STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–17.
Published online in International Academic Press (www.IAPress.org)

A Semi-Analytical Approach to Solving the Ornstein–Uhlenbeck Equation via Reproducing Kernel Hilbert Spaces (RKHS)

Erisbey Marin Cardona, Carlos A Ramirez V, Andres F. Valencia-López*

Departamento de Matemáticas, Universidad Tecnológica de Pereira, Pereira 660003, Colombia

Abstract This work proposes a semi-analytical method for approximating the solution of the Ornstein–Uhlenbeck (OU) stochastic differential equation by means of Reproducing Kernel Hilbert Spaces (RKHS). The OU process, widely used as a Gaussian mean-reverting model in physics and finance, is decomposed into a deterministic mean-reversion component and a purely stochastic component represented as an Itô integral. The deterministic part is reconstructed via regularized kernel regression, while the stochastic term is modeled through an autoregressive representation in an RKHS induced by a positive definite kernel. By combining the representer theorem with a kernel-based autoregressive scheme, we obtain a finite-dimensional linear system for the kernel coefficients together with a fixed-point algorithm for solving the preimage problem. The resulting framework provides a fully kernel-based approximation of the OU dynamics. The methodology is validated on synthetic Ornstein–Uhlenbeck trajectories and shows that the RKHS-based approach captures the mean-reversion behavior and stochastic variability of the process while preserving numerical stability and interpretability.

Keywords Ornstein–Uhlenbeck process, stochastic differential equations, reproducing kernel Hilbert spaces, kernel methods, autoregressive models, mean-reverting processes, semi-analytical methods

AMS 2010 subject classifications 60H10, 46E22, 62M10

DOI: 10.19139/soic-2310-5070-3469

1. Introduction

The Ornstein–Uhlenbeck (OU) process is one of the most fundamental continuous-time Gaussian Markov processes. It was originally introduced to describe the velocity of a Brownian particle subject to friction, and it has subsequently become a standard model for mean-reverting dynamics in diverse fields such as physics, biology and mathematical finance. In particular, OU-type processes and their variants underlie short-rate models (e.g., Vasicek-type dynamics), stochastic volatility specifications, and models of noisy mean-reverting signals.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space supporting a standard Brownian motion $\{W_t\}_{t\geq 0}$. The classical one-dimensional Ornstein–Uhlenbeck process $\{X_t\}_{t\geq 0}$ is defined as the solution of the stochastic differential equation (SDE)

$$dX_t = \theta(\mu - X_t)\, dt + \sigma\, dW_t, \qquad t \geq 0, \tag{1}$$

where

- $\theta > 0$ is the mean-reversion speed,
- $\mu \in \mathbb{R}$ is the long-term mean level,
- $\sigma > 0$ is the volatility parameter.

*Correspondence to: Erisbey Marín Cardona (Email: erics10-@utp.edu.co). Departamento de Matemáticas, Universidad Tecnológica de Pereira.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

The drift term $\theta(\mu - X_t)$ pulls the process towards the level μ, while the Brownian term $\sigma\, dW_t$ introduces random fluctuations around this mean-reverting behavior.

The OU process admits a closed-form solution and has Gaussian transition densities, which makes it analytically tractable and allows for explicit expressions of moments and stationary distributions. Nevertheless, in many practical situations one does not observe the full distributional structure: only a finite time series (possibly noisy or irregularly sampled) is available. In such cases, it is natural to seek data-driven reconstruction methods that respect the stochastic structure of the model and can exploit flexible function spaces.

In this work we adopt a semi-analytical perspective similar in spirit to previous RKHS-based approaches for stochastic models. The main idea is to decompose the OU solution into:

1. a *deterministic* component, governed by the mean-reversion dynamics in the absence of noise;
2. a *stochastic* component, given by an Itô integral with respect to Brownian motion.

We then embed these components into a Reproducing Kernel Hilbert Space associated with a positive definite kernel. The deterministic evolution is approximated via regularized kernel regression, whereas the stochastic term is modeled using an autoregressive structure in the RKHS. This leads to a hybrid scheme in which:

- the RKHS framework provides a non-parametric, data-driven representation of trajectories as linear combinations of kernel sections evaluated at observed time points;
- the autoregressive RKHS model captures temporal dependence and mean-reversion in the stochastic component;
- the representer theorem yields a finite-dimensional linear system for the kernel coefficients defining the reconstructed trajectory.

The main contributions of this paper are:

- We derive a decomposition of the Ornstein–Uhlenbeck solution into deterministic and stochastic parts, and show how each part can be embedded into a suitable RKHS.
- We develop a kernel-based autoregressive representation of the OU stochastic component and derive the associated normal equations for the autoregressive coefficients.
- We formulate and solve the preimage problem via a fixed-point iteration, recovering time-domain trajectories from their RKHS embeddings.

The remainder of the paper is organized as follows. In Section 2 we recall the explicit solution of the OU SDE and introduce its deterministic–stochastic decomposition. Section 3 presents the RKHS autoregressive modeling of the stochastic component and the associated normal equations. Section 3.2 discusses the preimage problem and provides a fixed-point algorithm for reconstructing trajectories. Section 4 outlines the numerical experiments with synthetic OU data. Finally, Section 5 summarizes the main findings and suggests future research directions.

2. Ornstein–Uhlenbeck Model and Deterministic–Stochastic Decomposition

We consider again the Ornstein–Uhlenbeck SDE

$$dX_t = \theta(\mu - X_t)\, dt + \sigma\, dW_t, \qquad X_0 = x_0, \tag{2}$$

with $\theta > 0$, $\mu \in \mathbb{R}$ and $\sigma > 0$. Equation (2) is linear in X_t, so it can be solved explicitly by standard methods. Multiplying both sides by the integrating factor $e^{\theta t}$ yields

$$e^{\theta t} dX_t + \theta e^{\theta t} X_t\, dt = \theta\mu e^{\theta t}\, dt + \sigma e^{\theta t}\, dW_t,$$

which can be written as

$$d\big(e^{\theta t}X_t\big) = \theta\mu e^{\theta t}\,dt + \sigma e^{\theta t}\,dW_t.$$

Integrating from 0 to t we obtain

$$e^{\theta t}X_t - X_0 = \theta\mu\int_0^t e^{\theta s}\,ds + \sigma\int_0^t e^{\theta s}\,dW_s.$$

Solving for X_t and using

$$\int_0^t e^{\theta s}\,ds = \frac{e^{\theta t}-1}{\theta},$$

we arrive at the well-known closed-form solution

$$X_t = \mu + (x_0-\mu)e^{-\theta t} + \sigma e^{-\theta t}\int_0^t e^{\theta s}\,dW_s. \tag{3}$$

Deterministic component

The first two terms in (3),

$$X^{\det}(t) := \mu + (x_0-\mu)e^{-\theta t}, \tag{4}$$

correspond to the solution of the *deterministic* ordinary differential equation obtained by formally setting $\sigma = 0$ in (2):

$$\frac{dX_t}{dt} = \theta(\mu - X_t), \qquad X_0 = x_0.$$

This solution describes pure mean-reversion towards the level μ with rate θ, in the absence of stochastic perturbations.

