Fox (2022) for a modern statistical review. Beyond classical linear VAR restrictions, a substantial literature has developed causality measures (not only tests), including horizon-specific and short-/long-run notions (Dufour and Taamouti, 2010), as well as nonlinear and distributional extensions based on conditional independence, copulas, conditional densities, and quantiles/expectiles (Bouezmarni et al., 2012; Taamouti et al., 2014; Song and Taamouti, 2018; Song and Taamouti, 2021; Bouezmarni et al., 2024). These contributions highlight that in many applications the goal is not merely to reject "no causality" but to quantify predictive importance in an interpretable way.

In functional settings, Granger-causality questions are often intrinsically functional rather than finite-dimensional. For instance, whether a yield curve $X_t(\cdot)$ helps forecast another curve $Y_t(\cdot)$ can depend on maturity; similarly, an intraday volatility profile may predict specific regions of an implied-volatility smile. Treating curves as high-dimensional vectors (e.g., via dense discretization) typically yields a large-parameter vector autoregression with exogenous variables (VARX), which can make inference unstable; by contrast, score-based VAR approaches built on functional principal component analysis (FPCA) require selecting a

truncation level that may mask localized predictive effects. A fully functional approach instead models lagged dependence through bounded linear operators on $\mathbb{H}$, retaining the infinite-dimensional structure and making the notion of "X helps forecast the entire curve Y" precise.

Related work on Granger causality for function-valued time series remains limited and has largely focused on point-null questions ("any non-zero effect") combined with projection-based implementation. Amblard and Michel (2013) formulate Granger causality for function-valued signals in a Hilbert-space framework, motivated partly by non-stationary signal-processing applications, but they do not develop an econometric testing theory for stationary functional Granger causality, which is the focus of this paper. Saumard (2017) formulates linear Granger non-causality for stationary functional time series, showing that under the null it can be characterized through restrictions on second-order moments, leading to a test based on comparisons of covariance operators. Implementation relies on existing procedures for testing equality of covariance operators for dependent functional data, such as Zhang and Shao (2015), typically carried out via projection methods. Beyond covariance-operator approaches, Shang et al. (2021) develop generalized correlation-type causality measures for bivariate stationary curve time series, while Sen et al. (2022) study Granger causality in multivariate functional autoregressions via Bayes factors. These approaches share limitations pertinent to the present work, as they aim to detect any non-zero predictive effect and rely on projection onto finite-dimensional subspaces. In contrast, we develop new procedures for relevant functional causality, based on operator-based effect sizes and fully functional/infinite-dimensional inference.

Our contribution departs from this literature along several dimensions. First, rather than testing whether cross-lag operators vanish (zero), we test whether an operator-based effect size exceeds a practically meaningful threshold, thereby targeting relevant functional Granger causality. Second, we formalize and analyze three functional analogues of classical Granger-causality measures: operator energy, forecast gain, and residual–predictor cross-covariance energy. Each captures distinct predictive content and requires separate asymptotic theory due to its unique limiting structure. To our knowledge, this unified operator-based framework does not appear elsewhere in the functional Granger causality literature. Third, methodologically, we combine ridge-regularized estimation of lag operators directly in the Hilbert space with self-normalized inference, yielding tuning-free tests that avoid long-run variance estimation and finite-dimensional truncation while delivering pivotal limit distributions under weak dependence. Taken together, these features extend the functional time-series literature beyond point-null non-causality testing toward fully infinite-dimensional, practically meaningful inference.

Classical Granger-causality tests target the point null "no causality" and can reject for effects that are statistically detectable but practically negligible. This concern is amplified for functional data: rich function spaces make it easy to pick up tiny localized dependencies (and, in turn, tiny forecast gains) that are statis-

tically significant but not substantively meaningful. In applications, causality is often actionable only above a materiality threshold. For example, when assessing whether option-trading activity, an implied-volatility curve, or an intraday order-flow profile helps forecast the subsequent volatility curve of the underlying asset, one may want to be conclusive only if the resulting gain is large enough to justify a statement such as "economically relevant predictability" (e.g. large enough to affect risk limits, hedging decisions, or regulatory conclusions about destabilization). This motivates a relevant hypothesis formulation:

$$H_0(\Delta) : d \leq \Delta \quad \text{vs} \quad H_1(\Delta) : d > \Delta,$$

where d is a nonnegative effect size measuring $X \to Y$ predictive strength and $\Delta > 0$ is a user-chosen relevance threshold reflecting scientific or economic significance. Relevant hypotheses avoid the classical "consistency against negligible effects" pitfall and provide a principled way to guard against spurious "discoveries" that can arise from the sheer richness of functional data. In functional time series, Dette et al. (2020) develop self-normalized inference for testing relevant hypotheses directly in a Hilbert space. Building on this line of work, Lin et al. (2026) propose an adjusted-range self-normalizer; in this paper we adopt both a quadratic-form and an adjusted-range SN to construct tuning-free relevant tests for functional Granger causality.

For inference under weak dependence, standard studentization typically requires estimating a long-run variance/covariance object and hence choosing bandwidths or lag truncation parameters. SN offers a tuning-free alternative by using subsample-path variation as an internal normalizer (Shao, 2010; Hong et al., 2024; Hong et al., 2025). A key advantage of SN for relevant hypotheses with $\Delta > 0$ is that, at the boundary case $d = \Delta$, the unknown long-run scale enters both the numerator and the self-normalizer and therefore cancels in the ratio, yielding a pivotal limit that depends only on Brownian motion functionals. Although the relevant framework nests the point null by letting $\Delta \downarrow 0$, this cancellation mechanism does not extend to the point-null $\Delta = 0$ case: there the limiting law is a weighted χ^2 distribution depending on the (unknown) spectrum of a covariance operator, so there is no single nuisance scale to eliminate. Accordingly, we treat $\Delta = 0$ as a degenerate benchmark and develop plug-in critical values (and local power analysis) for empirical comparison, while our main methodological focus is on $\Delta > 0$ where SN is most beneficial. Our implementation is fully functional: lag operators are estimated directly in $\mathbb{H}$ via regularized operator estimation (e.g. ridge/Tikhonov), so no fixed FPCA truncation level is part of the model; any basis/grid representation is only numerical.

In simulations, the proposed self-normalized tests exhibit accurate boundary size and strong power, with the adjusted-range normalizer performing particularly well under dependence and heavy-tailed innovations. In an empirical illustration using 2025 Bitcoin data (Deribit options and Bitfinex spot), implied-volatility

smiles predict next-day spot realized volatility at a modest relevance threshold, but the evidence weakens at stricter thresholds; reverse and flow-based spillovers are not economically material.

The remainder of the paper is organized as follows. Section 1 introduces the functional Granger-causality setup and the relevant effect sizes considered in this work. Section 2 develops self-normalized offline tests, stating the required assumptions and core large-sample results. Sections 3 and 4 report simulations and an empirical illustration. Section 5 concludes and discusses extensions.

1 Functional Granger causality

Let $\mathbb{H}$ be a separable Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$ (e.g. $\mathbb{H} = L^2(\mathcal{T})$ on a compact domain $\mathcal{T}$). We observe a paired functional time series $(X_t, Y_t) \in \mathbb{H}\times\mathbb{H}, t = 1,\dots,n$, and write $\mathcal{F}_{t-1}$ for the σ-field generated by $\{(X_s, Y_s) : s \le t-1\}$ and $\mathcal{F}^Y_{t-1}$ for the past of Y only. Here (X_t, Y_t) is a sequence of random elements in the product Hilbert space $\mathbb{H}\times\mathbb{H}$, with X_t and Y_t themselves random elements of $\mathbb{H}$ (i.e. random functions such as curves or surfaces). All equalities below are understood as equalities in $\mathbb{H}$ (e.g. in L^2 if $\mathbb{H} = L^2(\mathcal{T})$). This differs from classical multivariate time series, where (X_t, Y_t) takes values in $\mathbb{R}^d$ for fixed d: here the state space is infinite-dimensional, its geometry is governed by the Hilbert-space structure, and the index t labels sequentially observed functions rather than finite-dimensional vectors.

Let $(\mathbb{H}, \langle\cdot,\cdot\rangle)$ be a separable Hilbert space with norm $\|x\| \overset{\text{def}}{=} \sqrt{\langle x, x\rangle}$, and assume throughout that $\langle\cdot,\cdot\rangle$ is linear in its first argument. When $\mathbb{H} = L^2(\mathcal{T})$, we take $\langle f, g\rangle \overset{\text{def}}{=} \int_{\mathcal{T}} f(t)\, g(t)\, dt$ (or $\langle f, g\rangle \overset{\text{def}}{=} \int_{\mathcal{T}} f(t)\, \overline{g(t)}\, dt$ in the complex-valued case). For $x, y \in \mathbb{H}$, define the rank-one operator $x\otimes y : \mathbb{H}\to\mathbb{H}$ by $(x\otimes y)z \overset{\text{def}}{=} \langle z, y\rangle\, x$ for $z \in \mathbb{H}$. For square-integrable random elements $U, V \in \mathbb{H}$, define the (cross-)covariance operator $\Gamma_{UV} \overset{\text{def}}{=} \mathrm{Cov}(U, V) \overset{\text{def}}{=} \mathsf{E}\big[(U - \mathsf{E}\,U)\otimes(V - \mathsf{E}\,V)\big]$, and write $\Gamma_U \overset{\text{def}}{=} \Gamma_{UU}$. For a bounded linear operator $A : \mathbb{H}\to\mathbb{H}$, A^* denotes its Hilbert-space adjoint, i.e. $\langle Ax, y\rangle = \langle x, A^*y\rangle$ for all $x, y \in \mathbb{H}$; in particular, $\Gamma_{VU} = \Gamma^*_{UV}$.

A natural notion of functional Granger noncausality is

$$X \text{ does not Granger-cause } Y \text{ in mean if } \ \mathsf{E}[Y_t \mid \mathcal{F}_{t-1}] = \mathsf{E}\big[Y_t \mid \mathcal{F}^Y_{t-1}\big] \quad (\text{in } \mathbb{H}). \tag{1}$$

Here the conditional expectations are $\mathbb{H}$-valued random elements, so (1) requires equality of conditional mean functions, not merely equality of finitely many coordinates. Thus, functional Granger noncausality in mean means that the conditional mean function of the entire future curve $Y_t(\cdot)$ is the same whether or not past X-curves are included, once the past of Y is conditioned on.

In linear prediction classes, the statement in (1) corresponds to restrictions on cross-lag operators in a functional VARX model. A prototypical model is the functional autoregression with exogenous functional

regressors (FARX):

$$Y_t = \sum_{j=1}^{p} A_j Y_{t-j} + \sum_{j=1}^{p} B_j X_{t-j} + \varepsilon_t, \tag{2}$$

where $A_j, B_j : \mathbb{H} \to \mathbb{H}$ are bounded linear operators and ε_t is a mean-zero innovation in $\mathbb{H}$. Each operator A_j or B_j maps an entire past function to a new function. Based on the above model, functional Granger noncausality corresponds to

$$B_1 = \cdots = B_p = 0,$$

that is, the absence of any cross-lag operator mapping past X-curves into the conditional mean of Y_t.

We measure causality or predictive strength from $X \to Y$ using a scalar effect size d and test a relevant hypothesis. Three complementary fully functional choices follow. Importantly, each effect size below is defined directly in terms of operators or $\mathbb{H}$-norms, thereby capturing the causality or predictive strength of entire functions rather than selected coordinates.

(ES1) Operator energy. In (2), define

$$d_B \;=\; \sum_{j=1}^{p} \|B_j\|_{\mathrm{HS}}^2 \,, \tag{3}$$

where $\|\cdot\|_{\mathrm{HS}}$ denotes the Hilbert–Schmidt norm (defined for Hilbert–Schmidt operators, which are compact). When the B_j are Hilbert–Schmidt, $\|B_j\|_{\mathrm{HS}}^2$ equals the sum of squared singular values of B_j. Hence d_B measures the overall (global) strength of the mapping from past X-curves to the conditional mean function of Y_t across the domain $\mathcal{T}$.

(ES2) Forecast gain. Let $\hat{Y}_{t|t-1}^{(U)}$ and $\hat{Y}_{t|t-1}^{(R)}$ denote best linear one-step predictors of Y_t from $(X_{t-1:t-p}, Y_{t-1:t-p})$ (unrestricted) and from $Y_{t-1:t-p}$ only (restricted). Define

$$d_{\mathrm{PE}} \;=\; \mathsf{E}\left\|Y_t - \hat{Y}_{t|t-1}^{(R)}\right\|^2 - \mathsf{E}\left\|Y_t - \hat{Y}_{t|t-1}^{(U)}\right\|^2 \;\geq 0. \tag{4}$$

Here $\|\cdot\|$ is the $\mathbb{H}$-norm, so the prediction error is integrated over the full function domain. Hence d_{PE} measures the gain in integrated mean squared prediction accuracy of the entire curve $Y_t(\cdot)$ when past X-curves are included.

(ES3) Residual–predictor cross-covariance energy. Fit the restricted model (no X lags) and compute residuals $u_t = Y_t - \hat{Y}_{t|t-1}^{(R)}$. Let

$$C_{uX}^{(j)} = \mathsf{E}\left[u_t \otimes X_{t-j}\right], \qquad j = 1, \ldots, p,$$

and set

$$d_{uX} \;=\; \sum_{j=1}^{p} \left\| C_{uX}^{(j)} \right\|_{\mathrm{HS}}^{2}. \tag{5}$$

Each $C_{uX}^{(j)}$ is a Hilbert–Schmidt cross-covariance operator capturing linear dependence between the restricted-model residual $u_t(\cdot)$ and the lagged predictor $X_{t-j}(\cdot)$. If $C_{uX}^{(j)} = 0$, then u_t is uncorrelated with X_{t-j} in all directions, so after accounting for the restricted model there is no remaining linear predictive content from X_{t-j} for Y_t at lag j. Consequently, in the linear FARX/projection setting, $d_{uX} = 0$ (equivalently $C_{uX}^{(j)} = 0$ for all j) corresponds to functional Granger noncausality.

For any choice $d \in \{d_B, d_{\mathrm{PE}}, d_{uX}\}$ we test the relevant causality hypothesis

$$H_0(\Delta) : d \le \Delta \qquad \text{vs} \qquad H_1(\Delta) : d > \Delta, \tag{6}$$

where $\Delta > 0$ encodes economic/scientific relevance.

It is worth noting that, by formulating d in terms of operators and $\mathbb{H}$-norms, (6) concerns the predictive strength of entire functions and is invariant to discretization or basis representation. This distinguishes our framework from classical finite-dimensional Granger-causality tests applied to densely sampled curves.

