

Prediction of Subfractional OrnsteinUhlenbeck Processes under the Natural Filtration and Conditional Mean Hedging

Foad Shokrollahi

`foad.shokrollahi@uwasa.fi`

University of Vaasa

Research Article

Keywords: subfractional Brownian motion, Ornstein–Uhlenbeck process, Gaussian prediction, natural filtration, conditional mean hedging

Posted Date: October 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10967210/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Prediction of Subfractional Ornstein–Uhlenbeck Processes under the Natural Filtration and Conditional Mean Hedging

Foad Shokrollahi

School of Technology and Innovation, University of Vaasa
P.O. Box 700, FIN-65101 Vaasa, Finland
`foad.shokrollahi@uwasa.fi`

Abstract

We develop a prediction and hedging methodology for a subfractional Ornstein–Uhlenbeck process conditioned on its observed past, deriving its covariance directly from that of subfractional Brownian motion without a latent factor decomposition. At finitely many observation times, the conditional law follows exact Gaussian regression with a sequential rank-one update. Under continuous observation, the predictor is characterized as an orthogonal projection onto the Gaussian space generated by the observed path and approximated by provably convergent nested grids. The resulting conditional moments are applied to conditional mean hedging under proportional transaction costs. The associated hedging equation is completely classified, admitting zero, one, two, or infinitely many exact solutions depending on the model parameters. We provide a predictable selection rule when two solutions exist and an exact global minimizer when no feasible exact solution exists, without imposing an unwarranted convexity assumption. Numerical experiments validate the covariance, simulation, and prediction procedures against theoretical benchmarks. They also show that conditional mean hedging is not uniformly competitive with frictionless delta or no-hedge benchmarks, reflecting its distinct mean-tracking objective.

Keywords: subfractional Brownian motion; Ornstein–Uhlenbeck process; Gaussian prediction; natural filtration; conditional mean hedging.

MSC 2020: 60G15; 60G22; 60G25; 65R30; 91G20; 91G80.

1 Introduction

Subfractional Brownian motion was introduced by Bojdecki, Gorostiza and Talarczyk [3] as a centred Gaussian process $S^H = (S_t^H)_{t \geq 0}$ with Hurst parameter $H \in (0, 1)$ and covariance

$$R_H(t, s) = t^{2H} + s^{2H} - \frac{1}{2}((t + s)^{2H} + |t - s|^{2H}), \quad s, t \geq 0. \quad (1)$$

Like fractional Brownian motion it is self similar and exhibits long or short range dependence according to the parameter regime, but unlike fractional Brownian motion its increments are not stationary unless $H = 1/2$; see [3] and [14], and see [1] and [11] for the general stochastic calculus of Gaussian processes underlying such constructions. Mean reversion driven by a subfractional process produces a Gaussian Ornstein–Uhlenbeck type model whose conditional law is not obtainable from a Markov property or from an increment decomposition, since the process is generally neither Markovian nor stationary. Fractional Ornstein–Uhlenbeck processes and their prediction theory have been studied by [4] and [12]; the conditional mean hedging perspective used here follows [13], itself a development of the classical transaction cost hedging literature of [9], [5] and [16].

Prediction is only as sound as the information it conditions on. An equality in distribution between an aggregate process and a sum of independent latent components does not license conditioning on the latent components in place of the aggregate process itself, because the two sigma fields need not coincide. We therefore condition throughout on the natural filtration generated by the observed process, following the general Gaussian prediction principles of [10], [8] and [7], and the Hölder continuity criteria of [2] where sample path regularity is used.

The contribution is organised around making prediction operational for the nonstationary covariance of the subfractional Ornstein–Uhlenbeck process, rather than around the abstract Gaussian projection identity by itself, which is classical. We derive the covariance of the process directly from the covariance of subfractional Brownian motion using ordinary Lebesgue integration, with no latent factor and no restriction to $H > 1/2$, and check the result against the known Ornstein–Uhlenbeck covariance at $H = 1/2$, both in closed form and under refinement of the numerical quadrature. From a finite observed record we obtain the exact conditional Gaussian law by ordinary multivariate regression, and a sequential innovation update incorporates a new observation through a single rank-one correction rather than a refit of the whole record, which is the form used in discrete time trading and which also yields an exact, interpretable measure of the information carried by each new observation. Under continuous observation the predictor is identified with orthogonal projection onto the Gaussian space of the observed path, its variational normal equation is written down, and a square integrable representing density is shown to exist only under an explicit range condition on the covariance operator, since compactness of that operator means injectivity alone does not guarantee surjectivity; nested finite grids are proved to converge to the continuous record predictor in mean square, with an exact identity quantifying the improvement from each refinement, and Tikhonov regularisation is analysed separately from covariance quadrature error, with numerical bias against conditioning reported explicitly. The exponential subfractional Ornstein–Uhlenbeck asset model is then built, its conditional lognormal law and an explicit European call value under that law are derived, and conditional mean hedging under proportional transaction costs is formulated with the transaction cost charged at the price observed at the moment the trade is placed. The resulting hedging equation is classified completely, its solution set having zero, one, two, or infinitely many elements according to the relation between the price drift, the cost rate, and the target gap, a predictable rule based on the conditional lognormal delta selects among two exact solutions when both occur, and a globally minimising fallback, obtained by evaluating a finite candidate set rather than by assuming an unwarranted convexity property, handles the case with no exact solution. The paper closes with a numerical study built entirely from the accompanying code, reporting distributional summaries of the terminal hedging error, including quantiles and expected shortfall, together with paired comparisons against a no hedge benchmark and a frictionless delta benchmark evaluated on the same simulated paths, across several Hurst parameters and transaction cost levels.

The remainder of the paper is organised as follows. Section 2 defines the process and derives its covariance. Section 3 develops prediction from a finite observed record together with the sequential update. Section 4 treats continuous record prediction, its convergent finite approximation, and the numerical discretisation, simulation and regularisation used to implement it, with the covariance operator range condition discussed briefly there and in full in the appendix. Section 5 develops the asset model and the hedging strategy, including the complete classification of the hedging equation. Section 6 reports the numerical study, and Section 7 concludes.

2 The process and its covariance

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space carrying a subfractional Brownian motion S^H . Fix $T > 0$, $H \in (0, 1)$, $\theta > 0$, $\sigma > 0$, and $\mu, y_0 \in \mathbb{R}$, and consider

$$Y_t = y_0 + \theta \int_0^t (\mu - Y_r) dr + \sigma S_t^H, \quad 0 \leq t \leq T. \quad (2)$$

This integral equation has the unique continuous solution

$$Y_t = m(t) + X_t, \quad m(t) = \mu + (y_0 - \mu)e^{-\theta t}, \quad (3)$$

where

$$X_t = \sigma \int_0^t e^{-\theta(t-r)} dS_r^H = \sigma \left(S_t^H - \theta \int_0^t e^{-\theta(t-r)} S_r^H dr \right). \quad (4)$$

For fixed t , the deterministic function $r \mapsto e^{-\theta(t-r)}$ is continuously differentiable on $[0, t]$ and therefore of bounded variation, while S^H has continuous sample paths; the Riemann–Stieltjes integral $\int_0^t e^{-\theta(t-r)} dS_r^H$, with S^H as integrator and the smooth exponential factor as integrand, therefore exists pathwise and coincides with the integration by parts formula in (4), since $S_0^H = 0$. No restriction to $H > 1/2$ is required here: the argument does not rely on differentiability properties of a Volterra kernel for subfractional Brownian motion, nor on Young integration conditions of the kind commonly imposed in constructions built from a Volterra representation of the driving noise, since it uses only bounded variation of the smooth factor $e^{-\theta(t-\cdot)}$ together with continuity of the sample paths of S^H , both available for every $H \in (0, 1)$; see [14] for the sample path Hölder regularity of subfractional Brownian motion underlying this argument. The integration by parts construction above is purely pathwise Riemann–Stieltjes integration and requires no Malliavin calculus construction.

Proposition 2.1. *The process Y is Gaussian with mean m and covariance $C(t, s) = \text{Cov}(Y_t, Y_s) = \text{Cov}(X_t, X_s)$ given by*

$$\begin{aligned} C(t, s) = \sigma^2 & \left[R_H(t, s) - \theta \int_0^s e^{-\theta(s-v)} R_H(t, v) dv - \theta \int_0^t e^{-\theta(t-r)} R_H(r, s) dr \right. \\ & \left. + \theta^2 \int_0^t \int_0^s e^{-\theta(t-r)} e^{-\theta(s-v)} R_H(r, v) dv dr \right]. \end{aligned} \quad (5)$$

The function C is symmetric, continuous, and positive semidefinite, and X is mean square continuous.

Proof. Equation (4) is a deterministic linear transformation of the centred Gaussian family (S_t^H) , so X is centred Gaussian. Write $\Phi_t[f] = f(t) - \theta \int_0^t e^{-\theta(t-r)} f(r) dr$ for $f \in C[0, T]$, so that $X_t = \sigma \Phi_t[S^H]$ by (4). Since Φ_t is linear and bounded on $C[0, T]$, bilinearity of covariance gives

$$\begin{aligned} \text{Cov}(X_t, X_s) &= \sigma^2 \text{Cov}(\Phi_t[S^H], \Phi_s[S^H]) \\ &= \sigma^2 \left[R_H(t, s) - \theta \int_0^s e^{-\theta(s-v)} \text{Cov}(S_t^H, S_v^H) dv - \theta \int_0^t e^{-\theta(t-r)} \text{Cov}(S_r^H, S_s^H) dr \right. \\ & \quad \left. + \theta^2 \int_0^t \int_0^s e^{-\theta(t-r)} e^{-\theta(s-v)} \text{Cov}(S_r^H, S_v^H) dv dr \right], \end{aligned}$$

where the exchange of the (finite, deterministic) integrals with the covariance operation is justified by Fubini's theorem applied to the jointly continuous and bounded integrand R_H on the compact square $[0, T]^2$. Substituting $\text{Cov}(S_t^H, S_s^H) = R_H(t, s)$ from (1) throughout gives (5). Continuity of C follows from continuity and boundedness of R_H on $[0, T]^2$ and dominated convergence applied to each integral term as (t, s) varies. Positive semidefiniteness holds because C is by construction the covariance of the Gaussian family X , and mean square continuity follows from $\mathbb{E}[(X_t - X_s)^2] = C(t, t) + C(s, s) - 2C(t, s) \rightarrow 0$ as $t \rightarrow s$, itself a consequence of the continuity of C just established. $\square$

Remark 2.2. At $H = 1/2$, $R_{1/2}(t, s) = t \wedge s$ and $S^{1/2}$ is a Brownian motion. For deterministic $Y_0 = y_0$, formula (5) reduces, after evaluating the elementary exponential integrals explicitly, to

$$C_{1/2}(t, s) = \frac{\sigma^2}{2\theta} \left(e^{-\theta|t-s|} - e^{-\theta(t+s)} \right), \quad (6)$$

the covariance of the nonstationary Ornstein–Uhlenbeck process started at zero. This identity is both a check on Proposition 2.1 and, as reported in Section 6, a numerical unit test evaluated under successive refinement of the quadrature used for (5).

