Functional diffusion driven stochastic volatility model

Functional diffusion driven stochastic volatility model

PIOTR KOKOSZKA[1,a], NEDA MOHAMMADI[1,b], HAONAN WANG[1,c] and SHIXUAN WANG[2,d]

[1]*Department of Statistics, Colorado State University, USA,* [a]*piotr.kokoszka@colostate.edu,* [b]*neda.mohammadi@colostate.edu,* [c]*haonan.wang@colostate.edu*
[2]*Department of Economics, University of Reading, UK,* [d]*shixuan.wang@reading.ac.uk*

We propose a stochastic volatility model for time series of curves. It is motivated by dynamics of intraday price curves that exhibit both between days dependence and intraday price evolution. The curves are suitably normalized to stationary in a function space and are functional analogs of point-to-point daily returns. The between curves dependence is modeled by a latent autoregression. The within curves behavior is modeled by a diffusion process. We establish the properties of the model and propose several approaches to its estimation. These approaches are justified by asymptotic arguments that involve an interplay between the latent autoregression and the intraday diffusions. The asymptotic framework combines the increasing number of daily curves and the refinement of the discrete grid on which each daily curve is observed. Consistency rates for the estimators of the intraday volatility curves are derived as well as the asymptotic normality of the estimators of the latent autoregression. The estimation approaches are further explored and compared by an application to intraday price curves of over seven thousand U.S. stocks and an informative simulation study.

Keywords: Functional time series; intraday price curves; Itô diffusion process; stochastic volatility

1. Introduction

Time dependent volatility is one of the main features of financial time series. Diffusion models for price evolution have been employed for over one hundred years, starting perhaps with Bachelier (1900), with a robust development since the 1970s. This paper proposes a model that combines a diffusion model for the intraday price evolution with a stochastic volatility paradigm for day-to-day dependence. We combine the tools of time series analysis, functional data analysis and stochastic calculus. We develop estimation theory for a functional time series model of the general form $R_i(t) = g_i Y_i(t)$, where the curves Y_i are iid and the scalars g_i form a scalar time series. Multiplicative functional time series have been studied within the ARCH paradigm, references are given below, but not in the context of the stochastic volatility paradigm. Our objective is to develop theoretical foundations of the estimation of the functional stochastic volatility models.

The concept of univariate conditional heteroscedasticity traces back at least to Engle (1982) who proposed autoregressive conditional heteroscedastic (ARCH) models and the influential contribution of Bollerslev (1986) who proposed generalized autoregressive conditional heteroscedastic (GARCH) models. Later on, Bollerslev, Engle and Wooldridge (1988), Bollerslev (1990), Engle and Kroner (1995), Engle (2002), among many others, utilized GARCH models in analysing multivariate heteroscedastic time series. An overview of multivariate GARCH models is provided in Bauwens, Laurent and Rombouts (2006) and Silvennoinen and Teräsvirta (2009). Bayesian inference for multivariate GARCH models is addressed in Vrontos, Dellaportas and Politis (2003). Hörmann, Horváth and Reeder (2013) study conditional heteroscedasticity in the framework of functional data analysis (FDA). In particular, they propose a functional version of the ARCH model, which is extended to functional GARCH

models by Aue, Horváth and Pellatt (2017). Cerovecki et al. (2019) proposed a quasi–likelihood inference approach for functional GARCH models as an alternative to the least square method. The common feature of all above models is that the random volatility is measurable with respect to *past events*, i.e. conditionally on *past observations* the volatility process is deterministic. To illustrate this point, in the basic univariate ARCH(1), model

$$\begin{cases} r_i = g_i w_i, \quad w_i \sim \text{ iid } N(0,1); \\ g_i^2 = \eta_1 r_{i-1}^2 + \eta_0, \end{cases}$$

the current volatility g_i^2 is a function of the previous observation r_{i-1}. More complex functions lead to various models in the ARCH family. An alternative approach to model the randomness in volatility is the so called stochastic volatility (SV), where in contrast to ARCH-type models, conditionally on past events the volatility process is not fully observable. The simplest univariate formulation of stochastic volatility is