2 Self-normalized inference

This section develops (i) a regularized estimation strategy for the causality effect size d in a fully functional setting and (ii) tuning-free offline tests based on SN, focusing on inference that operates directly in the Hilbert-space framework and avoids finite-dimensional truncation or bandwidth selection.

2.1 Regularized estimation in functional autoregressions

To retain a fully functional framework, we do not impose a fixed FPCA truncation level as part of the model. Instead, we estimate lag operators in the functional autoregression model directly in the Hilbert space using regularization. Covariance operators in functional settings are typically compact and may not admit bounded inverses, motivating regularized estimation (Bosq, 2000; Horváth and Kokoszka, 2012; Aue et al., 2015). This approach keeps the estimation in the infinite-dimensional Hilbert space and avoids modeling choices tied to finite-dimensional representations.

Define the stacked lag vector

$$Z_{t-1} = (Y_{t-1}, \ldots, Y_{t-p},\, X_{t-1}, \ldots, X_{t-p}) \in \mathbb{H}^{2p},$$

collecting the relevant functional history of X and Y. This allows (2) to be written in operator form. Specifically, define the block operator $\mathcal{B} : \mathbb{H}^{2p} \to \mathbb{H}$ by

$$\mathcal{B}(z_1, \ldots, z_{2p}) = \sum_{j=1}^{p} A_j z_j + \sum_{j=1}^{p} B_j z_{p+j},$$

so that

$$Y_t = \mathcal{B} Z_{t-1} + \varepsilon_t.$$

This representation parallels finite-dimensional vector autoregressions but operates directly in Hilbert space, capturing dependence among entire functions rather than selected coordinates or basis coefficients.

Estimation of $\mathcal{B}$ is challenging because covariance operators in functional settings are compact and typically do not have bounded inverses. Regularization addresses this ill-posedness by stabilizing the inverse problem while preserving estimation in the infinite-dimensional space. We therefore employ Tikhonov (ridge) regularization via penalized least squares over Hilbert–Schmidt operators, yielding the estimator

$$\widehat{\mathcal{B}}_\lambda = \arg\min_{\mathcal{B}} \sum_{t=p+1}^{n} \|Y_t - \mathcal{B} Z_{t-1}\|^2 + \lambda \|\mathcal{B}\|_{\mathrm{HS}}^2, \qquad \lambda > 0. \tag{7}$$

Let

$$\widehat{\Gamma}_{ZZ} = \frac{1}{n-p} \sum_{t=p+1}^{n} Z_{t-1} \otimes Z_{t-1}, \qquad \widehat{\Gamma}_{YZ} = \frac{1}{n-p} \sum_{t=p+1}^{n} Y_t \otimes Z_{t-1},$$

denote the empirical second-moment and cross-moment operators. Then the normal equations yield the closed-form solution

$$\widehat{\mathcal{B}}_\lambda = \widehat{\Gamma}_{YZ} \left(\widehat{\Gamma}_{ZZ} + \lambda I\right)^{-1}, \tag{8}$$

where I is the identity operator on $\mathbb{H}^{2p}$. The estimated lag operators $\widehat{A}_{j,\lambda}$ and $\widehat{B}_{j,\lambda}$ correspond to the respective blocks of $\widehat{\mathcal{B}}_\lambda$, providing direct estimates of the dynamic relationships between past and future functions. These operators form the basis for the effect-size measures used for functional Granger-causality analysis. Thus, for the operator-based effect size (3), we use these estimators to compute

$$\widehat{d}_B = \sum_{j=1}^{p} \left\| \widehat{B}_{j,\lambda} \right\|_{\mathrm{HS}}^2,$$

which aggregates the strength of all cross-lag mappings.

For the forecast-improvement measure (4), we compare one-step prediction errors under restricted and unrestricted models, yielding an integrated assessment of predictive gain. For the cross-covariance measure

(5), we estimate residuals from the restricted model and form

$$\widehat{C}_{uX}^{(j)} = \frac{1}{n-p}\sum_{t=p+1}^{n} u_t \otimes X_{t-j}, \qquad \widehat{d}_{uX} = \sum_{j=1}^{p} \left\| \widehat{C}_{uX}^{(j)} \right\|_{\mathrm{HS}}^2,$$

which quantifies remaining cross-lag dependence after accounting for past Y dynamics.

All objects are defined in the infinite-dimensional space $\mathbb{H}$; discretization or basis expansion is used only for numerical implementation.

2.2 Self-normalized testing of functional Granger causality

We develop tuning-parameter-free inference for functional Granger causality, quantified by the effect size d, using SN, thereby avoiding the need to explicitly estimate long-run variance operators.

Let $\hat{d}_n$ denote an estimator of d based on the full sample $\{(X_t, Y_t)\}_{t=1}^n$. For $\lambda \in [0,1]$, let $\hat{d}_n(\lambda)$ be the red corresponding partial-sample statistic based on the first $\lfloor n\lambda \rfloor$ observations, constructed with the same normalization as $\hat{d}_n(1)$. Define the centered subsample path

$$G_n(\lambda) \stackrel{\text{def}}{=} \hat{d}_n(\lambda) - \lambda^2 \hat{d}_n(1), \qquad \lambda \in [0,1]. \tag{9}$$

The centering term $\lambda^2 \hat{d}_n(1)$ is tailored to these quadratic (squared-norm) effect sizes and ensures that $\sqrt{n}\, G_n(\cdot)$ has a non-degenerate limit. SN replaces the unknown dependence scale by a functional of $G_n(\cdot)$.

We consider two standard self-normalizers. First, the quadratic SN (Shao, 2010) aggregates subsample fluctuations via

$$\widehat{V}_n \stackrel{\text{def}}{=} \left(\int_0^1 G_n(\lambda)^2 \, \nu(\,\mathrm{d}\lambda) \right)^{1/2}, \tag{10}$$

where ν is a probability measure on $(0,1)$, often taken as discrete uniform on a grid away from 0 and 1. This automatically adapts to serial dependence. Second, the adjusted-range SN (Hong et al., 2024) emphasizes extremal deviations, defined as

$$\widehat{H}_n \stackrel{\text{def}}{=} \sup_{\lambda \in [0,1]} G_n(\lambda) - \inf_{\lambda \in [0,1]} G_n(\lambda), \tag{11}$$

with sup and inf evaluated on a fine grid. This normalization is particularly robust to local irregularities in the subsample path.

To establish the asymptotic validity of the tests based on the self-normalizers introduced above, we build on the self-normalized relevant-testing framework for functional time series developed in Dette et al. (2020) and tailor it to the present functional Granger-causality effect sizes. Specifically, we characterize the joint asymptotic behavior of the full-sample effect-size estimator $\hat{d}_n$ and the centered subsample path $G_n(\cdot)$ in

(9). This joint weak limit yields that the resulting self-normalized statistics (based on either $\widehat{V}_n$ or $\widehat{H}_n$) converge, under the boundary case $d = \Delta$, to functionals of standard Brownian motion and are therefore asymptotically pivotal in the sense that nuisance scale parameters (e.g., long-run variance terms) cancel out. This pivotality underpins the validity of the generic rejection rule stated in (13). Formalizing the argument requires strict stationarity and weak dependence conditions compatible with the Bernoulli-shift and m-approximability framework for functional time series (Hörmann and Kokoszka, 2010; Berkes et al., 2013), together with stability and regularity conditions for the FARX dynamics used to construct $\hat{d}_n$. For notational compactness, let $Z_t = (X_t, Y_t) \in \mathbb{H}^2$.

Assumption 1 (Stationarity and Bernoulli shift)**.** *$(Z_t)_{t\in\mathbb{Z}}$ is strictly stationary with mean $\mu \in \mathbb{H}^2$ and admits a Bernoulli-shift representation $Z_t = \mu + f(\varepsilon_t, \varepsilon_{t-1}, \ldots)$ for some measurable map f and i.i.d. innovations (ε_t).*

Assumption 2 (Weak dependence and moments)**.** *There exists $\psi \in (0, 1)$ such that $\mathsf{E}\,\|Z_0\|^{4+\psi} < \infty$ and Z_t is m-approximable in the sense of Hörmann and Kokoszka (2010) with summable approximation error of order $4 + \psi$.*

Assumption 3 (Well-posed FARX dynamics)**.** *Model (2) holds with bounded linear operators $A_j, B_j : \mathbb{H} \to \mathbb{H}$ and mean-zero innovations (ε_t). Let $\mathcal{A} : \mathbb{H}^p \to \mathbb{H}^p$ be the companion operator*

$$\mathcal{A}(y_1, \ldots, y_p) = \Big(\sum_{j=1}^{p} A_j y_j,\, y_1, \ldots, y_{p-1} \Big).$$

Assume geometric stability: $\|\mathcal{A}^k\|_{\mathrm{op}} \le C\rho^k$ for some $C < \infty$ and $\rho \in (0, 1)$. Moreover, $\mathsf{E}(\varepsilon_t \otimes X_{t-j}) = \mathsf{E}(\varepsilon_t \otimes Y_{t-j}) = 0$ for all $j \ge 1$.

Assumption 4 (Joint weak invariance principle for the effect-size estimator)**.** *Let d be one of (3)–(5) and let $\hat{d}_n$ be the corresponding estimator. Assume that in $\mathbb{R} \times C([0, 1])$,*

$$\Big(\sqrt{n}(\hat{d}_n - d),\, \sqrt{n} G_n(\cdot) \Big) \Rightarrow \Big(\tau B(1),\, \tau U(\cdot) \Big), \tag{12}$$

where B is standard Brownian motion, $\tau > 0$, and $U(\lambda) \overset{\mathrm{def}}{=} \lambda B(\lambda) - \lambda^2 B(1)$.

Remark 1 (On verifying Assumption 4)**.** *Assumptions 1–2 support functional invariance principles for sample means of $\mathbb{H}$-valued and Hilbert–Schmidt-valued processes. For d_{uX}, for instance, one may establish a Hilbert–Schmidt CLT for the sample mean of $u_t \otimes X_{t-j}$ and then apply the delta method to $C \mapsto \|C\|_{\mathrm{HS}}^2$ away from the origin; the relevance condition $\Delta > 0$ avoids degeneracy at the boundary. Regularization and residual-estimation errors must be controlled so that they are of smaller order than $n^{-1/2}$ in (12).*

Theorem 1 (Generic SN principle)**.** *Under Assumption 4, for any continuous positively homogeneous functional $\rho : C([0,1]) \to [0,\infty)$ with $\rho\{U(\cdot)\} > 0$ a.s.,*

$$\frac{\hat{d}_n - d}{\rho(G_n)} \Rightarrow \frac{B(1)}{\rho\{U(\cdot)\}},$$

and the limit law is pivotal (free of τ). In particular, (10) *and* (11) *yield the quadratic and adjusted-range pivots.*

Theorem 2 (Asymptotic level and consistency in the relevant regime)**.** *Assume $\Delta > 0$ and Assumption 4. Then the test* (13) *satisfies*

$$\lim_{n\to\infty} \mathbb{P}\{\text{reject}\} = \begin{cases} 0, & d < \Delta, \\ \alpha, & d = \Delta, \\ 1, & d > \Delta, \end{cases}$$

provided $\tau > 0$ at the boundary case $d = \Delta$.

Theorem 3 (Pitman local power)**.** *Assume $\Delta > 0$ and Assumption 4, and consider local alternatives $d_n = \Delta + c/\sqrt{n}$. Then the rejection probability of* (13) *converges to*

$$\mathbb{P}\left\{ \frac{B(1) + c/\tau}{\rho\{U(\cdot)\}} > q_{1-\alpha} \right\},$$

where ρ is the functional defining the chosen self-normalizer. In particular, the limit equals α at $c = 0$ and is nondecreasing in c (strictly increasing under continuity of the limit law).

Remark 2 (Degeneracy at $\Delta = 0$)**.** *When $\Delta = 0$ and $d = 0$, quadratic effect sizes are typically degenerate at first order. This is a principal reason for focusing on relevant hypotheses with $\Delta > 0$; see Dette et al. (2020).*

We now define the self-normalized rejection rule for testing relevant functional Granger causality. The rule compares the appropriately normalized effect-size estimator to the critical value from the corresponding pivotal limit, thereby avoiding nuisance parameters. Formally, the rejection rule is given by

$$\text{Reject } H_0 : d \le \Delta \quad \Longleftrightarrow \quad \frac{\hat{d}_n - \Delta}{\rho(G_n)} > q_{1-\alpha}, \tag{13}$$

where $\rho(G_n) = \widehat{V}_n$, defined in (10), for the quadratic SN and $\rho(G_n) = \widehat{H}_n$, defined in (11), for the adjusted-range SN, and $q_{1-\alpha}$ denotes the $(1-\alpha)$-quantile (hereafter critical value) of the associated pivotal limit.

The critical values for the quadratic SN follow Dette et al. (2020) and are obtained from simulated

quantiles based on 1000 replications using the grid $\lambda_i = i/20$, $i = 1, \ldots, 19$, yielding $q_{0.90} = 7.619$, $q_{0.95} = 10.530$, and $q_{0.99} = 16.081$. For the adjusted-range SN, the supremum and infimum are approximated on the same grid, and based on 10^6 Brownian-motion replications, we obtain critical values $q^*_{0.90}(\mathbb{S}) \approx 2.432$, $q^*_{0.95}(\mathbb{S}) \approx 3.269$, and $q^*_{0.99}(\mathbb{S}) \approx 5.159$.

2.3 Pitman local power under contiguous relevant alternatives

To quantify the sensitivity of our self-normalized tests to small deviations from the relevance boundary $d = \Delta$, we consider *Pitman local alternatives* of the form

$$H_{1,n}(c): \qquad d = d_n = \Delta + \frac{c}{\sqrt{n}}, \qquad c \in \mathbb{R}, \tag{14}$$

which exceed the relevance threshold by an amount of order $n^{-1/2}$, capturing asymptotic contiguous departures from the null. Theorem 4 below then establishes the corresponding local power of our self-normalized test.

Theorem 4 (Pitman local limit and local power)**.** *Assume $\Delta > 0$ and Assumption 4. Under $H_{1,n}(c)$,*

$$\frac{\hat{d}_n - \Delta}{\rho(G_n)} \Rightarrow \frac{B(1) + c/\tau}{\rho\{U(\cdot)\}}, \qquad U(\lambda) \stackrel{\text{def}}{=} \lambda B(\lambda) - \lambda^2 B(1). \tag{15}$$

where B is standard Brownian motion and τ is the asymptotic scale parameter from Assumption 4. Consequently, the asymptotic local power function is

$$\pi_\rho(c) \stackrel{\text{def}}{=} \lim_{n\to\infty} \mathbb{P}_{H_{1,n}(c)}\left\{ \frac{\hat{d}_n - \Delta}{\rho(G_n)} > q_{1-\alpha} \right\} = \mathbb{P}\left\{ \frac{B(1) + c/\tau}{\rho\{U(\cdot)\}} > q_{1-\alpha} \right\}. \tag{16}$$

In particular, $\pi_\rho(0) = \alpha$, $\pi_\rho(c)$ is nondecreasing in c, and $\lim_{c\to\infty} \pi_\rho(c) = 1$, $\lim_{c\to-\infty} \pi_\rho(c) = 0$.