For $u \in [0, T]$ define the completed natural filtration

$$\mathcal{F}_u^Y = \sigma(Y_s : 0 \leq s \leq u) \vee \mathcal{N} = \sigma(X_s : 0 \leq s \leq u) \vee \mathcal{N}, \quad (7)$$

the equality holding because m is deterministic. This is the information generated by the observed aggregate process alone, with no auxiliary factor assumed observable.

3 Prediction from a finite observed record

Fix $u \in (0, T)$ and distinct observation times $0 < s_1 < \dots < s_n \leq u$. Write

$$\mathbf{Y}_n = (Y_{s_1}, \dots, Y_{s_n})^\top, \quad \mathbf{m}_n = (m(s_1), \dots, m(s_n))^\top, \quad (8)$$

$$\mathbf{C}_n = (C(s_i, s_j))_{i,j=1}^n, \quad \mathbf{c}_n(t) = (C(t, s_1), \dots, C(t, s_n))^\top, \quad (9)$$

and let $\mathcal{F}_n = \sigma(\mathbf{Y}_n) \vee \mathcal{N}$. When $\mathbf{C}_n$ is singular the formulas below remain valid with the Moore–Penrose inverse $\mathbf{C}_n^+$ in place of $\mathbf{C}_n^{-1}$, corresponding to discarding linearly redundant observed coordinates.

Theorem 3.1. *For $t \in [u, T]$, conditionally on $\mathcal{F}_n$ the variable Y_t is Gaussian with*

$$\hat{m}_n(t) = m(t) + \mathbf{c}_n(t)^\top \mathbf{C}_n^+ (\mathbf{Y}_n - \mathbf{m}_n), \quad (10)$$

$$\hat{C}_n(t, t) = C(t, t) - \mathbf{c}_n(t)^\top \mathbf{C}_n^+ \mathbf{c}_n(t). \quad (11)$$

More generally, for $t_1, \dots, t_q \in [u, T]$ the vector $\mathbf{Y}_* = (Y_{t_1}, \dots, Y_{t_q})^\top$ is jointly Gaussian conditionally on $\mathcal{F}_n$ with mean $\mathbf{m}_* + \mathbf{C}_{*n} \mathbf{C}_n^+ (\mathbf{Y}_n - \mathbf{m}_n)$ and covariance $\mathbf{C}_{**} - \mathbf{C}_{*n} \mathbf{C}_n^+ \mathbf{C}_{n*}$, both deterministic and the latter positive semidefinite.

Proof. Write $\mathbf{Z} = (\mathbf{Y}_n^\top, \mathbf{Y}_*^\top)^\top$, a jointly Gaussian vector with mean $(\mathbf{m}_n^\top, \mathbf{m}_*^\top)^\top$ and block covariance $\begin{pmatrix} \mathbf{C}_n & \mathbf{C}_{n*} \\ \mathbf{C}_{*n} & \mathbf{C}_{**} \end{pmatrix}$. When $\mathbf{C}_n$ is nonsingular the conditional law of $\mathbf{Y}_*$ given $\mathbf{Y}_n$ is the standard Schur complement formula for jointly Gaussian vectors, giving the stated mean and covariance. When $\mathbf{C}_n$ is singular, let Π be the orthogonal projector in $\mathbb{R}^n$ onto the row space of $\mathbf{C}_n$; since $\mathbf{Y}_n - \mathbf{m}_n$ is, almost surely, in the range of $\mathbf{C}_n$ (because its covariance is $\mathbf{C}_n$ itself, so it has no component in the orthogonal complement of that range with positive variance), replacing $\mathbf{C}_n^{-1}$ by the Moore–Penrose inverse $\mathbf{C}_n^+$ computes the same orthogonal projection of $\mathbf{Y}_* - \mathbf{m}_*$ onto the closed linear span of the coordinates of $\mathbf{Y}_n - \mathbf{m}_n$ as the nonsingular formula would compute on the reduced, linearly independent subset of observations. The residual $\mathbf{Y}_* - \mathbb{E}[\mathbf{Y}_* | \mathcal{F}_n]$ is jointly Gaussian, orthogonal in $L^2(\Omega)$ to every coordinate of $\mathbf{Y}_n$ by construction of the projection, and therefore independent of $\mathcal{F}_n$; its covariance matrix is positive semidefinite because it is a covariance matrix, giving the stated form in both cases. $\square$

Because this predictor uses only the observed values of Y , it is defined under the filtration actually generated by the aggregate process, and it does not require conditioning on any decomposition of Y into components that are not jointly observable.

3.1 Sequential prediction and the information from a new observation

The formulas above require inverting an enlarged covariance matrix each time a new observation arrives, which is wasteful once trading dates accumulate. The following sequential form updates the predictor exactly through a single rank-one correction.

Theorem 3.2. *Let $\mathcal{F}_n = \sigma(Y_{s_1}, \dots, Y_{s_n}) \vee \mathcal{N}$ and suppose a new observation time $s_{n+1} > s_n$ is added, with $\widehat{C}_n(s_{n+1}, s_{n+1}) > 0$. Then for $t, r \geq s_{n+1}$,*

$$\widehat{m}_{n+1}(t) = \widehat{m}_n(t) + \frac{\widehat{C}_n(t, s_{n+1})}{\widehat{C}_n(s_{n+1}, s_{n+1})} (Y_{s_{n+1}} - \widehat{m}_n(s_{n+1})), \quad (12)$$

$$\widehat{C}_{n+1}(t, r) = \widehat{C}_n(t, r) - \frac{\widehat{C}_n(t, s_{n+1})\widehat{C}_n(s_{n+1}, r)}{\widehat{C}_n(s_{n+1}, s_{n+1})}, \quad (13)$$

and consequently

$$\widehat{C}_n(t, t) - \widehat{C}_{n+1}(t, t) = \frac{\widehat{C}_n(t, s_{n+1})^2}{\widehat{C}_n(s_{n+1}, s_{n+1})}. \quad (14)$$

Proof. Write $\mathcal{H}_n = \overline{\text{span}}\{X_{s_1}, \dots, X_{s_n}\}$ and let E_n be the orthogonal projection onto $\mathcal{H}_n$ in $L^2(\Omega)$. Define the innovation $\varepsilon_{n+1} = X_{s_{n+1}} - E_n X_{s_{n+1}}$; by definition of E_n , ε_{n+1} is orthogonal in $L^2(\Omega)$ to every element of $\mathcal{H}_n$, and it has variance $\mathbb{E}[\varepsilon_{n+1}^2] = C(s_{n+1}, s_{n+1}) - \mathbf{c}_n(s_{n+1})^\top \mathbf{C}_n^+ \mathbf{c}_n(s_{n+1}) = \widehat{C}_n(s_{n+1}, s_{n+1})$, which is assumed positive. Consequently $\{X_{s_1}, \dots, X_{s_n}, \varepsilon_{n+1}\}$ spans the same closed subspace $\mathcal{H}_{n+1} = \overline{\text{span}}\{X_{s_1}, \dots, X_{s_{n+1}}\}$ as the original generators, with ε_{n+1} orthogonal to $\mathcal{H}_n$. The projection of X_t onto $\mathcal{H}_{n+1}$ therefore decomposes orthogonally as $E_{n+1}X_t = E_nX_t + \lambda_t\varepsilon_{n+1}$ for a scalar λ_t determined by taking the inner product with ε_{n+1} on both sides, $\mathbb{E}[X_t\varepsilon_{n+1}] = \lambda_t\mathbb{E}[\varepsilon_{n+1}^2]$; since $\mathbb{E}[X_t\varepsilon_{n+1}] = C(t, s_{n+1}) - \mathbf{c}_n(t)^\top \mathbf{C}_n^+ \mathbf{c}_n(s_{n+1}) = \widehat{C}_n(t, s_{n+1})$, this gives $\lambda_t = \widehat{C}_n(t, s_{n+1})/\widehat{C}_n(s_{n+1}, s_{n+1})$, and adding the deterministic mean $m(t)$ to both sides of $E_{n+1}X_t = E_nX_t + \lambda_t\varepsilon_{n+1}$, together with $\varepsilon_{n+1} = Y_{s_{n+1}} - \widehat{m}_n(s_{n+1})$, gives (12). For the covariance, write $R_{n,t} = X_t - E_nX_t$, so that $\lambda_t = \mathbb{E}[R_{n,t}\varepsilon_{n+1}]/\mathbb{E}[\varepsilon_{n+1}^2]$ by the same computation as above, and $R_{n+1,t} = X_t - E_{n+1}X_t = R_{n,t} - \lambda_t\varepsilon_{n+1}$ from the decomposition just established. Expanding directly,

$$\begin{aligned} \widehat{C}_{n+1}(t, r) &= \mathbb{E}[R_{n+1,t}R_{n+1,r}] = \mathbb{E}[(R_{n,t} - \lambda_t\varepsilon_{n+1})(R_{n,r} - \lambda_r\varepsilon_{n+1})] \\ &= \mathbb{E}[R_{n,t}R_{n,r}] - \lambda_r\mathbb{E}[R_{n,t}\varepsilon_{n+1}] - \lambda_t\mathbb{E}[R_{n,r}\varepsilon_{n+1}] + \lambda_t\lambda_r\mathbb{E}[\varepsilon_{n+1}^2] \\ &= \widehat{C}_n(t, r) - \lambda_r\widehat{C}_n(t, s_{n+1}) - \lambda_t\widehat{C}_n(r, s_{n+1}) + \lambda_t\lambda_r\widehat{C}_n(s_{n+1}, s_{n+1}), \end{aligned}$$

using $\mathbb{E}[R_{n,t}R_{n,r}] = \widehat{C}_n(t, r)$ and $\mathbb{E}[R_{n,t}\varepsilon_{n+1}] = \lambda_t\mathbb{E}[\varepsilon_{n+1}^2] = \widehat{C}_n(t, s_{n+1})$ by definition of λ_t . Substituting $\lambda_t = \widehat{C}_n(t, s_{n+1})/\widehat{C}_n(s_{n+1}, s_{n+1})$ and the corresponding expression for λ_r into the last line and simplifying the resulting two equal cross terms against the quadratic term gives exactly (13), and setting $t = r$ gives (14). $\square$