$$\begin{cases} r_i = g_i w_i, \quad w_i \sim \text{ iid } N(0, \sigma_w^2); \\ g_i = \exp\{\varphi \log g_{i-1} + \varepsilon_i\}, \quad \varepsilon_i \sim \text{iid } N(0, \sigma_\varepsilon^2), \end{cases} \tag{1}$$

where φ is a parameter satisfying $|\varphi| < 1$. To draw an analogy to linear time series models, GARCH models are analogous to ARMA models, while the stochastic volatility models are analogous to state space models. The stochastic volatility model was introduced by Taylor (1982) and further developed by Taylor (1994), Shephard (1996) and Taylor (2008). Harvey, Ruiz and Shephard (1994), Danielsson (1998), and Asai, McAleer and Yu (2006) extend the stochastic volatility models to multivariate settings. A Bayesian approach to multivariate stochastic volatility models is investigated in Yu and Meyer (2006). There have been thousands of contributions to ARCH and SV modeling, and a number of excellent monographs have been published. We listed only selected papers emphasizing those dealing with multivariate models. We propose a *functional* SV model. Looking at our contribution from the angle of functional data analysis, we advance the theory of functional time series by studying a multiplicative model. Linear functional time series have been studied even before the publication of the monograph of Bosq (2000) and there is still great interest in them, see e.g. Kuenzer (2024) for a recent contribution. Nonlinear functional models with additive errors have also been studied, see e.g. Paparoditis and Shang (2023) and references therein.

In its simplest form, our model is given by equations analogous to equations (1):

$$\begin{cases} R_i(t) = g_i \int_0^t \sigma(u) dW_i(u), \quad t \in [0,1], \\ \log g_i = \varphi \log g_{i-1} + \varepsilon_i, \quad \varepsilon_i \sim \text{ iid } \mathcal{WN}(0, \sigma_\varepsilon^2), \quad i \in \mathbb{Z}, \end{cases} \tag{2}$$

where $\mathcal{WN}(0, \sigma_\varepsilon^2)$ denotes white noise with mean zero and finite second moment σ_ε^2. The $W_i(\cdot)$s are independent standard Brownian motions independent of the error sequence $\{\varepsilon_i\}$. The random coefficients g_i are positive with probability one, $\sigma(\cdot)$ is a nonparametric function with $\|\sigma\|_\infty < \infty$ and φ is a scalar with $|\varphi| < 1$, see Section 3 for a full discussion. Setting $t = 1$ in (2), we obtain

$$R_i(1) = g_i \int_0^1 \sigma(u) dW_i(u),$$

retrieving the univariate model (1) with $r_i = R_i(1)$, $w_i = \int_0^1 \sigma(u) dW_i(u)$ and $\sigma_w^2 = \int_0^1 \sigma^2(u) du$. This property shows that the proposed model (2) extends the well established univariate model (1).

The data that motivate model (2) are intraday price curves suitably transformed to form a stationary sequence of curves. Detailed definitions are given in Section 5. Basically, $R_i(t)$ is the cumulative return

on trading day i up to intraday time t. If the exchange opening time is rescaled to the unit interval, $R_i(1)$ is the return on day i, except that we compare the closing price to the opening price, rather than to the closing price on the previous day. The AR(1) formulation appearing in the second equation in (2) models the dependence structure between the daily curves. The strength of this dependence is quantified through φ, or more parameters, as explained in Section 4. Müller, Sen and Stadtmüller (2011) proposed a framework that uses diffusions to model volatility, but assumed i.i.d curves. Uncorrelated stochastic volatility curves are also studied in Jang, Jauch and Matteson (2022) who apply dimension reduction through basis expansions. This technique, however, ignores the roughness of the sample paths which is a crucial property for price processes; continuous time price models are diffusions with nowhere differentiable paths. Chong et al. (2022) consider the roughness to be a crucial feature of financial continuous time models and propose stochastic volatility models driven by fractional Brownian motion with the Hurst index $H < 1/2$. Their inference targets the Hurst index H and they establish a minimax theory for this parameter. Their work completes the results of Rosenbaum (2008) that focus on a similar inference problem with $H > 1/2$. The between curve dynamic dependence is not addressed in Rosenbaum (2008) and Chong et al. (2022) because they consider a single time interval. Our model incorporates the non-differentiability of the trajectories with temporal dependence between curves, which is widely believed to exist, see e.g. Christensen, Hounyo and Podolskij (2018). Our approach is somewhat related to the work of Visser (2011) who showed that the estimation of GARCH models can be improved if intraday volatility is employed.