Lemma 1 (A useful Brownian-bridge decomposition)**.** *Let B be standard Brownian motion on $[0,1]$ and define the Brownian bridge $\mathbb{B}(\lambda) \stackrel{\text{def}}{=} B(\lambda) - \lambda B(1)$. Then $\mathbb{B}(\cdot)$ is independent of $B(1)$ and*

$$U(\lambda) = \lambda B(\lambda) - \lambda^2 B(1) = \lambda\, \mathbb{B}(\lambda),$$

so that $\rho\{U(\cdot)\}$ is independent of $B(1)$ for any measurable functional ρ.

Corollary 5 (Closed-form representation of the local power curve)**.** *Under Theorem 4 and Lemma 1, let*

$R \overset{\text{def}}{=} \rho\{U(\cdot)\}$. *Then* $B(1) \sim N(0,1)$ *is independent of* R, *and*

$$\pi_\rho(c) = \mathbb{E}\Big[1 - \Phi\Big(q_{1-\alpha} R - \frac{c}{\tau}\Big)\Big], \tag{17}$$

where Φ *is the standard normal cdf. Moreover, if* φ *denotes the standard normal pdf,*

$$\pi'_\rho(0) = \frac{1}{\tau}\,\mathbb{E}[\varphi(q_{1-\alpha} R)] > 0. \tag{18}$$

The representation (17) *is convenient for plotting local-power curves and for comparing different self-normalizers via their Pitman slope* (18).

Corollary 6 (Special cases: quadratic SN and adjusted-range SN)**.** *(i) For the quadratic SN in* (10),

$$R = \rho\{U(\cdot)\} = \left(\int_0^1 U(\lambda)^2 \, \nu(\,\mathrm{d}\lambda)\right)^{1/2}.$$

(ii) For the adjusted-range SN in (11),

$$R = \rho\{U(\cdot)\} = \sup_{\lambda \in [0,1]} U(\lambda) - \inf_{\lambda \in [0,1]} U(\lambda).$$

Plugging the corresponding R *into* (17) *yields the asymptotic local-power curve* $\pi_\rho(c)$.

Note that the local power depends on c only through the standardized quantity c/τ; one may therefore plot $\pi_\rho(\kappa)$ with $\kappa = c/\tau$ to avoid nuisance-scale comparisons across designs.

The representation (17) also provides a convenient basis for numerical evaluation of the asymptotic local power function. Because the distribution of $R = \rho\{U(\cdot)\}$ depends only on functionals of Brownian motion, the limit power curve can be accurately approximated using Monte Carlo simulation, without requiring estimation of nuisance parameters or additional model-specific quantities. This approach allows one to visualize the sensitivity of the proposed test to contiguous relevant alternatives and to compare the performance of different self-normalizers. Algorithm 1 summarizes a simple simulation procedure for approximating the limit power function based on discretized Brownian bridge paths.

Algorithm 1 Monte Carlo approximation of $\pi_\rho(c)$ via (17)

Fix c, τ (or use the standardized drift c/τ), and the critical value $q_{1-\alpha}$.
Choose a fine grid $0 = \lambda_0 < \lambda_1 < \cdots < \lambda_m = 1$.
for $r = 1, \ldots, R$ **do**
  Simulate a Brownian bridge path $\{\mathbb{B}^{(r)}(\lambda_i)\}_{i=0}^m$ on the grid.
  Set $U^{(r)}(\lambda_i) = \lambda_i \, \mathbb{B}^{(r)}(\lambda_i)$ and compute $R^{(r)} = \rho\{U^{(r)}(\cdot)\}$.
  Compute $p^{(r)}(c) = 1 - \Phi\big(q_{1-\alpha} R^{(r)} - \frac{c}{\tau}\big)$.
end for
Output $\widehat{\pi}_\rho(c) = \frac{1}{R} \sum_{r=1}^R p^{(r)}(c)$.

2.4 Point-null comparison: the degenerate case and plug-in inference

The self-normalized procedures developed in Section 2.2 are designed for relevant hypotheses with strictly positive threshold $\Delta > 0$, where the effect size admits a standard $\sqrt{n}$ asymptotic expansion and yields pivotal self-normalized limits. When $\Delta = 0$, the relevant null $H_0 : d \leq \Delta$ reduces to the classical point null $H_0 : d = 0$, which corresponds to testing the absence of any causal effect. In this case, however, the asymptotic representation changes fundamentally. For the quadratic effect sizes considered in this paper, the point-null case is degenerate at first order, in the sense that $\hat{d}_n = O_p(n^{-1})$ under H_0. Consequently, the $\sqrt{n}$-expansion in Assumption 4 collapses, and the self-normalized statistic in (13) converges to zero rather than to a non-degenerate pivotal limit; see Dette et al. (2020).

To obtain a non-degenerate limit theory under the point null, it is therefore natural to work with the underlying linear Hilbert-space parameter whose squared norm defines the effect size. Specifically, many functional Granger-causality effect sizes can be written as

$$d = \|\theta\|_{\mathcal{H}}^2, \tag{19}$$

for some separable Hilbert space $\mathcal{H}$ and parameter $\theta \in \mathcal{H}$: for example, $\theta = (B_1, \ldots, B_p) \in \mathcal{L}_{\mathrm{HS}}(\mathbb{H})^p$ yields $d = d_B$ in (3), and $\theta = (C_{uX}^{(1)}, \ldots, C_{uX}^{(p)}) \in \mathcal{L}_{\mathrm{HS}}(\mathbb{H})^p$ yields $d = d_{uX}$ in (5). Let $\hat{\theta}_n$ denote the corresponding estimator (obtained from ridge/Tikhonov fits and sample cross-moments), and set

$$\hat{d}_n = \|\hat{\theta}_n\|_{\mathcal{H}}^2. \tag{20}$$

For the degenerate point-null case $\Delta = 0$, we develop a valid asymptotic theory based on the underlying linear Hilbert-space parameter, yielding non-degenerate limit distributions and reliable inference. Formalizing this point-null framework requires an additional assumption on the Hilbert-space CLT for the estimator $\hat{\theta}_n$ and on the covariance structure of its limiting Gaussian element, which we state next.

Assumption 5 (CLT for the linear object in the point-null case)**.** *In the Hilbert space $\mathcal{H}$, the estimator $\hat{\theta}_n$ satisfies*

$$\sqrt{n}\,(\hat{\theta}_n - \theta) \Rightarrow \mathcal{G},$$

where $\mathcal{G}$ is a mean-zero Gaussian element in $\mathcal{H}$ with (trace-class) covariance operator Σ. Let $\{(\lambda_k, e_k)\}_{k\geq 1}$ denote the (nonnegative) eigenvalues/eigenfunctions of Σ.

This assumption is standard in functional time-series econometrics and follows, for example, from functional central limit theorems for Hilbert–Schmidt operator-valued sample means.

Theorem 7 (Point-null limit distribution for $\Delta = 0$)**.** *Assume* (19)–(20) *and Assumption 5. Under the point null $H_0 : \theta = 0$ (equivalently $d = 0$), the scaled statistic*

$$T_n \overset{\text{def}}{=} n\hat{d}_n = \left\| \sqrt{n}\,\hat{\theta}_n \right\|_{\mathcal{H}}^2 \tag{21}$$

converges in distribution to

$$T_n \Rightarrow \|\mathcal{G}\|_{\mathcal{H}}^2 \overset{d}{=} \sum_{k=1}^{\infty} \lambda_k Z_k^2, \tag{22}$$

where $(Z_k)_{k\geq 1}$ are i.i.d. $N(0,1)$.

The limit distribution in (22) is a weighted chi-square law, reflecting the infinite-dimensional Gaussian fluctuations of the underlying operator estimator. This representation highlights the role of the long-run covariance operator Σ, whose spectrum determines the relative contribution of different functional directions to the test statistic.

In practice, critical values can be obtained using plug-in estimation of the covariance operator. Specifically, estimate Σ using a consistent long-run covariance estimator for the influence process driving $\hat{\theta}_n$, for example via HAC estimation after basis expansion or discretization. Let $\hat{\lambda}_k$ denote the eigenvalues of the estimated covariance operator. Then simulate

$$\hat{T}^{(r)} = \sum_{k=1}^{K} \hat{\lambda}_k Z_{k,r}^2, \qquad Z_{k,r} \overset{\text{i.i.d.}}{\sim} N(0,1),$$

for $r = 1, \ldots, R$, and approximate the critical value $c_{1-\alpha}$ by the empirical $(1-\alpha)$-quantile of $\{\hat{T}^{(r)}\}_{r=1}^{R}$.

As a computationally efficient alternative, one may use the Satterthwaite moment-matching approximation, which replaces the weighted chi-square distribution by a scaled chi-square law matching the first two

moments. Letting $m_1 = \sum_k \hat{\lambda}_k$ and $m_2 = 2\sum_k \hat{\lambda}_k^2$, define

$$\nu = \frac{2m_1^2}{m_2}, \qquad a = \frac{m_2}{2m_1}, \tag{23}$$

so that $\sum_k \hat{\lambda}_k Z_k^2$ is approximated by $a\,\chi_\nu^2$.

This plug-in procedure provides a feasible approach for classical point-null testing, but in contrast to the self-normalized relevant test, it requires explicit estimation of the covariance operator.

Theorem 8 (Point-null local power under $n^{-1/2}$ operator alternatives)**.** *Assume* (19)–(20) *and Assumption 5. Consider a sequence of local alternatives*

$$H_{1,n}(h): \qquad \theta = \theta_n = \frac{h}{\sqrt{n}}, \qquad h \in \mathcal{H} \text{ fixed.} \tag{24}$$

Then

$$T_n \Rightarrow \|\mathcal{G} + h\|_{\mathcal{H}}^2. \tag{25}$$

Equivalently, writing $h_k = \langle h, e_k \rangle$ and letting (Z_k) be i.i.d. $N(0,1)$,

$$\|\mathcal{G} + h\|_{\mathcal{H}}^2 \overset{d}{=} \sum_{k:\lambda_k>0} \lambda_k \Big(Z_k + \frac{h_k}{\sqrt{\lambda_k}}\Big)^2 + \sum_{k:\lambda_k=0} h_k^2. \tag{26}$$

If $c_{1-\alpha}$ denotes the $(1-\alpha)$-quantile of the null limit in (22)*, the corresponding asymptotic local power is*

$$\pi_0(h) \overset{\text{def}}{=} \lim_{n\to\infty} \mathbb{P}_{H_{1,n}(h)}\{T_n > c_{1-\alpha}\} = \mathbb{P}\{\|\mathcal{G} + h\|_{\mathcal{H}}^2 > c_{1-\alpha}\},$$

which equals α at $h = 0$ and increases as the signal $\|h\|_{\mathcal{H}}$ grows.

Overall, the above analysis highlights a fundamental distinction between relevant and point-null inference. While classical point-null testing requires explicit covariance estimation and second-order asymptotics, the self-normalized relevant test developed in this paper avoids these complications by focusing on practically meaningful thresholds $\Delta > 0$, yielding pivotal limits and tuning-free inference in fully infinite-dimensional settings.

3 Simulation study

This section investigates the finite-sample performance of the proposed tests for relevant functional Granger causality. We focus on the residual–predictor cross-covariance energy d_{uX} in (5), which yields a self-

normalized statistic with a tractable analytical structure.

3.1 Data generating processes

We generate paired functional series (X_t, Y_t) as random elements in $\mathbb{H}$ from a FARX(1) model

$$X_t = 0.5\, X_{t-1} + \varepsilon_t^X, \qquad Y_t = 0.4\, Y_{t-1} + B(\theta)\, X_{t-1} + \varepsilon_t^Y,$$

where $B(\theta) : \mathbb{H} \to \mathbb{H}$ is a bounded Hilbert–Schmidt operator governing the Granger causality from X to Y. We specify $B(\theta)$ to be diagonal in the Fourier basis $\{\phi_k\}_{k\geq 1}$,

$$B(\theta)\phi_k = \theta b_k\, \phi_k, \qquad b_k = 1/k,$$

so that higher-frequency components exhibit a weaker causal effect.

For numerical implementation, we truncate at dimension $D = 21$ and represent each curve in the span of $\{\phi_1, \ldots, \phi_D\}$:

$$X_t = \sum_{k=1}^{D} x_{t,k}\phi_k, \qquad Y_t = \sum_{k=1}^{D} y_{t,k}\phi_k.$$

Writing $\varepsilon_t^X = \sum_{k=1}^{D} e_{t,k}^X \phi_k$ and $\varepsilon_t^Y = \sum_{k=1}^{D} e_{t,k}^Y \phi_k$, the functional recursion reduces to diagonal coordinate dynamics

$$x_{t,k} = 0.5\, x_{t-1,k} + e_{t,k}^X, \qquad y_{t,k} = 0.4\, y_{t-1,k} + \theta b_k x_{t-1,k} + e_{t,k}^Y, \quad k = 1, \ldots, D.$$

We set $\mathrm{Var}(e_{t,k}^X) = \mathrm{Var}(e_{t,k}^Y) = \sigma_k^2$ with $\sigma_k = 1/k$ and generate innovations independently across processes and frequencies under three designs:

(i) *iid*: $e_{t,k} = z_{t,k}$ with $z_{t,k} \sim N(0, \sigma_k^2)$;

(ii) *fMA(1)*: $e_{t,k} = z_{t,k} + \psi z_{t-1,k}$ with $\psi = 0.7$ and $z_{t,k} \sim N(0, \sigma_k^2/(1+\psi^2))$, ensuring $\mathrm{Var}(e_{t,k}) = \sigma_k^2$;

(iii) t_5: Student-t innovations with five degrees of freedom, rescaled to variance σ_k^2.

Let $\theta_0 = 0.30$ and define $\Delta = d_{uX}(\theta_0)$ as the relevance boundary. We consider $n \in \{100, 200, 400\}$ and $\theta \in \{0, 0.10, 0.20, 0.25, 0.30, 0.35, 0.40, 0.50\}$. Rejection probabilities are based on $R = 1000$ Monte Carlo replications with burn-in length 200.

