Journal of Time Series Analysis

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Functional Vašiček Model

Piotr Kokoszka[1] | Neda Mohammadi[1] | Leonardo Santoro[2] | Hong Miao[3] | Haonan Wang[1]

[1]Department of Statistics, Colorado State University, Fort Collins, Colorado, USA | [2]Institute of Mathematics, EPFL, Lausanne, Switzerland | [3]Department of Finance and Real Estate, Colorado State University, Fort Collins, Colorado, USA

Correspondence: Piotr Kokoszka (piotr.kokoszka@colostate.edu)

Received: 23 December 2024 | **Revised:** 16 September 2025 | **Accepted:** 19 September 2025

Funding: This work was supported by the National Science Foundation (Grant No. DMS-2412408).

Keywords: functional data | intraday rates | Itô diffusion | Ornstein–Uhlenbeck process

ABSTRACT

We propose a new formulation of the Vašičekmodel within the framework of functional data analysis. We treat observations (continuous-time rates) within a suitably defined trading day as a single statistical object. We then consider a sequence of such objects, indexed by day. In addition to the common long-term rate, the objects are parametrized by two functional parameters, the volatility curve and the reversion curve, which replace analogous scalar parameters in the classical Vašičekmodel. Such a modeling paradigm allows us to estimate instantaneous reversion and volatility parameters within a trading day, thus allowing them to evolve with the time of day. The model is estimated within a new framework that combines techniques of functional data analysis with those of SDEs. In particular, large sample properties are derived as the number of days and the number of discrete time points at which the rate curves are observed tend to infinity.

MSC2020 Classification: 60H10, 60H30, 62M15, 62P05

1 | Introduction

The classical model of Vašiček (1977) for continuously observed interest rates $\{R(t), t \geq 0\}$ is given by the stochastic differential equation

$$dR(t) = \theta(\mu - R(t))dt + \sigma dW(t) \tag{1}$$

where the parameters θ and σ are positive constants, the *speed of reversion* and the *instantaneous volatility*, respectively, μ is a real constant, the *long-term mean level*, and $W(\cdot)$ is a Wiener process. The parameter μ is positive in practice and has the interpretation of long-term, or equilibrium, interest rate for the period under consideration. Theory behind the Vašiček model is explained in many textbooks, see, for example, Filipovic (2009), and the model, with it various extensions, has attracted a lot of attention in applied finance as well as theoretical research in continuous-time finance and stochastic processes, we review some most recent contributions later in this section.

The purpose of this paper is to introduce a new modeling paradigm and the requisite statistical theory that are suitable for fixed-term securities traded chiefly during the opening hours of an exchange. For example, while U.S. Treasury securities can be traded electronically at any time, the largest volume corresponds to U.S. business hours. The model we propose considers rates during those hours as daily curves and proposes the estimation of the continuous time term structure based on these curves. Model (2) does not account for a fundamentally different stochastic behavior at night or other periods of low trading volume. We therefore propose the model

$$dR_i(t) = \theta(t)\big(\mu - R_i(t)\big)dt + \sigma(t)dW_i(t), \quad i = 1, 2, \ldots, N \tag{2}$$

This article is part of the special issue "Peter J. Brockwell".

in which i refers to a trading day, defined by hours of high trading volume, and t to intraday time. We also extend model (2) by considering the speed of reversion and the instantaneous volatility as functions that can evolve with the time of the day, denoted t. Mathematically, the index i could refer to contiguous 24 h-long time intervals, but in the application we consider, t is in the interval 8:00–20:00 EST. Outside this interval, using an SDE to model the rates seems questionable due to the sparsity of the trades.

Before discussing our model further, it is useful to return to the standard model (2). For mathematical consistency, the initial value, $R(0)$, is assumed to be a random variable independent of the sigma algebra $\mathcal{F}_\infty = \sigma\{W(t), t \in [0,\infty)\}$ generated by the Wiener process. Applying Itô formula, see Øksendal (2003), Chapter 4, to the function $f(x,t) = e^{\theta t}x$, $t \in (0,\infty)$, $x \in \mathbb{R}$, we obtain the following closed-form solution to (2):

$$R(t) = e^{-\theta t}R(0) + \mu\left(1 - e^{-\theta t}\right) + \sigma\int_0^t e^{-\theta(t-u)}dW(u), \quad t \in [0,\infty)$$

Denote by $m(\cdot) = \mathbb{E}[R(\cdot)]$ and $D(\cdot) = \mathrm{Var}[R(\cdot)]$ the mean and the variance functions, respectively. The closed-form solution implies

$$\begin{aligned} m(t) &= e^{-\theta t}m(0) + \mu\left(1 - e^{-\theta t}\right), \\ D(t) &= e^{-2\theta t}D(0) + \sigma^2\int_0^t e^{-2\theta(t-u)}du \\ &= e^{-2\theta t}D(0) + \frac{\sigma^2}{2\theta}\left(1 - e^{-2\theta t}\right), \quad t \in [0,\infty) \end{aligned}$$

so

$$\lim_{t\to\infty} m(t) = \mu, \quad \lim_{t\to\infty} D(t) = \frac{\sigma^2}{2\theta}$$

The above, and similar, relations motivate a great deal of asymptotic estimation theory for the Vašičekmodel and its extensions. It is assumed that continuous-time observations are available on the interval $[0,T]$, and asymptotics as $T \to \infty$ are considered. In our setting, the asymptotic analysis of estimators is performed as N, the count of trading days, tends to infinity. Additionally, it accounts for the empirical fact that most available data sets consist of rates at discrete times within a day. These are often data products derived from more comprehensive data repositories. We thus introduce the parameter K, which is the count of observations within each trading day. Our asymptotic analysis assumes that $N \to \infty$ and $K \to \infty$. It is thus related to a general setting of Functional Data Analysis (FDA), for example, Horváth and Kokoszka (2012) or Hsing and Eubank (2015), where curves are treated as indivisible statistical objects whose number increases. A similar perspective is taken by Liu et al. (2016) who consider iid Ornstein–Uhlenbeck processes (solutions to (2) on [0,1]) as errors in their model. Our inference procedure establishes a connection between the unknown SDE objects μ, $\sigma(\cdot)$, and $\theta(\cdot)$ and nonparametric first- and second-order functions, $m(\cdot)$ and $D(\cdot)$. This approach offers two key advantages. Firstly, it enables nonparametric inference procedures without relying on within-curve assumptions like local stationarity, providing flexibility in capturing the underlying dynamics. Secondly, it allows for inference that does not rely on an expanding window, $[0,T]$, which is unrealistic if the structure of the data is not stationary in some sense.

We now list a few recent contributions related to the Vašiček model, without aiming at completeness, due to a very large volume of related publications. We begin with contributions with a theoretical emphasis. Jiang et al. (2023) study deviation inequalities for quadratic functionals in the explosive Vašičekmodel. They obtain the self-normalized Cramér-type moderate deviations and joint moderate deviations for the maximum likelihood estimators via asymptotic analysis techniques. Xiao et al. (2014) propose a fractional Vašičekmodel to describe the dynamics of the short rate in the pricing environment of equity warrants. Xiao and Yu (2019) investigate a fractional Vašičekmodel with long-range dependence assumed to be driven by a fractional Brownian motion with the Hurst parameter greater than or equal to one-half. They develop an asymptotic theory for estimators of two parameters in the drift function when a continuous record of observations is available. Nourdin and Diu Tran (2019) propose the Vašiček model driven by a Hermite process and study its estimation, also in continuous time. Estimation of nonnegative Lévy-driven Ornstein–Uhlenbeck process is studied in Brockwell et al. (2007). Properties of such processes are studied in depth by Lindner and Sato (2009). More general ARMA-type continuous-time models driven by the Lévy motion are considered by Brockwell and Lindner (2009) and Brockwell et al. (2011). Turning to applied research, Cai et al. (2022) study pricing of American put options in a market with a stochastic interest rate and finite-time maturity. They present a numerical study of the option price and the optimal exercise boundary for the Vašičekmodel. Chen et al. (2020) study an ambiguity-averse investor with a Cramer–Lundberg surplus to be allocated into a mean-reverting asset representing a commodity and a bond following the Vašičekmodel. Turning to papers with more applied angle, Orlando et al. (2020) propose a new methodology for forecasting the future expected interest rate through the Vašičekand Cox–Ingersoll–Ross (CIR) models based on rolling windows from observed financial market data. Kharrat and Arfaoui (2023) solve the time-fractional Vašičekmodel for European options. Nowman (2010) estimate the Generalized Vašičekterm structure model using the United Kingdom and the Euro panel data. Wu et al. (2019) develop a framework for assessing the optimal refinancing strategy in continuous time when the interest rate is stochastic and follows a Vašičekmodel. Rodrigo and Mamon (2014) suggest a new method to calibrate the CIR interest rate models from bond prices by defining an appropriate generating function and derive recursive relations between the derivatives of the generating function and the bond prices. The parameters of the Vašičekand CIR models are then obtained by solving a system of linearly independent equations arising from the recursive relations. Mehrdoust and Najafi (2020) consider a fractional version of the Vašičekmodel, where the noise part of the model is modeled as fractional Brownian motion and apply the model to price a European option on the zero-coupon bond.

There has also been recent interest in inference for functional replications akin to (2) outside the context of the Vašičekmodel. Comte and Genon-Catalot (2020) consider the model given by the equations $dX_i(t) = b(X_i(t)) + \sigma(X_i(t))dW_i(t)$, and focus on the estimation of the drift function $b(\cdot)$. Mohammadi et al. (2024) consider the equations $dX_i(t) = \mu(t)X_i(t)^\alpha + \sigma(t)X_i(t)^\beta dW_i(t)$, and estimate the functions $\mu(\cdot)$ and $\sigma(\cdot)$ in a setting where sparse, error-contaminated observations are available. Zhou and

Müller (2025) consider the most general formulation given by (3) in Section 2 and propose an algorithm for simulating sample paths based on estimation of the conditional mean and conditional variance.

The remainder of the paper is organized as follows. For ease of reference, we collect in Section 2 the requisite mathematical results related to the SDEs we study. Our model is developed in Section 3. Its estimation together with the large-sample theory, is presented in Section 4. An application to U.S. treasuries and an analysis of finite sample properties of the estimators are presented in Section 5. The proofs of all theoretical results are collected in Section 6.