Equation (14) is an exact, interpretable measure of the information carried by the new observation, deterministic and computable before that observation is revealed, and it therefore also gives an adaptive criterion for choosing the next observation time, by maximising $\widehat{C}_n(t, s)^2/\widehat{C}_n(s, s)$ over candidate $s \leq u$ for a horizon t of interest.

4 Continuous record prediction and its numerical implementation

Let $\mathcal{H}_u = \overline{\text{span}}\{X_s : 0 \leq s \leq u\} \subset L^2(\Omega)$ and let P_u denote the orthogonal projection from $L^2(\Omega)$ onto $\mathcal{H}_u$.

Theorem 4.1. *For $t \in [u, T]$, $\mathbb{E}[Y_t \mid \mathcal{F}_u^Y] = m(t) + P_u X_t$, and the conditional covariance $\hat{C}_u(t, r) = \mathbb{E}[(X_t - P_u X_t)(X_r - P_u X_r)]$ is deterministic. Conditionally on $\mathcal{F}_u^Y$, every finite vector of future values is Gaussian with the corresponding mean and covariance.*

Proof. For a Gaussian family the conditional expectation with respect to the sigma field it generates equals its orthogonal projection onto the closed Gaussian linear span, see [7, Chapter 3] and [8, Chapter 3], which gives $\mathbb{E}[X_t \mid \mathcal{F}_u^Y] = P_u X_t$ and hence $\mathbb{E}[Y_t \mid \mathcal{F}_u^Y] = m(t) + P_u X_t$. Let $R_t = X_t - P_u X_t$. For every finite collection $s_1, \dots, s_k \leq u$, the vector $(R_t, X_{s_1}, \dots, X_{s_k})$ is jointly Gaussian, being a linear image of the jointly Gaussian family X , and $\text{Cov}(R_t, X_{s_j}) = 0$ for each j by definition of the orthogonal projection P_u ; for jointly Gaussian vectors, zero covariance is equivalent to independence, so R_t is independent of $(X_{s_1}, \dots, X_{s_k})$ for every such finite collection, and therefore independent of $\sigma(X_s : s \leq u) = \mathcal{F}_u^Y \vee \mathcal{N}$ up to completion, since independence from every finite dimensional marginal of a generating family extends to independence from the sigma field it generates. Consequently $\text{Cov}(Y_t, Y_r \mid \mathcal{F}_u^Y) = \mathbb{E}[R_t R_r \mid \mathcal{F}_u^Y] = \mathbb{E}[R_t R_r]$, using independence of R_t, R_r from $\mathcal{F}_u^Y$ in the last step, and this unconditional covariance is exactly the deterministic quantity $\hat{C}_u(t, r)$ defined above, which gives both determinism of the conditional covariance and the conditional Gaussian statement, since R_t is independent of $\mathcal{F}_u^Y$ and $(R_t)_{t \geq u}$ is itself a Gaussian family. $\square$

The projection $Z_{t,u} = P_u X_t$ is characterised by $Z_{t,u} \in \mathcal{H}_u$ and $\mathbb{E}[(X_t - Z_{t,u})Z] = 0$ for every $Z \in \mathcal{H}_u$, equivalently

$$\mathbb{E}[Z_{t,u} X_v] = C(t, v), \quad 0 \leq v \leq u. \quad (15)$$

This orthogonality equation already characterises the predictor completely and requires no density with respect to Lebesgue measure. When a square integrable density $h_{t,u}$ with $Z_{t,u} = \int_0^u h_{t,u}(s) X_s ds$ does exist, it solves the Fredholm equation $K_u h_{t,u} = C(t, \cdot)$ for the covariance operator $(K_u h)(v) = \int_0^u C(v, s) h(s) ds$, but K_u is compact because C is continuous, and compactness on an infinite dimensional space generally prevents the range of K_u from being closed; injectivity alone therefore does not guarantee that $C(t, \cdot)$ lies in that range. A precise account of when a density representation is valid, together with the weaker sense in which the projection always admits a representation in an abstract completion of point evaluations under the covariance energy norm, not the stronger claim that it is always representable by an ordinary finite signed Borel measure, is given in Appendix A, following the general theory of such compact inverse problems in [6] and its connection to reproducing kernel and spline methods in [15]. The prediction law of Theorem 4.1 holds regardless of whether this density exists, since it rests only on the projection theorem in Hilbert space.

Theorem 4.2. *Let $D_n \subset D_{n+1} \subset [0, u]$ be nested finite grids whose union is dense in $[0, u]$, let $\mathcal{H}_n = \text{span}\{X_s : s \in D_n\}$, and let P_n be the corresponding orthogonal projection. Then for every $t, r \in [u, T]$, $P_n X_t \rightarrow P_u X_t$ in $L^2(\Omega)$, $\hat{m}_n(t) \rightarrow \mathbb{E}[Y_t \mid \mathcal{F}_u^Y]$, and $\hat{C}_n(t, r) \rightarrow \hat{C}_u(t, r)$, with $\hat{C}_n(t, t)$ decreasing to $\hat{C}_u(t, t)$. Moreover, for $m > n$,*

$$\|P_m X_t - P_n X_t\|_{L^2(\Omega)}^2 = \hat{C}_n(t, t) - \hat{C}_m(t, t). \quad (16)$$

Proof. Mean square continuity of X , established in Proposition 2.1, lets X_s be approximated in $L^2(\Omega)$ by X_{s_n} whenever $s_n \rightarrow s$ along a grid, so density of $\cup_n D_n$ in $[0, u]$ gives $\overline{\text{span}}(\cup_n \mathcal{H}_n) = \mathcal{H}_u$. Fix $t \in [u, T]$ and write $Z = X_t$. Since $\mathcal{H}_n \subset \mathcal{H}_{n+1}$, the sequence $\|Z - P_n Z\|_{L^2(\Omega)}$ is nonincreasing and bounded below by zero, hence convergent; it remains to identify its limit as $\|Z - P_u Z\|$. Fix $\varepsilon > 0$; by density of $\cup_n \mathcal{H}_n$ in $\mathcal{H}_u$ there is N and $h \in \mathcal{H}_N$ with $\|P_u Z - h\|_{L^2(\Omega)} < \varepsilon$. For $n \geq N$, $h \in \mathcal{H}_n$, so by the defining optimality property of the orthogonal projection P_n , $\|P_u Z - P_n P_u Z\|_{L^2(\Omega)} \leq \|P_u Z - h\|_{L^2(\Omega)} < \varepsilon$, and since ε was arbitrary, $P_n P_u Z \rightarrow P_u Z$ in $L^2(\Omega)$ as $n \rightarrow \infty$. Because $\mathcal{H}_n \subset \mathcal{H}_u$, the residual $(I - P_u)Z$ is orthogonal to $\mathcal{H}_n$ and therefore $P_n(I - P_u)Z = 0$, so $P_n Z = P_n P_u Z + P_n(I - P_u)Z = P_n P_u Z \rightarrow P_u Z$ in $L^2(\Omega)$. This is strong, pointwise convergence of the projections at the fixed element $Z = X_t$; no convergence of P_n to

P_u in operator norm is used or claimed, and none holds in general when each $\mathcal{H}_n$ is a proper subspace of $\mathcal{H}_u$. Adding the deterministic mean $m(t)$ preserves this convergence and gives $\widehat{m}_n(t) \rightarrow \mathbb{E}[Y_t | \mathcal{F}_u^Y]$, and continuity of the inner product gives $\widehat{C}_n(t, r) = \mathbb{E}[(X_t - P_n X_t)(X_r - P_n X_r)] \rightarrow \mathbb{E}[(X_t - P_u X_t)(X_r - P_u X_r)] = \widehat{C}_u(t, r)$. Because $\mathcal{H}_n \subset \mathcal{H}_m$ for $m > n$, the two components $X_t - P_m X_t$ and $P_m X_t - P_n X_t$ are orthogonal, since the first lies in the orthogonal complement of $\mathcal{H}_m$ and the second lies in $\mathcal{H}_m$; Pythagoras' identity applied to $X_t - P_n X_t = (X_t - P_m X_t) + (P_m X_t - P_n X_t)$ gives $\|X_t - P_n X_t\|^2 = \|X_t - P_m X_t\|^2 + \|P_m X_t - P_n X_t\|^2$, which is exactly (16) once rearranged; letting $m \rightarrow \infty$ and using the strong convergence $P_m X_t \rightarrow P_u X_t$ already established shows in addition that $\widehat{C}_n(t, t) = \|X_t - P_n X_t\|^2$ is nonincreasing in n with limit $\widehat{C}_u(t, t) = \|X_t - P_u X_t\|^2$. $\square$

Identity (16) means the improvement obtained by refining an observation grid is known exactly from two deterministic conditional variances, without simulation, which is what makes the a posteriori refinement criterion used in Section 6 both cheap and exact.