We model the dependence between the curves through a latent autoregression, roughness of sample paths through a stochastic integral, and intraday volatility through the product $g_i\sigma(\cdot)$, where the g_i exhibit day-to-day dependence. The intensity of between–curve dependence, governed by the autoregressive model, is tuned by the parameter φ. Independently of this component, the intensity of the within curve dependence, modeled through the Itô integral, is quantified by the function $\sigma(\cdot)$. From now on, we call the deterministic component $\sigma(\cdot)$ the diffusion function while keeping in mind that the stochastic diffusion function is $g_i\sigma(\cdot)$. Our main objective is to establish conditions for the existence of solutions to model equations and develop inference for model parameters, which include the autoregressive parameters and the function $\sigma(\cdot)$. It is thus a parametric-nonparametric estimation problem that involves challenges not encountered in previous research, as explained in Section 3. Estimation must take into account the fact that the daily trajectories are observed at discrete time points within a day. The availability of replications indexed by i (day) suggests to approach the problem from a functional data analysis perspective. However, the roughness of the trajectories makes established FDA approaches that assume smoothness less attractive. There is comprehensive research on rough trajectories, with a typical model assuming that smooth trajectories are observed with i.i.d. randomly scattered measurement errors, see Yao, Müller and Wang (2005), Hall, Müller and Wang (2006) and Li and Hsing (2010), among many others. These approaches are effective for biomedical data. For continuous time price data, roughness is however a fundamental property modeled by diffusions rather than caused by completely random errors. Our strategy is built on approximating the (latent) quadratic variation process that usually satisfies some desired regularity property. Utilizing the approximating quadratic variation processes in lieu of the hidden curves paves the way to employing FDA techniques. Consistency of the realized quadratic variation process and hence estimated objects are addressed under decay of step size, Δ say, and growth of sample size N. Galbraith, Zinde-Walsh and Zhu (2015) study GARCH models and propose an argument based on the realized quadratic variation process. Due to the nature of the GARCH model, Galbraith, Zinde-Walsh and Zhu (2015) do not require the consistency of the realized quadratic variation process and treat the induced discrepancy as a noise term. This substantially differentiates their approach from the current study.

The remainder of the paper is organized as follows. In Section 2, for ease of reference and to fix notation and terminology, we present relevant information related to stochastic integrals. Section 3 is

dedicated to the explanation of the proposed model and estimation approaches in the simplest case of the latent autoregression of order 1, which already contains key model and estimation features, and allows us to focus on them. Extension to AR(p) latent autoregression is developed in Section 4. An Application to U.S. intraday stock prices and a simulation study are presented in Section 5. We conclude in Section 6 with a brief summary and discussion of directions for future research. Online Supplementary Material, Kokoszka et al. (2024), contains the proofs of the results of Sections 3 and 4, as well as additional information related to the empirical analysis in Section 5.

2. Preliminaries

For ease of reference, we collect in this section useful facts related to Itô diffusion processes. More details are provided in the monographs of Øksendal (2003), Aït-Sahalia and Jacod (2014), and Karatzas and Shreve (1991), among others. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space. We define the $\mathbb{R}$–valued Itô diffusion (diffusion in short) process by

$$\begin{aligned} dX(t) =& \mu(t, X(t))dt + \sigma(t, X(t))dW(t), \qquad 0 < t, \\ X(t) =& X(0), \qquad t = 0, \end{aligned} \tag{3}$$

where W denotes a standard Brownian motion. Integrals with respect to W should be understood in the sense of the stochastic Itô integral. We assume that the initial distribution is independent of the σ-algebra $\mathcal{F}_\infty$ generated by $\{W(t)\}_{t\geq 0}$. The function $\mu(\cdot,\cdot)$, the so called drift (viscosity), and the function $\sigma(\cdot,\cdot)$, the so called diffusion (volatility), are Borel measurable. The following theorem, a consequence of Theorem 5.2.1 in Øksendal (2003), provides sufficient conditions for 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 functions $\mu(\cdot,\cdot) : [0,T] \times \mathbb{R} \longmapsto \mathbb{R}$ and $\sigma(\cdot,\cdot) : [0,T] \times \mathbb{R} \longmapsto \mathbb{R}$ 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 Lipschitz continuity in the space variable i.e.