2 | Preliminaries

Consider the $\mathbb{R}$-valued Itô diffusion (diffusion in short) process

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t), \qquad t > 0 \tag{3}$$

where W denotes a standard Brownian motion. We assume that the initial distribution, $X(0)$, is independent of the σ-algebra $\mathcal{F}_\infty$ generated by $\{W(t)\}_{t\geq 0}$. Integrals with respect to W should be understood in the sense of the stochastic Itô integral. The functions $\mu(\cdot,\cdot)$ (drift or viscosity), and $\sigma(\cdot,\cdot)$, (diffusion or volatility), are Borel measurable. The following theorem, a consequence of Theorem 5.2.1 in Øksendal (2003), provides sufficient conditions for the existence and uniqueness of the process X satisfying the stochastic differential Equation (3).

Theorem 2.1 (Existence and Uniqueness). *Let T be a positive number and the functions $\mu(\cdot,\cdot) : [0,T]\times\mathbb{R} \mapsto \mathbb{R}$ and $\sigma(\cdot,\cdot) : [0,T]\times\mathbb{R} \mapsto \mathbb{R}$ be measurable functions satisfying the linear growth condition*

$$|\mu(t,x)| + |\sigma(t,x)| \leq L(1+|x|), \quad x\in\mathbb{R},\ t\in[0,T] \tag{4}$$

and the Lipschitz continuity in the space variable, that is,

$$|\mu(t,x)| - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq L|x-y|, \quad x,y\in\mathbb{R},\ t\in[0,T] \tag{5}$$

for some constant $L > 0$. Let moreover $X(0)$ be a random variable independent of the σ-algebra $\mathcal{F}_\infty$ generated by $\{W(t)\}_{t\geq 0}$ and such that $\mathbb{E}[X(0)^2] < \infty$. Then the stochastic differential Equation (3) admits a unique solution with time-continuous trajectories and adapted to the filtration $\{\mathcal{F}_t^{X_0}\}$ generated by $X(0)$ and the standard Brownian motion $\{W(s)\}_{s\leq t}$. Moreover,

$$\mathbb{E}\left[\int_0^T |X(t)|^2 dt\right] < \infty$$

In the sequel, we employ a corollary of Dambis–Dubins–Schwarz theorem which expresses *any* continuous local martingale as a time change of a Brownian motion, see, for example, Section 5.3.2 in Le Gall (2016).

Corollary 2.1. *Assume the setting of Theorem 2.1 and set the drift function equal to zero, $\mu(\cdot,\cdot) = 0$. Then,*

$$\left\{\int_0^t \sigma(u, X(u))dW(u),\ t\in[0,1]\right\} \stackrel{\mathcal{L}aw}{=} \left\{W\left(\int_0^t \sigma^2(u,X(u))du\right),\ t\in[0,1]\right\}$$

where the equality in distribution is in the space $C([0,1])$ of continuous functions equipped with the supremum norm.

3 | Vašiček Model for Daily Curves and Its Estimation

Leading up to the subsequent model development, consider the following extension of the Vašiček model (2):

$$dR(t) = \theta(t)(\mu - R(t))dt + \sigma(t)dW(t), \quad t\in[0,\infty) \tag{6}$$

where $\theta(\cdot)$ and $\sigma(\cdot)$ are positive functions that we call the *instantaneous speed of reversion* and *instantaneous volatility*, respectively, while μ remains constant with the same interpretation as the long-term mean level. We impose the following assumptions to ensure existence and uniqueness of a solution to the stochastic differential Equation (6), see Theorem 3.1 below.

Assumption 3.1. The functions θ and σ are continuous and bounded on $[0,\infty)$.

Assumption 3.2. The random variable $R(0)$ is independent of $W(\cdot)$ and square integrable, that is, $\mathbb{E}|R(0)|^2 < \infty$.

Theorem 3.1. *Suppose Assumptions 3.1 and 3.2 hold. Then the stochastic differential Equation (6) admits a unique solution in the form*

$$R(t) = \exp\left\{-\int_0^t \theta(u)du\right\}R(0) + \mu\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\theta(s)ds + \int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\sigma(s)dW(s), \quad \times t\in[0,\infty)$$

Moreover, the mean function $m(\cdot) = \mathbb{E}[R(\cdot)]$ and variance function $D(\cdot) = \mathrm{Var}[R(\cdot)]$ are given by

$$m(t) = \exp\left\{-\int_0^t \theta(u)du\right\}m(0) + \mu\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\theta(s)ds, \quad t\in[0,\infty) \tag{8}$$

$$D(t) = \exp\left\{-2\int_0^t \theta(u)du\right\}D(0) + \int_0^t \exp\left\{-2\int_s^t \theta(u)du\right\}\sigma^2(s)ds, \quad t\in[0,\infty) \tag{9}$$

which satisfy the ordinary differential equations

$$\partial m(t) = -\theta(t)m(t) + \mu\theta(t), \quad t\in[0,\infty) \tag{10}$$

$$\partial D(t) = -2\theta(t)D(t) + \sigma^2(t), \quad t\in[0,\infty) \tag{11}$$

Now, assume we observe N i.i.d. replications $R_i(\cdot)$, $i = 1,2,\ldots,N$, satisfying (6) or equivalently (7), limited to the unit time interval, $t \in [0,1]$, that is,

$$dR_i(t) = \theta(t)\big(\mu - R_i(t)\big)dt + \sigma(t)dW_i(t)$$
$$t \in [0,1], \quad i = 1, 2, \ldots, N \tag{12}$$

For each i, the initial error $R_i(0) - \mu$ is independent from $W_i(\cdot)$. The pairs $(R_i(\cdot), W_i(\cdot))$ are independent across i. The parameters $(\mu, \theta(\cdot), \sigma(\cdot))$ quantify the shared random structure of all curves. In particular, μ determines their long-term level. We propose an inference procedure for the parameters $(\mu, \theta(\cdot), \sigma(\cdot))$ based on N replications of the curves $R_i(\cdot)$ observed at discrete time points $t_k = \Delta k$, $k = 0, 1, \ldots, K$, with $\Delta = \frac{1}{K}$. We thus observe the $N \times (K+1)$ matrix

$$\left[R_i\left(\frac{k}{K}\right),\ 1 \le i \le N,\ 0 \le k \le K\right] \tag{13}$$

With continuous observations, the quadratic variation process is observable, and the diffusion function is the derivative of the quadratic variation process, see, for example, Cialenco (2018), p. 311, for a more detailed discussion. An entirely different approach to the estimation of model (12) is required. It uses the availability of the discretely observed replications indexed by i rather that the limiting behavior as $t \to \infty$ and a continuously observed trajectory.

Recall the ordinary differential Equations (10) and (11) where $m(\cdot)$ and $D(\cdot)$ denote, respectively, the common mean and variance functions of the random curves $R_i(\cdot)$. Setting

$$Q(t) = \int_0^t \sigma^2(u)du, \qquad t \in [0,1] \tag{14}$$

Equation (11) entails

$$\sigma^2(t) = \partial Q(t), \qquad t \in [0,1] \tag{15}$$

$$\theta(t) = \frac{\sigma^2(t) - \partial D(t)}{2D(t)}, \qquad t \in [0,1] \tag{16}$$

Note that Equation (10) implies

$$\mu = \frac{\partial m(t) + \theta(t)m(t)}{\theta(t)}, \qquad t \in [0,1] \tag{17}$$

so that, in particular

$$\mu = \text{median}\left(\frac{\partial m(t) + \theta(t)m(t)}{\theta(t)}\ :\ t \in [0,1]\right) \tag{18}$$

Equations (15–17) reduce inference for $\sigma(\cdot)$, $\theta(\cdot)$ and μ to inference for $\partial Q(\cdot)$, $D(\cdot)$, $\partial D(\cdot)$, $m(\cdot)$ and $\partial m(\cdot)$. The latter functions are more directly accessible from the observations (13) than the model parameters in Equation (12).

For each i, define the step function

$$\tilde{Q}_i(t) = \sum_{k=1}^{K} |R_i(t_k) - R_i(t_{k-1})|^2 \mathbb{1}\{t_k \le t\}, \quad t \in [0,1] \tag{19}$$

This function is called the realized or empirical quadratic variation process. We now propose a nonparametric approach to estimate the derivative of the quadratic variation process $\partial Q(t)$ based on local linear estimation, see Fan and Gijbels (1996) for a general exposition and Mohammadi et al. (2024) for a recent study in the context of SDEs. We first define

$$\overline{Q}(t) = \frac{1}{N}\sum_{i=1}^{N} \tilde{Q}_i(t), \quad t \in [0,1] \tag{20}$$

We then apply a local linear smoother on $\overline{Q}(t_k)$, $k = 0, 1, \ldots, K$, that is, we compute

$$\left(\hat{\gamma}_0(t), \hat{\gamma}_1(t)\right)^\top = \underset{\{\gamma_0,\gamma_1\}}{\text{argmin}} \sum_{k=0}^{K}\left\{\overline{Q}(t_k) - \gamma_0 - \gamma_1 h_Q^{-1}(t_k - t)\right\}^2 \times h_Q^{-1} K_Q\left(h_Q^{-1}(t_k - t)\right) \tag{21}$$

where h_Q is a bandwidth parameter depending on N and Δ, and $K_Q(\cdot)$ an integrable kernel function. This leads to the estimators

$$\left(\hat{Q}(t), \widehat{\partial Q}(t)\right) = \left(\hat{\gamma}_0(t), h_Q^{-1}\hat{\gamma}_1(t)\right), \quad t \in [0,1] \tag{22}$$

We need only the second component, that is, the estimator of the derivative, $\widehat{\partial Q}(t), t \in [0,1]$.

Remark 3.1. Equation (22) gives $\hat{Q}(\cdot)$ as an estimate for the quadratic variation process, different from the simple average

$$\overline{Q}(t) = \frac{1}{N}\sum_{i=1}^{N} \tilde{Q}_i(t), \quad t \in [0,1]$$

Even though we do not use $\hat{Q}(\cdot)$, such a smooth estimator may be useful in other contexts.

We proceed in the same way to estimate the mean function $m(\cdot)$ and its derivative $\partial m(\cdot)$. For each $t \in [0,1]$, define

$$\left(\hat{m}(t), \widehat{\partial m}(t)\right) = \left(\hat{\alpha}_0(t), h_m^{-1}\hat{\alpha}_1(t)\right), \quad t \in [0,1] \tag{23}$$

where

$$\left(\hat{\alpha}_0(t), \hat{\alpha}_1(t)\right)^\top = \underset{\{\alpha_0,\alpha_1\}}{\text{argmin}} \sum_{k=0}^{K}\left\{\overline{R}(t_k) - \alpha_0 - \alpha_1 h_m^{-1}(t_k - t)\right\}^2 \times h_m^{-1} K_m\left(h_m^{-1}(t_k - t)\right) \tag{24}$$

with $\overline{R}(t_k) = \frac{1}{N}\sum_{i=1}^{N} R_i(t_k)$, h_m a bandwidth depending on N and Δ, and $K_m(\cdot)$ an integrable kernel function.