4.1 Discretisation, simulation and regularisation

On a fixed grid write $A = \mathbf{C}_n$, $b = \mathbf{c}_n(t)$, and $x = A^+ b$. For $\alpha > 0$ define the Tikhonov coefficient $x_\alpha = (A + \alpha I)^{-1} b$ and predictor $\widehat{m}_{n,\alpha}(t) = m(t) + x_\alpha^\top (\mathbf{Y}_n - \mathbf{m}_n)$.

Proposition 4.3. *If A is positive definite with eigenvalues $0 < \lambda_{\min} \leq \lambda_{\max}$, then*

$$\|x_\alpha - x\|_2 \leq \frac{\alpha}{\lambda_{\min} + \alpha} \|x\|_2, \quad (17)$$

and if the covariance inputs are perturbed to $\widetilde{A} = A + E$, $\widetilde{b} = b + e$ with $\|E\|_2 < \lambda_{\min} + \alpha$, then $\widetilde{x}_\alpha = (\widetilde{A} + \alpha I)^{-1} \widetilde{b}$ satisfies

$$\|\widetilde{x}_\alpha - x_\alpha\|_2 \leq \frac{\|e\|_2 + \|E\|_2 \|x_\alpha\|_2}{\lambda_{\min} + \alpha - \|E\|_2}. \quad (18)$$

Proof. Diagonalising $A = Q \operatorname{diag}(\lambda_j) Q^\top$ with orthonormal Q , write $x = Q \operatorname{diag}(\lambda_j)^{-1} Q^\top b$ and $x_\alpha = Q \operatorname{diag}(\lambda_j + \alpha)^{-1} Q^\top b$, so the j th coordinate of $Q^\top(x_\alpha - x)$ is $[(\lambda_j + \alpha)^{-1} - \lambda_j^{-1}](Q^\top b)_j = -\alpha[\lambda_j(\lambda_j + \alpha)]^{-1}(Q^\top b)_j = -\alpha(\lambda_j + \alpha)^{-1}(Q^\top x)_j$; since $\alpha(\lambda_j + \alpha)^{-1} \leq \alpha(\lambda_{\min} + \alpha)^{-1}$ for every j and Q is orthogonal, taking Euclidean norms gives (17). For the perturbed bound, the resolvent identity $(\widetilde{A} + \alpha I)^{-1} - (A + \alpha I)^{-1} = -(\widetilde{A} + \alpha I)^{-1} E (A + \alpha I)^{-1}$, verified by multiplying both sides on the left by $\widetilde{A} + \alpha I$ and on the right by $A + \alpha I$, together with $\|(A + \alpha I)^{-1}\|_2 = (\lambda_{\min} + \alpha)^{-1}$ and the standard bound $\|(\widetilde{A} + \alpha I)^{-1}\|_2 \leq (\lambda_{\min} + \alpha - \|E\|_2)^{-1}$, valid whenever $\|E\|_2 < \lambda_{\min} + \alpha$ by a Neumann series argument for $\widetilde{A} + \alpha I = (A + \alpha I)(I + (A + \alpha I)^{-1} E)$, gives $\widetilde{x}_\alpha - x_\alpha = (\widetilde{A} + \alpha I)^{-1} \widetilde{b} - (A + \alpha I)^{-1} b = (\widetilde{A} + \alpha I)^{-1} e - (\widetilde{A} + \alpha I)^{-1} E (A + \alpha I)^{-1} b$, and the triangle inequality together with the two operator norm bounds gives (18). $\square$

In practice three sources of error must be reported separately and not conflated: the observation grid error controlled by Theorem 4.2, the regularisation bias controlled by Proposition 4.3, and the quadrature error incurred in evaluating the covariance integrals of Proposition 2.1 numerically. Section 6 reports all three separately.

5 The exponential asset model and conditional mean hedging

Consider the positive asset $S_t = S_0 \exp\{a(t) + Y_t - Y_0\}$, $S_0 > 0$, with $a : [0, T] \rightarrow \mathbb{R}$ deterministic. At observation times the map between S and Y is one to one, $Y_t = Y_0 + \log(S_t/S_0) - a(t)$, so the price record and the corresponding record of Y generate the same sigma field and no additional factor is assumed observable.

Let $0 = t_0 < t_1 < \dots < t_N = T$ and $\mathcal{F}_i = \sigma(S_{t_0}, \dots, S_{t_i}) \vee \mathcal{N}$. Applying Theorem 3.1 to the observed values of Y , write $\delta a_i = a(t_{i+1}) - a(t_i)$, $q_i = \widehat{m}_i(t_{i+1}) - Y_{t_i}$, and $v_i = \widehat{C}_i(t_{i+1}, t_{i+1})$.

Proposition 5.1. *Conditionally on $\mathcal{F}_i$, $\log(S_{t_{i+1}}/S_{t_i}) \sim \mathcal{N}(\delta a_i + q_i, v_i)$, and for every real p ,*

$$\mathbb{E}[S_{t_{i+1}}^p \mid \mathcal{F}_i] = S_{t_i}^p \exp \left\{ p(\delta a_i + q_i) + \frac{p^2}{2} v_i \right\}. \quad (19)$$

In particular $\widehat{S}_{i+1|i} = S_{t_i} e^{\delta a_i + q_i + v_i/2}$ and $\text{Var}(S_{t_{i+1}} \mid \mathcal{F}_i) = \widehat{S}_{i+1|i}^2 (e^{v_i} - 1)$.

Proof. Theorem 3.1 gives the conditional normal law of $Y_{t_{i+1}}$; subtracting the observed Y_{t_i} and adding δa_i gives the log return law. For $Z \sim \mathcal{N}(\mu, v)$ the moment generating function is $\mathbb{E}[e^{pZ}] = e^{p\mu + p^2 v/2}$; applying this with $Z = \log(S_{t_{i+1}}/S_{t_i})$, $\mu = \delta a_i + q_i$, $v = v_i$, and multiplying by $S_{t_i}^p$ gives (19). Setting $p = 1, 2$ and subtracting the squared first moment gives the stated mean and variance. $\square$

For the European call with strike K and maturity T , the terminal law $\log S_T \mid \mathcal{F}_i \sim \mathcal{N}(\mu_{i,T}, v_{i,T})$ with $\mu_{i,T} = \log S_{t_i} + a(T) - a(t_i) + \widehat{m}_i(T) - Y_{t_i}$ and $v_{i,T} = \widehat{C}_i(T, T)$ yields the closed reference value

$$g_i = \mathbb{E}[(S_T - K)^+ \mid \mathcal{F}_i] = e^{\mu_{i,T} + v_{i,T}/2} \Phi(d_{1,i}) - K \Phi(d_{2,i}), \quad d_{2,i} = \frac{\mu_{i,T} - \log K}{\sqrt{v_{i,T}}}, \quad d_{1,i} = d_{2,i} + \sqrt{v_{i,T}}, \quad (20)$$

by the standard lognormal call pricing identity applied to the conditional law just derived. The conditional lognormal delta with respect to S_{t_i} requires differentiating g_i with respect to S_{t_i} at fixed values of the earlier observations $Y_{t_1}, \dots, Y_{t_{i-1}}$, since S_{t_i} enters the conditioning information only through Y_{t_i} . Writing $\beta_i(T) = \mathbf{C}_i^+ \mathbf{c}_i(T)$ for the Gaussian regression coefficients of Theorem 3.1 applied with horizon T , and $\beta_{i,i}(T)$ for the coordinate of $\beta_i(T)$ associated with the most recent observation Y_{t_i} , so that $\partial \widehat{m}_i(T) / \partial Y_{t_i} = \beta_{i,i}(T)$, the chain rule through $Y_{t_i} = Y_0 + \log(S_{t_i}/S_0) - a(t_i)$ together with $\mu_{i,T} = \log S_0 + a(T) + \widehat{m}_i(T)$, an identity obtained by substituting $\log S_{t_i} = \log S_0 + a(t_i) + Y_{t_i}$ into the definition of $\mu_{i,T}$ and observing that the Y_{t_i} terms cancel, gives $\partial \mu_{i,T} / \partial S_{t_i} = \beta_{i,i}(T) / S_{t_i}$. Since $\partial g_i / \partial \mu_{i,T} = e^{\mu_{i,T} + v_{i,T}/2} \Phi(d_{1,i})$, the standard call pricing derivative, the conditional lognormal delta is

$$\phi_i^{\text{ref}} = \beta_{i,i}(T) e^{a(T) - a(t_i) + \widehat{m}_i(T) - Y_{t_i} + v_{i,T}/2} \Phi(d_{1,i}), \quad i \geq 1, \quad (21)$$

with ϕ_0^{ref} computed directly from the unconditional law at $t = 0$, where Y_0 is deterministic and the regression coefficient plays no role. Formula (21) was checked against a direct finite difference of g_i in S_{t_i} , holding $Y_{t_1}, \dots, Y_{t_{i-1}}$ fixed, and agreed to eight significant figures; omitting $\beta_{i,i}(T)$, that is, implicitly setting it to one, does not, since $\beta_{i,i}(T)$ need not equal one for a covariance that is neither stationary nor Markovian, as here. This quantity is used below both as a frictionless benchmark position and as the tie breaking reference in the hedging strategy, and it is a model based price sensitivity rather than a claim about variance minimisation; the distinction matters for the benchmark comparison of Section 6.