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

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,*

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

We now present the Itô isometry that is one of the most useful results in the context of stochastic calculus. Define $\mathcal{V} = \mathcal{V}(S,T)$ to be the class of functions

$$f : [0,\infty) \times \Omega \longrightarrow \mathbb{R},$$

satisfying

(i) $(t,\omega) \mapsto f(t,\omega)$ is $\mathcal{B}\times\mathcal{F}$ measurable, where $\mathcal{B}$ denotes the Borel σ–algebra on $[0,\infty)$,
(ii) $f(t,\cdot)$ is $\mathcal{F}_t$ adapted, where $\mathcal{F}_t$ is the σ–algebra generated by $\{W(s)\}_{s\le t}$,
(iii) $\mathbb{E}[\int_S^T |f(t,\omega)|^2 dt] < \infty$.

If $f \in \mathcal{V}$, then, according to Corollary 3.1.7 in Øksendal (2003),

$$\mathbb{E}\left[\int_S^T f(t,\omega)dW(t)\right]^2 = \mathbb{E}\left[\int_S^T |f(t,\omega)|^2 dt\right].$$

The quadratic variation process and its empirical counterpart, also known as realized quadratic variation process, play a fundamental role in the study of Itô semimartingales. Our inferential procedure and consequently our consistency results heavily rely on calculation of realized quadratic variation and its convergence to the true process. Let $X(\cdot)$ satisfy model (3) and $\{t_k\}$ be an equispaced partition of the unit interval with step size Δ. The realized quadratic variation process at point t is defined through the following sum of squared increments:

$$\sum_k |X(t_k) - X(t_{k-1})|^2 \mathbb{I}\{t_k \le t\}, \quad t \in [0,1]. \tag{6}$$

The above sum is tightly related to the quadratic variation process which will be denoted by $\langle X,X\rangle_t$, $t \in [0,1]$. According to Proposition 3.2.17 in Karatzas and Shreve (1991), under the conditions of Theorem 2.1, the process $\langle X,X\rangle_t$ can be defined by

$$\langle X,X\rangle_t = \int_0^t \sigma^2(u,X(u))du, \quad t \in [0,1]. \tag{7}$$

We are now ready to present Theorem 2.2 which examines convergence of the realized quadratic variation to the theoretical counterpart $\langle X,X\rangle_t$, see Theorem 1.14 and relation (3.23) in Aït-Sahalia and Jacod (2014) for a more general statement.

Theorem 2.2. *Assume the conditions of Theorem 2.1, and let $\Pi_N = \{t_{N,k}\}$ be a sequence of partitions of the unit interval with step size $\Delta(N)$ that tends to zero, as N increases. Then the realized quadratic variation process tends to the quadratic variation process uniformly in probability, i.e. as $\Delta(N) \longrightarrow 0$*

$$\sup_{0\le t\le 1}\left|\sum_k |X(t_{N,k}) - X(t_{N,k-1})|^2 \mathbb{I}\{t_{N,k} \le t\} - \int_0^t \sigma^2(u,X(u))du\right| \xrightarrow{\mathbb{P}} 0, \tag{8}$$

where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability.

We close this section by stating a time change result as a corollary of Dambis–Dubins–Schwarz theorem which expresses *any* continuous local martingale as a time change of a Brownian motion, see e.g. 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 $X(t)$ has the same distribution as $W(\langle X,X\rangle_t)$, i.e.