Stochastic component

The third term in (3),

$$\Gamma_t := \sigma e^{-\theta t}\int_0^t e^{\theta s}\,dW_s, \tag{5}$$

is an Itô integral and thus defines a Gaussian martingale with zero mean. This term captures the random fluctuations around the deterministic mean-reversion path. By construction, Γ_t has

$$\mathbb{E}[\Gamma_t] = 0, \qquad \mathrm{Var}(\Gamma_t) = \sigma^2 e^{-2\theta t}\int_0^t e^{2\theta s}\,ds = \frac{\sigma^2}{2\theta}\big(1-e^{-2\theta t}\big),$$

and the full process $X_t = X^{\det}(t) + \Gamma_t$ is Gaussian with mean

$$\mathbb{E}[X_t] = \mu + (x_0-\mu)e^{-\theta t}$$

and variance $\mathrm{Var}(X_t) = \mathrm{Var}(\Gamma_t)$. As $t\to\infty$, the distribution of X_t converges to the stationary normal law

$$X_\infty \sim \mathcal{N}\left(\mu,\,\frac{\sigma^2}{2\theta}\right).$$

Discrete-time representation

For numerical and data-driven purposes, it is convenient to sample the process at discrete times $t_n = n\Delta t$, with step size $\Delta t > 0$. Using either the exact solution (3) or a standard Euler–Maruyama discretization, one obtains a recursion of the form

$$X_{n+1} = \phi X_n + c + \varepsilon_n, \qquad n = 0, 1, \ldots, \tag{6}$$

where

$$\phi = e^{-\theta \Delta t}, \qquad c = \mu(1 - e^{-\theta \Delta t}),$$

and $\{\varepsilon_n\}$ is a zero-mean Gaussian noise sequence with variance determined by σ, θ and Δt. Equation (6) exhibits a discrete-time AR(1)-type structure, making the OU process particularly suitable for autoregressive modeling.

In the proposed RKHS-based framework, we interpret either the centered process $X_t - X^{\det}(t)$ or the stochastic component Γ_t as the primary stochastic signal to be embedded in a Reproducing Kernel Hilbert Space. The discrete-time representation (6) then motivates an RKHS autoregressive model for the stochastic part of the dynamics, which will be developed in detail in Section 3.

3. RKHS Embedding of the Stochastic Ornstein–Uhlenbeck Component

In this section we embed the stochastic part of the Ornstein–Uhlenbeck dynamics into a Reproducing Kernel Hilbert Space (RKHS), following the methodology previously developed for the Black–Scholes equation in our earlier work [7], and adapting it to the mean-reverting structure of the OU process.

Recall the decomposition

$$X_t = X^{\det}(t) + \Gamma_t,$$

where $X^{\det}(t)$ is the deterministic mean-reversion component and Γ_t is the stochastic Itô integral term given in (5). For numerical and data-driven purposes, we work with the discretized version of the centered process

$$Z_t := X_t - X^{\det}(t),$$

sampled at times $t_n = n\Delta t$. Motivated by the AR(1)-type recursion (6), we approximate $\{Z_t\}$ by an autoregressive model of order p:

$$Z_t = \sum_{j=1}^{p} a_j Z_{t-j} + \eta_t, \tag{7}$$

where $a_j \in \mathbb{R}$ are scalar coefficients and η_t is a zero-mean noise term.

3.1. Autoregressive Structure in RKHS

Embedding the AR process (7) into a Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}$ with feature map $\Phi : \mathcal{X} \to \mathcal{H}$ allows us to describe the dynamics in a linear form in $\mathcal{H}$. As in [7], we consider the embedded process

$$\Phi(Z_t) \in \mathcal{H},$$

and postulate the autoregressive representation

$$\Phi(Z_t) = \sum_{j=1}^{p} \alpha_j \Phi(Z_{t-j}) + \varepsilon_t, \tag{8}$$

where the coefficients $\alpha_j \in \mathbb{R}$ characterize the embedded AR dynamics and $\varepsilon_t \in \mathcal{H}$ is a residual term.

Equation (8) can be seen as the natural lifting of the scalar AR model (7) to the feature space induced by the kernel $k(\cdot, \cdot)$, with

$$\langle \Phi(x), \Phi(y) \rangle_{\mathcal{H}} = k(x, y).$$

Physical Interpretation of the OU Model. Although the representation in RKHS is abstract, it has a clear physical meaning when applied to the Ornstein–Uhlenbeck process, which classically models, for instance, the velocity of a Brownian particle under friction, temperature fluctuations around equilibrium, or other mean-reverting physical quantities:

- *Preservation of Zero-Mean Fluctuations:* The stochastic component Γ_t (and thus Z_t) is a zero-mean Gaussian process. The RKHS embedding is constructed so that the regression in feature space does not introduce spurious deterministic drift; in expectation, the centered nature of the fluctuations is preserved. This is consistent with the physical picture of fluctuations around an equilibrium state determined by $X^{\det}(t)$ and the parameters (θ, μ, σ).
- *Temporal Correlations and Relaxation Time:* The OU process is characterized by an exponential decay of correlations with relaxation time $1/\theta$. The AR structure in RKHS captures these temporal correlations in a flexible way: past fluctuations influence the current state through the coefficients α_j and the kernel $k(\cdot,\cdot)$. In particular, the embedded dynamics approximate how perturbations at earlier times relax towards equilibrium.
- *Kernel Hyperparameter as Correlation Scale:* For kernels such as the Gaussian kernel with bandwidth σ_k, the hyperparameter acts as a correlation length (or time) for the fluctuations: large σ_k corresponds to smoother, long-range dependence in the feature space, whereas small σ_k emphasizes short-range, rapidly decorrelating fluctuations. This links directly the kernel choice with physical properties such as correlation time and characteristic scale of the noise.

Estimation of AR Coefficients. As in the Black–Scholes setting [7], the coefficients α_j are obtained by minimizing the mean squared error functional:

$$J(\boldsymbol{\alpha}) = \mathbb{E}\left[\left\|\Phi(Z_t) - \sum_{j=1}^{p} \alpha_j \Phi(Z_{t-j})\right\|_{\mathcal{H}}^2\right], \tag{9}$$

with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)^\top$.

Expanding the squared norm and using the reproducing property $\langle \Phi(x), \Phi(y)\rangle_{\mathcal{H}} = k(x,y)$ leads to:

$$J(\boldsymbol{\alpha}) = \mathbb{E}\big[k(Z_t, Z_t)\big] - 2\sum_{j=1}^{p} \alpha_j \mathbb{E}\big[k(Z_t, Z_{t-j})\big] + \sum_{i,j=1}^{p} \alpha_i \alpha_j \mathbb{E}\big[k(Z_{t-i}, Z_{t-j})\big].$$

Setting the derivative with respect to α_m equal to zero yields the linear system:

$$\sum_{j=1}^{p} \alpha_j \mathbb{E}[k(Z_{t-m}, Z_{t-j})] = \mathbb{E}[k(Z_t, Z_{t-m})], \qquad m = 1, \ldots, p. \tag{10}$$

Defining the *RKHS covariance matrix*:

$$\mathbf{K} = \begin{bmatrix} \mathbb{E}[k(Z_{t-1}, Z_{t-1})] & \ldots & \mathbb{E}[k(Z_{t-1}, Z_{t-p})] \\ \vdots & \ddots & \vdots \\ \mathbb{E}[k(Z_{t-p}, Z_{t-1})] & \ldots & \mathbb{E}[k(Z_{t-p}, Z_{t-p})] \end{bmatrix}_{p\times p}, \tag{11}$$

the *coefficient vector*:

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix}_{p\times 1}, \tag{12}$$

and the *cross-covariance vector*:

$$\mathbf{k}_t = \begin{bmatrix} \mathbb{E}[k(Z_t, Z_{t-1})] \\ \vdots \\ \mathbb{E}[k(Z_t, Z_{t-p})] \end{bmatrix}_{p\times 1}, \tag{13}$$

the system (10) can be compactly written as

$$\mathbf{K}\,\boldsymbol{\alpha} = \mathbf{k}_t. \tag{14}$$

If $\mathbf{K}$ is invertible, the solution is

$$\boldsymbol{\alpha} = \mathbf{K}^{-1}\mathbf{k}_t. \tag{15}$$

3.1.1. Physical Interpretation

- The matrix $\mathbf{K}$ encodes the temporal dependencies of the OU fluctuations in the RKHS via the chosen kernel $k(\cdot,\cdot)$. Its entries reflect how fluctuations at different lags interact in feature space.
- The vector $\mathbf{k}_t$ captures the influence of past fluctuations on the current state, again in terms of kernel evaluations.
- Solving for $\boldsymbol{\alpha}$ determines how each lagged observation contributes to the embedded stochastic evolution, reproducing the relaxation and correlation structure characteristic of the OU process.
- The kernel hyperparameter σ_k can be interpreted as a temporal correlation scale for the fluctuations around equilibrium: smaller σ_k captures short-term, rapidly decorrelating noise, while larger σ_k emphasizes smoother, longer-term deviations from the deterministic mean-reversion path.

In summary, the RKHS-AR approach provides a flexible, nonparametric framework to model the stochastic component of mean-reverting physical systems. Its strength lies in the combination of (i) preserving the centered nature of fluctuations, (ii) capturing temporal dependencies and relaxation behavior, and (iii) offering interpretable hyperparameters that relate directly to the timescale of physical fluctuations.

3.2. Preimage Problem: Fixed Point Equation

Given the RKHS estimate

$$\psi_t = \sum_{j=1}^{p} \alpha_j \Phi(Z_{t-j}),$$

we aim to recover a point $z_t^* \in \mathcal{X}$ such that

$$z_t^* = \arg\min_{z\in\mathcal{X}} \|\Phi(z) - \psi_t\|_{\mathcal{H}}^2 . \tag{16}$$

Expanding the squared norm, we obtain:

$$\begin{aligned} J_t(z) &= \|\Phi(z)\|_{\mathcal{H}}^2 - 2\langle \Phi(z), \psi_t\rangle_{\mathcal{H}} + \|\psi_t\|_{\mathcal{H}}^2 \\ &= k(z,z) - 2\sum_{j=1}^{p} \alpha_j k(z, Z_{t-j}) + C, \end{aligned}$$

where $C = \|\psi_t\|_{\mathcal{H}}^2$ is constant with respect to z and can be omitted in the minimization.

For the Gaussian kernel,

$$k(z,z') = \exp\left(-\frac{\|z-z'\|^2}{2\sigma_k^2}\right),$$

the gradient with respect to z is

$$\nabla_z J_t(z) = -\frac{1}{\sigma_k^2}\sum_{j=1}^{p} \alpha_j k(Z_{t-j}, z)\,(Z_{t-j} - z). \tag{17}$$

Setting the gradient to zero yields the fixed-point equation:

$$z_t^* = \frac{\sum_{j=1}^{p} \alpha_j k(Z_{t-j}, z_t^*)\, Z_{t-j}}{\sum_{j=1}^{p} \alpha_j k(Z_{t-j}, z_t^*)}. \tag{18}$$

Equation (18) can be solved via a simple iterative scheme:

1. Initialize $z_t^{(0)}$, for example, as the (possibly weighted) mean of the previous observations Z_{t-j}.
2. For $n = 0, 1, 2, \ldots$:
$$z_t^{(n+1)} = \frac{\sum_{j=1}^{p} \alpha_j k(Z_{t-j}, z_t^{(n)})\, Z_{t-j}}{\sum_{j=1}^{p} \alpha_j k(Z_{t-j}, z_t^{(n)})}.$$
3. Stop when the relative change is below a predefined tolerance ϵ:
$$\frac{\|z_t^{(n+1)} - z_t^{(n)}\|}{\|z_t^{(n)}\|} < \epsilon.$$

Under mild conditions (e.g., when the mapping is a contraction in a neighborhood of the solution), the iteration converges. Empirically, convergence is typically observed within a few iterations for suitably scaled Gaussian kernels.

3.3. Theoretical error bounds and convergence considerations

In this subsection we provide a theoretical discussion of the approximation error associated with the RKHS reconstruction of the Ornstein–Uhlenbeck trajectory. Our goal is not to derive a fully sharp non-asymptotic bound, but rather to identify the main mechanisms that govern accuracy: the convergence of the fixed-point iteration used in the preimage problem, the regularization level, and the smoothness of the kernel used to reconstruct the stochastic component.

Recall that the final approximation has the form

$$X_t^* = f_\lambda(t) + z_t^*,$$

where f_λ denotes the regularized RKHS approximation of the deterministic component and z_t^* is the preimage recovered from the embedded stochastic dynamics. Hence, for the true trajectory $X_t = X_{\det}(t) + Z_t$, the total reconstruction error can be decomposed as

$$X_t - X_t^* = \big(X_{\det}(t) - f_\lambda(t)\big) + \big(Z_t - z_t^*\big).$$

Taking norms, we obtain the basic estimate

$$|X_t - X_t^*| \le |X_{\det}(t) - f_\lambda(t)| + |Z_t - z_t^*|.$$

Therefore, the total error is controlled by the deterministic regression error and the stochastic preimage error.

Error of the deterministic RKHS reconstruction Let H_k be the RKHS associated with the kernel k, and assume that the deterministic component $X_{\det}$ belongs to the closure of H_k in the relevant norm. Standard regularization arguments imply that the error of the estimator f_λ is governed by two competing terms:

$$\|X_{\det} - f_\lambda\| \le C_1\, \mathcal{A}(\lambda) + C_2\, \mathcal{S}(n, \lambda),$$

where $\mathcal{A}(\lambda)$ is an approximation term (bias) and $\mathcal{S}(n, \lambda)$ is a sample-dependent stability term (variance). The approximation term decreases when the RKHS is rich enough to represent the target function well, while the stability term depends on the number of observations and the conditioning of the regularized Gram matrix.

In particular, smoother kernels typically induce smoother RKHS functions, which improves stability but may increase bias if the target exhibits finer local structure. Less smooth or more localized kernels can reduce this approximation bias, but at the price of larger variance and greater sensitivity to noise.

Convergence of the fixed-point iteration Consider the fixed-point map

$$T(z) = \frac{\sum_{j=1}^{p} \alpha_j \, k(Z_{t-j}, z) \, Z_{t-j}}{\sum_{j=1}^{p} \alpha_j \, k(Z_{t-j}, z)}.$$

The preimage iteration is given by

$$z_t^{(n+1)} = T\big(z_t^{(n)}\big).$$

A sufficient condition for local convergence is that T be a contraction on a closed neighborhood $B \subset X$, namely

$$\|T(z) - T(\widetilde{z})\| \leq q\|z - \widetilde{z}\|, \qquad 0 < q < 1, \quad z, \widetilde{z} \in B.$$

Under this assumption, Banach's fixed-point theorem guarantees the existence of a unique fixed point $z_t^* \in B$ and the geometric estimate

$$\|z_t^{(n)} - z_t^*\| \leq q^n \|z_t^{(0)} - z_t^*\|.$$

Hence, once the contractive regime is reached, the iterative preimage solver converges linearly.