It remains to estimate the variance function $D(\cdot)$ and its derivative $\partial D(\cdot)$. For this purpose, we first estimate the second-moment function $G(t) = \mathbb{E}[R_i^2(t)]$ and its derivative $\partial G(t) = \frac{\partial G}{\partial t}$. Define

$$\left(\hat{G}(t), \widehat{\partial G}(t)\right) = \left(\hat{\beta}_0(t), h_G^{-1}\hat{\beta}_1(t)\right), \quad t \in [0,1] \tag{25}$$

where

$$\left(\hat{\beta}_0(t), \hat{\beta}_1(t)\right)^\top = \underset{\{\beta_0,\beta_1\}}{\text{argmin}} \sum_{k=0}^{K}\left\{\overline{R^2}(t_k) - \beta_0 - \beta_1 h_G^{-1}(t_k - t)\right\}^2 \times h_G^{-1} K_G\left(h_G^{-1}(t_k - t)\right) \tag{26}$$

with $\overline{R^2}(t_k) = \frac{1}{N}\sum_{i=1}^{N} R_i^2(t_k)$, h_G a bandwidth parameter depending on N and Δ, and $K_G(\cdot)$ an integrable kernel function. We set

$$\left(\widehat{D}(t), \widehat{\partial D}(t)\right) = \left(\widehat{G}(t) - \widehat{m}^2(t), \widehat{\partial G}(t) - 2\widehat{m}(t)\widehat{\partial m}(t)\right) \quad t \in [0,1] \tag{27}$$

Using (22) we define

$$\widehat{\sigma}^2(t) = \widehat{\partial Q}(t), \quad t \in [0,1] \tag{28}$$

Combining (28) with (23) and (27) and the system of Equations (15), (16), and (18), we propose the estimators

$$\widehat{\theta}(t) = \frac{\widehat{\sigma}^2(t) - \widehat{\partial D}(t)}{2\widehat{D}(t)}, \quad t \in [0,1] \tag{29}$$

$$\widehat{\mu} = \text{median}\left(\frac{\widehat{\partial m}(t) + \widehat{\theta}(t)\widehat{m}(t)}{\widehat{\theta}(t)} \;:\; t \in [0,1]\right) \tag{30}$$

In Section 4, we establish a number of asymptotic results that lead to consistency rates for the estimators $\widehat{\sigma}^2(\cdot)$, $\widehat{\theta}(\cdot)$ and $\widehat{\mu}$, respectively in Equations (28), (29), and (30). In Section 5, we examine their finite sample properties.

Remark 3.2. The estimates $\widehat{\theta}$ and $\widehat{\sigma}^2$ are not guaranteed to be strictly positive. To ensure positivity, one could use the maximum of the proposed estimate and a small constant (e.g., 0.001). However, this modification does not affect the rate of convergence or any of the theoretical results studied in the next section. In practice, we have not encountered this issue.

We conclude this section with a useful observation that has an important bearing on both the theory and implementation of the above estimators. As an alternative to (21, 22), we can propose the following optimization

$$\left(\check{\eta}_0(t), \check{\eta}_1(t)\right)^\top = \underset{\{\eta_0,\eta_1\}}{\text{argmin}} \sum_{i=1}^{N}\sum_{k=0}^{K}\left\{\widetilde{Q}_i(t_k) - \eta_0 - \eta_1 h_Q^{-1}(t_k - t)\right\}^2 \times h_Q^{-1} K_Q\left(h_Q^{-1}(t_k - t)\right) \tag{31}$$

with all objects the same as in Equation (21). This leads to the estimator

$$\left(\check{Q}(t), \partial\check{Q}(t)\right) = \left(\check{\eta}_0(t), h_Q^{-1}\check{\eta}_1(t)\right), \quad t \in [0,1] \tag{32}$$

Proposition 3.1 below shows that estimators (22) and (32) are equal. However, the computational cost of (32) is evidently much higher, as it requires to find the minimizer of a sum of $N \cdot (K+1)$ terms, whereas the optimization problem (21) requires the simple computation of an average of N terms and the minimization of a sum of $K+1$ terms. It is clear that (22) is more convenient to implement in practice. On the other hand, (32) is mathematically more convenient to work with as it allows the summation of independent random variables over the index i, enabling the application of concentration inequalities. See proof of Proposition (3.1) and Theorem 4.3 for more details. Obviously, one may propose such alternatives to (23–24) and (25–26).

Proposition 3.1. *The optimization problems (21, 22) and (31, 32) admit the same solutions, that is,*

$$\left(\widehat{Q}(t), \widehat{\partial Q}(t)\right) = \left(\check{Q}(t), \partial\check{Q}(t)\right), \quad t \in [0,1]$$

4 | Asymptotic Properties of the Estimators

We work under the following assumptions.

Assumption 4.1. The volatility function is continuously differentiable, that is, $\sigma(\cdot) \in C^1([0,1])$. Equivalently, the quadratic variation process $Q(\cdot) = \int_0^{\cdot} \sigma^2(u)du$ is twice continuously differentiable, that is, $Q(\cdot) \in C^2([0,1])$.

Assumption 4.2. The speed of reversion function $\theta(\cdot)$ is continuously differentiable, that is, $\theta(\cdot) \in C^1([0,1])$.

Assumption 4.3. The kernel function $K(\cdot)$ is supported on the compact interval $[-1,1]$ with finite integral, $\int |K(u)|du < \infty$. Its Fourier transform $K^\dagger(\cdot) = \frac{1}{2\pi}\int \exp\{-\mathrm{i}ut\}K(u)du$, where i is the imaginary unit, satisfies $\int \left|K^\dagger(u)\right|du < \infty$.

Assumption 4.4. The bandwidth parameter $h_Q \to 0$ satisfies

$$\frac{K}{h_Q^2}\left(\frac{\log N}{N}\right)^{1-\frac{2}{\alpha}} = O(1)$$

Assumption 4.5. The bandwidth parameter $h_m \to 0$ satisfies

$$\frac{1}{h_m^2}\left(\frac{\log N}{N}\right)^{1-\frac{2}{\alpha}} = O(1)$$

Assumption 4.6. The bandwidth parameter $h_G \to 0$ satisfies

$$\frac{1}{h_G^2}\left(\frac{\log N}{N}\right)^{1-\frac{2}{\alpha}} = O(1)$$

Assumption 4.7. $\mathbb{E}\left|R_1(0)\right|^\alpha < \infty$.

We begin with two results that establish uniform bounds on moments of the realizations of model (12) and its empirical variation process. These results are used in subsequent proofs, but are also of independent interest, as they could be applied in other contexts.

Theorem 4.1. *In addition to the assumptions of Theorem 3.1, suppose that Assumption 4.7 holds for some $\alpha > 0$. Then, for the same value(s) of α,*

$$\sup_{0 \le t \le 1} \mathbb{E}\left|R_1(t)\right|^\alpha < \infty$$

where $R_1(\cdot)$ is defined in Equation (12).

Theorem 4.2. *Suppose Assumptions of Theorem 3.1 hold. Then, for any $\alpha > 0$,*

$$\sup_{0 \le t \le 1} \mathbb{E}\left|\widetilde{Q}_1(t)\right|^\alpha = \mathbb{E}\left|\widetilde{Q}_1(1)\right|^\alpha \le M_\alpha < \infty \tag{33}$$

with $\widetilde{Q}_1(\cdot)$ defined in Equation (19). The constant M_α depends only on α and $\|\sigma\|_\infty := \sup_{t\in[0,1]} |\sigma(t)|$.

The remaining theorems establish uniform convergence rates on all estimators defined in Section 3.

Theorem 4.3. *In addition to the assumptions of Theorem 3.1, suppose that Assumptions 4.1, 4.3, and 4.4 hold true for some $\alpha > 2$. Then, regardless of the interplay between K and N, with probability 1,*

$$\sup_{t\in[0,1]} \left|\widehat{Q}(t) - Q(t)\right| = O(\mathcal{R}_Q(N,K)),$$
$$\sup_{t\in[0,1]} \left|\widehat{\partial Q}(t) - \partial Q(t)\right| = h_Q^{-1} O(\mathcal{R}_Q(N,K))$$

where $\mathcal{R}_Q(N,K) = \left(\frac{\log N}{KN}\right)^{\frac{1}{2}} + h_Q^2$.

Theorem 4.4. *In addition to the assumptions of Theorem 3.1, suppose that Assumptions 4.2, 4.3, 4.5, and 4.7, with some $\alpha > 2$, hold true. Then, with probability 1,*

$$\sup_{t\in[0,1]} |\hat{m}(t) - m(t)| = O(\mathcal{R}_m(N)),$$
$$\sup_{t\in[0,1]} \left|\widehat{\partial m}(t) - \partial m(t)\right| = h_m^{-1} O(\mathcal{R}_m(N))$$

where $\mathcal{R}_m(N) = \left(\frac{\log N}{N}\right)^{\frac{1}{2}} + h_m^2$.

Theorem 4.5. *Suppose Assumptions 4.1, 4.2, 4.3, 4.6, and 4.7, with some $\alpha > 4$, hold true. Then, with probability 1,*

$$\sup_{t\in[0,1]} \left|\widehat{G}(t) - G(t)\right| = O(\mathcal{R}_G(N)),$$
$$\sup_{t\in[0,1]} \left|\widehat{\partial G}(t) - \partial G(t)\right| = h_G^{-1} O(\mathcal{R}_G(N))$$

where $\mathcal{R}_G(N) = \left(\frac{\log N}{N}\right)^{\frac{1}{2}} + h_G^2$.