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

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

3. Development and estimation of order 1 model

In this section, we focus on the model defined in the Introduction by equations (2). It already contains the most essential elements of the proposed framework and the important issues related to its properties and estimation are easier to explain. An extension to higher order latent autoregressions is presented in Section 4. For ease of reference, we display equations (2) as

$$R_i(t) = g_i \int_0^t \sigma(u) dW_i(u), \quad t \in [0,1], \quad i \in \mathbb{Z}, \tag{9}$$

$$\log g_i = \varphi \log g_{i-1} + \varepsilon_i, \quad \varepsilon_i \sim \text{ iid } \mathcal{WN}(0, \sigma_\varepsilon^2). \tag{10}$$

We will use the following assumptions. Not all of them are needed for every result, as specified in the following, but all results are valid if all conditions listed below hold.

1. The function $\sigma(\cdot)$ is nonnegative and deterministic with $\sup_{0 \le t \le 1} \sigma(t) = \|\sigma\|_\infty < \infty$,
2. the scalar coefficients g_i are nonnegative with probability one. This assumption is equivalent to setting $g_i = \exp(x_i)$, for a real–valued random sequence $\{x_i\}$, and formulating the AR(1) model (10) in terms of $\{x_i\}$,
3. the random processes $W_i(\cdot)$ are independent standard Brownian motions (Wiener processes),
4. the sequences $\{W_i\}$ and $\{g_i\}$ are independent,
5. the autoregressive coefficient φ in (10) satisfies $|\varphi| < 1$,
6. The function $\sigma(t)$, $t \in [0,1]$, is non–zero almost everywhere with respect to Lebesgue measure, i.e. $\mathcal{L}eb\{t : \sigma(t) = 0\} = 0$, where $\mathcal{L}eb$ denotes the Lebesgue measure restricted to the unit interval $[0,1]$.
7. $\mathbb{E}(\varepsilon_0^4) = \eta \sigma_\varepsilon^2 < \infty$.

REMARK 3.1. Assumptions 1 and 6 together imply that there is no subinterval of $[0,1]$ on which the volatility $\sigma(\cdot)$ is infinity or zero. Assumptions 6 also implies that for any fixed $\alpha \in (0,1)$, function $G(t) = \int_0^t \sigma^2(u)du$ is bounded away from zero on the restricted domain $t \in [\alpha, 1]$:

$$\inf_{\alpha \le t \le 1} \int_0^t \sigma^2(u)du \ge \int_0^\alpha \sigma^2(u)du > 0. \tag{11}$$

The proofs of Lemma A.1. and Corollary A.1. show that a larger α may lead to more precise estimators, at the expense of reducing the range of estimation to the interval $[\alpha, 1]$. We comment on this point further in Remark 3.2.

The proposed model (9)–(10) decomposes the full random behavior of the curves R_i into the between curves dynamics quantified by the g_i and the within curve dynamics described by the stochastic integrals $\int_0^t \sigma(u)dW_i(u)$, $t \in [0,1]$. The between curves dynamics is regulated by the dependence between the g_is that replaces the independence of curves assumption used in previous research discussed in the Introduction. The within curve dynamics is expressed in terms of the diffusion $\int_0^t \sigma(u)dW_i(u)$, which models the roughness of the trajectories.

Our first theorem establishes the existence and uniqueness of a strictly stationary functional sequence R_i satisfying (9)–(10), as well as the identifiability of model components. We present the proof here because it is short and we use in the following the relations it contains. For the sake of compactness,

we denote $\langle R_i, R_i\rangle_t$ by $Q_i(t)$ from now on. According to (7),

$$Q_i(t) := \langle R_i, R_i\rangle_t = g_i^2 \int_0^t \sigma^2(u)du = g_i^2 \exp\{H(t)\}, \quad t \in (0,1], \quad i = 1,2,\ldots,N, \tag{12}$$

where

$$H(t) = \log \int_0^t \sigma^2(u)du = \log G(t), \qquad t \in (0,1]. \tag{13}$$

Theorem 3.1. *Suppose conditions 1–5 hold. Then, there exists a unique strictly stationary functional sequence R_i satisfying* (9)–(10). *Moreover, the scalar g_i is distinguishable from the function σ in* (10).

Proof. First observe that condition 5, the finiteness of the second moment of the ε_i and their iid property imply the existence, the uniqueness and the strict stationarity of the sequence $\{\log g_i\}$ satisfying (10) and hence the same properties of the sequence $\{g_i\}$, cf. condition 2.