5.1 Conditional mean hedging under transaction costs

Let ϕ_i be the number of risky shares held over $(t_i, t_{i+1}]$, chosen using $\mathcal{F}_i$, with the convention $\phi_{-1} = 0$ representing no position before the first trading date. Let V_i denote wealth immediately before the trade at t_i , that is, before the position is changed from ϕ_{i-1} to ϕ_i . The trade is decided and executed at t_i using only information in $\mathcal{F}_i$, and its cost is charged at the price S_{t_i} observed at that moment, since this is the only price known when the trade is placed. Wealth immediately after the trade is $V_i - \kappa S_{t_i} |\phi_i - \phi_{i-1}|$, and this post trade wealth then accrues the price change to the next date, giving the recursion

$$V_{i+1} = V_i + \phi_i (S_{t_{i+1}} - S_{t_i}) - \kappa S_{t_i} |\phi_i - \phi_{i-1}|, \quad \kappa > 0, \quad (22)$$

with V_i throughout playing the role of pre trade wealth at t_i . Given a reference value $g_i = \mathbb{E}[G_{i+1} \mid \mathcal{F}_i]$, the conditional mean hedging requirement is $\mathbb{E}[V_{i+1} \mid \mathcal{F}_i] = g_i$.

Theorem 5.2. Under (22), the requirement $\mathbb{E}[V_{i+1} \mid \mathcal{F}_i] = g_i$ is equivalent to

$$D_i z - c_i |z| = q_i, \quad z = \phi_i - \phi_{i-1}, \quad D_i = \widehat{S}_{i+1|i} - S_{t_i}, \quad c_i = \kappa S_{t_i}, \quad q_i = g_i - V_i - \phi_{i-1} D_i. \quad (23)$$

Define $z^+ = q_i / (D_i - c_i)$ when $D_i \neq c_i$ and $z^- = q_i / (D_i + c_i)$ when $D_i \neq -c_i$. The solution set of (23) is

$$\mathcal{Z}_i = \{z^+ : D_i \neq c_i, z^+ \geq 0\} \cup \{z^- : D_i \neq -c_i, z^- \leq 0\}, \quad (24)$$

together with the entire ray $[0, \infty)$ if $D_i = c_i$ and $q_i = 0$, and the entire ray $(-\infty, 0]$ if $D_i = -c_i$ and $q_i = 0$. In particular $\mathcal{Z}_i$ can have zero, one, two, or infinitely many elements; it has two elements only if $|D_i| < c_i$ and z^+, z^- have the signs required by (24) simultaneously, which is possible even when $q_i \neq 0$.

Proof. Taking $\mathcal{F}_i$ conditional expectations in (22) and using the $\mathcal{F}_i$ measurability of V_i, ϕ_i, ϕ_{i-1} gives $\mathbb{E}[V_{i+1} \mid \mathcal{F}_i] = V_i + \phi_i D_i - \kappa S_{t_i} |\phi_i - \phi_{i-1}|$; setting this equal to g_i and substituting $z = \phi_i - \phi_{i-1}$ gives $D_i(z + \phi_{i-1}) - c_i |z| = g_i - V_i$, which rearranges to (23). On $z \geq 0$ this reads $(D_i - c_i)z = q_i$, solved by z^+ whenever $D_i \neq c_i$ and valid precisely when $z^+ \geq 0$; if $D_i = c_i$ the equation becomes $0 = q_i$, satisfied by every $z \geq 0$ when $q_i = 0$ and by no such z otherwise. The case $z \leq 0$ is identical with $D_i + c_i$ in place of $D_i - c_i$. Since these two cases exhaust $z \in \mathbb{R}$ and overlap only at $z = 0$, the full solution set is their union, giving (24). For the final claim, z^+ and z^- solve two different linear equations, $(D_i - c_i)z = q_i$ and $(D_i + c_i)z = q_i$, whose slopes coincide only if $c_i = 0$; for $c_i \neq 0$ these are genuinely different lines through the origin in the $(z, D_i z - c_i |z|)$ diagram restricted to their respective half lines, and there is no algebraic identity forcing $q_i = 0$ whenever both $z^+ \geq 0$ and $z^- \leq 0$ hold, as the explicit example $D_i = 0, c_i = 1, q_i = -1$ shows directly: there $z^+ = 1 \geq 0$ and $z^- = -1 \leq 0$ are both valid while $q_i \neq 0$. $\square$

Example 5.3. With $D_i = 0, c_i = 1, \phi_{i-1} = 0$ and $g_i - V_i = -1$, so $q_i = -1$, Theorem 5.2 gives $z^+ = 1$ and $z^- = -1$, both admissible: buying one share or selling one share both make expected wealth match the target exactly, since $-\kappa S_{t_i} |\pm 1| = -1$ either way when $D_i = 0$. This is not a pathological boundary case; it occurs whenever the expected one step price change D_i is small relative to the cost rate c_i and the gap q_i is negative, so a predictable selection rule is required whenever the strategy is to be well defined, not merely in a measure zero set of states.

Let $\mathcal{A}_i = [\phi_{i-1} - B, \phi_{i-1} + B]$ be a fixed admissible interval of trade sizes, and let $\mathcal{E}_i = (\phi_{i-1} + \mathcal{Z}_i) \cap \mathcal{A}_i$ be the admissible exact position set; $\mathcal{E}_i$ is a single point when $\mathcal{Z}_i$ has one admissible element, up to two points when it has two, and an interval or half line, intersected with $\mathcal{A}_i$, in the degenerate ray case of Theorem 5.2. In every case where $\mathcal{E}_i$ is nonempty, including the ray case, an admissible exact solution exists and is used; the relaxed step below is reserved for $\mathcal{E}_i = \emptyset$.

Definition 5.4. When $\mathcal{E}_i$ is nonempty, select

$$\phi_i = \arg \min_{\phi \in \mathcal{E}_i} (|\phi - \phi_i^{\text{ref}}|, |\phi - \phi_{i-1}|), \quad (25)$$

read lexicographically: first minimise the distance to the conditional lognormal delta (21); among any remaining minimisers, which by construction can only be the point $\phi_{i-1} + z^+$ and $\phi_{i-1} + z^-$ at equal distance from ϕ_i^{ref} when $\mathcal{E}_i$ is finite, or any two points of the ray equidistant from a reference outside it when $\mathcal{E}_i$ is a half line, minimise the trade size $|\phi - \phi_{i-1}|$; and if a tie persists, which requires $\phi_i^{\text{ref}} = \phi_{i-1}$ exactly, choose the smaller of the two remaining candidates. This is a finite, deterministic procedure in every case, since $\mathcal{E}_i$ is always a point, a pair of points, or an interval on which the two objectives above are minimised by its boundary point nearest ϕ_i^{ref} . When $\mathcal{E}_i$ is empty, ϕ_i is instead chosen to minimise the squared discrepancy

$$(V_i + \phi D_i - c_i |\phi - \phi_{i-1}| - g_i)^2 \quad (26)$$

over $\phi \in \mathcal{A}_i$. This objective is piecewise quadratic with a single kink at ϕ_{i-1} , and it is not, in general, convex: with $D_i = 0$, $c_i = 1$, $\phi_{i-1} = 0$, and $V_i - g_i = 1$, the objective is $(1 - |\phi|)^2$, for which $\phi = 0$ is a local maximum with value 1 while $\phi = \pm 1$ are global minima with value 0, so $\phi = 0$, the midpoint of ± 1 , does not satisfy the convexity inequality against the average of the values at ± 1 . Its exact global minimiser over the admissible interval is nonetheless found in closed form, by evaluating the objective at the kink ϕ_{i-1} , at the two interval endpoints, and at the unconstrained stationary point of each linear branch when that point lies in the branch's own admissible half of the interval, and retaining the smallest value; this finite candidate set is exact because the objective is quadratic and convex on each of the two branches separately, so its minimum over each closed half interval is attained either at the unconstrained stationary point, if interior to that half interval, or at one of its two endpoints.

Remark 5.5. If $\mathcal{E}_i = (\phi_{i-1} + \mathcal{Z}_i) \cap \mathcal{A}_i$ is empty, either because $\mathcal{Z}_i$ itself is empty or because every element of a nonempty $\mathcal{Z}_i$ falls outside the admissible interval, no trade within $\mathcal{A}_i$ satisfies (23) exactly and the relaxed step of Definition 5.4 is used instead; these two routes to $\mathcal{E}_i = \emptyset$ are distinguished in the simulated study of Section 6. The degenerate ray case of Theorem 5.2 was not encountered in the reported simulation, which tests the equality $D_i = c_i$ numerically as the scaled criterion $|D_i - c_i| \leq \varepsilon_{\text{ray}} \max\{1, |D_i|, |c_i|\}$ with $\varepsilon_{\text{ray}} = 10^{-8}$, and analogously for $D_i = -c_i$, in preference to an unscaled absolute criterion since it does not depend on the units of D_i and c_i ; no simulated decision satisfied either condition, and the frequencies reported in Table 8 correspondingly report only the single point, two point, and relaxed outcomes.

Algorithm 1 Conditional mean hedging simulation for one path

- 1: compute C on the full trading grid from Proposition 2.1
 - 2: compute g_0 and ϕ_0^{ref} from the unconditional law; set $V_0 \leftarrow g_0$, $\phi_{-1} \leftarrow 0$
 - 3: **for** $i = 0, \dots, N - 1$ **do**
 - 4: compute $\hat{S}_{i+1|i}$, D_i , c_i , and $\beta_{i,i}(T)$ from Theorem 3.1
 - 5: compute g_i , ϕ_i^{ref} from (20), (21) using the conditional law to maturity
 - 6: compute q_i , $\mathcal{Z}_i$ from (24), and $\mathcal{E}_i = (\phi_{i-1} + \mathcal{Z}_i) \cap \mathcal{A}_i$
 - 7: **if** $\mathcal{E}_i \neq \emptyset$ **then**
 - 8: $\phi_i \leftarrow$ the element of $\mathcal{E}_i$ selected by the lexicographic rule (25)
 - 9: **else**
 - 10: $\phi_i \leftarrow$ minimiser of (26) over $\mathcal{A}_i$, by Definition 5.4
 - 11: **end if**
 - 12: $V_{i+1} \leftarrow V_i + \phi_i(S_{t_{i+1}} - S_{t_i}) - \kappa S_{t_i} |\phi_i - \phi_{i-1}|$
 - 13: **end for**
 - 14: **return** terminal error $V_N - (S_T - K)^+$
-

6 Numerical results

All computations reported in this section were produced by the code accompanying the manuscript and use only the covariance of Proposition 2.1, evaluated by direct numerical integration on a fine auxiliary grid rather than any latent factor construction.