Theorem 4.6. *Suppose Assumptions 4.1, 4.2, 4.3, 4.5, 4.6, and 4.7, with some $\alpha > 4$, hold. Then, with probability 1,*

$$\sup_{t\in[0,1]} \left|\widehat{D}(t) - D(t)\right| = O\left(\left(\frac{\log N}{N}\right)^{\frac{1}{2}} + h_G^2 + h_m^2\right),$$
$$\sup_{t\in[0,1]} \left|\widehat{\partial D}(t) - \partial D(t)\right| = O\left(\left(h_G^{-1} + h_m^{-1}\right)\left(\frac{\log N}{N}\right)^{\frac{1}{2}} + h_G + h_m\right)$$

Theorem 4.7. *In addition to the assumptions of Theorem 3.1, suppose that Assumptions 4.1, 4.3, and 4.4 hold true for some $\alpha > 2$ then, regardless of the interplay between K and N, with probability 1,*

$$\sup_{t\in[0,1]} |\hat{\sigma}(t) - \sigma(t)| = O\left(h_Q^{-1}\left(\frac{\log N}{KN}\right)^{\frac{1}{2}} + h_Q\right)$$

Theorem 4.8. *Suppose Assumptions 4.1–4.7, with $\alpha > 4$ hold true then, regardless of the interplay between K and N, with probability 1,*

$$\sup_{t\in[0,1]} \left|\hat{\theta}(t) - \theta(t)\right| = O(\mathcal{R}_\theta(N,K))$$

with $\mathcal{R}_\theta(N,K) = h_Q^{-1}\left(\frac{\log N}{KN}\right)^{\frac{1}{2}} + \left(h_G^{-1} + h_m^{-1}\right)\left(\frac{\log N}{N}\right)^{\frac{1}{2}} + h_Q + h_G + h_m$.

Theorem 4.9. *Assume conditions of Theorem 4.8 hold true then, with probability 1,*

$$|\hat{\mu} - \mu| = O(\mathcal{R}_\mu(N,K))$$

with $\mathcal{R}_\mu(N,K) = h_Q^{-1}\left(\frac{\log N}{KN}\right)^{\frac{1}{2}} + \left(h_G^{-1} + h_m^{-1}\right)\left(\frac{\log N}{N}\right)^{\frac{1}{2}} + h_Q + h_G + h_m$.

The rates in Theorems 4.7, 4.8, and 4.9 could be used to derive optimal bandwidths, but we do not pursue this problem in this paper. As with all nonparametric problems of this type, the selection of optimal bandwidths does not reduce to deriving optimal rates, but we can expect that in finite samples, σ can be estimated more precisely than θ. This difference is borne out in simulations in Section 5. This can be explained as follows. To minimize the error in the estimation of the function σ, we need $h_Q \sim (\log N)^{1/4}(NK)^{-1/4}$. For θ and μ, the optimal rates are $h_Q, h_m \sim (\log N)^{1/4}(N)^{-1/4}$. With these rates, the error in the estimation of the function σ is roughly of the order $(NK)^{-1/4}$ while for θ and μ the errors are roughly of the order $N^{-1/4}$. Finite sample errors will be impacted by the kernel and the bandwidth selection rule.

5 | Application to U.S. Treasuries and a Simulation Study

We begin with an application to specific data that will motivate the simulations that follow. We consider the following futures: FV (TNote 5 Yr), TY (TNote 10 Yr), US (TBond 30 Yr). The 5-year and 10-year T-note, and 30-year T-Bill futures contracts are traded at the Chicago Mercantile Exchange (CME Group). The trade data were obtained from TickData. The continuous contracts were constructed using front contracts and rolled over at 28 days before expiration dates. The data are sampled at the 5-min frequency in our application, but practically any other frequency is possible. Examples of curves we work with are shown in Figure 1.

The data are available from 2003 to 2014 on regular trading days. We treat each calendar year separately, so $N \approx 250$. We use the values at the 5-minute discretization over the time interval from 8:00 to 20:00 EST, that is, $K = 144$. Even though there is an overnight separation between the trading days we consider, the assumption of independence may not hold perfectly. We estimated the Vašičekmodel for each of the 11 calendar years separately. In all linear smoothing formulas ((21), (24), and (26)) we used the Epanechnikov kernel with the bandwidth $h = 0.5 \cdot K^{-1/5}$.

The results are presented in Figure 2. The rates in the third column are in basis points (1/100th of a percent), so the equilibrium rates vary between 1.1 and 1.7 percent, depending on the year, for the 30-year bond, and are lower for the T-notes. The volatilities $\sigma(t)$ are highest around noon EST. The reversion

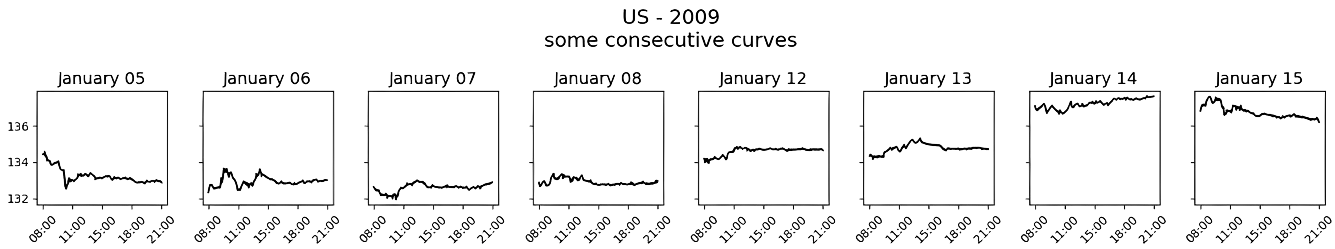

FIGURE 1 | First 8 consecutive available days in 2009, US 30-year bond.

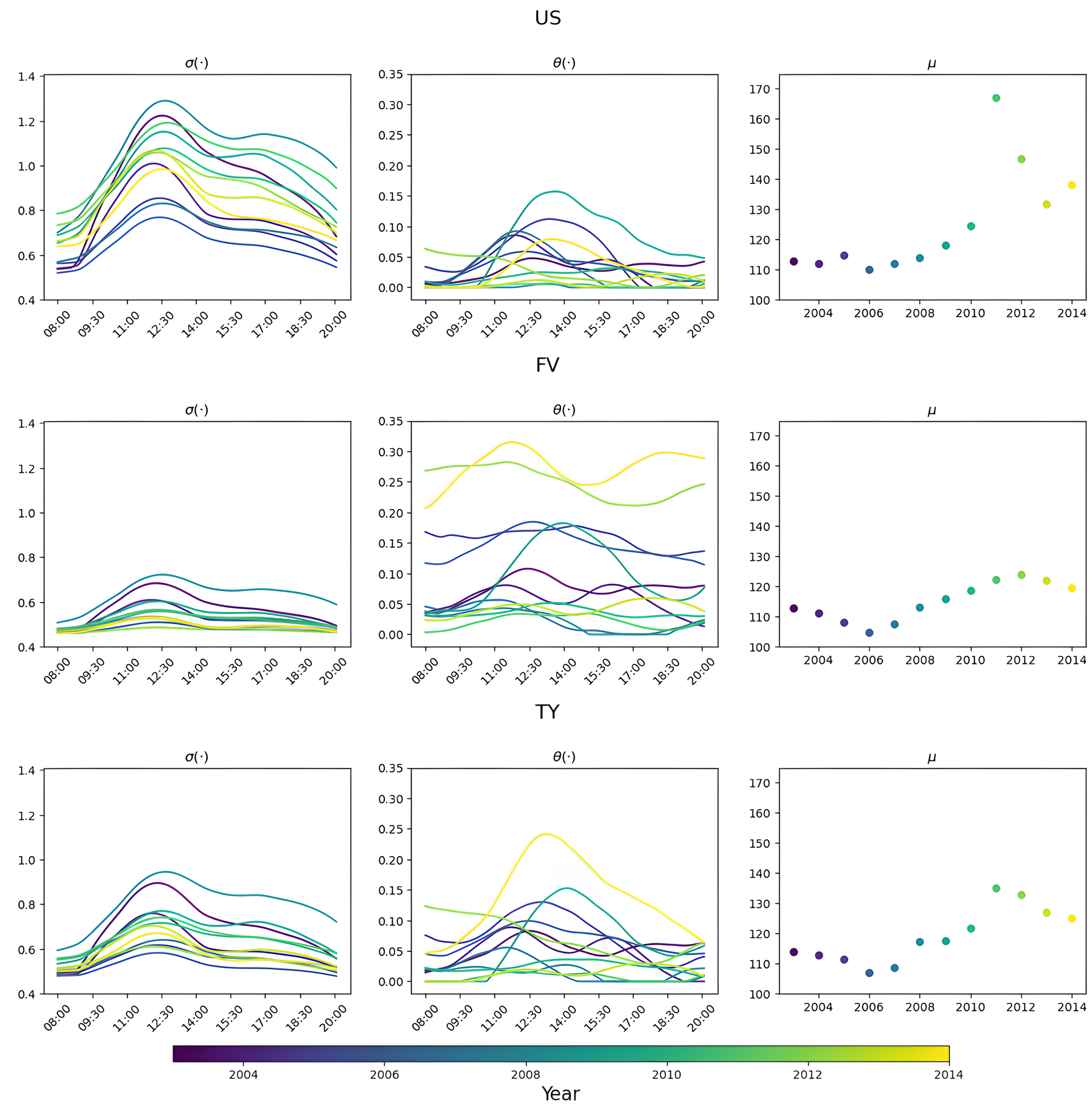

FIGURE 2 | Estimated values of instant volatility, speed of reversion, and the long-term mean for each treasury future over the years considered.

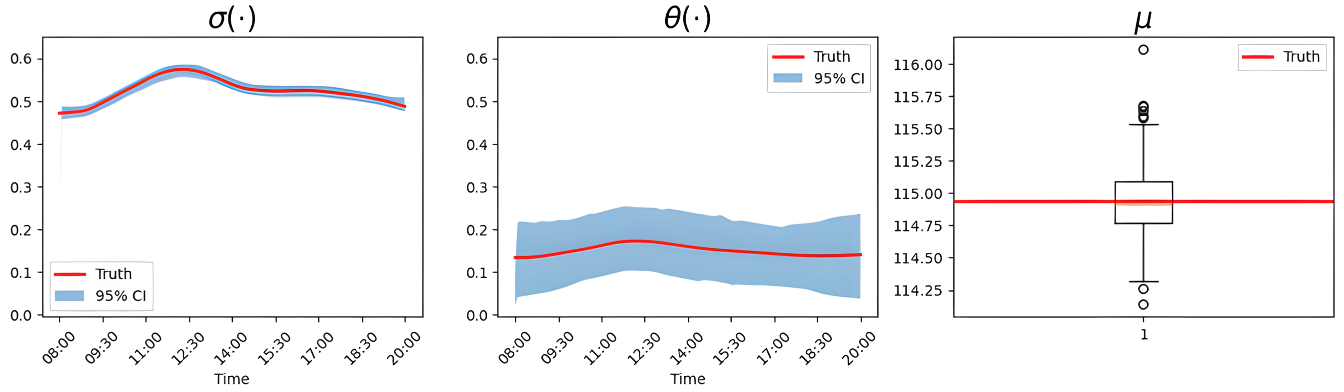

FIGURE 3 | 95% prediction bands and a boxplot obtained over 500 independent experiments based on data resampling FV futures for an "average year".

coefficients $\theta(t)$ appear to follow a similar pattern with a larger spread between years for the T-notes. The last conclusion must be taken with some caution because, as we noted at the end of Section 4 the function θ is estimated less accurately than the function σ. Figure 2 does strongly suggest that neither of these functions can be assumed to be constant throughout a trading day.