We now argue that condition 1 implies existence and uniqueness of the Itô integral $\int_0^t \sigma(u)dW_i(u)$, for each i. This follows from Theorem 2.1 because conditions (4) and (5) hold. Equation (9) thus defines the functional sequence $\{R_i\}$ directly.

Observe that the integrated volatility $\int_0^t \sigma^2(u)du$ appearing in (12) is identifiable from g_i^2 if and only if the term $\log \int_0^t \sigma^2(u)du$ is identifiable in the sum

$$\log Q_i(t) = 2\log g_i + \log \int_0^t \sigma^2(u)du = 2\log g_i + H(t). \tag{14}$$

The summands appearing in (14) are identifiable due to the assumption $\mathbb{E}(\varepsilon_i) = 0$ which implies $\mathbb{E}(\log g_i) = 0$. Thus, $H(t)$ is identifiable via

$$\mathbb{E}\log Q_i(t) = H(t). \tag{15}$$

□

We now turn to estimation. We assume one has access to discrete observations $R_i(t_{i,k})$, $i = 1,2,\ldots,N$, $k = 0,1,\ldots,m$. We assume that the points $t_{i,k}$ are the same for all curves, that is $t_{i,k} = t_k = \Delta k$, for some positive Δ that decays to zero. We also assume $t_0 = 0$ and $t_m = 1$. We aim to develop inference for the vector

$$\boldsymbol{\theta} = [G(\cdot), \varphi, \sigma_\varepsilon^2], \tag{16}$$

with the function $G(\cdot)$ is defined in (13). We focus on the function G because it is sufficiently smooth, whereas under our general assumptions, σ can be basically any measurable function. To recover σ from G we need to assume that σ is continuous. It can then be computed by differentiating G numerically.

The dense (high frequency) sampling regime of observations suggest working with the quadratic variation processes $Q_i(t)$, $t \in [0,1]$, which can be estimated in such a framework. Representation (14) paves the way to explaining the idea behind our inferential procedure. Assume one has access to the latent quadratic variation processes $\{Q_i(t), t \in (0,1]\}$, for $i = 1,2,\ldots,N$. Equation (14) together with (10) leads to the family of AR(1) models, indexed by time t,

$$\log Q_i(t) - H(t) = \varphi\,[\log Q_{i-1}(t) - H(t)] + 2\varepsilon_i. \tag{17}$$

Each AR(1) model is defined by the time index i, and we have a family of such models indexed by continuous time t. However, by (10), these models share the common innovation terms ε_i which establish a connection between them that is explored in our estimation procedure. Now, one can apply any of the well–known estimation techniques, see for example Chapter 8 of Brockwell and Davis (1991), to obtain the oracle estimates (oracle only because the $Q_i(t)$ are not observable). To focus on a specific simple approach, we set, cf. (15),

$$\widetilde{H}(t) = \frac{1}{N}\sum_{i=1}^{N} \log Q_i(t), \qquad t \in (0,1], \tag{18}$$

or equivalently

$$\widetilde{G}(t) = \exp\left(\frac{1}{N}\sum_{i=1}^{N} \log Q_i(t)\right), \qquad t \in (0,1].$$

Next, we define, respectively, the lag zero and lag one oracle empirical autocovariances

$$\begin{aligned}\gamma_{0,N} =& \frac{1}{4N}\sum_{i=1}^{N}\left(\log Q_i(t) - \frac{1}{N}\sum_{i=1}^{N}\log Q_i(t)\right)^2 \\ =& \frac{1}{N}\sum_{i=1}^{N}\left(\log g_i - \frac{1}{N}\sum_{i=1}^{N}\log g_i\right)^2.\end{aligned} \tag{19}$$

and

$$\begin{aligned}\gamma_{1,N} =& \frac{1}{4N}\sum_{i=1}^{N-1}\left(\log Q_i(t) - \frac{1}{N}\sum_{i=1}^{N}\log Q_i(t)\right)\left(\log Q_{i+1}(t) - \frac{1}{N}\sum_{i=1}^{N}\log Q_i(t)\right) \\ =& \frac{1}{N}\sum_{i=1}^{N-1}\left(\log g_i - \frac{1}{N}\sum_{i=1}^{N}\log g_i\right)\left(\log g_{i+1} - \frac{1}{N}\sum_{i=1}^{N}\log g_i\right).\end{aligned} \tag{20}$$