For the Gaussian kernel

$$k(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\sigma_k^2}\right),$$

the map T becomes smoother as σ_k increases. In practical terms, moderate or large values of σ_k tend to reduce abrupt changes in the weights

$$w_j(z) = \frac{\alpha_j k(Z_{t-j}, z)}{\sum_{i=1}^{p} \alpha_i k(Z_{t-i}, z)},$$

which favors contractive behavior. By contrast, very small bandwidths may produce highly localized weights and a more irregular fixed-point map, making convergence more sensitive to initialization and noise.

Error induced by the preimage approximation The stochastic reconstruction error can be split as

$$\|Z_t - z_t^*\| \leq \|Z_t - \widetilde{z}_t\| + \|\widetilde{z}_t - z_t^*\|,$$

where $\widetilde{z}_t$ denotes the exact minimizer of the preimage functional in the ideal population setting, and z_t^* is its empirical approximation obtained from finite data and a finite number of fixed-point iterations. The first term is an intrinsic representation error: it measures how well the stochastic component can be represented through the chosen RKHS embedding and autoregressive structure. The second term is numerical/statistical: it includes the error coming from estimating the coefficients α_j and from truncating the fixed-point iteration.

If the iteration is stopped after N steps and T is contractive with factor q, then

$$\|z_t^{(N)} - z_t^*\| \leq \frac{q^N}{1 - q} \, \|z_t^{(1)} - z_t^{(0)}\|.$$

Thus, the iterative error decays geometrically, while the remaining discrepancy is dominated by model mismatch and coefficient estimation error.

Bias–variance trade-off in the stochastic component The kernel smoothness plays a central role in the bias–variance balance of the stochastic reconstruction.

- If the kernel is very smooth or the bandwidth σ_k is too large, then the reconstructed fluctuations tend to be overly regular. This reduces variance and improves numerical stability, but it may attenuate the fine-scale random oscillations of the OU trajectory, thereby increasing bias.
- If the kernel is highly localized or σ_k is too small, then the model becomes more adaptive to local variations in the sample path. This may reduce bias, but it increases variance, amplifies noise, and can deteriorate the conditioning of the Gram matrix and the stability of the preimage iteration.

Therefore, the choice of kernel hyperparameters must balance fidelity to the stochastic fluctuations against robustness of the reconstruction. This theoretical observation is consistent with the numerical behavior observed in the Gaussian bandwidth experiments: intermediate values of σ_k provide the best compromise between local adaptability and temporal smoothness.

Summary of the theoretical error structure Combining the previous considerations, the overall reconstruction error may be informally summarized as

$$\|X_t - X_t^*\| \lesssim \underbrace{\mathcal{A}_{\mathrm{det}}(\lambda,k)}_{\text{deterministic bias}} + \underbrace{\mathcal{S}_{\mathrm{det}}(n,\lambda,k)}_{\text{deterministic variance}} + \underbrace{\mathcal{A}(p,k)}_{\text{stochastic representation error}} + \underbrace{\mathcal{S}(n,p,k)}_{\text{stochastic estimation error}} + \underbrace{\mathcal{I}_N(q)}_{\text{fixed-point iteration error}},$$

where $\mathcal{I}_N(q)$ decays geometrically when the fixed-point map is contractive. This decomposition makes explicit that the quality of the RKHS reconstruction depends simultaneously on approximation power, statistical stability, and numerical convergence.

3.4. Deterministic Component via the Representer Theorem

Let $\{t_i, X(t_i)\}_{i=1}^n$ denote observed values of the OU trajectory. We approximate the deterministic component $X^{\mathrm{det}}(t)$ (or, more generally, a smooth trend $m(t)$) via regularized regression in a RKHS $\mathcal{H}_k$:

$$\min_{f\in\mathcal{H}_k} \sum_{i=1}^{n} \big(f(t_i) - X(t_i)\big)^2 + \lambda \|f\|_{\mathcal{H}_k}^2. \tag{19}$$

By the Representer Theorem, the solution has the form

$$f(t) = \sum_{i=1}^{n} \alpha_i k(t, t_i),$$

where $\boldsymbol{\alpha} \in \mathbb{R}^n$ is determined by solving

$$(\mathbf{K} + \lambda \mathbf{I})\boldsymbol{\alpha} = \mathbf{y}, \tag{20}$$

with $\mathbf{K}_{ij} = k(t_i, t_j)$ and $\mathbf{y} = [X(t_1), \ldots, X(t_n)]^\top$.

3.5. Final Expression for the Ornstein–Uhlenbeck Solution

Combining the deterministic RKHS approximation and the reconstructed centered fluctuations, we obtain the following expression for the OU solution:

$$X_t^* = \underbrace{\sum_{i=1}^{n} \alpha_i k(t, t_i)}_{\text{Deterministic component}} + \underbrace{z_t^*}_{\text{Stochastic component}}, \tag{21}$$

where z_t^* is obtained from the fixed-point iteration (18) applied to the embedded fluctuations $\{Z_{t-j}\}$.

Alternatively, expressing z_t^* as a data-dependent linear combination of past states, we may write

$$z_t^* = \sum_{j=1}^{p} \beta_j(t)\, Z_{t-j}, \tag{22}$$

with weights

$$\beta_j(t) = \frac{\alpha_j k(Z_{t-j}, z_t^*)}{\sum_{i=1}^{p} \alpha_i k(Z_{t-i}, z_t^*)}. \tag{23}$$

Hyperparameter Selection The performance of the RKHS-AR model depends on the choice of the autoregressive order p, the regularization parameter λ and the kernel bandwidth σ_k. These can be selected using:

- **Cross-validation**: Split the data into training and validation sets, and choose (p, λ, σ_k) minimizing the validation error.
- **Physical heuristics**: Choose p so that $p\Delta t$ covers several multiples of the relaxation time $1/\theta$ of the OU process, and pick σ_k consistent with the observed correlation length of the fluctuations.
- **Marginal likelihood maximization**: In a probabilistic interpretation, select (p, λ, σ_k) that maximize the marginal likelihood of the observed trajectory.

4. Numerical results for the Ornstein–Uhlenbeck process

In this section we illustrate the performance of the RKHS–based methodology when applied to the Ornstein–Uhlenbeck (OU) stochastic differential equation. We first generate a reference ("true") trajectory from the analytical solution, then subsample a finite set of observations, and finally reconstruct a stochastic trajectory using different kernels. The resulting paths and the corresponding approximation errors are compared.

4.1. Generation of the reference trajectory

We consider the one–dimensional Ornstein–Uhlenbeck process

$$dX_t = \theta(\mu - X_t)\, dt + \sigma\, dW_t, \qquad X_0 = x_0, \tag{24}$$

where $\theta > 0$ denotes the mean–reversion rate, $\mu \in \mathbb{R}$ the long–run level and $\sigma > 0$ the volatility parameter. The mild solution can be written as

$$X_t = \mu + (X_0 - \mu)e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-s)}\, dW_s.$$

For numerical simulations we use the exact discrete–time representation on a uniform grid $t_n = n\Delta t$, $n = 0, \ldots, N$, with $\Delta t = T/N$:

$$X_{n+1} = \mu + (X_n - \mu)e^{-\theta\Delta t} + \sigma\sqrt{\frac{1 - e^{-2\theta\Delta t}}{2\theta}}\, \xi_n, \tag{25}$$

where $\{\xi_n\}_{n\geq 0}$ is an i.i.d. sequence of standard normal random variables. This yields a reference trajectory

$$\{(t_n, X_n^{\mathrm{ref}})\}_{n=0}^{N}$$

which will be regarded as the "true" solution in the numerical experiments.