Because u_t denotes the population one-step-ahead prediction error from the restricted predictor based on past Y only, the residual–predictor cross-covariance is the partial (projection) covariance operator:

$$C_{uX} = \mathrm{Cov}(u_t, X_{t-1}) = B(\theta)\, \Gamma_{X|Y}, \qquad \Gamma_{X|Y} \overset{\mathrm{def}}{=} \Gamma_X - \Gamma_{XY}\Gamma_Y^{-1}\Gamma_{YX},$$

where $\Gamma_X = \mathrm{Cov}(X_{t-1})$, $\Gamma_Y = \mathrm{Cov}(Y_{t-1})$, $\Gamma_{XY} = \mathrm{Cov}(X_{t-1}, Y_{t-1})$, and $\Gamma_{YX} = \Gamma_{XY}^*$.

Under our diagonal Fourier design, all covariance operators are diagonal in $\{\phi_k\}$. Let $\gamma_{X,k} \stackrel{\text{def}}{=} \mathrm{Var}(x_{t-1,k})$, $\gamma_{Y,k} \stackrel{\text{def}}{=} \mathrm{Var}(y_{t-1,k})$, and $\gamma_{XY,k} \stackrel{\text{def}}{=} \mathrm{Cov}(x_{t-1,k}, y_{t-1,k})$. Then the eigenvalues of $\Gamma_{X|Y}$ on the simulated D-dimensional subspace are

$$\gamma_{X|Y,k} = \gamma_{X,k} - \frac{\gamma_{XY,k}^2}{\gamma_{Y,k}}, \qquad k = 1, \ldots, D,$$

which equals $\mathrm{Var}(x_{t-1,k} \mid y_{t-1,k})$ under Gaussianity. Consequently,

$$d_{uX}(\theta) = \|C_{uX}\|_{\mathrm{HS}}^2 = \theta^2 \sum_{k=1}^{D} \big(b_k\, \gamma_{X|Y,k}\big)^2.$$

We compute $\Delta = d_{uX}(\theta_0)$ using stationary second moments (or via a long Monte Carlo pre-run).

Within each replication, we estimate the restricted FAR(1) model by ridge (Tikhonov) regularization on the Fourier coefficient representation, using penalty $\lambda_n = n^{-1/2}$. Residual coefficient vectors $\hat{u}_t$ are used to form

$$\widehat{C}_{uX,n}(1) = \frac{1}{n} \sum_{t=2}^{n} \hat{u}_t \otimes X_{t-1}, \qquad \hat{d}_{uX,n}(1) = \|\widehat{C}_{uX,n}(1)\|_{\mathrm{HS}}^2.$$

In the orthonormal basis representation, the Hilbert–Schmidt norm coincides with the Frobenius norm, enabling implementation via standard matrix norm calculations. For the self-normalizer, we re-estimate the restricted model on subsamples $\{1, \ldots, \lfloor n\lambda \rfloor\}$ and compute $\hat{d}_{uX,n}(\lambda)$ on the grid $\lambda_i = i/20$, $i = 1, \ldots, 19$, i.e. $m = 20$. This yields the centered process

$$G_n(\lambda) = \hat{d}_{uX,n}(\lambda) - \lambda^2 \hat{d}_{uX,n}(1),$$

which enters both the quadratic and adjusted-range self-normalizers.

3.2 Simulation results

Figures 1 and 2 report empirical rejection probabilities as functions of θ. Across all designs, rejection probabilities are close to zero in the interior of the null ($\theta < \theta_0$), increase as θ moves into the alternative, and improve with sample size. At the boundary $\theta = \theta_0$, the SN procedures are mildly conservative under i.i.d. and t_5 innovations, while the fMA(1) design induces slight over-rejection for $n \geq 200$.

[Figure 1 about here.]

[Figure 2 about here.]

Figure 3 compares the quadratic and adjusted-range self-normalizers directly. The adjusted-range version

delivers uniformly higher power for $\theta > \theta_0$ while maintaining comparable boundary behavior at $\theta = \theta_0$, with gains particularly pronounced under dependent (fMA(1)) and heavy-tailed innovations (t_5). This aligns with the intuition that range-based self-normalization is less sensitive to irregularities that can inflate second-moment–based normalizers.

[Figure 3 about here.]

Table 1 evaluates the tests performance at the relevance boundary

$$H_0(\Delta): \; d_{uX}(\theta) \leq \Delta,$$

at $\theta = \theta_0$, where $\Delta = d_{uX}(\theta_0)$. This configuration is least favorable under the one-sided alternative $H_1 : d_{uX}(\theta) > \Delta$, since within the strict null ($d_{uX}(\theta) < \Delta$) the statistic exhibits negative drift away from the rejection region. Accordingly, rejection frequencies at $\theta = \theta_0$ provide a finite-sample approximation to the maximal type-I error (size) over $H_0(\Delta)$.

The last column of Table 1 does not assess size for $H_0(\Delta)$. It reports a classical plug-in test corresponding to the point null $H_0 : d_{uX}(\theta) = 0$ obtained by setting $\Delta = 0$. Since $d_{uX}(\theta_0) = \Delta > 0$, this null is false at $\theta = \theta_0$; the reported rejection frequencies therefore measure power, which converges to one as n increases.

[Table 1 about here.]

3.3 Point-null comparison ($\Delta = 0$)

To contrast relevance-based functional Granger-causality inference with classical detection of any nonzero causal effect, we also report a plug-in test of the point null $H_0 : C_{uX} = 0$ (equivalently $d_{uX} = 0$). The test statistic is

$$T = (n-1)\big\|\widehat{C}_{uX}^{\mathrm{std}}\big\|_{\mathrm{HS}}^2, \qquad \widehat{C}_{uX}^{\mathrm{std}} = \frac{1}{n-1}\sum_{t=2}^{n} \hat{u}_t \otimes X_{t-1},$$

where $\hat{u}_t$ denotes residuals from the restricted model. Under H_0, $\sqrt{n-1}\,\mathrm{vec}(\widehat{C}_{uX}^{\mathrm{std}})$ is asymptotically Gaussian, implying that T converges to a quadratic form in a high-dimensional normal vector. We estimate the diagonal long-run variance componentwise using the Newey–West estimator (Bartlett kernel, bandwidth $L = \lfloor (n-1)^{1/4} \rfloor$) and approximate the resulting quadratic-form distribution by a scaled chi-square distribution via Satterthwaite moment matching.

Figure 4 reports rejection probabilities for the point-null test. As expected, this test exhibits power against relatively small departures from zero, reflecting that asymptotic power against local alternatives does not imply that the corresponding effect is relevant in magnitude. This illustrates the conceptual

distinction between point-null inference—which assesses the existence of any nonzero effect—and relevance-based testing, which targets effects that exceed a pre-specified threshold and are therefore economically meaningful.

[Figure 4 about here.]

Example 1. *To illustrate the distinction between statistical and economic significance, consider a volatility-forecasting setting. Let Y_t denote a scalar target such as log realized volatility of the underlying asset on day t, and let $X_{t-1}(\cdot)$ be a functional predictor, for example the previous day's option-implied volatility smile at fixed maturity, represented as a function of log-moneyness and sampled on $D_X = 21$ grid points. Denote by $\hat{u}_t$ the residual from the restricted model that excludes X and conditions only on lags of Y and other controls.*

After standardization, suppose each coordinate of X_{t-1} exhibits a small but nonzero lagged correlation ρ_j with $\hat{u}_t$, with $\rho_j \approx 0.01$ for all j. Although each coordinate effect is weak, the induced cross-covariance energy aggregates across dimensions,

$$\|C_{uX}^{\mathrm{std}}\|_{\mathrm{HS}}^2 \approx \sum_{j=1}^{D_X} \rho_j^2 \approx 21 \times (0.01)^2 = 0.0021.$$

A point-null test targeting $C_{uX}^{\mathrm{std}} = 0$ therefore exhibits increasing rejection probability as n grows, since the null is violated by small deviations that accumulate in high dimensions. However, the corresponding root mean square,

$$\sqrt{\|C_{uX}^{\mathrm{std}}\|_{\mathrm{HS}}^2 / D_X} \approx 0.01,$$

is economically negligible and would alter volatility forecasts by approximately one percent of a standard deviation, a magnitude unlikely to influence hedging or portfolio decisions once transaction costs and estimation error are considered.

Relevance-based testing addresses this issue by imposing a threshold r_0 for economically meaningful effects. With $r_0 = 0.05$ (equivalently $\Delta = D_X D_Y r_0^2 = 0.0525$ when $D_Y = 1$), the above configuration lies within the null region and the test does not reject. Inference therefore focuses on causal relations that exceed a pre-specified relevance threshold and are capable of affecting economic decisions. The same dimensionality-driven aggregation of small coordinate effects applies to other functional predictors, such as intraday return or volume curves sampled over many time bins.

4 Empirical analysis

This section provides an empirical illustration of the proposed relevance-based functional Granger causality framework in the Bitcoin market. The analysis addresses a central debate in financial economics concerning the extent to which option markets primarily enhance price discovery and informational efficiency of the underlying asset or instead whether speculative trading in derivatives markets contributes to excess volatility in the spot market. We operationalize these competing mechanisms using daily option-market functionals and test for economically relevant volatility spillovers in both directions between options and spot, thereby focusing on causal relations that are capable of influencing economic decision-making rather than statistical significance.

4.1 Data sources and sample

We use transaction-level BTC–USD spot data from Bitfinex and BTC option trades from Deribit. The sample spans January 1, 2025 to December 31, 2025 (UTC), providing a high-frequency window for assessing causal dynamics in cryptocurrency markets. Bitfinex spot transactions are aggregated into a daily realized volatility series, while Deribit option trades are transformed into daily functional representations of the implied-volatility smile.

[Figure 5 about here.]

Figure 5 summarizes the resulting spot volatility series together with trading activity, and Figure 6 characterizes the option-market functionals used in the analysis. These data sources allow us to investigate economically relevant functional Granger causality between derivatives and spot markets using rich functional information rather than scalar summaries.

4.2 Constructing spot volatility and option-market functionals

Let $P_{t,i}$ denote the transaction price of the i-th Bitfinex BTC–USD trade on day t. We compute intraday log-returns $r_{t,i} = \log P_{t,i} - \log P_{t,i-1}$ and form daily realized variance $\mathrm{RV}_t = \sum_i r_{t,i}^2$, with realized volatility $\sigma_t = \sqrt{\mathrm{RV}_t}$. Days with missing transactions are retained in a full calendar and aligned with option-market observations to preserve temporal consistency.

[Figure 6 about here.]

For each day t, Deribit option trades are filtered to a near-one-month maturity window $\tau \in [\tau_0 - \delta_\tau, \tau_0 + \delta_\tau]$ with $\tau_0 = 30$ days and $\delta_\tau = 10$ days. Log-moneyness is defined as $u = \log(K/S_t)$, where K is the strike price

and S_t the contemporaneous spot index. We construct two functional objects on a fixed grid $u \in [-0.25, 0.25]$ of size $D = 21$: (i) an implied-volatility smile $X_t^{\mathrm{IV}}(u)$, obtained by smoothing implied volatility as a function of log-moneyness and evaluating it on the grid; and (ii) an amount (flow) smile $X_t^{\mathrm{Amt}}(u)$, constructed analogously using $\log(1+\text{amount})$ as the response variable. The implied-volatility smile proxies information about future volatility and price discovery, whereas the flow-based smile captures cross-sectional trading pressure and speculative activity. Figure 6 summarizes the resulting functional objects and their features.

4.3 Full-sample relevant Granger-causality results

We align the daily spot and option-functional series by retaining dates for which (i) Bitfinex RV is available and (ii) a complete daily smile can be constructed. This yields $n = 239$ matched observations in 2025. The analysis focuses on the residual–predictor cross-covariance energy d_{uX} defined in (5), and we test the relevance hypothesis $H_0(\Delta) : d_{uX} \leq \Delta$ against $H_1(\Delta) : d_{uX} > \Delta$ at significance level $\alpha = 0.05$ using both the quadratic SN and adjusted-range SN tests. One-day lagged variables ($p = 1$) are considered throughout.

To render the relevance threshold economically interpretable, we parameterize Δ via a root-mean-square (RMS) correlation benchmark r_0. After column-wise standardization (z-scoring each coordinate) of X and Y, we set

$$\Delta = D_X D_Y \, r_0^2,$$

so that r_0 represents a benchmark magnitude for the average (RMS) correlation strength between the restricted residual and the predictor. We report results for $r_0 \in \{0.05, 0.10, 0.15\}$, ranging from modest to strict relevance.

[Table 2 about here.]

Table 2 yields two main findings. First, for the price-discovery proxy (implied-volatility smile), we reject the relevance null for the Granger causality from Options to SpotVol (Options $\to$ SpotVol) at the modest threshold $r_0 = 0.05$ under both self-normalizers. This indicates that option-market information contains economically meaningful predictive content for next-day spot volatility at that level of relevance. Second, we find no evidence of economically relevant Granger causality for the reverse direction (SpotVol $\to$ Options) and no relevant causality (predictive) content from the flow-based smile to spot volatility over the full sample. When the relevance requirement is strengthened to $r_0 \in \{0.10, 0.15\}$, the null cannot be rejected, highlighting that statistical and economic significance are distinct in causal inference.

4.4 Rolling-window analysis and lead–lag regimes

Causal relationships in cryptocurrency markets may evolve over time. To examine stability, we conduct a rolling-window analysis using windows of length $W = 90$ matched observations, advanced in increments of five observations. Because option data are not available on every calendar day, the calendar span of a window may exceed 90 days, but each window contains 90 aligned (X_t, Y_t) pairs; windows with fewer than 60 observations are omitted to ensure reliable inference.

For visualization, we report excess statistics defined as the self-normalized test statistic minus its 5% critical value, so that positive values correspond to rejection of the relevance null. Figure 7 plots excess statistics for $r_0 = 0.05$, and Figure 8 reports rolling estimates of the RMS size $\hat{r}_{\mathrm{RMS}} = \sqrt{\max(\hat{d}, 0)/(D_X D_Y)}$ together with reference horizontal lines at $r_0 \in \{0.05, 0.10, 0.15\}$. While the magnitude of excess statistics differs across normalizers—reflecting scale differences in self-normalization—their rejection patterns are broadly aligned. The adjusted-range self-normalizer exhibits slightly higher sensitivity in borderline windows, consistent with its superior local-power properties documented in the simulation analysis.

[Figure 7 about here.]

[Figure 8 about here.]

[Table 3 about here.]

Table 3 reports rejection frequencies across rolling windows. With 90 matched observations per window, economically relevant Granger causalities are detected only episodically and primarily at the lowest relevance threshold $r_0 = 0.05$. At this level, the adjusted-range self-normalizer rejects more frequently than the quadratic alternative in the amount-smile channel, consistent with its higher finite-sample sensitivity documented by the local-power analysis in Section 2.3.