We now report the results of a simulation study motivated by the above data analysis, which is also meant to validate our conclusions. As the parameters $\sigma(\cdot), \theta(\cdot)$ and μ we chose the averages of the curves in Figure 2 across the years 2003 and 2014 for the 10-year treasury futures (FV). To mimic this instance of data, since there are approximately 250 trading days each year, we generated $N = 250$ independent, identically distributed curves, each with the starting point drawn from a Gaussian distribution $\mathcal{N}(\nu, \tau)$, where ν is the average of all $R(0)$ in the FV dataset and τ their standard deviation. The time interval, the kernels, the bandwidths, and the discretization ($K = 144$) are the same as in the data analysis. The results of the estimation based on 500 Monte Carlo runs are shown in Figure 3. We see that the function $\sigma(\cdot)$ is estimated very precisely, as suggested by the rate in Theorem 4.3. Consequently, we can say with high confidence that for some years the estimated curves in the left panel of Figure 2 correspond to statistically significantly different functional parameters. The spread of the estimators of the reversion curves $\theta(\cdot)$ is around 0.20, so we still can be fairly confident the some differences in the middle panel of Figure 2 are significant. The spread of the long-term rates is about 0.3 and definitely less than one basis point, accounting for a slightly leptokurtic distribution of the initial values. The differences in the right panel of Figure 2 can thus also be interpreted as statistically significant.

6 | Proofs of Theoretical Results

6.1 | Proofs of Theorem 3.1 and Proposition 3.1

Proof of Theorem 3.1. The existence of a unique solution is a consequence of the Theorem 2.1. We now obtain the representation of this unique solution. Applying Itô formula (Theorem 4.1.2 in Øksendal (2003)) to the function $f(t,x) = x\exp\left\{\int_0^t \theta(u)du\right\}$, we obtain

$$
\begin{aligned}
d&\left(R(t)\exp\left\{\int_0^t \theta(u)du\right\}\right)\\
&= \frac{\partial f}{\partial t}(t,R(t))dt + \frac{\partial f}{\partial x}(t,R(t))dR(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,R(t))(dR(t))^2\\
&= R(t)\exp\left\{\int_0^t \theta(u)du\right\}\theta(t)dt + \exp\left\{\int_0^t \theta(u)du\right\}dR(t) + 0\\
&= R(t)\exp\left\{\int_0^t \theta(u)du\right\}\theta(t)dt + \exp\left\{\int_0^t \theta(u)du\right\}\\
&\quad\times \theta(t)(\mu - R(t))dt + \exp\left\{\int_0^t \theta(u)du\right\}\sigma(t)dW(t)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
d\left(R(t)\exp\left\{\int_0^t \theta(u)du\right\}\right) &= \mu\exp\left\{\int_0^t \theta(u)du\right\}\theta(t)dt\\
&\quad + \exp\left\{\int_0^t \theta(u)du\right\}\sigma(t)dW(t)
\end{aligned}
$$

This leads to

$$
\begin{aligned}
&R(t)\exp\left\{\int_0^t \theta(u)du\right\}\\
&\quad= R(0) + \mu\int_0^t \exp\left\{\int_0^s \theta(u)du\right\}\theta(s)ds\\
&\qquad+ \int_0^t \exp\left\{\int_0^s \theta(u)du\right\}\sigma(s)dW(s)
\end{aligned}
$$

which entails the desired closed-form representation (7), that is,

$$
\begin{aligned}
R(t) &= \exp\left\{-\int_0^t \theta(u)du\right\}R(0)\\
&\quad+ \mu\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\theta(s)ds\\
&\quad+ \int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\sigma(s)dW(s), \qquad t\in[0,\infty) \qquad (34)
\end{aligned}
$$

Taking expectation of both sides of (34), we have

$$
\begin{aligned}
m(t) &= \exp\left\{-\int_0^t \theta(u)du\right\}m(0)\\
&\quad+ \mu\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\theta(s)ds \qquad (35)
\end{aligned}
$$

that gives (8). Representation (35) can be rewritten in the form

$$m(t) = \exp\left\{-\int_0^t \theta(u)du\right\} m(0) + \mu \exp\left\{-\int_0^t \theta(u)du\right\} \times \int_0^t \exp\left\{\int_0^s \theta(u)du\right\}\theta(s)ds + 0$$

This leads to

$$\begin{aligned}\partial m(t) &= -\theta(t)\exp\left\{-\int_0^t \theta(u)du\right\} m(0) \\ &\quad - \mu\theta(t)\exp\left\{-\int_0^t \theta(u)du\right\}\int_0^t \exp\left\{\int_0^s \theta(u)du\right\}\theta(s)ds \\ &\quad + \mu\exp\left\{-\int_0^t \theta(u)du\right\}\exp\left\{\int_0^t \theta(u)du\right\}\theta(t) \\ &= -\theta(t)\exp\left\{-\int_0^t \theta(u)du\right\} m(0) \\ &\quad - \mu\theta(t)\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\theta(s)ds + \mu\theta(t)\end{aligned}$$

which together with (35) implies

$$\partial m(t) = -\theta(t)m(t) + \mu\theta(t)$$

as claimed in Equation (10). Applying Corollary 2.1, we see that the variance function $D(\cdot)$ of the solution (34) satisfies

$$D(t) = \exp\left\{-2\int_0^t \theta(u)du\right\} D(0) + \int_0^t \exp\left\{-2\int_s^t \theta(u)du\right\}\sigma^2(s)ds$$

This entails (9). It will be convenient to work with the representation

$$D(t) = \exp\left\{-2\int_0^t \theta(u)du\right\} D(0) + \exp\left\{-2\int_0^t \theta(u)du\right\}\int_0^t \exp\left\{2\int_0^s \theta(u)du\right\}\sigma^2(s)ds \tag{36}$$

Taking the derivative, we obtain

$$\begin{aligned}\partial D(t) &= -2\theta(t)\exp\left\{-2\int_0^t \theta(u)du\right\} D(0) \\ &\quad - 2\theta(t)\exp\left\{-2\int_0^t \theta(u)du\right\} \\ &\quad \times \int_0^t \exp\left\{2\int_0^s \theta(u)du\right\}\sigma^2(s)ds \\ &\quad + \exp\left\{-2\int_0^t \theta(u)du\right\}\exp\left\{2\int_0^t \theta(u)du\right\}\sigma^2(t) \\ &= -2\theta(t)\exp\left\{-2\int_0^t \theta(u)du\right\} D(0) \\ &\quad - 2\theta(t)\int_0^t \exp\left\{-2\int_s^t \theta(u)du\right\}\sigma^2(s)ds + \sigma^2(t)\end{aligned}$$

that together with (36) leads to (11). This completes the proof. □

Proof of Proposition 3.1. Both optimization problems (21) and (31) admit a closed-form solution. So we first present these closed-form solutions, and then we justify their equality. Observe that the solution (22) has the form

$$\begin{bmatrix}\hat{Q}(t) \\ h_Q\widehat{\partial Q}(t)\end{bmatrix} = \begin{bmatrix}\mathcal{A}_0 & \mathcal{A}_1 \\ \mathcal{A}_1 & \mathcal{A}_2\end{bmatrix}^{-1}\begin{bmatrix}S_0 \\ S_1\end{bmatrix}(t) \tag{37}$$

where for $0 \le t \le 1$,

$$\mathcal{A}_p(t) = \frac{1}{K+1}\sum_{k=0}^{K}\left(\frac{t_k - t}{h_Q}\right)^p h_Q^{-1}K_Q\left(h_Q^{-1}(t_k - t)\right), \quad p = 0,1,2 \tag{38}$$

and

$$S_p(t) = \frac{1}{K+1}\sum_{k=0}^{K}\overline{Q}(t_k)\left(\frac{t_k - t}{h_Q}\right)^p h_Q^{-1}K_Q\left(h_Q^{-1}(t_k - t)\right) \quad p = 0,1 \tag{39}$$

On the other hand, the solution (32) admits the closed from representation

$$\begin{bmatrix}\check{Q}(t) \\ h_Q\partial\check{Q}(t)\end{bmatrix} = \begin{bmatrix}\mathcal{A}_0^* & \mathcal{A}_1^* \\ \mathcal{A}_1^* & \mathcal{A}_2^*\end{bmatrix}^{-1}\begin{bmatrix}S_0^* \\ S_1^*\end{bmatrix}(t) \tag{40}$$

where for $0 \le t \le 1$,

$$\begin{aligned}\mathcal{A}_p^*(t) &= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left(\frac{t_k - t}{h_Q}\right)^p h_Q^{-1}K_Q\left(h_Q^{-1}(t_k - t)\right) \qquad (41) \\ &= \frac{1}{K+1}\sum_{k=0}^{K}\left(\frac{t_k - t}{h_Q}\right)^p h_Q^{-1}K_Q\left(h_Q^{-1}(t_k - t)\right), \quad p = 0,1,2\end{aligned}$$

and

$$\begin{aligned}S_p^*(t) &= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\widetilde{Q}_i(t_k)\left(\frac{t_k - t}{h_Q}\right)^p h_Q^{-1}K_Q\left(h_Q^{-1}(t_k - t)\right) \\ &= \frac{1}{K+1}\sum_{k=0}^{K}\overline{Q}(t_k)\left(\frac{t_k - t}{h_Q}\right)^p h_Q^{-1}K_Q\left(h_Q^{-1}(t_k - t)\right) \\ &\qquad p = 0,1\end{aligned} \tag{42}$$

Comparing (37), (38), and (39) to (40), (41), and (42) completes the proof. □

6.2 | Proof of Theorems 4.1 and 4.2

Proof of Theorem 4.1. Recall the closed-form representation (7) limited to the unit compact interval [0, 1]:

$$R(t) = \exp\left\{-\int_0^t \theta(u)du\right\}R(0) + \mu\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\theta(s)ds + \int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\sigma(s)dW(s), \quad t \in [0,1]$$

This implies that the desired result, $\sup_{0\le t\le 1}\mathbb{E}|R(t)|^\alpha < \infty$, holds if

$$\mathbb{E}|R(0)|^\alpha < \infty \tag{43}$$

$$\sup_{0\le t\le 1}\left|\mu\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\theta(s)ds\right|^\alpha < \infty \tag{44}$$

and

$$\sup_{0\le t\le 1}\mathbb{E}\left|\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\sigma(s)dW(s)\right|^\alpha < \infty \tag{45}$$