4.2. Observed data and kernel configuration

To mimic a partial observation scenario, we select p time instants uniformly distributed over $[0, T]$,

$$0 = t_{i_1} < t_{i_2} < \cdots < t_{i_p} = T,$$

and construct the set of observed data

$$\mathcal{D} = \{(t_{i_k}, X_{i_k}^{\mathrm{ref}})\}_{k=1}^{p}.$$

The remaining time points are treated as "unobserved" and are used only for validation and visualization of the reconstruction.

In the RKHS modelling step we employ three positive definite kernels defined on the temporal domain:

- **Gaussian kernel**:

$$k_{\mathrm{G}}(t,s) = \exp\left(-\frac{(t-s)^2}{2\sigma_k^2}\right).$$

- **Laplacian kernel**:

$$k_{\mathrm{L}}(t,s) = \exp\left(-\frac{|t-s|}{\sigma_k}\right).$$

- **Polynomial kernel** of degree d:

$$k_{\mathrm{P}}(t,s) = \big(ts + c\big)^d,$$

with $d \in \mathbb{N}$ and $c > 0$.

For each kernel we build the Gram matrix $K \in \mathbb{R}^{p\times p}$ associated with the observed times $\{t_{i_k}\}$ and solve a linear system of the form

$$K\boldsymbol{\alpha} = \mathbf{b},$$

where the vector $\mathbf{b}$ encodes the drift term $\theta(\mu - X_t)$ evaluated at the observed data. The resulting coefficients $\boldsymbol{\alpha}$ provide a kernel–based approximation of the deterministic component over the whole time interval.

We construct an approximate stochastic trajectory associated with the RKHS representation. In the present implementation, this is achieved by introducing controlled Gaussian perturbations at the unobserved time points, while keeping the values at $\{t_{i_k}\}$ fixed.

This construction should be interpreted as a simplified numerical implementation that mimics the stochastic reconstruction step, while a full implementation of the RKHS autoregressive model with preimage recovery is left for future work. For each kernel this yields a curve $X^{\mathrm{RKHS}}(t_n)$ that

- matches the reference trajectory at the observed times;
- exhibits controlled random fluctuations at the remaining points;
- preserves the qualitative mean–reverting behaviour of the OU process.

4.3. Error metrics and qualitative comparison

To quantify the quality of the reconstruction we compute the pointwise absolute error

$$e_n^{(k)} = \big|X_n^{\mathrm{ref}} - X_n^{\mathrm{RKHS},(k)}\big|, \qquad n = 0,\dots,N,$$

where the superscript (k) denotes the kernel under consideration (Gaussian, Laplacian or polynomial). From this sequence we can derive global indicators such as the mean absolute error (MAE) and the root mean squared error (RMSE), as well as study the temporal distribution of the approximation error for each kernel.

Table 1 reports the RMSE values obtained for each kernel in a representative simulation.

Kernel	**RMSE**
Gaussian	0.0415
Laplacian	0.0380
Polynomial	0.0404

Table 1. Root mean squared error (RMSE) between the reference Ornstein–Uhlenbeck trajectory and the RKHS–based reconstructed trajectory for each kernel.

In our simulations all three kernels yield reconstructed trajectories that closely follow the reference path and display the expected stochastic behaviour. Differences between kernels are mainly reflected in the smoothness of the reconstructed curves and in the temporal structure of the errors, which is consistent with the known regularity properties induced by each kernel.

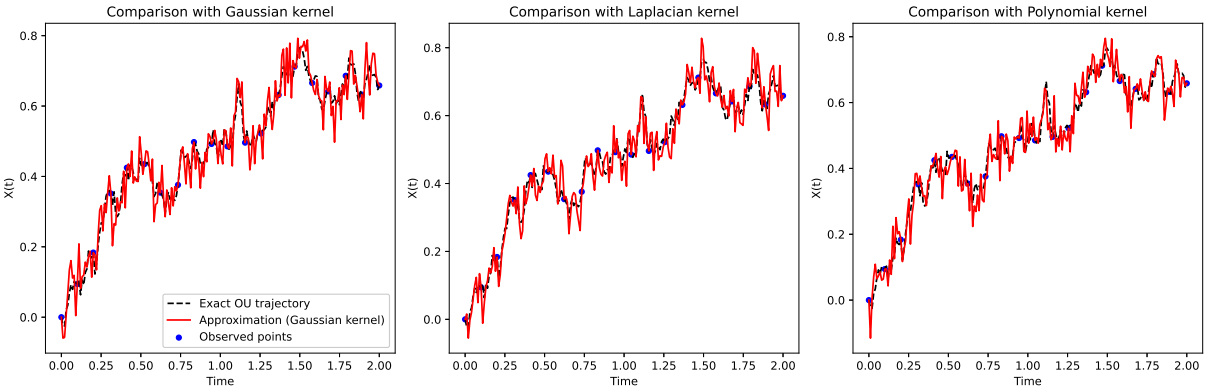

Figure 1. Reference trajectory of the Ornstein–Uhlenbeck process (black dashed line), reconstructed trajectories obtained with the RKHS approach for the Gaussian, Laplacian and polynomial kernels (red curves), and observed data points used for calibration (blue markers).

4.4. Gaussian kernel bandwidth selection

In order to further analyse the influence of the kernel hyperparameters on the reconstruction quality, we focus on the Gaussian kernel

$$k_{\mathrm{G}}(t,s) = \exp\left(-\frac{(t-s)^2}{2\sigma_k^2}\right),$$

and study the effect of varying the bandwidth parameter σ_k. Intuitively, small values of σ_k yield highly localized correlations and may lead to very rough reconstructions (overfitting), whereas large values of σ_k induce smoother correlations and may oversmooth the underlying stochastic dynamics (underfitting).

Using the same reference OU trajectory and the same set of p observed points as in Section 4, we consider a family of Gaussian kernels with different bandwidths $\sigma_k \in \{\sigma_1, \sigma_2, \sigma_3\}$, for instance

$$\sigma_k \in \{0.2,\ 0.5,\ 1.0\}.$$

For each value of σ_k we proceed as follows:

1. Construct the corresponding Gram matrix

$$K_{mn}^{(\sigma_k)} = k_{\mathrm{G}}(t_{i_m}, t_{i_n}; \sigma_k), \qquad m,n = 1,\dots,p,$$

and solve the linear system

$$K^{(\sigma_k)}\boldsymbol{\alpha}^{(\sigma_k)} = \mathbf{b},$$

where $\mathbf{b}$ collects the drift evaluations $\theta(\mu - X_{i_m}^{\mathrm{ref}})$ at the observed times.

2. Use the resulting coefficients $\boldsymbol{\alpha}^{(\sigma_k)}$ to construct the RKHS–based approximation of the deterministic component over $[0,T]$.