To characterize the lead–lag causality structure, each window is classified into four regimes: Option-led (rejection of Options $\to$ SpotVol only), Spot-led (rejection of SpotVol $\to$ Options only), Feedback (rejections in both directions), and None (no rejections). Tables 4–5 report regime shares for quadratic and adjusted-range self-normalizers, revealing that economically relevant lead–lag activity is episodic and primarily confined to the amount-smile channel at low relevance thresholds, whereas higher thresholds and the implied-volatility smile channel display negligible causal regimes.

[Table 4 about here.]

[Table 5 about here.]

The rolling analysis indicates that economically relevant Granger causalities are episodic rather than pervasive in 2025. This finding underscores the importance of relevance-based inference since conventional point-null tests of Granger causality may report widespread statistical significance in large samples even when effects are too small to influence economic decisions. By contrast, the relevance framework identifies periods in which predictive strength exceeds a pre-specified economic threshold, providing insights into when causal relations are consequential for decision-making.

5 Conclusion

This paper develops relevance-based tests for functional Granger causality formulated directly in a separable Hilbert space. In contrast to conventional point-null procedures that test for any nonzero causal effect, we focus on economically meaningful causality by evaluating whether the predictive effect of one functional process on another exceeds a pre-specified relevance threshold. This conceptual shift from statistical significance to economic relevance is particularly important in high-dimensional functional environments, where aggregation across infinitely many coordinates can yield statistical significance even when incremental predictive content is economically negligible.

We construct fully functional effect-size measures based on residual–predictor cross-covariance operators, yielding interpretable metrics of causal strength. Estimation proceeds via regularized operator methods that preserve the infinite-dimensional structure of the data; discretization and basis expansions are employed purely for numerical implementation rather than as modeling restrictions. For inference, we develop tuning-free self-normalized statistics that eliminate the need for long-run variance estimation and bandwidth selection. Simulation evidence demonstrates reliable size control at the relevance boundary and strong power under serial dependence and heavy-tailed innovations. An adjusted-range normalization further enhances sensitivity in near-boundary configurations while maintaining comparable finite-sample stability.

An empirical application to Bitcoin option and spot markets illustrates the practical relevance of the proposed tests. Functional predictors extracted from implied-volatility surfaces exhibit economically meaningful predictive effects for next-day realized volatility at moderate relevance thresholds. In contrast, reverse-direction effects and flow-based functional proxies rarely exceed economically interpretable magnitudes. Rolling-window evidence indicates that economically relevant causal linkages are episodic rather than persistent, underscoring the importance of distinguishing economically meaningful causal effects from statistically significant but quantitatively minor ones.

Taken together, the proposed framework provides a rigorous and economically disciplined approach to functional causal inference. By embedding effect-size thresholds directly into the null hypothesis, the method-

ology aligns hypothesis testing with substantive economic interpretation while remaining fully coherent in infinite-dimensional settings. The resulting procedures deliver inference that is theoretically rigorous, computationally feasible, and directly informative for decision-making in high-dimensional time series applications.

Figure Legends

Figure 1. Empirical rejection probabilities for the relevant $X \to Y$ test based on quadratic SN ($m = 20$, $\alpha = 0.05$), under i.i.d., fMA(1), and t_5 innovations. The vertical dashed line marks $\theta_0 = 0.30$ and the horizontal line marks $\alpha = 0.05$.

Figure 2. Empirical rejection probabilities for the relevant $X \to Y$ test based on adjusted-range SN ($m = 20$, $\alpha = 0.05$), under i.i.d., fMA(1), and t_5 innovations. The vertical dashed line marks $\theta_0 = 0.30$ and the horizontal line marks $\alpha = 0.05$.

Figure 3. Empirical rejection probabilities for the relevant $X \to Y$ test comparing quadratic SN and adjusted-range SN ($m = 20$, $\alpha = 0.05$), under i.i.d., fMA(1), and t_5 innovations. The vertical dashed line marks $\theta_0 = 0.30$ and the horizontal line marks $\alpha = 0.05$.

Figure 4. Empirical rejection probabilities for the plug-in point-null test of $H_0 : C_{uX} = 0$ (equivalently $\Delta = 0$) under i.i.d., fMA(1), and t_5 innovations. The horizontal line marks $\alpha = 0.05$.

Figure 5. Daily BTC spot RV from Bitfinex (2025, UTC) and the corresponding number of spot trades per day.

Figure 6. Summary of daily Deribit option-market functionals in 2025. The panels report the cross-day mean together with dispersion bands (quantiles) over the sample.

Figure 7. Rolling-window excess statistics for the relevant SN tests with $r_0 = 0.05$ (window=90 aligned observations, step=5). Positive values indicate rejection.

Figure 8. Rolling-window estimated effect size in RMS-correlation units $\hat{r}_{\mathrm{RMS}}$ for each channel and direction (window=90 aligned observations, step=5). Horizontal lines indicate the relevance thresholds $r_0 \in \{0.05, 0.10, 0.15\}$.

References

Amblard, P.-O. and O. Michel (2013). "Causalité de Granger pour des signaux à valeurs fonctionnelles". In: *XXIVème Colloque francophone de traitement du signal et des images (GRETSI 2013)*. Brest, France, pp. 809–812.

Aue, A., D. Dubart Norinho, and S. Hörmann (2015). "On the prediction of stationary functional time series". In: *Journal of the American Statistical Association* 110.509, pp. 378–392. DOI: 10.1080/01621459.2014.909317.

Bardsley, P., L. Horváth, P. Kokoszka, and G. Young (2017). "Change point tests in functional factor models with application to yield curves". In: *The Econometrics Journal* 20.1, pp. 86–117. DOI: 10.1111/ectj.12075.

Berkes, I., L. Horváth, and G. Rice (2013). "Weak invariance principles for sums of dependent random functions". In: *Stochastic Processes and their Applications* 123.2, pp. 385–403. DOI: 10.1016/j.spa.2012.10.003.

Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications*. Vol. 149. Lecture Notes in Statistics. New York, NY: Springer. DOI: 10.1007/978-1-4612-1154-9.

Bouezmarni, T., M. Doukali, and A. Taamouti (2024). "Testing Granger Non-Causality in Expectiles". In: *Econometric Reviews* 43.1, pp. 30–51. DOI: 10.1080/07474938.2023.2246823.

Bouezmarni, T., J. V. K. Rombouts, and A. Taamouti (2012). "Nonparametric Copula-Based Test for Conditional Independence with Applications to Granger Causality". In: *Journal of Business & Economic Statistics* 30.2, pp. 275–287. DOI: 10.1080/07350015.2011.638831.

Dette, H., K. Kokot, and S. Volgushev (2020). "Testing relevant hypotheses in functional time series via self-normalization". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 82.3, pp. 629–660. DOI: 10.1111/rssb.12370.

Dufour, J.-M. and A. Taamouti (2010). "Short and Long Run Causality Measures: Theory and Inference". In: *Journal of Econometrics* 154.1, pp. 42–58. DOI: 10.1016/j.jeconom.2009.06.008.

Ferraty, F. and P. Vieu (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. New York, NY: Springer. DOI: 10.1007/0-387-36620-2.

Granger, C. W. J. (1969). "Investigating Causal Relations by Econometric Models and Cross-spectral Methods". In: *Econometrica* 37.3, pp. 424–438.

Hong, Y., Z. Lin, O. Linton, W. Newey, and J. Sun (2025). "Confidence interval estimation and hypothesis testing using the adjusted-range-based self-normalization approach". Working paper.

Hong, Y., O. Linton, B. McCabe, J. Sun, and S. Wang (2024). "Kolmogorov–Smirnov type testing for structural breaks: A new adjusted-range based self-normalization approach". In: *Journal of Econometrics* 238.2, p. 105603. DOI: 10.1016/j.jeconom.2023.105603.

Hörmann, S. and P. Kokoszka (2010). "Weakly dependent functional data". In: *The Annals of Statistics* 38.3, pp. 1845–1884. DOI: 10.1214/09-AOS768.

Horváth, L. and P. Kokoszka (2012). *Inference for Functional Data with Applications*. Springer Series in Statistics. New York, NY: Springer. DOI: 10.1007/978-1-4614-3655-3.

Lengwiler, Y. and C. Lenz (2010). "Intelligible factors for the yield curve". In: *Journal of Econometrics* 157.2, pp. 481–491. DOI: 10.1016/j.jeconom.2010.04.001.

Lin, Z., J. Sun, W. K. Härdle, and M. Zhu (2026). "Monitoring Relevant Shifts in Functional Time Series". Working paper.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin Heidelberg: Springer. DOI: 10.1007/978-3-540-27752-1.

Müller, H.-G., R. Sen, and U. Stadtmüller (2011). "Functional data analysis for volatility". In: *Journal of Econometrics* 165.2, pp. 233–245. DOI: 10.1016/j.jeconom.2011.08.002.

Ramsay, J. O. and B. W. Silverman (2005). *Functional Data Analysis*. 2nd ed. New York, NY: Springer. DOI: 10.1007/b98888.

Saumard, M. (2017). "Linear causality in the sense of Granger with stationary functional time series". In: *Functional Statistics and Related Fields*. Ed. by G. Aneiros, E. G. Bongiorno, R. Cao, and P. Vieu. Contributions to Statistics. Cham: Springer, pp. 225–231. DOI: 10.1007/978-3-319-55846-2_30.

Sen, R., A. Majumdar, and S. Sikaria (2022). "Bayesian Testing of Granger Causality in Functional Time Series". In: *Journal of Quantitative Economics* 20.Suppl 1, pp. 191–210. DOI: 10.1007/s40953-022-00306-x.

Shang, H., K. Ji, and U. Beyaztas (July 2021). "Granger causality of bivariate stationary curve time series". In: *Journal of Forecasting* 40.4, pp. 626–635. DOI: 10.1002/for.2732.

Shao, X. (2010). "A self-normalized approach to confidence interval construction in time series". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 72.3, pp. 343–366. DOI: 10.1111/j.1467-9868.2009.00737.x.

Shojaie, A. and E. B. Fox (2022). "Granger Causality: A Review and Recent Advances". In: *Annual Review of Statistics and Its Application* 9, pp. 289–319. DOI: 10.1146/annurev-statistics-040120-010930.

Song, X. and A. Taamouti (2018). "Measuring Nonlinear Granger Causality in Mean". In: *Journal of Business & Economic Statistics* 36.2, pp. 321–333. DOI: 10.1080/07350015.2016.1166118.

Song, X. and A. Taamouti (2021). "Measuring Granger Causality in Quantiles". In: *Journal of Business & Economic Statistics* 39.4, pp. 937–952. DOI: 10.1080/07350015.2020.1739531.

Taamouti, A., T. Bouezmarni, and A. El Ghouch (2014). "Nonparametric Estimation and Inference for Conditional Density Based Granger Causality Measures". In: *Journal of Econometrics* 180.2, pp. 251–264. DOI: 10.1016/j.jeconom.2014.03.001.

Zhang, X. and X. Shao (2015). "Two sample inference for the second-order property of temporally dependent functional data". In: *Bernoulli* 21.2, pp. 909–929.

Tables

Table 1: Boundary rejection probabilities at the relevance threshold $\theta = \theta_0 = 0.30$ (Monte Carlo s.e. in parentheses).

Errors	n	Quadratic SN	Adjusted-range SN	Plug-in ($\Delta = 0$)
iid	100	0.030 (0.005)	0.032 (0.006)	0.900 (0.009)
	200	0.030 (0.005)	0.034 (0.006)	0.995 (0.002)
	400	0.028 (0.005)	0.035 (0.006)	1.000 (0.000)
fMA(1)	100	0.036 (0.006)	0.038 (0.006)	0.828 (0.012)
	200	0.060 (0.008)	0.064 (0.008)	0.987 (0.004)
	400	0.078 (0.008)	0.076 (0.008)	1.000 (0.000)
t_5	100	0.007 (0.003)	0.007 (0.003)	0.848 (0.011)
	200	0.013 (0.004)	0.020 (0.004)	0.995 (0.002)
	400	0.017 (0.004)	0.022 (0.005)	0.998 (0.001)

Notes: Nominal level $\alpha = 0.05$ and $R = 1000$ Monte Carlo replications. The relevance threshold is set to $\Delta = d_{uX}(\theta_0)$ with $\theta_0 = 0.30$. The SN normalizers are computed on the grid $\lambda_i = i/20$ ($m = 20$). Innovation designs: iid = Gaussian i.i.d.; fMA(1) = functional MA(1) innovations $e_{t,k} = z_{t,k} + 0.7z_{t-1,k}$ scaled to match $\mathrm{Var}(e_{t,k}) = \sigma_k^2$; t_5 = Student-t with 5 degrees of freedom rescaled to variance σ_k^2. Monte Carlo s.e. are $\sqrt{\hat{p}(1-\hat{p})/R}$. The first two SN columns correspond to the proposed relevant test of $H_0(\Delta) : d_{uX} \leq \Delta$; therefore the boundary row $\theta = \theta_0$ is a size diagnostic at the least favorable null point. The last column (Plug-in with $\Delta = 0$) tests the classical point null $H_0 : d_{uX} = 0$ and is evaluated at $\theta = \theta_0 > 0$; hence it should be interpreted as power rather than size.

Table 2: Relevant SN Granger-causality tests for BTC (2025), using Bitfinex spot RV and Deribit option-market functionals.

Channel	Direction	r_0	n	Quad. stat	Adj.-range stat
IV smile (price discovery)	Options → SpotVol	0.05	239	12.186*	3.617*
IV smile (price discovery)	SpotVol → Options	0.05	239	-43.850	-10.468
IV smile (price discovery)	Options → SpotVol	0.10	239	8.113	2.408
IV smile (price discovery)	SpotVol → Options	0.10	239	-197.434	-47.133
IV smile (price discovery)	Options → SpotVol	0.15	239	1.326	0.394
IV smile (price discovery)	SpotVol → Options	0.15	239	-453.408	-108.242
Amount smile (flow/speculation)	Options → SpotVol	0.05	239	-6.397	-1.813
Amount smile (flow/speculation)	SpotVol → Options	0.05	239	-2.468	-1.239
Amount smile (flow/speculation)	Options → SpotVol	0.10	239	-45.354	-12.855
Amount smile (flow/speculation)	SpotVol → Options	0.10	239	-20.429	-10.258
Amount smile (flow/speculation)	Options → SpotVol	0.15	239	-110.282	-31.259
Amount smile (flow/speculation)	SpotVol → Options	0.15	239	-50.364	-25.289

Notes: A star indicates rejection of the relevant null $H_0(\Delta)$ at level $\alpha = 0.05$. The statistics are $(\hat{d}_n - \Delta)/V$ (quadratic) and $(\hat{d}_n - \Delta)/H$ (adjusted-range).