Condition (43) holds by the assumptions of this Theorem. Inequality (44) is a consequence of the continuity of $\theta(\cdot)$. To justify (45), observe that

$$\begin{aligned}
&\sup_{0\le t\le 1}\mathbb{E}\left|\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\}\sigma(s)dW(s)\right|^\alpha \\
&\le \sup_{0\le t\le 1}\exp\left\{-\alpha\int_0^t \theta(u)du\right\} \\
&\sup_{0\le t\le 1}\mathbb{E}\left|\int_0^t \exp\left\{\int_0^s \theta(u)du\right\}\sigma(s)dW(s)\right|^\alpha \\
&= \sup_{0\le t\le 1}\exp\left\{-\alpha\int_0^t \theta(u)du\right\} \\
&\sup_{0\le t\le 1}\mathbb{E}\left|W\left(\int_0^t \exp\left\{2\int_0^s \theta(u)du\right\}\sigma^2(s)d(s)\right)\right|^\alpha \qquad (46)\\
&< \infty \qquad (47)
\end{aligned}$$

Equality (46) is a consequence of Dambis–Dubins–Schwarz theorem, see Corollary 2.1, and (47) follows from the continuity of $\sigma(\cdot)$ and $\theta(\cdot)$. □

Proof of Theorem 4.2. Recall that

$$Q(t)=\int_0^t \sigma^2(u)du, \quad \widetilde{Q}_i(t)=\sum_{k=1}^K |R_i(t_k)-R_i(t_{k-1})|^2\mathbb{1}\{t_k\le t\},$$
$$t\in[0,1]$$

which implies the equality $\sup_{0\le t\le 1}\mathbb{E}\left|\widetilde{Q}_1(t)\right|^\alpha=\mathbb{E}\left|\widetilde{Q}_1(1)\right|^\alpha$. We now obtain an upper bound for $\mathbb{E}\left|\widetilde{Q}_1(1)-Q_1(1)\right|^\alpha$. Observe that

$$\begin{aligned}
\left|\widetilde{Q}_1(1)-Q_1(1)\right|^\alpha &= \left|\sum_{k=1}^K\left|R_1(t_j)-R_1(t_{j-1})\right|^2-\int_0^1\sigma^2(u)du\right|^\alpha \\
&= \left|\sum_{k=1}^K\left(\left|R_1(t_k)-R_1(t_{k-1})\right|^2-\int_{t_{k-1}}^{t_k}\sigma^2(u)du\right)\right|^\alpha \\
&= \left|\sum_{k=1}^K\left(\frac{\left|R_1(t_k)-R_1(t_{k-1})\right|^2}{\int_{t_{k-1}}^{t_k}\sigma^2(u)du}-1\right)\int_{t_{k-1}}^{t_k}\sigma^2(u)du\right|^\alpha \\
&=: \left|\sum_{k=1}^K \xi_k\int_{t_{k-1}}^{t_k}\sigma^2(u)du\right|^\alpha
\end{aligned}$$

where ξ_k form a sequence of independent random variables with $\xi_k+1\sim\chi^2(1)$. Hence, $\{\xi_k\}$ is a ϕ-mixing sequence with bounded moments of any order. Applying Theorem 1 in Yoshihara (1978), for any even positive integer α we have

$$\mathbb{E}\left|\sum_{k=1}^K \xi_k\int_{t_{k-1}}^{t_k}\sigma^2(u)du\right|^\alpha \le c_\alpha\Delta^{-\alpha/2}\|\sigma\|_\infty^{2\alpha}\Delta^\alpha < c_\alpha\|\sigma\|_\infty^{2\alpha}$$

where c_α is an absolute constant depending on α only. This gives the result for positive even integers. For general $\alpha>0$, there is always a positive even integer larger than α. This completes the proof. □

6.3 | Proof of Theorem 4.3

Lemma 1. *Suppose Assumptions of Theorem 3.1*

$$\sup_{t\in[0,1]}\mathbb{E}\left|\widetilde{Q}_1(t)-Q(t)\right|^2 \le 6\left\|\sigma^2\right\|_\infty^2\frac{1}{K}+2\left\|\sigma^2\right\|_\infty^2\frac{1}{K^2}=O\left(\frac{1}{K}\right)$$

Proof of Lemma 1. This proof follows the lines of the proof of Lemma A.1 in Kokoszka et al. (2025). □

Proof of Theorem 4.3. Recall that solution (22) has the from

$$\begin{bmatrix}\widehat{Q}(t)\\ h_Q\widehat{\partial Q}(t)\end{bmatrix}=\begin{bmatrix}\mathcal{A}_0 & \mathcal{A}_1\\ \mathcal{A}_1 & \mathcal{A}_2\end{bmatrix}^{-1}\begin{bmatrix}S_0\\ S_1\end{bmatrix}(t) \tag{48}$$

where for $0\le t\le 1$,

$$\mathcal{A}_p(t)=\frac{1}{N(K+1)}\sum_{i=1}^N\sum_{k=0}^K\left(\frac{t_k-t}{h_Q}\right)^p h_Q^{-1}K_Q\left(h_Q^{-1}(t_k-t)\right)$$
$$p=0,1,2 \tag{49}$$

and

$$S_p(t)=\frac{1}{N(K+1)}\sum_{i=1}^N\sum_{k=0}^K\widetilde{Q}_i(t_k)\left(\frac{t_k-t}{h_Q}\right)^p h_Q^{-1}$$
$$\times K_Q\left(h_Q^{-1}(t_k-t)\right), \quad p=0,1 \tag{50}$$

Equation (48) implies

$$\begin{aligned}
\begin{bmatrix}\widehat{Q}(t)-Q(t)\\ h_Q\widehat{\partial Q}(t)-h_Q\partial Q(t)\end{bmatrix} &= \begin{bmatrix}\mathcal{A}_0 & \mathcal{A}_1\\ \mathcal{A}_1 & \mathcal{A}_2\end{bmatrix}^{-1} \\
&\times\left(\begin{bmatrix}S_0\\ S_1\end{bmatrix}-\begin{bmatrix}\mathcal{A}_0 & \mathcal{A}_1\\ \mathcal{A}_1 & \mathcal{A}_2\end{bmatrix}\begin{bmatrix}Q\\ h_Q\partial Q\end{bmatrix}\right)(t) \\
&= \begin{bmatrix}\mathcal{A}_0 & \mathcal{A}_1\\ \mathcal{A}_1 & \mathcal{A}_2\end{bmatrix}^{-1}\begin{bmatrix}\widetilde{S}_0\\ \widetilde{S}_1\end{bmatrix}(t) \qquad (51)
\end{aligned}$$

where

$$\widetilde{S}_0(t)=S_0(t)-\mathcal{A}_0Q(t)-\mathcal{A}_1h_Q\partial Q(t)=:\ S_0(t)-\Lambda_0(t),$$
$$\widetilde{S}_1(t)=S_1(t)-\mathcal{A}_1Q(t)-\mathcal{A}_2h_Q\partial Q(t)=:\ S_1(t)-\Lambda_1(t)$$

Hence, (51) can be rewritten in the form

$$\begin{bmatrix}\widehat{Q}(t)-Q(t)\\ h_Q\widehat{\partial Q}(t)-h_Q\partial Q(t)\end{bmatrix}=\frac{1}{\mathcal{A}_0\mathcal{A}_2-\mathcal{A}_1^2}\begin{bmatrix}\mathcal{A}_2\widetilde{S}_0-\mathcal{A}_1\widetilde{S}_1\\ -\mathcal{A}_1\widetilde{S}_0+\mathcal{A}_0\widetilde{S}_1\end{bmatrix}(t) \tag{52}$$

We investigate $\widetilde{S}_0(t)$ in detail; the other terms are similar. Note that

$$
\begin{aligned}
\widetilde{S}_0(t) &= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\widetilde{Q}_i(t_k)h_Q^{-1}K_Q\left(h_Q^{-1}(t_k-t)\right)-\Lambda_0(t)\\
&= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}Q(t_k)h_Q^{-1}K_Q\left(h_Q^{-1}(t_k-t)\right)-\Lambda_0(t)\\
&\quad+\frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left(\widetilde{Q}_i(t_k)-Q(t_k)\right)h_Q^{-1}K_Q\left(h_Q^{-1}(t_k-t)\right)\\
&=: A_1(t)+A_2(t)
\end{aligned}
$$

We first consider $A_1(\cdot)$. Observe that

$$
\begin{aligned}
A_1(t) &= \frac{1}{K+1}\sum_{k=0}^{K}\left[Q(t_k)-Q(t)-(t_k-t)\partial Q(t)\right]\\
&\quad\times h_Q^{-1}K_Q\left(h_Q^{-1}(t_k-t)\right)\\
&= O(h_Q^2),\quad \text{(w.p.1) uniformly on } 0\le t\le 1 \qquad (53)
\end{aligned}
$$

where Assumptions 4.1 and 4.3 allow Taylor's expansion to be used to obtain (53). We now investigate $A_2(\cdot)$. First observe that $\mathbb{E}A_2(t)=0$, for all $t\in[0,1]$. This implies

$$
\begin{aligned}
&\left|A_2(t)-\mathbb{E}A_2(t)\right|\\
&\le \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|h_Q^{-1}K_Q\left(h_Q^{-1}(t_k-t)\right)\\
&\le \int\left|K^{\dagger}(h_Q u)\right|du\times\frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|\\
&\quad\times\mathbb{I}\{t_k\in[t-h_Q,t+h_Q]\}\\
&= O(h_Q^{-1})\times\frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|\\
&\quad\times\mathbb{I}\{t_k\in[t-h_Q,t+h_Q]\}
\end{aligned}
$$

where $K^{\dagger}(\cdot)$ denotes the Fourier transform of the kernel function $K(\cdot)$, see Assumption 4.3. We now prove that

$$
\begin{aligned}
S(N,K) &:= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|\mathbb{I}\\
&\quad\times\{t_k\in[t-h_Q,t+h_Q]\}\\
&= O\left(\frac{h_Q^2\log N}{KN}\right)^{\frac{1}{2}},\text{uniformly on } 0\le t\le 1
\end{aligned}
$$