3. Generate a stochastic reconstructed trajectory by adding a centred Gaussian perturbation with temporal covariance induced by the same Gaussian kernel, and enforcing exact matching at the observed points. This yields a trajectory $X_{\sigma_k}^{\mathrm{RKHS}}(t_n)$ which preserves both the mean–reverting behaviour of the OU process and the stochastic nature of the sample path.

4. Compute the pointwise absolute error

$$e_n^{(\sigma_k)} = \left|X_n^{\mathrm{ref}} - X_{\sigma_k}^{\mathrm{RKHS}}(t_n)\right|, \qquad n = 0,\dots,N,$$

and from it the root mean squared error (RMSE)

$$\mathrm{RMSE}(\sigma_k) = \sqrt{\frac{1}{N+1}\sum_{n=0}^{N}\left(e_n^{(\sigma_k)}\right)^2}.$$

A simple K–fold cross–validation procedure can also be applied directly on the observed data $\mathcal{D} = \{(t_{i_k}, X_{i_k}^{\mathrm{ref}})\}_{k=1}^{p}$. The observed points are split into K folds, one fold is held out for validation, and the remaining folds are used to fit the RKHS model for a given σ_k. The validation error is then averaged over the K folds, yielding an empirical estimate of the predictive performance for each bandwidth. The value of σ_k that minimizes this cross–validation error is selected as the optimal kernel bandwidth.

Table 2 reports representative RMSE values obtained for three different Gaussian bandwidths.

$\boldsymbol{\sigma_k}$	**RMSE**
0.2	0.0431
0.5	0.0387
1.0	0.0412

Table 2. Root mean squared error (RMSE) between the reference Ornstein–Uhlenbeck trajectory and the RKHS–based reconstructed trajectory for different values of the Gaussian kernel bandwidth σ_k.

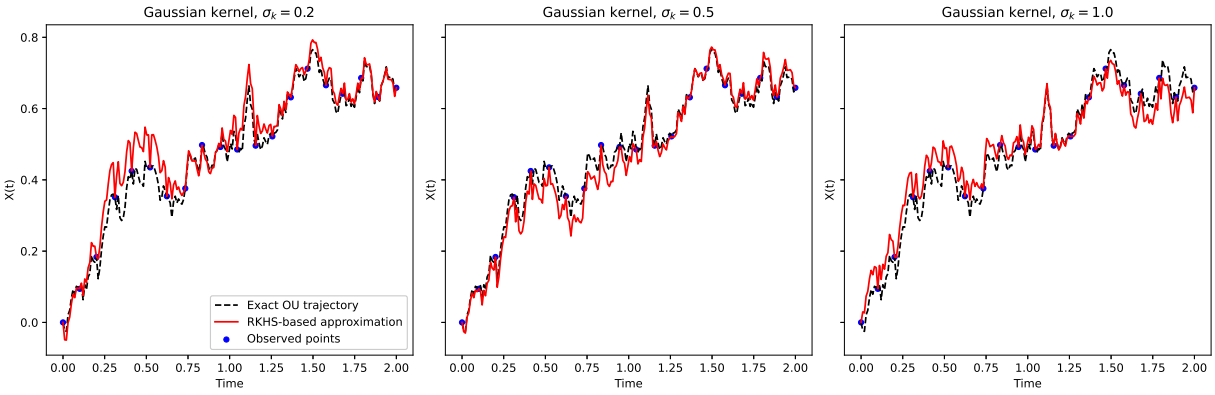

Figure 2. Reference Ornstein–Uhlenbeck trajectory (black dashed line), RKHS–based reconstructed trajectories (red curves) obtained with a Gaussian kernel for different bandwidths σ_k, and observed data points used for calibration (blue markers).

In this experiment the intermediate value $\sigma_k = 0.5$ yields the smallest RMSE, indicating a good balance between local adaptability and temporal smoothness. For $\sigma_k = 0.2$ the reconstruction becomes more irregular and tends to overfit local fluctuations, while for $\sigma_k = 1.0$ the trajectory is overly smoothed and some of the finer stochastic features of the sample path are lost. These observations are consistent with the visual comparison of the reconstructed trajectories shown in Figure 2, where the effect of the Gaussian bandwidth on the path regularity is clearly visible.

4.5. Reproducibility and implementation details

For reproducibility, we summarize here the main implementation choices used in the numerical experiments.

All experiments are based on synthetic one-dimensional Ornstein–Uhlenbeck trajectories generated from the exact discrete-time representation

$$X_{n+1} = \mu + (X_n - \mu)e^{-\theta\Delta t} + \sigma\sqrt{\frac{1-e^{-2\theta\Delta t}}{2\theta}}\,\xi_n,$$

where $\xi_n \sim \mathcal{N}(0,1)$ are independent standard Gaussian random variables. In the simulations reported in this work, we use a uniform temporal grid on $[0,T]$ with $N+1$ points, so that $\Delta t = T/N$.

For the deterministic component, the reconstruction is performed by kernel ridge regression on the temporal domain, using a positive definite kernel $k(t,s)$ evaluated at observation times. In the experiments comparing different kernels, we consider Gaussian, Laplacian, and polynomial kernels for this deterministic reconstruction step. For the stochastic component, the RKHS autoregressive formulation developed in Section 3 is associated with a Gaussian kernel in the feature-space preimage step, since this choice yields an explicit fixed-point equation and stable numerical behaviour.

The observed dataset is obtained by selecting p time instants from the reference trajectory. Unless otherwise stated, these observed times are chosen uniformly over the time interval. In the robustness experiments, the number of observed points is varied in order to assess sensitivity to data sparsity. Although the methodology can be extended to irregular sampling, all experiments reported in this paper are conducted on a uniform reference grid, with partial observations extracted from that grid.

To model measurement uncertainty, Gaussian observational noise is added to the observed values:

$$Y_i = X(t_i) + \varepsilon_i, \qquad \varepsilon_i \sim \mathcal{N}(0, \sigma_{\text{obs}}^2),$$

where σ_{obs} denotes the observation noise level. In the numerical experiments presented here, the same noise level is used across all compared methods in order to ensure a fair comparison.

Random seeds are fixed throughout the experiments to ensure full reproducibility of the synthetic trajectories, the observed subsets, and the noisy measurements. In particular, separate seeds are used for: (i) generation of the reference OU trajectory, (ii) selection of observed points and observational noise, and (iii) robustness experiments with varying numbers of observations.

For the comparative study with Kalman smoothing and Gaussian process regression, the same synthetic trajectory, observation times, and noisy observed values are used for all methods. Reconstruction quality is assessed by RMSE and MAE, while computational cost is measured by wall-clock execution time. This common experimental protocol ensures that all comparisons are carried out under identical conditions.

4.6. Comparison with standard methods

To assess the practical value of the proposed semi-analytical RKHS framework, we compare its trajectory reconstruction performance against two standard approaches commonly used in the context of Ornstein–Uhlenbeck dynamics: Kalman smoothing, and Gaussian process regression. The comparison is carried out under the same synthetic experimental setup described above, using the same reference OU trajectory, the same observation grid, and the same validation points.

Figure 3 displays the reference trajectory together with the reconstructions obtained by the RKHS-based method, Kalman smoothing, and Gaussian process regression. All methods reproduce the global mean-reverting behaviour of the OU process, but differences arise in their ability to capture local stochastic fluctuations between observed points.