Table 3: Rolling-window summary of relevant SN Granger-causality rejections (window=90, step=5, 2025).

Channel	Direction	r_0	#Windows	Quad. rej.%	Adj.-range rej.%
IV smile (price discovery)	Options → SpotVol	0.05	30	0.0	0.0
IV smile (price discovery)	SpotVol → Options	0.05	30	0.0	0.0
Amount smile (flow/speculation)	Options → SpotVol	0.05	30	3.3	10.0
Amount smile (flow/speculation)	SpotVol → Options	0.05	30	20.0	23.3
IV smile (price discovery)	Options → SpotVol	0.10	30	0.0	0.0
IV smile (price discovery)	SpotVol → Options	0.10	30	0.0	0.0
Amount smile (flow/speculation)	Options → SpotVol	0.10	30	0.0	0.0
Amount smile (flow/speculation)	SpotVol → Options	0.10	30	0.0	0.0
IV smile (price discovery)	Options → SpotVol	0.15	30	0.0	0.0
IV smile (price discovery)	SpotVol → Options	0.15	30	0.0	0.0
Amount smile (flow/speculation)	Options → SpotVol	0.15	30	0.0	0.0
Amount smile (flow/speculation)	SpotVol → Options	0.15	30	0.0	0.0

Notes: A window is counted as a rejection if the corresponding rolling test rejects $H_0(\Delta)$ at level $\alpha = 0.05$.

Table 4: Rolling-window lead–lag regimes based on the quadratic SN test (window=90, step=5, 2025).

Channel	r_0	Option-led(%)	Spot-led(%)	Feedback(%)	None(%)
Amount smile (flow/speculation)	0.05	3.3	20.0	0.0	76.7
Amount smile (flow/speculation)	0.10	0.0	0.0	0.0	100.0
Amount smile (flow/speculation)	0.15	0.0	0.0	0.0	100.0
IV smile (price discovery)	0.05	0.0	0.0	0.0	100.0
IV smile (price discovery)	0.10	0.0	0.0	0.0	100.0
IV smile (price discovery)	0.15	0.0	0.0	0.0	100.0

Table 5: Rolling-window lead–lag regimes based on the adjusted-range SN test (window=90, step=5, 2025).

Channel	r_0	Option-led(%)	Spot-led(%)	Feedback(%)	None(%)
Amount smile (flow/speculation)	0.05	6.7	20.0	3.3	70.0
Amount smile (flow/speculation)	0.10	0.0	0.0	0.0	100.0
Amount smile (flow/speculation)	0.15	0.0	0.0	0.0	100.0
IV smile (price discovery)	0.05	0.0	0.0	0.0	100.0
IV smile (price discovery)	0.10	0.0	0.0	0.0	100.0
IV smile (price discovery)	0.15	0.0	0.0	0.0	100.0

Figures

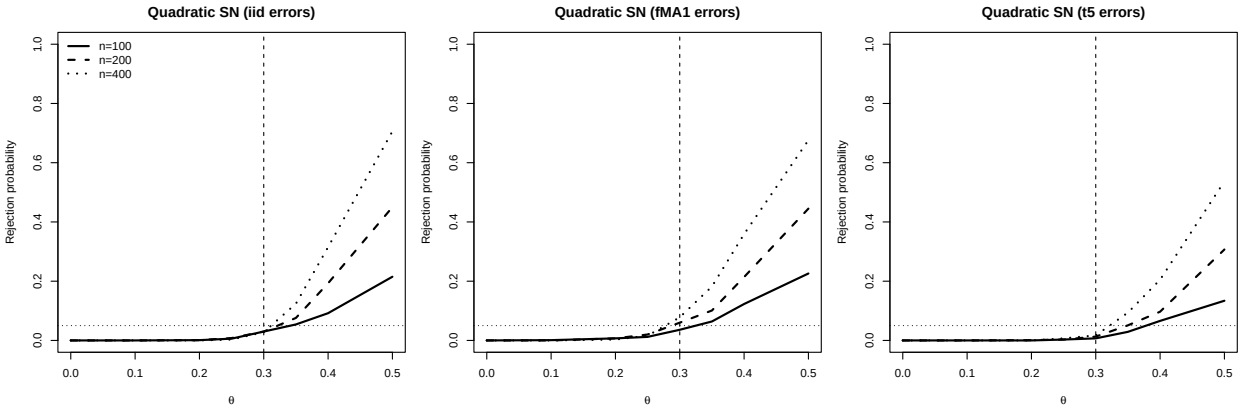

Figure 1: Empirical rejection probabilities for the relevant $X \to Y$ test based on quadratic SN ($m = 20$, $\alpha = 0.05$), under i.i.d., fMA(1), and t_5 innovations. The vertical dashed line marks $\theta_0 = 0.30$ and the horizontal line marks $\alpha = 0.05$.

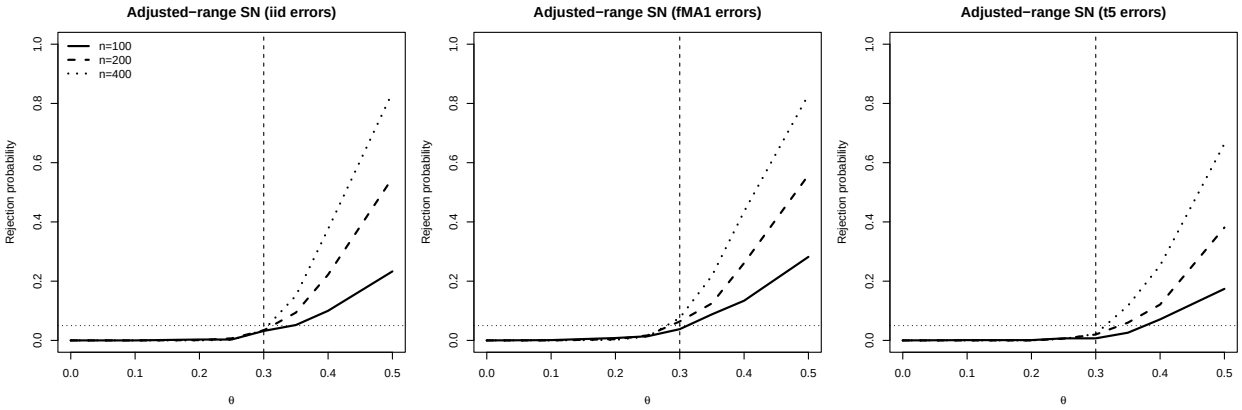

Figure 2: Empirical rejection probabilities for the relevant $X \to Y$ test based on adjusted-range SN ($m = 20$, $\alpha = 0.05$), under i.i.d., fMA(1), and t_5 innovations. The vertical dashed line marks $\theta_0 = 0.30$ and the horizontal line marks $\alpha = 0.05$.

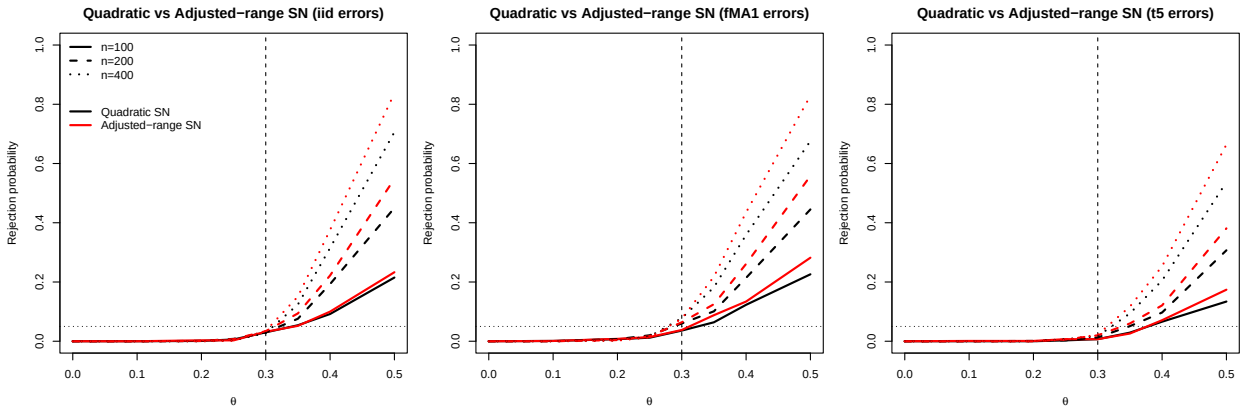

Figure 3: Empirical rejection probabilities for the relevant $X \to Y$ test comparing quadratic SN and adjusted-range SN ($m = 20$, $\alpha = 0.05$), under i.i.d., fMA(1), and t_5 innovations. The vertical dashed line marks $\theta_0 = 0.30$ and the horizontal line marks $\alpha = 0.05$.

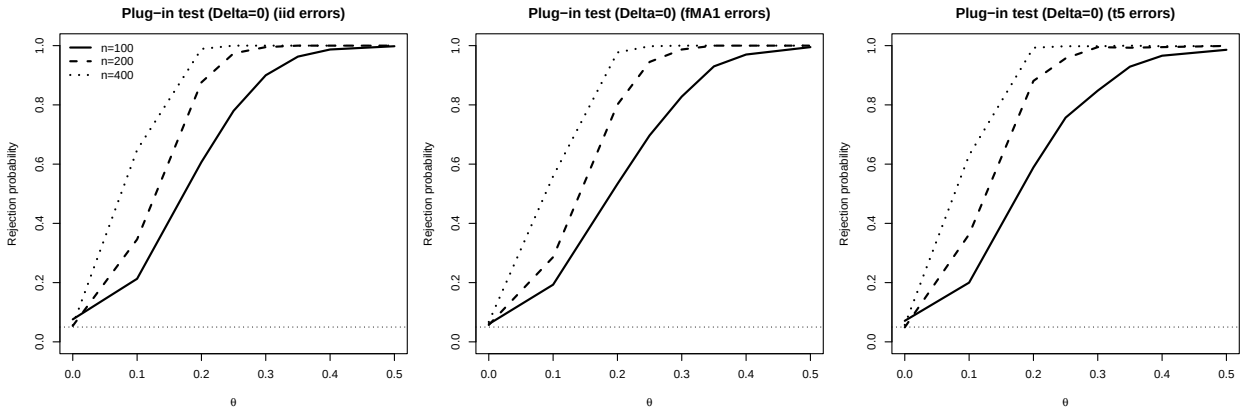

Figure 4: Empirical rejection probabilities for the plug-in point-null test of $H_0 : C_{uX} = 0$ (equivalently $\Delta = 0$) under i.i.d., fMA(1), and t_5 innovations. The horizontal line marks $\alpha = 0.05$.

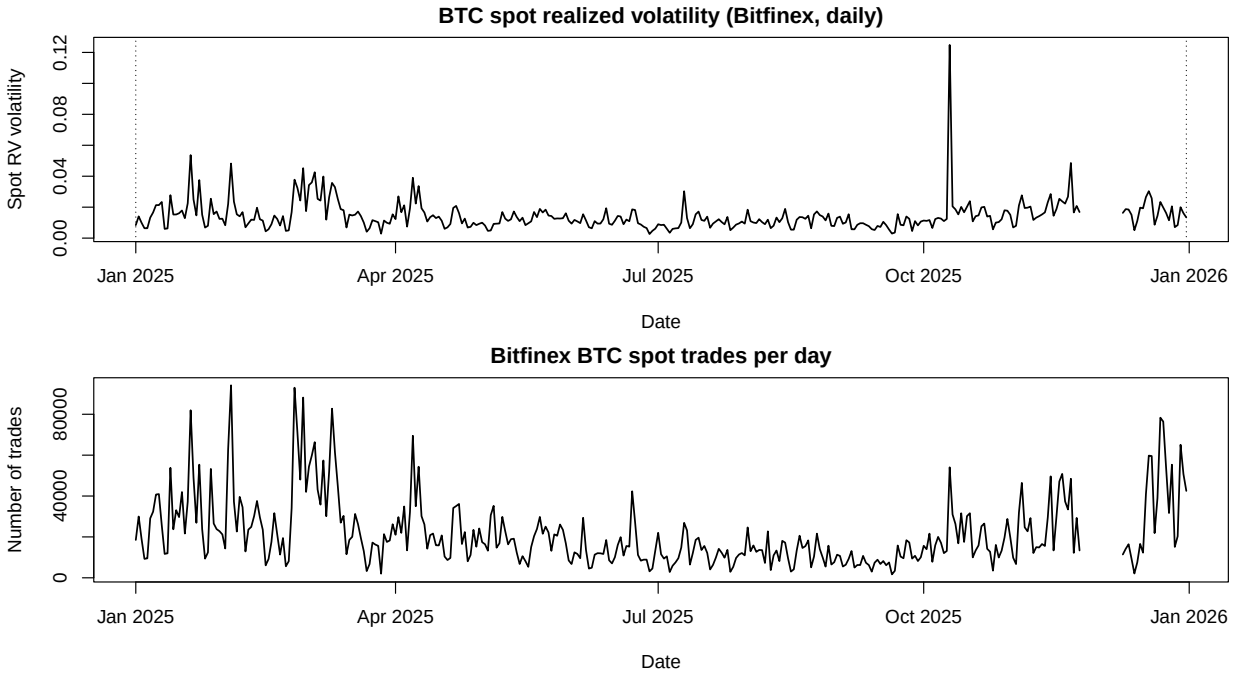

Figure 5: Daily BTC spot RV from Bitfinex (2025, UTC) and the corresponding number of spot trades per day.