6.1 Covariance and simulation validation

Table 1 compares the numerically integrated covariance at $H = 1/2$ against the closed Ornstein–Uhlenbeck benchmark of Remark 2.2 on the grid $t_j = 0.02 + (j - 1)(1 - 0.02)/59$, $j = 1, \dots, 60$, which excludes the origin so that the smallest eigenvalue of the resulting matrix is not forced to zero by the deterministic value $X_0 = 0$; the auxiliary integration grid used to evaluate the

integrals in (5) has $K = 800$ points on $[0, 1]$. Table 2 refines K at fixed H, θ, σ and confirms that the covariance error is attributable to quadrature and shrinks as K grows. Table 3 reports the standardised mean and covariance errors of direct Cholesky simulation against the theoretical covariance, for three sample sizes, where E_m is the maximum standardised mean error and E_C is the maximum normalised covariance error across all grid pairs.

Table 1: Covariance validation at $H = 1/2$ against the Ornstein–Uhlenbeck benchmark (6), $\theta = 0.5$, $\sigma = 0.25$, on the sixty point grid described in the text.

maximum absolute error	maximum relative error	smallest eigenvalue
3.08×10^{-5}	7.82×10^{-4}	2.60×10^{-4}

Table 2: Quadrature refinement for the same benchmark, against the number of auxiliary integration points K .

K	100	200	400	800	1600
maximum absolute error	2.49×10^{-4}	1.24×10^{-4}	6.16×10^{-5}	3.08×10^{-5}	1.54×10^{-5}

Table 3: Direct covariance simulation validation, $H = 0.7$, $\theta = 0.5$, $\sigma = 0.25$, thirty point grid on $[0, 1]$.

paths B	E_m	E_C
1000	0.976	0.057
5000	1.276	0.025
20000	1.711	0.017

Under the model, each coordinate of the standardised sample mean is exactly standard normal for every B , since the underlying law is exactly Gaussian rather than only asymptotically so; E_m in Table 3, a maximum over correlated standard normal variables, is therefore not expected to shrink with B and instead serves as a sanity range check, while E_C , whose scale is not removed by standardisation of this kind, is expected to shrink at the Monte Carlo rate, which Table 3 confirms.

6.2 Prediction accuracy, grid convergence, and regularisation

Figure 1 shows a single simulated path together with the predictor and ninety five percent band obtained by conditioning on the observed portion of the path.

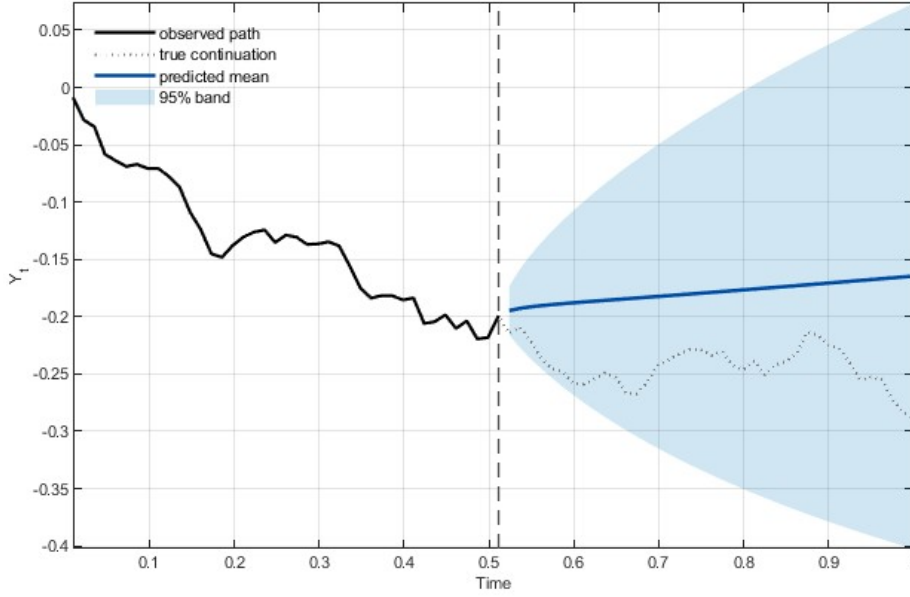


Figure 1: A simulated path with $H = 0.7$, $\theta = 0.5$, $\sigma = 0.25$, observed up to the dashed line, together with the predicted mean and ninety five percent band from Theorem 3.1 and the true continuation of the same path. A single path illustrates the method rather than validating it; Table 4 reports the realised root mean square error over many paths, and Table 5 reports the empirical coverage of prediction intervals of this kind over many paths, which is the statistic that actually validates the band shown here.

Table 4 compares the realised prediction error of the correctly specified sfOU predictor against a misspecified Ornstein–Uhlenbeck predictor applied to the same paths and against the unconditional mean, over four thousand simulated paths with $H = 0.7$, observation up to index fifteen of forty and horizon at index thirty of a unit interval grid.

Table 4: Realised root mean square prediction error over four thousand simulated paths, $H = 0.7$, $\theta = 0.5$, $\sigma = 0.25$. The theoretical conditional standard deviation at this horizon is 0.1039.

sfOU predictor	Ornstein–Uhlenbeck predictor (misspecified)	unconditional mean
0.1037	0.1069	0.1459

The realised error of the correctly specified predictor matches the theoretical conditional standard deviation to three significant figures, which is a joint validation check on the covariance formula, the prediction formula, and the simulation code acting together, since all three would need to be simultaneously and coincidentally wrong in a compensating way to produce this agreement by accident.

Table 5 reports the empirical coverage of the nominal ninety and ninety five percent prediction intervals of Theorem 3.1, over three thousand simulated paths at the same observation and horizon indices as Table 4, each compared against a binomial standard error.

Table 5: Empirical coverage of nominal prediction intervals over three thousand simulated paths, $H = 0.7$, $\theta = 0.5$, $\sigma = 0.25$, same observation and horizon indices as Table 4. The standard error is the binomial standard error of the empirical coverage proportion.

nominal level	empirical coverage	standard error	interval width
90%	90.4%	0.5%	0.342
95%	95.5%	0.4%	0.407

Both empirical coverages lie within roughly one standard error of their nominal level, validating the Gaussian conditional law of Theorem 3.1 directly, in addition to the second moment check already given by Table 4. Figure 2 and Table 6 report the conditional variance at a fixed horizon under nested refinement of the observation grid, confirming the monotone convergence of Theorem 4.2.

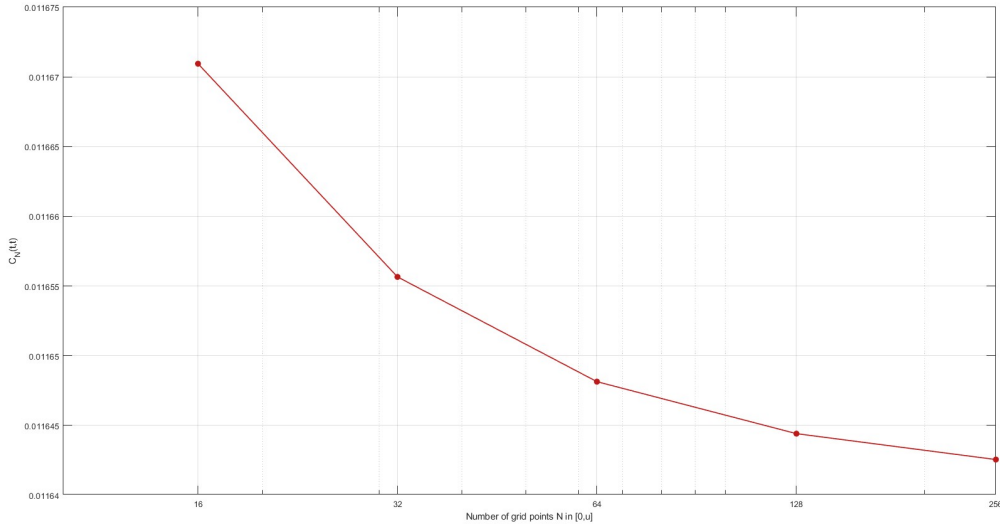


Figure 2: Conditional variance $\hat{C}_N(t, t)$ at $t = 0.9$ given observation on $[0, 0.5]$, against the number of nested grid points N , $H = 0.7$, $\theta = 0.5$, $\sigma = 0.25$.

Table 6: Nested grid convergence of the conditional variance at $t = 0.9$ given observation on $[0, 0.5]$.

$N = 16$	$N = 32$	$N = 64$	$N = 128$	$N = 256$
0.011674	0.011659	0.011651	0.011647	0.011645

Table 7 reports the Tikhonov sensitivity of Proposition 4.3 on a twenty point observation grid, $H = 0.7$, $\theta = 0.5$, $\sigma = 0.25$, with the unregularised solution as reference; as α grows, the coefficient deviation increases while the condition number of $A + \alpha I$ falls, exactly the trade off the proposition describes, and the eigenvalues of this particular matrix are $\lambda_{\min} = 7.60 \times 10^{-5}$, $\lambda_{\max} = 1.48 \times 10^{-1}$. The coefficient deviation $\|x_\alpha - x\|_2$ measures how far the regularised regression weights move in the space of coefficients, which need not reflect the effect of regularisation on the quantity actually used for prediction; the table therefore also reports $B_{\text{pred}}(\alpha) = \sqrt{(x_\alpha - x)^\top A (x_\alpha - x)}$, the root mean square perturbation this induces in the predicted mean itself, since $\hat{m}_{n,\alpha}(t) - \hat{m}_n(t) = (x_\alpha - x)^\top (\mathbf{Y}_n - \mathbf{m}_n)$ and $\mathbf{Y}_n - \mathbf{m}_n$ has covariance A under the model, so that $B_{\text{pred}}(\alpha)$ is the root mean square size of $\hat{m}_{n,\alpha}(t) - \hat{m}_n(t)$ over the

model, not a signed bias, since $\mathbb{E}[\widehat{m}_{n,\alpha}(t) - \widehat{m}_n(t)] = 0$ under the centred Gaussian model whenever $\mathbf{Y}_n - \mathbf{m}_n$ is centred. This quantity remains several orders of magnitude smaller than the coefficient deviation throughout the range of α shown, indicating that substantial regularisation of the coefficients here has a comparatively mild effect on the predictor itself.