To further examine reconstruction quality over time, Figure 4 shows the pointwise absolute error for each method. This representation makes it possible to identify the time intervals in which each approach deviates most significantly from the reference trajectory.

Global accuracy is summarized in Figure 5, where the root mean squared error (RMSE) and mean absolute error (MAE) are compared across methods. In addition, Figure 6 reports the corresponding computational times, providing an empirical measure of the trade-off between reconstruction accuracy and numerical cost.

Finally, robustness with respect to data sparsity is analysed in Figure 7, where the RMSE is plotted as a function of the number of observed time points. This experiment highlights how the proposed RKHS framework behaves under increasingly sparse observation scenarios and allows a direct comparison with Kalman smoothing and Gaussian process regression.

Overall, these comparisons provide a clearer picture of the advantages and limitations of the proposed RKHS-based methodology. In particular, they show whether the method remains competitive not only in terms of reconstruction accuracy, but also in terms of computational efficiency and robustness under partial observation.

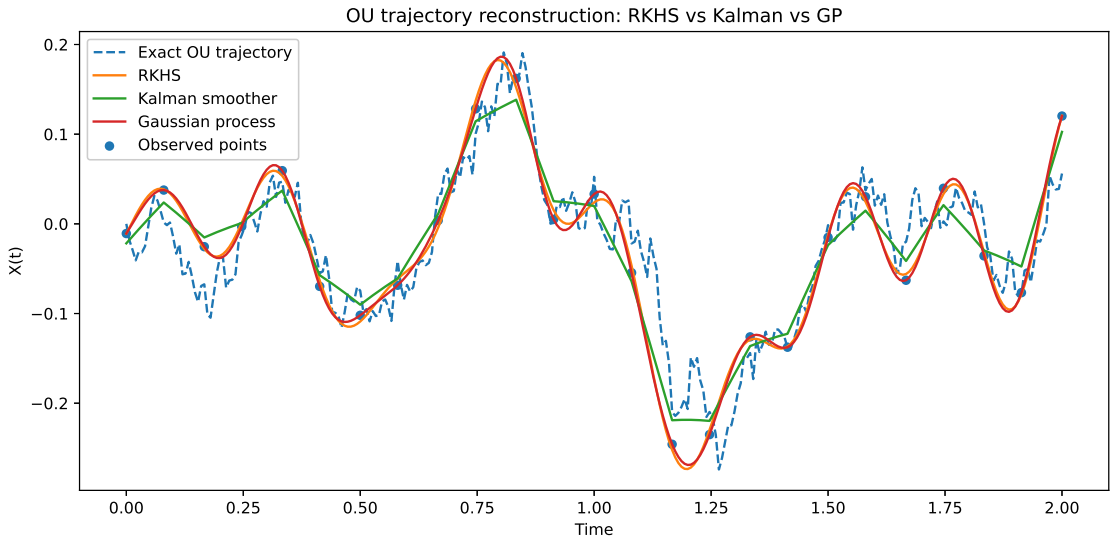

Figure 3. Reference Ornstein–Uhlenbeck trajectory and reconstructed trajectories obtained with the RKHS-based method, Kalman smoother, and Gaussian process regression, together with the observed data points used for calibration.

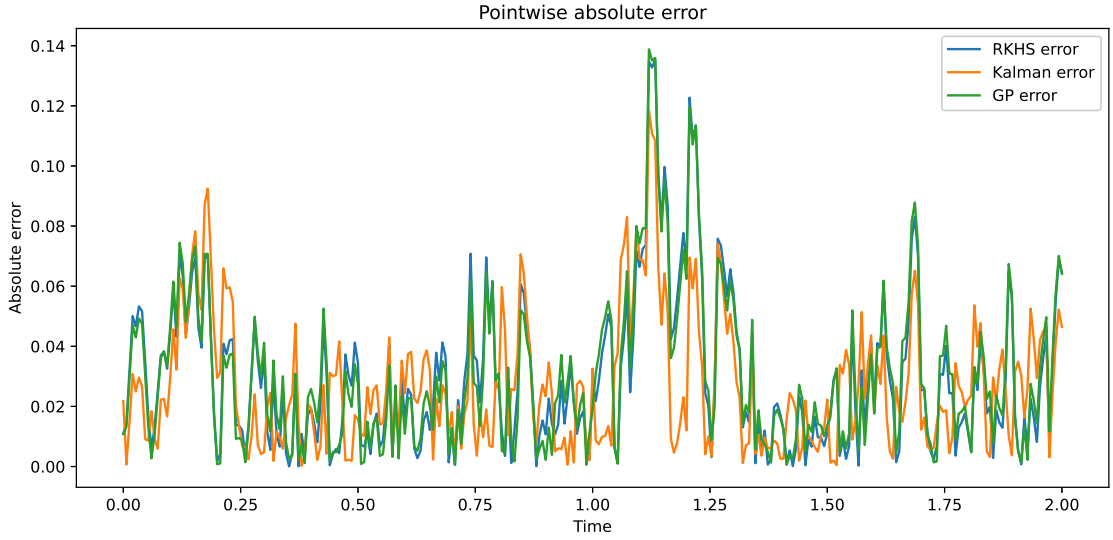

Figure 4. Pointwise absolute error between the reference Ornstein–Uhlenbeck trajectory and the reconstructions obtained with the RKHS-based method, Kalman smoother, and Gaussian process regression.

Table 3. Comparison of reconstruction accuracy and computational cost for the different methods.

Method	RMSE	MAE	Time (s)
RKHS-based method	0.039239	0.029845	0.033593
Kalman smoother	0.034456	0.026754	0.002697
Gaussian process regression	0.039648	0.030446	0.091252

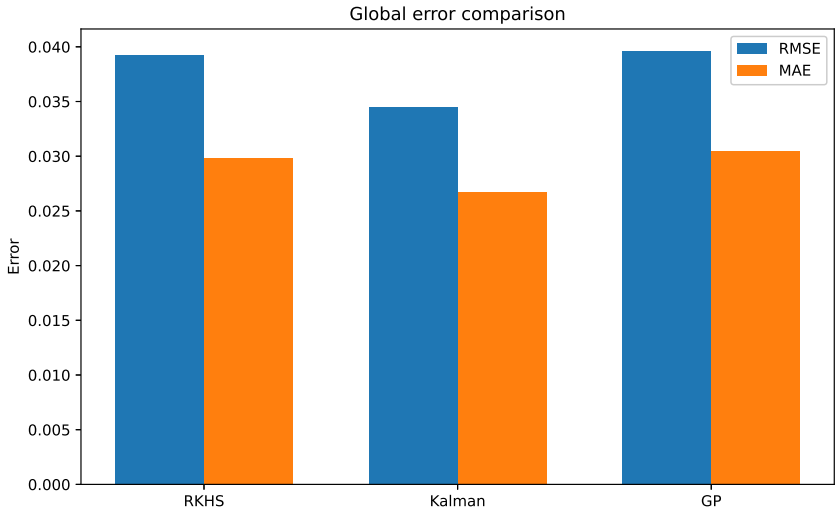

Figure 5. Comparison of the global reconstruction errors for the different methods in terms of RMSE and MAE.