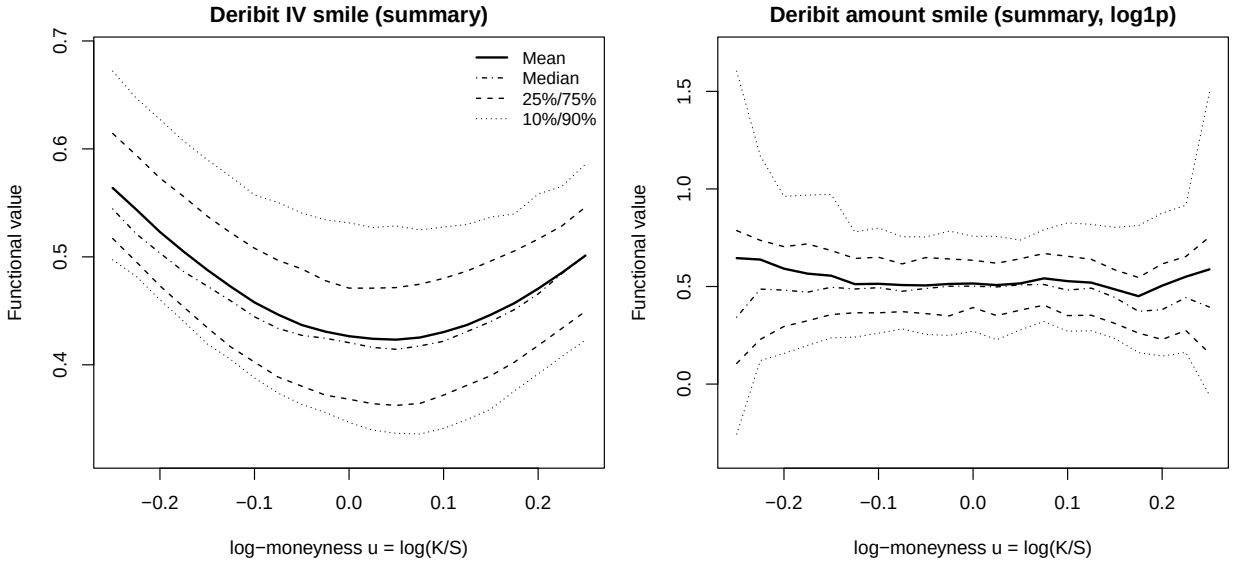

Figure 6: Summary of daily Deribit option-market functionals in 2025. The panels report the cross-day mean together with dispersion bands (quantiles) over the sample.

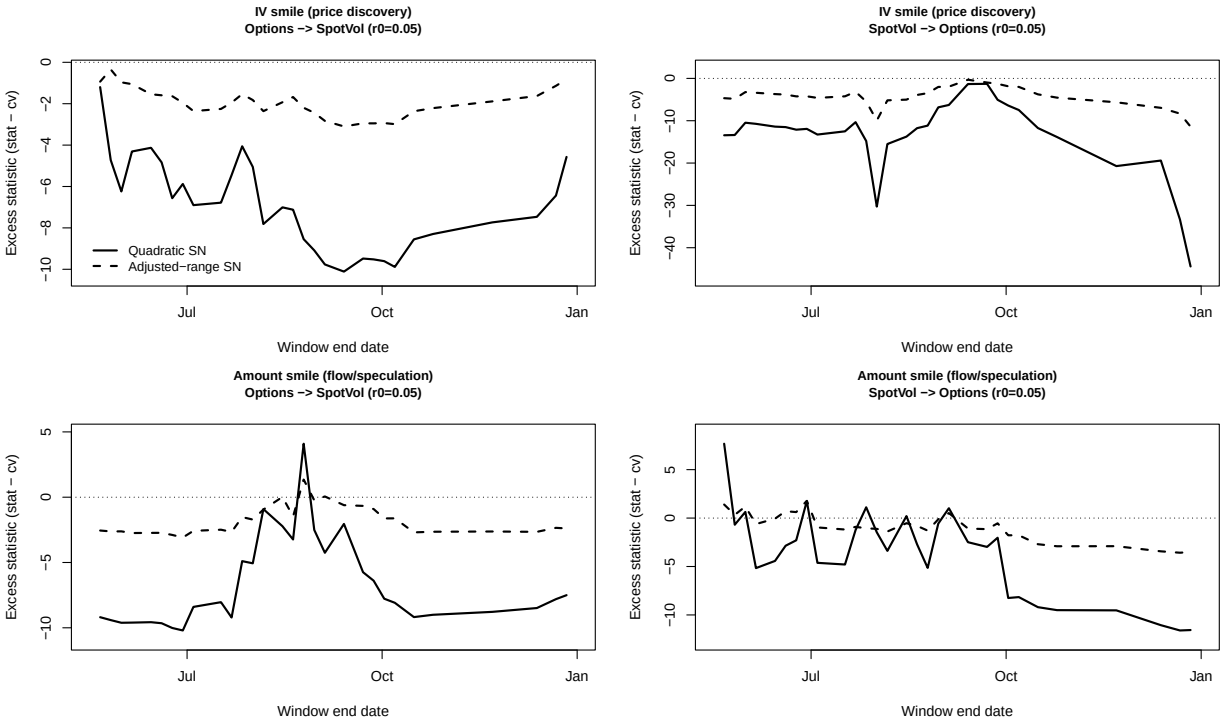

Figure 7: Rolling-window excess statistics for the relevant SN tests with $r_0 = 0.05$ (window=90 aligned observations, step=5). Positive values indicate rejection.

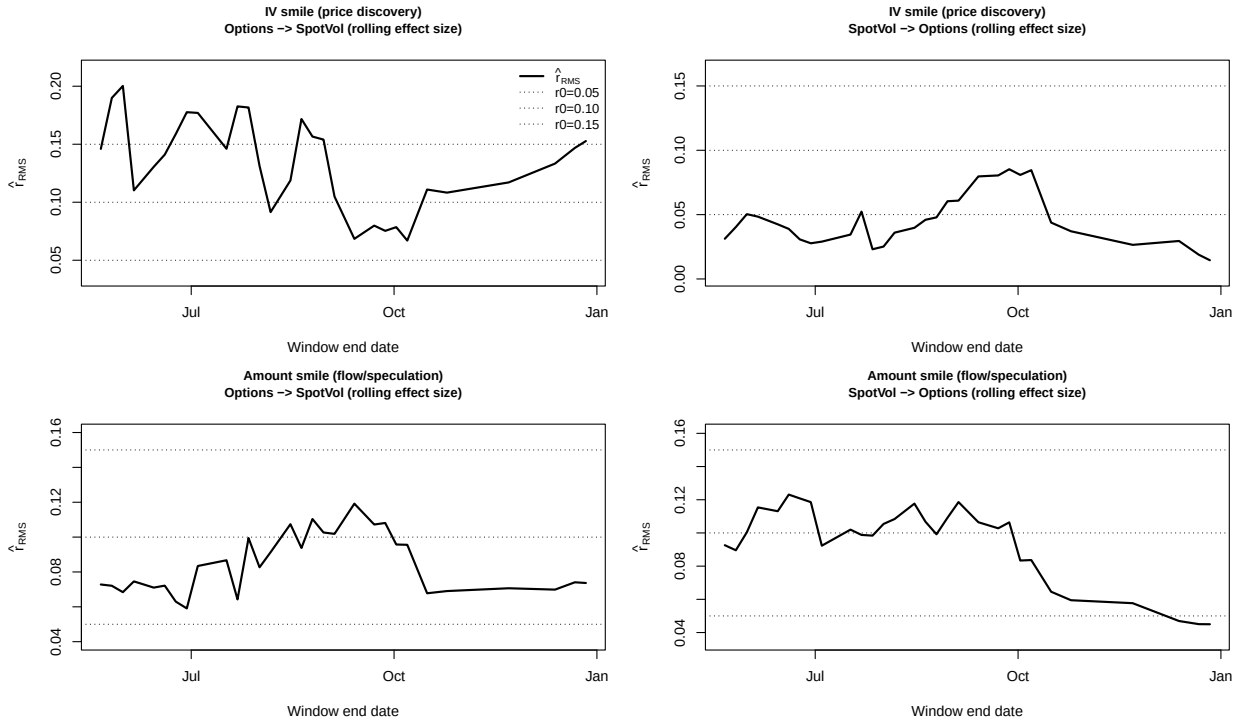

Figure 8: Rolling-window estimated effect size in RMS-correlation units $\hat{r}_{\mathrm{RMS}}$ for each channel and direction (window=90 aligned observations, step=5). Horizontal lines indicate the relevance thresholds $r_0 \in \{0.05, 0.10, 0.15\}$.

Appendix

A Proofs

This appendix gives more detailed proofs of the asymptotic results stated in Section 2. All arguments follow the standard invariance-principle route for SN in relevant hypothesis testing (see, e.g., Dette et al., 2020): once the joint weak convergence in Assumption 4 is available, the self-normalized statistics are obtained by continuous mappings and Slutsky-type arguments. For convenience, throughout the appendix we write

$$A_n \stackrel{\text{def}}{=} \sqrt{n}(\hat{d}_n - d), \qquad H_n(\lambda) \stackrel{\text{def}}{=} \sqrt{n} G_n(\lambda), \qquad \lambda \in [0,1].$$

Assumption 4 is then equivalently stated as

$$(A_n,\, H_n) \Rightarrow (\tau B(1),\, \tau U) \qquad \text{in } \mathbb{R} \times C([0,1]), \tag{A.1}$$

where $U(\lambda) = \lambda B(\lambda) - \lambda^2 B(1)$.

A.1 Preliminary lemmas

Lemma A.1 (A useful identity for homogeneous functionals)**.** *Let $\rho : C([0,1]) \to [0,\infty)$ be positively homogeneous, i.e. $\rho(cf) = c\,\rho(f)$ for all $c \geq 0$ and all $f \in C([0,1])$. Then for any random element $G_n \in C([0,1])$,*

$$\rho\big(\sqrt{n}\, G_n\big) = \sqrt{n}\, \rho(G_n) \qquad \textit{and hence} \qquad \frac{\hat{d}_n - d}{\rho(G_n)} = \frac{A_n}{\rho(H_n)}.$$

Proof. The first identity is immediate from positive homogeneity with $c = \sqrt{n}$. For the ratio, note that $A_n = \sqrt{n}(\hat{d}_n - d)$ and $H_n = \sqrt{n} G_n$, hence

$$\frac{A_n}{\rho(H_n)} = \frac{\sqrt{n}(\hat{d}_n - d)}{\rho(\sqrt{n} G_n)} = \frac{\sqrt{n}(\hat{d}_n - d)}{\sqrt{n}\rho(G_n)} = \frac{\hat{d}_n - d}{\rho(G_n)}.$$

□

Lemma A.2 (Positivity of the pivotal denominators)**.** *Let B be standard Brownian motion on $[0,1]$ and $U(\lambda) = \lambda B(\lambda) - \lambda^2 B(1)$. Then the following hold.*

(i) (Quadratic SN.) If ν is a finite measure on $[0,1]$ such that $\nu(\{\lambda : \lambda \in (0,1)\}) > 0$ and ν is not supported

on a single point, then

$$\int_0^1 U(\lambda)^2 \, \nu(\,\mathrm{d}\lambda) > 0 \qquad a.s.$$

In particular, for the discrete grid $\lambda_i = i/m$, $i = 1, \ldots, m-1$, with ν the discrete uniform distribution on $\{\lambda_i\}$, the above integral is positive a.s.

(ii) (Adjusted-range SN.) We have

$$\sup_{\lambda \in [0,1]} U(\lambda) - \inf_{\lambda \in [0,1]} U(\lambda) > 0 \qquad a.s.$$

Proof. Write $U(\lambda) = \lambda\{B(\lambda) - \lambda B(1)\} = \lambda\,\mathbb{B}(\lambda)$, where $\mathbb{B}$ is the Brownian bridge. In particular, U is a continuous Gaussian process with $U(0) = U(1) = 0$.

(i) If $\int_0^1 U(\lambda)^2 \nu(\,\mathrm{d}\lambda) = 0$, then $U(\lambda) = 0$ for ν-a.e. λ. If ν is discrete with support $\{\lambda_1, \ldots, \lambda_r\} \subset (0,1)$, then this implies $U(\lambda_j) = 0$ for all j. But $(U(\lambda_1), \ldots, U(\lambda_r))$ is a non-degenerate Gaussian vector (its covariance matrix is positive definite whenever the points are distinct), hence

$$\mathbb{P}\big(U(\lambda_1) = \cdots = U(\lambda_r) = 0\big) = 0,$$

which proves the claim for the grids used in this paper. If ν charges an interval, then $\int U^2 \, \mathrm{d}\nu = 0$ implies $U(\lambda) = 0$ on a set having an accumulation point. By continuity this forces $U(\lambda) = 0$ for all λ in that interval. Equivalently, $\mathbb{B}(\lambda) = 0$ on an interval, which implies $B(\lambda) = \lambda B(1)$ on that interval. Pick two distinct points $0 < \lambda_1 < \lambda_2 < 1$ in that interval. The event implies $U(\lambda_1) = U(\lambda_2) = 0$, that is, the two linear constraints $B(\lambda_i) = \lambda_i B(1)$, $i = 1, 2$, on the Gaussian vector $(B(\lambda_1), B(\lambda_2), B(1))$. Because this vector has a nonsingular covariance matrix, these constraints hold with probability zero. Thus the denominator is positive a.s.

(ii) If $\sup U - \inf U = 0$, then U is constant on $[0,1]$. Because $U(0) = 0$, this constant must be 0, hence $U(\lambda) \equiv 0$. As in part (i), $U \equiv 0$ is equivalent to $\mathbb{B} \equiv 0$, i.e. $B(\lambda) = \lambda B(1)$ for all λ, an event of probability zero. □

A.2 Proof of Theorem 1

Proof. We proceed in three steps.

Assumption 4 gives (A.1). Because ρ is continuous on $C([0,1])$ (under the supremum norm), the mapping

$$\Phi : \mathbb{R} \times C([0,1]) \to \mathbb{R} \times \mathbb{R}, \qquad \Phi(x, f) = (x, \rho(f)),$$

is continuous. By the continuous mapping theorem applied to (A.1),

$$\big(A_n, \rho(H_n)\big) \Rightarrow \big(\tau B(1),\, \rho(\tau U)\big). \tag{A.2}$$

By positive homogeneity, $\rho(\tau U) = \tau\rho(U)$.

By Lemma A.1,

$$\frac{\hat{d}_n - d}{\rho(G_n)} = \frac{A_n}{\rho(H_n)}.$$

Define the ratio map $\Psi : \mathbb{R} \times (0, \infty) \to \mathbb{R}$ by $\Psi(x, y) = x/y$. This map is continuous. By assumption, $\rho\{U(\cdot)\} > 0$ a.s. (and for the quadratic and range functionals this follows from Lemma A.2).

From (A.2) and the continuity of Ψ at all points with $y > 0$, another application of the continuous mapping theorem yields

$$\frac{A_n}{\rho(H_n)} \Rightarrow \frac{\tau B(1)}{\tau\rho\{U(\cdot)\}} = \frac{B(1)}{\rho\{U(\cdot)\}}.$$

This proves the claimed convergence. The limit is pivotal because the nuisance scale τ cancels. □

A.3 Proof of Theorem 2

Proof. Write $\widehat{N}_n \stackrel{\text{def}}{=} \rho(G_n)$ for the self-normalizer. We first record its stochastic order.

By Lemma A.1 and the continuous mapping theorem as in (A.2),

$$\sqrt{n}\,\widehat{N}_n = \rho(H_n) \Rightarrow \rho(\tau U) = \tau\rho(U),$$

hence $\widehat{N}_n = O_p(n^{-1/2})$ and, in particular, $\widehat{N}_n = o_p(1)$. Moreover, $A_n = O_p(1)$ implies $\hat{d}_n \to d$ in probability.