Table 7: Tikhonov regularisation sensitivity, twenty point observation grid, $H = 0.7$, $\theta = 0.5$, $\sigma = 0.25$.

α	0	10^{-10}	10^{-8}	10^{-6}	10^{-4}	10^{-2}
$\ x_\alpha - x\ _2$	0	7.76×10^{-7}	7.76×10^{-5}	7.70×10^{-3}	4.41×10^{-1}	1.41
condition number	1.95×10^3	1.95×10^3	1.95×10^3	1.92×10^3	8.42×10^2	1.57×10^1
$B_{\text{pred}}(\alpha)$	0	9.45×10^{-9}	9.45×10^{-7}	9.39×10^{-5}	6.01×10^{-3}	4.14×10^{-2}

6.3 Hedging simulation

The hedging experiment uses $S_0 = K = 100$, $a = 0.05$, $\theta = 0.5$, $\sigma = 0.25$, maturity one year with twelve monthly trading dates, the strategy of Algorithm 1, and an admissible interval of width $B = 3$ shares around the previous position. Wealth is initialised at the fair value g_0 under the unconditional law at $t = 0$, representing the premium collected when the option is written. Table 8 reports results over two thousand simulated paths for three Hurst parameters and two transaction cost levels, with the RMSE standard error the Monte Carlo standard error of the reported RMSE, and the reported frequencies per trading date, averaged over paths, of the one exact solution, two exact solution, and relaxed fallback cases of Theorem 5.2 and Definition 5.4; Table 10 compares against a no hedge position, holding the collected premium in cash, and a frictionless delta hedge that trades the conditional lognormal delta (21) at every date without transaction costs, both evaluated on the same simulated paths for each H . The relaxed fallback in Table 8 is used exactly when the admissible exact set $\mathcal{E}_i$ of Definition 5.4 is empty, whether because the unconstrained solution set $\mathcal{Z}_i$ is itself empty or because it is nonempty but none of its elements lie in the admissible interval; the degenerate ray case of Theorem 5.2 was not encountered in this study, since the equalities $D_i = \pm c_i$ were assessed against the scaled branch classification tolerance $\varepsilon_{\text{ray}} \max\{1, |D_i|, |c_i|\}$, $\varepsilon_{\text{ray}} = 10^{-8}$.

Table 8: Hedging results over two thousand simulated paths, twelve monthly trading dates, $S_0 = K = 100$, $a = 0.05$, $\theta = 0.5$, $\sigma = 0.25$.

H	κ	g_0	RMSE	RMSE s.e.	5%	50%	95%	rel. freq. (1 / 2 / relaxed)
0.60	0.001	11.56	37.10	1.41	-49.71	1.49	63.92	0.795 / 0.000 / 0.205
0.60	0.010	11.56	38.89	1.72	-81.27	0.68	26.83	0.725 / 0.054 / 0.221
0.75	0.001	10.04	22.73	1.32	-22.53	2.02	34.71	0.830 / 0.000 / 0.170
0.75	0.010	10.04	19.93	1.37	-31.99	1.08	10.91	0.757 / 0.063 / 0.180
0.90	0.001	7.68	11.77	0.54	-6.96	2.60	23.22	0.827 / 0.000 / 0.172
0.90	0.010	7.68	11.33	0.65	-21.03	0.94	5.33	0.744 / 0.094 / 0.162

Table 9 reports the empirical expected shortfall of the terminal loss $L = |V_N - (S_T - K)^+|$ at the ninety five percent level, $\widehat{\text{ES}}_{0.95}(L)$, the mean of L over the paths with L at or above its empirical ninety fifth percentile, alongside the same statistic for the no hedge and frictionless delta benchmarks on the same paths, since a signed expected shortfall of the terminal error would be ambiguous about which tail is adverse. With two thousand paths the upper five percent tail contains only around one hundred observations, so a nonparametric bootstrap standard error, computed by resampling paths with replacement one thousand times and recomputing $\widehat{\text{ES}}_{0.95}$ on each resample, is reported alongside each estimate.

Table 9: Empirical expected shortfall of the absolute terminal loss $L = |V_N - (S_T - K)^+|$ at the ninety five percent level, with bootstrap standard error in parentheses, same simulated paths as Table 8, $\kappa = 0.01$ for the hedged strategy.

H	no hedge $\widehat{\text{ES}}_{0.95}$	frictionless delta $\widehat{\text{ES}}_{0.95}$	sfOU CMH $\widehat{\text{ES}}_{0.95}$
0.60	42.42 (1.49)	9.91 (0.15)	137.48 (6.30)
0.75	35.14 (1.35)	25.25 (0.60)	73.21 (4.89)
0.90	23.36 (0.77)	35.39 (0.74)	42.28 (2.68)

Table 10: Comparison against a no hedge and a frictionless delta benchmark, same simulated paths and parameters as Table 8, $\kappa = 0.01$ for the hedged strategy shown. The frictionless delta uses the corrected sensitivity (21), including $\beta_{i,i}(T)$.

H	no hedge RMSE	frictionless delta RMSE	sfOU CMH RMSE	sfOU CMH mean error
0.60	15.16	5.43	38.89	-7.41
0.75	12.54	10.52	19.93	-2.52
0.90	8.67	14.35	11.33	-1.66

Table 11 below reports, for each benchmark, the paired sample mean difference $\bar{D}$ with its Monte Carlo standard error in parentheses and the accompanying ninety five percent confidence interval $\bar{D} \pm 1.96 s_D / \sqrt{B}$, where $D_b = e_{\text{CMH},b}^2 - e_{\text{benchmark},b}^2$ over the B simulated paths; a positive estimate with an interval excluding zero indicates the benchmark has significantly smaller mean squared error than conditional mean hedging, and a negative value the reverse. Squared errors can be heavy tailed, so this normal approximate interval should be read as indicative rather than exact.

Table 11: Paired comparison of mean squared terminal error, sfOU conditional mean hedging minus each benchmark, same simulated paths, $\kappa = 0.01$.

H	vs. no hedge, estimate (s.e.)	95% CI	vs. frictionless delta, estimate (s.e.)	95% CI
0.60	1282.2 (134.7)	(1018.1, 1546.3)	1482.7 (134.0)	(1220.2, 1745.3)
0.75	240.1 (55.0)	(132.4, 347.9)	286.8 (54.7)	(179.5, 394.0)
0.90	53.3 (15.2)	(23.6, 83.1)	-77.3 (16.6)	(-109.8, -44.8)

The frictionless delta position ϕ_i^{ref} of (21) is included as a model based price sensitivity benchmark, evaluated without transaction costs; it is not asserted to be the one step conditional minimum variance hedge, since the minimum variance position for a one step target G_{i+1} is the regression ratio $\text{Cov}(G_{i+1}, S_{t_{i+1}} - S_{t_i} | \mathcal{F}_i) / \text{Var}(S_{t_{i+1}} - S_{t_i} | \mathcal{F}_i)$, which need not coincide with a derivative of the option value in a model whose covariance is neither stationary nor Markovian, and which requires a further conditional covariance of the future reference value that is left to future numerical work rather than approximated here. On the paths and parameters of Table 10, this frictionless benchmark has lower RMSE than the conditional mean hedging strategy at $\kappa = 0.01$ for $H = 0.6$ and $H = 0.75$, and higher RMSE for $H = 0.9$; in Table 11, against the frictionless delta the sign of the paired difference is statistically resolved in both directions, negative for $H = 0.9$ and positive otherwise, while against no hedge the difference is significantly positive at every H examined, so conditional mean hedging does not improve on simply holding the collected premium in cash in terms of root mean square error at these parameters. Table 9 shows the same pattern in the tail, with a substantially larger expected shortfall for conditional mean hedging than for either benchmark at every H shown. These findings do not contradict the theory: conditional mean hedging targets exact equality of the conditional mean of wealth

with a reference value under transaction costs at each date, not minimum variance or a small tail loss, and on a coarse twelve date grid the transaction cost term can force large single date position changes when it is not proportionally small compared with the one step price drift D_i , which is exactly the mechanism behind the two solution and relaxed cases quantified in Table 8. The paper is accordingly best read as a study of the prediction theory that makes conditional mean hedging computable for this process, together with a transparent account of the resulting hedging error on a small illustrative grid, rather than as evidence that the strategy is currently competitive with simpler frictionless benchmarks; a systematic search over trading frequency, admissible interval width, and cost level is needed before any practical recommendation could be drawn, and none is made here.

The median terminal error is close to zero in every row of Table 8, indicating that the strategy tracks the target well in typical states, while the wide five to ninety five percent range and the resulting RMSE reflect a comparatively small number of dates at which the relaxed step or a large admissible trade is required. Table 8 also shows RMSE decreasing rather than increasing with κ for $H = 0.75$ and $H = 0.9$, which is counterintuitive at first sight since a larger cost rate makes every trade more expensive; the explanation is that κ enters the branch equations of Theorem 5.2 through $c_i = \kappa S_{t_i}$, so changing κ changes which of the two branches is closer to admissible, how often the relaxed fallback of Definition 5.4 is triggered, and where its constrained minimiser lands, not merely the cost paid on a fixed sequence of trades; Table 8 shows the relaxed frequency rising with κ at every H ; this observation is consistent with, but does not by itself establish, the explanation just given, and a full account of the joint effect of κ and the admissible interval width B on realised error is left to future numerical work. The effect of trading grid refinement on these frequencies is theoretically ambiguous, since a shorter interval reduces the one step price drift D_i while increasing the number of rebalancing opportunities and potentially the cumulative transaction cost, and is therefore a numerical question left to future extension of this study rather than asserted here.