To do it, define

$$
U_{N,K} := \left(\frac{h_Q^2}{K}\right)^{\frac{1}{2}}\left(\frac{\log N}{N}\right)^{-\frac{1}{2}}
$$

and choose $h_Q\to 0$ such that

$$
U_{N,K}^{1-\alpha}=O\left(\frac{h_Q^2\log N}{KN}\right)^{\frac{1}{2}}\quad\text{that is,}\quad \frac{K}{h_Q^2}\left(\frac{\log N}{N}\right)^{1-\frac{2}{\alpha}}=O(1) \qquad (54)
$$

for some $\alpha>2$. Then we have

$$
\begin{aligned}
&S(N,K)\\
&= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|\mathbb{I}\left\{\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|>U_{N,K}\right\}\\
&\quad\times\mathbb{I}\{t_k\in[t-h_Q,t+h_Q]\}\\
&\quad+\frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|\mathbb{I}\left\{\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|\le U_{N,K}\right\}\\
&\quad\times\mathbb{I}\{t_k\in[t-h_Q,t+h_Q]\}\\
&=: B_1+B_2
\end{aligned}
$$

Considering B_1, we have

$$
\begin{aligned}
B_1 &\le \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|^{1-\alpha+\alpha}\\
&\quad\times\mathbb{I}\left\{\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|>U_{N,K}\right\}\\
&\le \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|^{\alpha}U_{N,K}^{1-\alpha}\\
&= O(1)U_{N,K}^{1-\alpha},\qquad \text{w.p.1 (uniformly on } 0\le t\le 1) \qquad (55)\\
&= O\left(\frac{h_Q^2\log N}{KN}\right)^{\frac{1}{2}},\qquad \text{w.p.1 (uniformly on } 0\le t\le 1) \qquad (56)
\end{aligned}
$$

where (55) is a consequence of continuity of the function $\sigma(\cdot)$, which in particular implies that (33) holds for some $\alpha>2$, see Theorem 4.2. The last Equation (56) is a consequence of (54).

We now turn to B_2. First observe that

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{1}{K+1}\sum_{k=0}^{K}\left|\widetilde{Q}_i(t_k)-Q(t_k)\right|\mathbb{I}\{t_k\in[t-h_Q,t+h_Q]\}\right)\\
&\le \frac{1}{(K+1)^2}\sup_{t\in[0,1]}\mathbb{E}\left[\left|\widetilde{Q}_i(t)-Q(t)\right|\right]^2\left(2h_Q(K+1)\right)^2\\
&\le c\frac{4h_Q^2}{K} \qquad (57)
\end{aligned}
$$

where $c=6\|\sigma^2\|_\infty^2$. The upper bound (57) is a consequence of Lemma 1. The above, together with the Bernstein inequality, implies, for any positive number η,

$$
\begin{aligned}
\mathbb{P}\left(B_2\ge\eta\left(\frac{h_Q^2\log N}{KN}\right)^{\frac{1}{2}}\right) &= \mathbb{P}\left(NB_2\ge N\eta\left(\frac{h_Q^2\log N}{KN}\right)^{\frac{1}{2}}\right)\\
&\le \exp\left\{-\frac{\eta^2N^2\left(\frac{h_Q^2\log N}{KN}\right)}{8Nc\frac{h_Q^2}{K}+\frac{2}{3}\eta N\frac{h_Q^2}{K}}\right\}\\
&= \exp\left\{-\frac{\eta^2\log N}{8c+\frac{2}{3}\eta}\right\}\\
&= N^{-\frac{\eta^2}{8c+\frac{2}{3}\eta}} \qquad (58)
\end{aligned}
$$

Choosing η large enough entails summability of (58) over index N. Applying the Borel–Cantelli lemma, we conclude that there is a subset $\Omega_0 \subset \Omega$ with full probability measure such that for each $\omega \in \Omega_0$ there is $N_0(\omega)$ with

$$B_2 \leq \eta\left(\frac{h_Q^2 \log N}{KN}\right)^{\frac{1}{2}}, \quad N \geq N_0(\omega)$$

This completes the proof for $\widetilde{S}_0(t)$. The other terms appearing in Equation (52) can be treated similarly. This completes the proof. □

6.4 | Proofs of the Remaining Theorems of Section 4

Lemma 2. *If, in addition to the assumptions of Theorem 3.1, Assumption 4.2 holds, then the mean function $m(\cdot)$ is twice continuously differentiable, that is, $m(\cdot) \in C^2([0,1])$.*

Proof of Lemma 2. The lemma follows from the representation (8), which can be rewritten as

$$\begin{aligned} m(t) &= \exp\left\{-\int_0^t \theta(u)du\right\}m(0) + \mu\int_0^t \exp\left\{-\int_s^t \theta(u)du\right\} \\ &\quad\times \theta(s)ds \\ &= \exp\left\{-\int_0^t \theta(u)du\right\}m(0) + \mu\exp\left\{-\int_0^t \theta(u)du\right\} \\ &\quad\times \int_0^t \exp\left\{\int_0^s \theta(u)du\right\}\theta(s)ds \end{aligned}$$

□

Proof of Theorem 4.4. The solution (23) admits the closed from representation

$$\begin{bmatrix} \hat{m}(t) \\ h_Q\widehat{\partial m}(t) \end{bmatrix} = \begin{bmatrix} \mathcal{A}_0 & \mathcal{A}_1 \\ \mathcal{A}_1 & \mathcal{A}_2 \end{bmatrix}^{-1} \begin{bmatrix} S_0 \\ S_1 \end{bmatrix}(t)$$

where for $0 \leq t \leq 1$,

$$\mathcal{A}_p(t) = \frac{1}{K+1}\sum_{k=0}^{K}\left(\frac{t_k - t}{h_m}\right)^p h_m^{-1}K_m\left(h_m^{-1}(t_k - t)\right), \quad p = 0,1,2$$

and

$$\begin{aligned} S_p(t) &= \frac{1}{N(K+1)}\sum_{k=0}^{K}\overline{R}(t_k)\left(\frac{t_k - t}{h_m}\right)^p h_m^{-1}K_m\left(h_m^{-1}(t_k - t)\right) \\ &= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}R_i(t_k)\left(\frac{t_k - t}{h_m}\right)^p h_m^{-1}K_m\left(h_m^{-1}(t_k - t)\right), \\ &\qquad p = 0,1 \end{aligned}$$

Following the lines of the proof of Theorem 4.3, it is enough to investigate the terms

$$\begin{aligned} A_1(t) + A_2(t) &:= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}m(t_k)h_m^{-1}K_m\left(h_m^{-1}(t_k - t)\right) \\ &\quad - \Lambda_0(t) + \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left(R_i(t_k) - m(t_k)\right)h_m^{-1}K_m \\ &\quad\times\left(h_m^{-1}(t_k - t)\right) \end{aligned}$$

where

$$\Lambda_0(t) := \mathcal{A}_0 m(t) + \mathcal{A}_1 h_m \partial m(t)$$

The bias term $A_1(t)$ satisfies

$$\begin{aligned} A_1(t) &= \frac{1}{K+1}\sum_{k=0}^{K}\left[m(t_k) - m(t) - (t_k - t)\partial m(t)\right]h_m^{-1}K_m \\ &\quad\times\left(h_m^{-1}(t_k - t)\right) \\ &= O(h_m^2), \quad \text{(w.p.1) uniformly on } 0 \leq \text{t} \leq 1 \end{aligned}$$

where Assumption 4.2, Lemma 2 and Assumption 4.3 allows the application of Taylor's expansion. We now investigate $A_2(\cdot)$. Similarly to the proof of Theorem 4.3, we have $\mathbb{E}A_2(t) = 0$, for all $t \in [0,1]$. This implies

$$\begin{aligned} \left|A_2(t) - \mathbb{E}A_2(t)\right| &\leq O(h_Q^{-1}) \times \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|R_i(t_k) - m(t_k)\right| \\ &\quad\times \mathbb{I}\left\{t_k \in [t - h_m, t + h_m]\right\} \end{aligned}$$

So, it is enough to obtain

$$\begin{aligned} S(N,K) &:= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|R_i(t_k) - m(t_k)\right| \\ &\quad\times \mathbb{I}\left\{t_k \in [t - h_m, t + h_m]\right\} \\ &= O\left(\frac{h_m^2 \log N}{N}\right)^{\frac{1}{2}}, \text{ uniformly on } 0 \leq t \leq 1 \end{aligned}$$

For this purpose, we decompose $S(N,L)$ into two terms:

$$\begin{aligned} S(N,K) &= \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|R_i(t_k) - m(t_k)\right| \\ &\quad\times \mathbb{I}\left\{\left|R_i(t_k) - m(t_k)\right| > U_{N,K}\right\}\mathbb{I}\left\{t_k \in [t - h_m, t + h_m]\right\} \\ &\quad+ \frac{1}{N(K+1)}\sum_{i=1}^{N}\sum_{k=0}^{K}\left|R_i(t_k) - m(t_k)\right| \\ &\quad\times \mathbb{I}\left\{\left|R_i(t_k) - m(t_k)\right| \leq U_{N,K}\right\} \\ &\quad\times \mathbb{I}\left\{t_k \in [t - h_m, t + h_m]\right\} \\ &=: B_1 + B_2 \end{aligned}$$

where

$$U_{N,K} := h_m\left(\frac{\log N}{N}\right)^{-\frac{1}{2}}$$

and choose $h_m \to 0$ such that

$$U_{N,K}^{1-\alpha} = O\left(\frac{h_m^2 \log N}{N}\right)^{\frac{1}{2}} \qquad \text{that is} \qquad \frac{1}{h_m^2}\left(\frac{\log N}{N}\right)^{1-\frac{2}{\alpha}} = O(1) \tag{59}$$

for some $\alpha > 2$ satisfying the statement of Theorem 4.4. This implies

$$
\begin{aligned}
B_1 &\le \frac{1}{N(K+1)} \sum_{i=1}^{N} \sum_{k=0}^{K} \left| R_i(t_k) - m(t_k) \right|^{1-\alpha+\alpha} \\
&\quad \times \mathbb{I}\left\{ \left| R_i(t_k) - m(t_k) \right| > U_{N,K} \right\} \\
&\le \frac{1}{N(K+1)} \sum_{i=1}^{N} \sum_{k=0}^{K} \left| R_i(t_k) - m(t_k) \right|^{\alpha} U_{N,K}^{1-\alpha} \\
&= O(1) U_{N,K}^{1-\alpha}, \qquad \text{w.p.1 (uniformly on } 0 \le t \le 1) \qquad (60) \\
&= O\left(\frac{h_m^2 \log N}{N} \right)^{\frac{1}{2}}, \quad \text{w.p.1 (uniformly on } 0 \le t \le 1) \quad (61)
\end{aligned}
$$

where (60) is a consequence of Assumption 4.7, for some $\alpha > 2$, combined with Theorem 4.1. Equality (61) is a consequence of (59). We now turn to B_2. We first obtain an upper bound on the variance of the inner sum:

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{1}{K+1} \sum_{k=0}^{K} \left| R_i(t_k) - m(t_k) \right| \mathbb{I}\left\{ t_k \in [t - h_m, t + h_m] \right\} \right) \\
&\le \frac{1}{(K+1)^2} \sup_{t \in [0,1]} \mathbb{E} \left| R_i(t) - m(t) \right|^2 \left(2h_m(K+1) \right)^2 \\
&\le c h_m^2 \qquad (62)
\end{aligned}
$$

where (62) is a consequence of Theorem 4.1, which is in turn a consequence of Assumption 4.7, for some $\alpha > 2$. Applying Bernstein's inequality, we conclude, for any positive number η,

$$
\begin{aligned}
\mathbb{P}\left(B_2 \ge \eta \left(\frac{h_m^2 \log N}{N} \right)^{\frac{1}{2}} \right) &= \mathbb{P}\left(N B_2 \ge N\eta \left(\frac{h_m^2 \log N}{N} \right)^{\frac{1}{2}} \right) \\
&\le \exp\left\{ - \frac{\eta^2 N^2 \left(\frac{h_m^2 \log N}{N} \right)}{2Nch_m^2 + \frac{2}{3}\eta N h_m^2} \right\} \\
&= \exp\left\{ - \frac{\eta^2 \log N}{2c + \frac{2}{3}\eta} \right\} \\
&= N^{-\frac{\eta^2}{8c + \frac{2}{3}\eta}}
\end{aligned}
$$

The rest of the argument follows the lines of the proof of Theorem 4.3. □

Lemma 3. *If, in addition to the assumptions of Theorem 3.1, Assumptions 4.1 and 4.2 hold, then the variance function $D(\cdot)$ is twice continuously differentiable, that is, $D(\cdot) \in C^2([0,1])$.*

Proof of Lemma 3. The lemma is a consequence of representation (9) that can be rewritten as

$$
\begin{aligned}
D(t) &= \exp\left\{ -2 \int_0^t \theta(u) du \right\} D(0) \\
&\quad + \int_0^t \exp\left\{ -2 \int_s^t \theta(u) du \right\} \sigma^2(s) ds \\
&= \exp\left\{ -2 \int_0^t \theta(u) du \right\} D(0) + \exp\left\{ -2 \int_0^t \theta(u) du \right\} \\
&\quad \times \int_0^t \exp\left\{ 2 \int_0^s \theta(u) du \right\} \sigma^2(s) ds
\end{aligned}
$$

□

Proof of Theorem 4.5. A slight modification of the proof of Theorem 4.4 entails the desired result. The solution (25) admits the closed from representation

$$
\begin{bmatrix} \hat{G}(t) \\ h_G \widehat{\partial G}(t) \end{bmatrix} = \begin{bmatrix} \mathcal{A}_0 & \mathcal{A}_1 \\ \mathcal{A}_1 & \mathcal{A}_2 \end{bmatrix}^{-1} \begin{bmatrix} S_0 \\ S_1 \end{bmatrix} (t)
$$

where for $0 \le t \le 1$,

$$
\mathcal{A}_p(t) = \frac{1}{K+1} \sum_{k=0}^{K} \left(\frac{t_k - t}{h_G} \right)^p h_G^{-1} K_G \left(h_G^{-1}(t_k - t) \right), \quad p = 0, 1, 2
$$

and

$$
\begin{aligned}
S_p(t) &= \frac{1}{N(K+1)} \sum_{k=0}^{K} \overline{R^2}(t_k) \left(\frac{t_k - t}{h_G} \right)^p h_G^{-1} K_G \left(h_G^{-1}(t_k - t) \right) \\
&= \frac{1}{N(K+1)} \sum_{i=1}^{N} \sum_{k=0}^{K} R_i^2(t_k) \left(\frac{t_k - t}{h_G} \right)^p h_G^{-1} K_G \left(h_G^{-1}(t_k - t) \right), \\
&\qquad p = 0, 1
\end{aligned}
$$

Compared to Theorem 4.4, the higher-order moment condition $\alpha > 4$ is required to guarantee the inequality

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{1}{K+1} \sum_{k=0}^{K} \left| R_i^2(t_k) - m(t_k) \right| \mathbb{I}\left\{ t_k \in [t - h_m, t + h_m] \right\} \right) \\
&\le \frac{1}{(K+1)^2} \sup_{t \in [0,1]} \mathbb{E} \left| R_i^2(t) - G(t) \right|^2 \left(2h_m(K+1) \right)^2 \le c h_m^2
\end{aligned}
$$

as a counterpart of (62). The rest of the argument is similar. □

Proof of Theorem 4.6. Combining Theorems 4.4 and 4.5, we obtain the claim. □

Proofs of Theorems 4.7, 4.8 and 4.9. These theorems are consequences of Theorems 4.3–4.6. □

Acknowledgments

This research was partially supported by the United States National Science Foundation grant DMS–2412408. We thank an anonymous reviewer for reading the original submission carefully and helping us improve the paper.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings will be available in Blumberg Data at https://data.bloomberg.com/ following an embargo from the date of publication to allow for commercialization of research findings.

References

Brockwell, P., R. Davis, and Y. Yang. 2007. “Estimation for Nonnegative Lévy-Driven Ornstein-Uhlenbeck Processes.” *Journal of Applied Probability* 44: 977–989.

Brockwell, P., R. Davis, and Y. Yang. 2011. “Estimation for Non-Negative Lévy-Driven CARMA Processes.” *Journal of Business & Economic Statistics* 29: 250–259.

Brockwell, P., and A. Lindner. 2009. “Existence and Uniqueness of Stationary Lévy-Driven CARMA Processes.” *Stochastic Processes and their Applications* 119: 2660–2681.

Cai, C., T. De Angelis, and J. Palczewski. 2022. “The American Put With Finite-Time Maturity and Stochastic Interest Rate.” *Mathematical Finance* 32: 1170–1213.

Chen, J., M. Davison, M. Escobar-Anel, and G. Zafari. 2020. “Robust Portfolios With Commodities and Stochastic Interest Rates.” *Quantitative Finance* 21: 991–1010.

Cialenco, I. 2018. “Statistical Inference for SPDEs: An Overview.” *Statistical Inference for Stochastic Processes* 21: 309–329.

Comte, F., and V. Genon-Catalot. 2020. “Nonparametric Drift Estimation for i.i.d. Paths of Stochastic Differential Equations.” *Annals of Statistics* 48: 3336–3365.

Fan, J., and I. Gijbels. 1996. *Local Polynomial Modelling and Its Applications*. Chapman & Hall/CRC.

Filipovic, D. 2009. *Term-Structure Models: A Graduate Course*. Springer.

Horváth, L., and P. Kokoszka. 2012. *Inference for Functional Data With Applications*. Springer.

Hsing, T., and R. Eubank. 2015. *Theoretical Foundations of Functional Data Analysis, With an Introduction to Linear Operators*. Wiley.

Jiang, H., Y. Pan, and X. Wei. 2023. “Self-Normalized Cramér-Type Moderate Deviations for Explosive Vasicek Model.” *Journal of Theoretical Probability* 37, no. 1: 223–228.

Kharrat, M., and H. Arfaoui. 2023. “A New Stabled Relaxation Method for Pricing European Options Under the Time-Fractional Vasicek Model.” *Computational Economics* 61: 1745–1763.

Kokoszka, P., N. Mohammadi, H. Wang, and S. Wang. 2025. “Functional Diffusion Driven Stochastic Volatility Model.” *Bernoulli* 31: 922–947.

Le Gall, J.-F. 2016. *Brownian Motion, Martingales, and Stochastic Calculus*. Springer.

Lindner, A., and K.-I. Sato. 2009. “Continuity Properties and Infinite Divisibility of Stationary Distributions of Some Generalized Ornstein-Uhlenbeck Processes.” *Annals of Probability* 37: 250–274.

Liu, X., H. Xiao, and R. Chen. 2016. “Convolutional Autoregressive Models for Functional Time Series.” *Journal of Econometrics* 194: 263–282.

Mehrdoust, F., and A. R. Najafi. 2020. “A Short Memory Version of the Vasicek and Evaluating European Options on Zero-Coupon Bonds.” *Journal of Computational and Applied Mathematics* 375, no. 112: 796.

Mohammadi, N., L. Santoro, and V. Panaretos. 2024. “Nonparametric Estimation for SDE With Sparsely Sampled Paths: An FDA Perspective.” *Stochastic Processes and their Applications* 167, no. 104: 239.

Nourdin, I., and T. Diu Tran. 2019. “Statistical Inference for Vasicek-Type Model Driven by Hermite Processes.” *Stochastic Processes and their Applications* 129: 3774–3791.

Nowman, K. B. 2010. “Modelling the UK and Euro Yield Curves Using the Generalized Vasicek Model: Empirical Results From Panel Data for One and Two Factor Models.” *International Review of Financial Analysis* 19: 334–341.

Øksendal, B. 2003. *Stochastic Differential Equations*. Springer.

Orlando, G., R. M. Mininni, and M. Bufalo. 2020. “Forecasting Interest Rates Through Vasicek and CIR Models: A Partitioning Approach.” *Journal of Forecasting* 39: 569–579.

Rodrigo, M. R., and R. S. Mamon. 2014. “An Alternative Approach to the Calibration of the Vasicek and CIR Interest Rate Models via Generating Functions.” *Quantitative Finance* 14: 1961–1970.

Vašiček, O. 1977. “An Equilibrium Characterization of the Term Structure.” *Journal of Financial Economics* 5: 177–188.

Wu, X., D. Xie, and D. A. Edwards. 2019. “An Optimal Mortgage Refinancing Strategy With Stochastic Interest Rate.” *Computational Economics* 53: 1353–1375.

Xiao, W., and J. Yu. 2019. “Asymptotic Theory for Estimating Drift Parameters in the Fractional Vasicek Model.” *Econometric Theory* 35: 198–231.

Xiao, W., W. Zhang, X. Zhang, and X. Chen. 2014. “The Valuation of Equity Warrants Under the Fractional Vasicek Process of the Short-Term Interest Rate.” *Physica A* 394: 320–327.

Yoshihara, K. 1978. “Moment Inequalities for Mixing Sequences.” *Kodai Mathematical Journal* 1: 316–328.

Zhou, Y., and H.-G. Müller. 2025. “Dynamic Modelling of Sparse Longitudinal Data and Functional Snippets With Stochastic Differential Equations.” *Journal of the Royal Statistical Society (B)* 87: 833–849.