Pick $\varepsilon = (\Delta - d)/2 > 0$. Because $\widehat{N}_n = o_p(1)$, we have $\Delta + q_{1-\alpha}\widehat{N}_n \le \Delta + \varepsilon$ with probability tending to one. Therefore,

$$\mathbb{P}(\text{reject}) = \mathbb{P}\{\hat{d}_n > \Delta + q_{1-\alpha}\widehat{N}_n\} \le \mathbb{P}\{\hat{d}_n > \Delta + \varepsilon\} + o(1) \le \mathbb{P}\{\hat{d}_n > d + \varepsilon\} + o(1) \to 0,$$

since $\hat{d}_n \to d$ in probability.

Pick $\varepsilon = (d - \Delta)/2 > 0$. Again $\widehat{N}_n = o_p(1)$ implies $q_{1-\alpha}\widehat{N}_n = o_p(1)$ and therefore $\Delta + q_{1-\alpha}\widehat{N}_n \le \Delta + \varepsilon$ with probability tending to one. Hence

$$\mathbb{P}(\text{not reject}) = \mathbb{P}\{\hat{d}_n \le \Delta + q_{1-\alpha}\widehat{N}_n\} \le \mathbb{P}\{\hat{d}_n \le \Delta + \varepsilon\} + o(1) \le \mathbb{P}\{\hat{d}_n \le d - \varepsilon\} + o(1) \to 0,$$

so $\mathbb{P}(\text{reject}) \to 1$.

On the event $\{\widehat{N}_n > 0\}$ the rejection rule (13) is equivalent to

$$\frac{\hat{d}_n - d}{\widehat{N}_n} > q_{1-\alpha}.$$

By Theorem 1,

$$\frac{\hat{d}_n - d}{\widehat{N}_n} \Rightarrow T \stackrel{\text{def}}{=} \frac{B(1)}{\rho\{U(\cdot)\}}.$$

Because $B(1)$ has a continuous density and $\rho\{U(\cdot)\} > 0$ a.s., the distribution of T is continuous: conditionally on $R = \rho\{U(\cdot)\}$, $T \mid R = r \sim N(0, r^{-2})$, and mixtures of continuous distributions are continuous. Therefore, if $q_{1-\alpha}$ is the $(1-\alpha)$-quantile of T, then $\mathbb{P}(T > q_{1-\alpha}) = \alpha$.

It remains to note that $\mathbb{P}(\widehat{N}_n > 0) \to 1$. Indeed, $\sqrt{n}\widehat{N}_n \Rightarrow \tau\rho(U)$ and $\tau > 0$; since $\rho(U) > 0$ a.s., we have $\mathbb{P}(\tau\rho(U) > 0) = 1$ and hence $\mathbb{P}(\widehat{N}_n > 0) \to 1$. Combining these facts gives $\mathbb{P}(\text{reject}) \to \alpha$ when $d = \Delta$. □

A.4 Proof of Theorem 3

Proof. Let $d_n = \Delta + c/\sqrt{n}$ be a Pitman sequence of local alternatives. Assumption 4 is assumed to hold along this sequence with the same limiting scale parameter $\tau > 0$.

Write

$$\sqrt{n}(\hat{d}_n - \Delta) = \sqrt{n}(\hat{d}_n - d_n) + \sqrt{n}(d_n - \Delta) = \sqrt{n}(\hat{d}_n - d_n) + c.$$

By Assumption 4 with d replaced by d_n,

$$\sqrt{n}(\hat{d}_n - d_n) \Rightarrow \tau B(1).$$

Therefore,

$$\sqrt{n}(\hat{d}_n - \Delta) \Rightarrow \tau B(1) + c. \tag{A.3}$$

As in the proof of Theorem 1,

$$\sqrt{n}\,\widehat{N}_n = \rho(H_n) \Rightarrow \rho(\tau U) = \tau\rho(U).$$

On $\{\widehat{N}_n > 0\}$,

$$\frac{\hat{d}_n - \Delta}{\widehat{N}_n} = \frac{\sqrt{n}(\hat{d}_n - \Delta)}{\sqrt{n}\widehat{N}_n}.$$

Combining (A.3) with the denominator convergence and applying Slutsky's theorem yields

$$\frac{\hat{d}_n - \Delta}{\widehat{N}_n} \Rightarrow \frac{\tau B(1) + c}{\tau \rho(U)} = \frac{B(1) + c/\tau}{\rho(U)}.$$

Since $\mathbb{P}(\widehat{N}_n > 0) \to 1$ by the argument in the proof of Theorem 2, the rejection probability converges to

$$\mathbb{P}\left\{ \frac{B(1) + c/\tau}{\rho(U)} > q_{1-\alpha} \right\},$$

which is the stated limit. □

A.5 Proof of Theorem 4

Proof. The statement is an immediate re-parametrization of Theorem 3. Indeed, Theorem 3 shows that the local-power curve depends on the Pitman drift c only through the standardized effect size $\kappa = c/\tau$. Replacing c/τ by κ in the limit in Theorem 3 yields (16). □

A.6 Proof of Lemma 1

Proof. By definition,

$$U(\lambda) = \lambda B(\lambda) - \lambda^2 B(1) = \lambda\{B(\lambda) - \lambda B(1)\} = \lambda\, \mathbb{B}(\lambda),$$

where $\mathbb{B}(\lambda) \stackrel{\text{def}}{=} B(\lambda) - \lambda B(1)$ is the Brownian bridge. For any fixed $\lambda \in [0, 1]$,

$$\operatorname{Cov}\big(B(1), \mathbb{B}(\lambda)\big) = \operatorname{Cov}\big(B(1), B(\lambda)\big) - \lambda \operatorname{Var}\big(B(1)\big) = \lambda - \lambda = 0.$$

More generally, for any finite set $\lambda_1, \ldots, \lambda_k$, the vector

$$\Big(B(1),\, \mathbb{B}(\lambda_1), \ldots, \mathbb{B}(\lambda_k)\Big)$$

is jointly Gaussian with zero cross-covariances between the first coordinate and the remaining coordinates. Hence $B(1)$ is independent of $(\mathbb{B}(\lambda_1), \ldots, \mathbb{B}(\lambda_k))$. Because $(B(1), \mathbb{B}(\cdot))$ is a Gaussian element in $\mathbb{R} \times C([0, 1])$, independence of all finite-dimensional marginals implies independence of $B(1)$ and the entire bridge process $\mathbb{B}(\cdot)$. Therefore $\rho\{U(\cdot)\} = \rho\{\lambda\, \mathbb{B}(\lambda)\}$ depends only on the bridge path and is independent of $B(1)$. □

A.7 Proof of Corollary 5

Proof. Let $R = \rho\{U(\cdot)\}$. By Lemma 1, R is independent of $B(1) \sim N(0,1)$ and satisfies $R > 0$ a.s. By Theorem 4,

$$\pi_\rho(c) = \mathbb{P}\left\{ \frac{B(1) + c/\tau}{R} > q_{1-\alpha} \right\}.$$

Conditioning on R yields

$$\pi_\rho(c) = \mathsf{E}\left[\mathbb{P}\left(B(1) > q_{1-\alpha} R - \frac{c}{\tau} \,\middle|\, R \right) \right] = \mathsf{E}\left[1 - \Phi\left(q_{1-\alpha} R - \frac{c}{\tau} \right) \right],$$

which is (17). For the derivative, note that the integrand is bounded between 0 and 1 and is differentiable in c with derivative

$$\frac{\partial}{\partial c} \left\{ 1 - \Phi\left(q_{1-\alpha} R - \frac{c}{\tau} \right) \right\} = \frac{1}{\tau}\, \varphi\left(q_{1-\alpha} R - \frac{c}{\tau} \right),$$

where φ is the standard normal density. Since φ is bounded, dominated convergence justifies differentiation under the expectation, giving

$$\pi'_\rho(c) = \frac{1}{\tau}\, \mathsf{E}\left[\varphi\left(q_{1-\alpha} R - \frac{c}{\tau} \right) \right],$$

and evaluating at $c = 0$ yields (18). □

A.8 Proof of Corollary 6

Proof. Corollary 5 holds for any admissible functional ρ. For the quadratic self-normalizer (10), $\rho(f) = \rho_{2,\nu}(f) = \{\int_0^1 f(\lambda)^2\, \nu(\mathrm{d}\lambda)\}^{1/2}$, so $R = \rho\{U(\cdot)\}$ equals $\{\int_0^1 U(\lambda)^2\, \nu(\mathrm{d}\lambda)\}^{1/2}$. For the adjusted-range self-normalizer (11), $\rho(f) = \rho_{\mathrm{range}}(f) = \sup_\lambda f(\lambda) - \inf_\lambda f(\lambda)$, so $R = \sup_\lambda U(\lambda) - \inf_\lambda U(\lambda)$. Substituting these expressions into (17) and (18) gives the stated formulas. □

A.9 Proof of Theorem 7

Proof. We give a slightly more explicit argument.

Under $H_0 : \theta = 0$, Assumption 5 implies

$$\sqrt{n}\, \hat{\theta}_n = \sqrt{n}(\hat{\theta}_n - \theta) \Rightarrow \mathcal{G} \qquad \text{in } \mathcal{H}.$$

The map $\phi : \mathcal{H} \to \mathbb{R}$, $\phi(x) = \|x\|_{\mathcal{H}}^2$, is continuous. Therefore the continuous mapping theorem yields

$$T_n = \left\| \sqrt{n}\, \hat{\theta}_n \right\|_{\mathcal{H}}^2 \Rightarrow \|\mathcal{G}\|_{\mathcal{H}}^2.$$

Since Σ is trace class, there exists an orthonormal basis $\{e_k\}_{k\geq 1}$ of $\mathcal{H}$ consisting of eigenfunctions of Σ with eigenvalues $\lambda_k \geq 0$ and $\sum_{k\geq 1}\lambda_k < \infty$. Define the scalar coordinates $G_k = \langle \mathcal{G}, e_k\rangle$. Then $(G_k)_{k\geq 1}$ is a centered Gaussian sequence with $\mathrm{Var}(G_k) = \lambda_k$ and $\mathrm{Cov}(G_k, G_\ell) = 0$ for $k \neq \ell$, hence the standardized variables $Z_k = G_k/\sqrt{\lambda_k}$ (for $\lambda_k > 0$) are independent $N(0,1)$. Consequently,

$$\|\mathcal{G}\|_{\mathcal{H}}^2 = \sum_{k\geq 1}\langle \mathcal{G}, e_k\rangle^2 = \sum_{k\geq 1}\lambda_k Z_k^2,$$

with the convention that terms with $\lambda_k = 0$ contribute zero. This proves (22). □

A.10 Proof of Theorem 8

Proof. Under the local alternatives $H_{1,n}(h)$ in (24), we have $\theta_n = h/\sqrt{n}$ and hence

$$\sqrt{n}\,\hat{\theta}_n = \sqrt{n}(\hat{\theta}_n - \theta_n) + h.$$

Assumption 5 yields $\sqrt{n}(\hat{\theta}_n - \theta_n) \Rightarrow \mathcal{G}$ in $\mathcal{H}$, and Slutsky's theorem gives

$$\sqrt{n}\,\hat{\theta}_n \Rightarrow \mathcal{G} + h.$$

Applying the continuous mapping theorem to $\phi(x) = \|x\|_{\mathcal{H}}^2$ yields (25).

For the series representation, expand $h = \sum_{k\geq 1} h_k e_k$ with $h_k = \langle h, e_k\rangle$ and use the representation $\mathcal{G} = \sum_{k\geq 1}\sqrt{\lambda_k}\, Z_k e_k$ with i.i.d. $Z_k \sim N(0,1)$. Then

$$\|\mathcal{G} + h\|_{\mathcal{H}}^2 = \sum_{k\geq 1}\big(\sqrt{\lambda_k}\, Z_k + h_k\big)^2 = \sum_{k:\lambda_k>0}\lambda_k\Big(Z_k + \frac{h_k}{\sqrt{\lambda_k}}\Big)^2 + \sum_{k:\lambda_k=0} h_k^2,$$

which gives (26). □

Remark A.1 (Pitman local-power comparison and "Pitman-optimal" self-normalizers)**.** *Theorem 3 yields a fully pivotal Pitman criterion for comparing candidate self-normalizers: for a fixed significance level α and a fixed functional ρ, the entire local-power curve depends only on Brownian motion through (B, U). Consequently, for any finite collection of admissible functionals $\{\rho_r\}$, one may define a Pitman-optimal choice at a standardized local effect size $\kappa = c/\tau$ as any maximizer of*

$$\pi_\rho(\kappa) \;=\; \mathbb{P}\left\{\frac{B(1)+\kappa}{\rho\{U(\cdot)\}} > q_{1-\alpha}(\rho)\right\}, \qquad q_{1-\alpha}(\rho) \;=\; \inf\left\{x : \mathbb{P}\Big(\frac{B(1)}{\rho\{U(\cdot)\}} \leq x\Big) \geq 1-\alpha\right\}.$$

This comparison can be carried out once by Monte Carlo simulation of Brownian motion, without simulating any functional time series.

Corollary A.1 (Quadratic and adjusted-range pivots)**.** *Let* $U(\lambda) = \lambda B(\lambda) - \lambda^2 B(1)$*. For the quadratic functional* $\rho_{2,\nu}(f) = \{\int_0^1 f(\lambda)^2 \nu(\,\mathrm{d}\lambda)\}^{1/2}$ *and the adjusted-range functional* $\rho_{\mathrm{range}}(f) = \sup_{\lambda\in[0,1]} f(\lambda) - \inf_{\lambda\in[0,1]} f(\lambda)$*, the corresponding pivotal limits in Theorem 1 are*

$$W = \frac{B(1)}{\big(\int_0^1 U(\lambda)^2\,\nu(\,\mathrm{d}\lambda)\big)^{1/2}}, \qquad S = \frac{B(1)}{\sup_{\lambda\in[0,1]} U(\lambda) - \inf_{\lambda\in[0,1]} U(\lambda)}.$$

Under Pitman drifts $d_n = \Delta + c/\sqrt{n}$*, the limiting rejection probabilities are obtained by replacing* $B(1)$ *with* $B(1) + c/\tau$ *in* W *and* S*, respectively.*

Proof. The expressions for W and S are direct consequences of Theorem 1 upon choosing $\rho = \rho_{2,\nu}$ and $\rho = \rho_{\mathrm{range}}$, respectively, since $\rho\{U(\cdot)\}$ equals the corresponding denominator. Under Pitman drifts $d_n = \Delta + c/\sqrt{n}$, the numerator shift follows from Theorem 3, which replaces $B(1)$ by $B(1) + c/\tau$ in the pivotal limit. □