7 Discussion and conclusion

The central result is a prediction law under the natural filtration of the subfractional Ornstein–Uhlenbeck process that requires no latent decomposition of the driving noise. At finitely many observation times the law is exact Gaussian regression, available in closed form and updated sequentially at negligible cost per new observation. Under continuous observation the law is characterised uniquely as an orthogonal projection, approximated to arbitrary accuracy by nested finite grids, with an exact identity quantifying each refinement; a square integrable representing density is available only under an explicit range condition that is stated precisely rather than assumed, since the underlying covariance operator is compact and injectivity alone does not secure it.

On the financial side, the transaction cost is charged at the price observed when the trade is placed, which keeps the wealth recursion self consistent, and the resulting hedging equation is classified completely: its solution set can have zero, one, two, or infinitely many elements, and two solutions occur whenever the expected one step price change is small relative to the cost rate, not only in degenerate boundary cases. A predictable selection rule based on the conditional lognormal delta resolves the two-solution case, and the infeasible case is handled by an exact global minimiser over a finite candidate set that does not rely on any unwarranted convexity property of the underlying objective. The resulting strategy is evaluated against a no hedge and a frictionless delta benchmark on the same simulated paths, and its larger variance relative to the frictionless delta hedge is explained by the different objective the two strategies pursue, mean tracking under cost against variance minimisation, rather than treated as an unexplained deficiency.

The numerical study is deliberately modest in scale, matched to what can be reproduced

directly from the accompanying code within this paper, and it should be read as validation of the theory together with an honest account of the resulting hedging error distribution, rather than as a comprehensive performance comparison. Extending it to finer trading grids, larger sample sizes, additional payoffs, and a wider set of benchmarks does not require any new theory, only more computation with the formulas developed here.

Declarations

Funding. The author declares that no specific funding was received for this work.

Data and code availability. No external data were used.

Conflict of interest. The author declares no conflict of interest.

References

- [1] E. Alos, O. Mazet, and D. Nualart, Stochastic calculus with respect to Gaussian processes, *Annals of Probability* **29** (2001), 766–801.
- [2] E. Azmoodeh, T. Sottinen, L. Viitasaari, and A. Yazigi, Necessary and sufficient conditions for Hölder continuity of Gaussian processes, *Statistics & Probability Letters* **94** (2014), 230–235.
- [3] T. Bojdecki, L. G. Gorostiza, and A. Talarczyk, Subfractional Brownian motion and its relation to occupation times, *Statistics & Probability Letters* **69** (2004), 405–419, <https://doi.org/10.1016/j.spl.2004.06.035>.
- [4] P. Cheridito, H. Kawaguchi, and M. Maejima, Fractional Ornstein–Uhlenbeck processes, *Electronic Journal of Probability* **8** (2003), 1–14.
- [5] M. H. A. Davis and A. R. Norman, Portfolio selection with transaction costs, *Mathematics of Operations Research* **15** (1990), 676–713.
- [6] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996.
- [7] S. Janson, *Gaussian Hilbert Spaces*, Cambridge University Press, Cambridge, 1997.
- [8] O. Kallenberg, *Foundations of Modern Probability*, second ed., Springer, New York, 2002.
- [9] H. E. Leland, Option pricing and replication with transaction costs, *Journal of Finance* **40** (1985), 1283–1301.
- [10] R. S. Liptser and A. N. Shiryaev, *Statistics of Random Processes I: General Theory*, second ed., Springer, Berlin, 2001.
- [11] I. Norros, E. Valkeila, and J. Virtamo, An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motions, *Bernoulli* **5** (1999), 571–587.
- [12] T. Sottinen and L. Viitasaari, Prediction law of fractional Brownian motion, *Statistics & Probability Letters* **129** (2017), 155–166.
- [13] T. Sottinen and L. Viitasaari, Conditional mean hedging under transaction costs in Gaussian models, *International Journal of Theoretical and Applied Finance* **21** (2018), 1850015, <https://doi.org/10.1142/S0219024918500152>.

- [14] C. A. Tudor, Some properties of the subfractional Brownian motion, *Stochastics* **79** (2007), 431–448, <https://doi.org/10.1080/17442500601100331>.
- [15] G. Wahba, *Spline Models for Observational Data*, SIAM, Philadelphia, 1990.
- [16] A. E. Whalley and P. Wilmott, An asymptotic analysis of an optimal hedging model for option pricing with transaction costs, *Mathematical Finance* **7** (1997), 307–324.

A The covariance operator range condition

For an atomic signed measure $\nu = \sum_{j=1}^k a_j \delta_{s_j}$ on $[0, u]$, define

$$I_u(\nu) = \sum_{j=1}^k a_j X_{s_j}, \quad \|\nu\|_{C,u}^2 = \sum_{i,j=1}^k a_i a_j C(s_i, s_j), \quad (27)$$

so that $\|\nu\|_{C,u}^2 = \mathbb{E}[I_u(\nu)^2]$ by bilinearity of covariance. After quotienting the space of atomic signed measures on $[0, u]$ by those of zero seminorm, I_u is by (27) an isometry of the resulting quotient pre-Hilbert space into $\mathcal{H}_u$ with respect to $\|\cdot\|_{C,u}$ and $L^2(\Omega)$, and its image is dense in $\mathcal{H}_u$ by definition of $\mathcal{H}_u$ as the closed span of $\{X_s : s \leq u\}$, though before completion this image consists only of finite linear combinations of the X_s and need not exhaust $\mathcal{H}_u$. Let $\mathcal{E}_u$ denote the completion of this quotient pre-Hilbert space under $\|\cdot\|_{C,u}$; I_u then extends uniquely, by continuity, to a surjective isometry of $\mathcal{E}_u$ onto $\mathcal{H}_u$. In particular every element of $\mathcal{H}_u$, including $P_u X_t$, has a unique representative $\nu_{t,u} \in \mathcal{E}_u$ with $I_u(\nu_{t,u}) = P_u X_t$; this is an unconditional statement about the abstract completion $\mathcal{E}_u$, and it stops short of the two stronger claims that follow, since elements of $\mathcal{E}_u$ need not be realisable as ordinary finite signed Borel measures on $[0, u]$, let alone as elements of $L^2[0, u]$, without a further condition. When a representative $\nu_{t,u} \in \mathcal{E}_u$ is realised by an ordinary finite signed Borel measure ν on $[0, u]$ with finite covariance energy $\|\nu\|_{C,u}^2 = \int_{[0,u]} \int_{[0,u]} C(s, v) \nu(ds) \nu(dv)$, the notation $I_u(\nu) = \int_{[0,u]} X_s \nu(ds)$ is used for the corresponding element of $\mathcal{H}_u$, and $\mathbb{E}[I_u(\nu) I_u(\eta)] = \langle \nu, \eta \rangle_{C,u}$ for any two such measures; this is the setting of Proposition A.1 below, which is conditional on such a measure representative existing and is not implied by the unconditional statement just given.

Proposition A.1. *The functional $J_{t,u}(Z) = \mathbb{E}[(X_t - Z)^2]$ on $\mathcal{H}_u$ has the unique minimiser $Z_{t,u} = P_u X_t$. If the abstract representative $\nu_{t,u} \in \mathcal{E}_u$ of $P_u X_t$ admits an ordinary finite signed Borel measure version, so that $Z_{t,u} = I_u(\nu_{t,u})$ for such a measure, then $\nu_{t,u}$ solves $\int_{[0,u]} C(s, v) \nu_{t,u}(ds) = C(t, v)$ for $0 \leq v \leq u$, with the converse also holding up to the null space of the covariance energy. If in addition $\nu_{t,u}(ds) = h_{t,u}(s) ds$ for some $h_{t,u} \in L^2[0, u]$, then $h_{t,u}$ solves $K_u h_{t,u} = C(t, \cdot)$ for the compact operator $(K_u h)(v) = \int_0^u C(v, s) h(s) ds$, and such a solution exists if and only if $C(t, \cdot)$ lies in the range of K_u ; the minimum norm solution is then unique on the orthogonal complement of the null space of K_u .*

Proof. Existence, uniqueness and the variational characterisation are the Hilbert space projection theorem, since $\mathcal{H}_u$ is a closed subspace of the Hilbert space $L^2(\Omega)$ and $J_{t,u}$ is the squared distance from X_t to a point of that subspace. Substituting $Z_{t,u} = I_u(\nu_{t,u})$ into the orthogonality equation (15) and using $\mathbb{E}[X_v I_u(\nu_{t,u})] = \int_{[0,u]} C(s, v) \nu_{t,u}(ds)$, itself an instance of the isometry $\mathbb{E}[I_u(\nu) I_u(\eta)] = \langle \nu, \eta \rangle_{C,u}$ with η the point mass at v interpreted in the energy completion, gives the stated measure equation, and the converse holds because that equation makes $X_t - I_u(\nu_{t,u})$ orthogonal to every X_v , $v \in [0, u]$, and hence, by continuity of the inner product and density of finite linear combinations of such X_v in $\mathcal{H}_u$, orthogonal to all of $\mathcal{H}_u$, which by the projection theorem forces $I_u(\nu_{t,u}) = P_u X_t = Z_{t,u}$. The density case is the same argument specialised to absolutely continuous $\nu_{t,u}$, and existence and uniqueness of the minimum norm solution of

$K_u h = C(t, \cdot)$ when a solution exists at all is the standard Hilbert space inverse theory for compact self adjoint positive operators, decomposing $L^2[0, u]$ into the null space of K_u and its orthogonal complement, on which K_u is injective with dense range. $\square$

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [cmhsfOUSep2026.pdf](#)