We now define the oracle Yule-Walker estimators

$$\tilde{\varphi} = \gamma_{0,N}^{-1}\gamma_{1,N}, \qquad \tilde{\sigma}_{\varepsilon}^2 = \gamma_{0,N} - \tilde{\varphi}\gamma_{1,N}.$$

Altogether, we propose the following vector of oracle estimators

$$\tilde{\boldsymbol{\theta}} = \left[\widetilde{G}(t), t \in [0,1], \quad \tilde{\varphi}, \quad \tilde{\sigma}_{\varepsilon}^2\right]. \tag{21}$$

Notice, however, that the $\log g_i$ and the quadratic variation processes $\{Q_i(t), t \in [0,1]\}$, for $i = 1, 2, \ldots, N$, are unobservable. Motivated by (8), we replace $\{Q_i(t), t \in [0,1]\}$ by their realized counterparts

$$\begin{aligned}\widehat{Q}_i(t) =& \sum_{k=1}^{m} |R_i(t_k) - R_i(t_{k-1})|^2 \mathbb{I}\{t_k \le t\} \\ =& g_i^2 \sum_{k}\left|\int_{t_{k-1}}^{t_k} \sigma(u)dW_i(u)\right|^2 \mathbb{I}\{t_k \le t\}, \qquad t \in [0,1], \quad i = 1, 2, \ldots N.\end{aligned} \tag{22}$$

We can compute

$$\widehat{H}(t) = \frac{1}{N}\sum_{i=1}^{N} \log \widehat{Q}_i(t), \qquad t \in (0,1], \tag{23}$$

or equivalently

$$\widehat{G}(t) = \exp\left(\frac{1}{N}\sum_{i=1}^{N} \log \widehat{Q}_i(t)\right), \qquad t \in (0,1]. \tag{24}$$

Analogously to (19) and (20), we define

$$\Gamma_{0,N}(t) = \frac{1}{4N}\sum_{i=1}^{N}\left(\log \widehat{Q}_i(t) - \frac{1}{N}\sum_{i=1}^{N}\log \widehat{Q}_i(t)\right)^2, \tag{25}$$

and

$$\Gamma_{1,N}(t) = \frac{1}{4N}\sum_{i=1}^{N-1}\left(\log \widehat{Q}_i(t) - \frac{1}{N}\sum_{i=1}^{N}\log \widehat{Q}_i(t)\right)\left(\log \widehat{Q}_{i+1}(t) - \frac{1}{N}\sum_{i=1}^{N}\log \widehat{Q}_i(t)\right). \tag{26}$$

Notice that in contrast to $\gamma_{0,N}$ and $\gamma_{1,N}$, $\Gamma_{0,N}(t)$ and $\Gamma_{1,N}(t)$ are not necessarily independent of t. Plugging them in the Yule-Walker equations defines a family of estimators, indexed by time t, for the scalars φ and σ_ε^2. In order to address this issue and remove the dependence on t, we propose the following three estimation procedures. Procedures B and C involve averaging over the interval $[\alpha, 1] \subset [0,1]$, for some $0 < \alpha < 1$. Our theory explains that averaging over the whole interval $[0,1]$ is not possible. This is due to the structure of the model and is further elaborated on in Remark 3.2. It turns out, cf. Section 5, that averaging over $[0,1]$ is not possible in practice either. Theorems stated at the end of this section show that all three procedures lead to consistent estimators with the $N^{-1/2}$ convergence rate (asymptotic normality). In Section 5, we investigate which approach works best in finite samples, as well as the effect of the truncation parameter α.

PROCEDURE A: We define

$$\check{\boldsymbol{\theta}} = \left[\widehat{G}(t), t \in (0,1], \quad \check{\varphi}, \quad \check{\sigma}_\varepsilon^2\right], \tag{27}$$

where

$$\check{\varphi} = \Gamma_{0,N}^{-1}(1)\Gamma_{1,N}(1), \qquad \check{\sigma}_\varepsilon^2 = \Gamma_{0,N}(1) - \check{\varphi}\Gamma_{1,N}(1). \tag{28}$$

The estimators (28) are motivated by the fact that the *total* variability in the stochastic volatility model is accumulated at point $t = 1$. The curve $\widehat{G}$ is defined by (24).