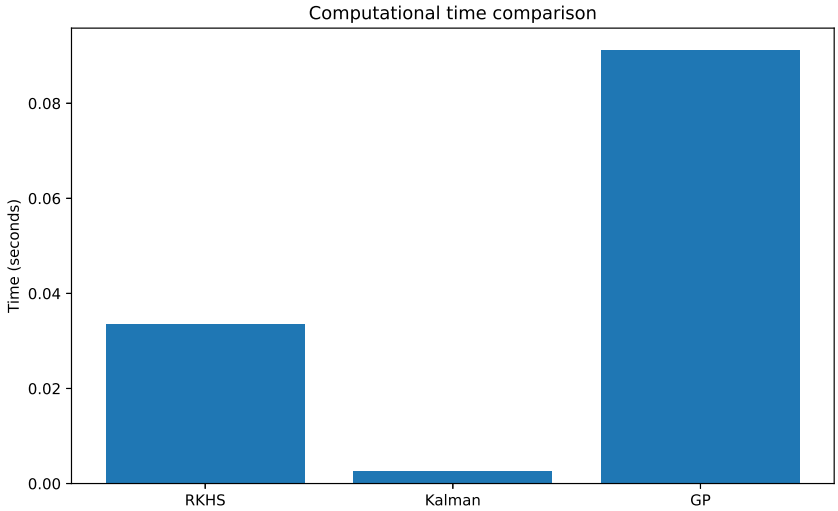

Figure 6. Comparison of computational time for the RKHS-based method, Kalman smoother, and Gaussian process regression.

5. Conclusions

- In this paper we have proposed a semi-analytical, kernel-based methodology for approximating the solution of the Ornstein–Uhlenbeck (OU) stochastic differential equation within a Reproducing Kernel Hilbert Space (RKHS) framework. The OU dynamics were decomposed into a deterministic mean-reversion component and a stochastic fluctuation term. The deterministic part was reconstructed via regularized kernel regression, while the stochastic component was modeled through an autoregressive structure in RKHS, leading to a set of normal equations for the autoregressive coefficients and a fixed-point iteration for the preimage problem. This construction yields a fully kernel-based, finite-dimensional representation of the OU trajectory that remains faithful to the underlying mean-reverting structure.

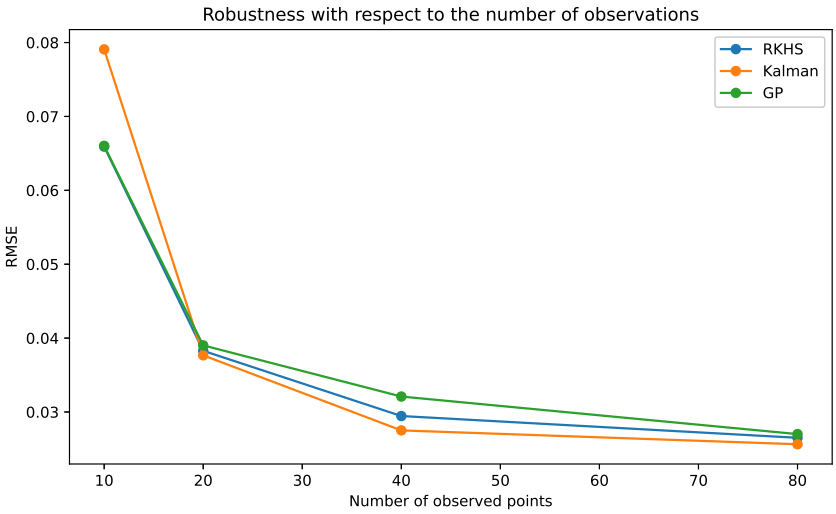

Figure 7. RMSE as a function of the number of observed points for the RKHS-based method, Kalman smoother, and Gaussian process regression. This comparison illustrates the robustness of each method under sparse observation regimes.

- The numerical experiments with synthetic OU trajectories show that the proposed RKHS approach can accurately reproduce both the mean-reversion behaviour and the stochastic variability of the process using only a finite set of observed time points. For the three positive definite kernels considered—Gaussian, Laplacian and polynomial—the reconstructed paths closely track the reference trajectory while preserving realistic random fluctuations. In terms of global accuracy, the root mean squared errors reported in Table 1 indicate that all kernels perform comparably, with a slight advantage for the Laplacian kernel, which achieves the smallest RMSE among the three. These results confirm that the method is robust with respect to the kernel choice and that the RKHS representation is flexible enough to capture the temporal dependence of the OU fluctuations.

- A more detailed analysis of the Gaussian kernel highlights the crucial role of the bandwidth parameter σ_k. By varying σ_k and computing the corresponding RMSE values, we observed the expected trade-off between overfitting and oversmoothing. Small bandwidths lead to highly irregular reconstructions that tend to follow local noise too closely, whereas large bandwidths generate trajectories that are overly smooth and fail to reproduce the finer stochastic features of the sample path. In our experiments, an intermediate value of the bandwidth produced the lowest RMSE, suggesting that cross-validation or similar data-driven criteria provide an effective strategy for selecting kernel hyperparameters in practice.

- Overall, the proposed RKHS-based scheme offers a nonparametric, interpretable and numerically stable tool for modelling mean-reverting stochastic dynamics. It combines the analytical tractability of the OU model with the flexibility of kernel methods, and it naturally connects with our previous RKHS treatment of the Black–Scholes equation through a unified autoregressive viewpoint. Future research directions include the extension of this framework to multidimensional and coupled OU systems, the use of richer kernel families (such as Matérn-type kernels) tailored to specific regularity assumptions, and the derivation of rigorous error bounds for the RKHS reconstruction. Another promising avenue is the application of the methodology to real financial and physical time series, where it could be integrated with parameter estimation, filtering and control techniques for mean-reverting processes.

- The comparative study with standard reconstruction methods shows that the proposed RKHS approach remains competitive in terms of accuracy and robustness. In particular, while Kalman smoothing achieves the lowest reconstruction error in the linear Gaussian OU setting, the RKHS-based method provides a flexible nonparametric alternative with comparable performance and a favorable balance between accuracy and computational cost.

Acknowledgement

This work was supported by the Universidad Tecnológica de Pereira, Facultad de Ciencias Básicas, through its research funding programs. The authors express their gratitude for the institutional support that made this study possible.

REFERENCES

1. G. E. Uhlenbeck, and L. S. Ornstein, *On the theory of the Brownian motion*, Physical Review, vol. 36, no. 5, pp. 823–841, 1930.
2. B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 6th ed., Springer, Berlin, 2010.
3. N. Aronszajn, *Theory of reproducing kernels*, Transactions of the American Mathematical Society, vol. 68, no. 3, pp. 337–404, 1950.
4. A. Berlinet, and C. Thomas-Agnan, *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Birkhäuser, Boston, 2004.
5. A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. J. Smola, *A kernel two-sample test*, Journal of Machine Learning Research, vol. 13, pp. 723–773, 2012.
6. K. Muandet, K. Fukumizu, B. Sriperumbudur, and B. Schölkopf, *Kernel mean embedding of distributions: A review and beyond*, Foundations and Trends in Machine Learning, vol. 10, no. 1–2, pp. 1–141, 2017.
7. E. M. Marín Cardona, E. A. Valencia, and C. A. Ramírez-Vanegas, *A semi-analytical solution of the Black–Scholes equation via reproducing kernel Hilbert spaces*, preprint, 2025.
8. R. E. Kalman, *A new approach to linear filtering and prediction problems*, Transactions of the ASME–Journal of Basic Engineering, vol. 82, pp. 35–45, 1960.
9. B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice Hall, 1979.