PROCEDURE B: In the second procedure, we first average, over $t \in [\alpha, 1]$, the autocovariance functions $\Gamma_{0,N}(t)$ and $\Gamma_{1,N}(t)$, that is for fixed $\alpha \in (0,1)$, we define

$$\bar{\Gamma}_{0,N} := \frac{1}{1-\alpha}\int_\alpha^1 \Gamma_{0,N}(t)dt, \qquad \bar{\Gamma}_{1,N} := \frac{1}{1-\alpha}\int_\alpha^1 \Gamma_{1,N}(t)dt.$$

We then plug the integrated autocovariances in the Yule-Walker equations and propose

$$\bar{\varphi} = \bar{\Gamma}_{0,N}^{-1}\bar{\Gamma}_{1,N}, \qquad \bar{\sigma}_{\varepsilon}^2 = \bar{\Gamma}_{0,N} - \bar{\varphi}\bar{\Gamma}_{1,N}. \tag{29}$$

The above, together with (24) gives the vector of estimates

$$\bar{\boldsymbol{\theta}} = \left[\widehat{G}(t), t \in (0,1], \quad \bar{\varphi}, \quad \bar{\sigma}_{\varepsilon}^2\right]. \tag{30}$$

PROCEDURE C: This procedure is motivated by Yao, Müller and Wang (2005) who, in a different context, obtain a family of estimates, $\{\hat{\sigma}^2(t)\}_{t\in\mathcal{T}}$ say, for a scalar parameter σ^2 and propose $\hat{\sigma}^2 = \frac{1}{|\mathcal{T}|}\int_{\mathcal{T}}\hat{\sigma}^2(t)dt$. Here, we first plug the autocovariance functions $\Gamma_{0,N}(t)$ and $\Gamma_{1,N}(t)$ in the Yule–Walker estimates and then take average over t i.e. we define the integrated estimates

$$\hat{\varphi} = \frac{1}{1-\alpha}\int_{\alpha}^{1}\Gamma_{0,N}^{-1}(t)\Gamma_{1,N}(t)dt, \qquad \hat{\sigma}_{\varepsilon}^2 = \frac{1}{1-\alpha}\int_{\alpha}^{1}\left(\Gamma_{0,N}(t) - \hat{\varphi}\Gamma_{1,N}(t)\right)dt, \tag{31}$$

where α is any fixed positive number. The above, together with (24), produce the final estimator vector

$$\hat{\boldsymbol{\theta}} = \left[\widehat{G}(t), t \in (0,1], \quad \hat{\varphi}, \quad \hat{\sigma}_{\varepsilon}^2\right]. \tag{32}$$

REMARK 3.2. Model (9)–(10) links the parameters φ and σ_{ε}^2 to the process $\{g_i\}$ only, while the observational scheme provides the product $g_i\int_0^t\sigma(u)dW_i(u)$, at discrete times, and does not distinguish g_i and $\int_0^t\sigma(u)dW_i(u)$. This fact is more apparent in the oracle identity (12). To perform inference, we apply the logarithmic function $\log(\cdot)$ to the processes $Q_i(\cdot)$ (in practice to the realized processes $\widehat{Q}_i(\cdot)$) and impose the condition $\mathbb{E}\log g_i = 0$. To establish consistency of this method, we require the function $G(t) = \int_0^t\sigma^2(u)du$ to be bounded away from zero. This forces us to remove the subinterval $[0,\alpha)$ from our analysis when we want to draw inference about $G(\cdot)$ or $H(\cdot)$ or when we want to apply Procedures B or C, see the proofs of Theorems 3.2, 3.4 and 3.5. Procedure A uses $t = 1$ only, where boundedness away from zero holds true due to $G(1) = \int_0^1\sigma^2(u)du > 0$.

The asymptotic properties of the estimates (27), (30) and (32) are addressed under assumptions on the growth of sample size N (the number of curves) and the decay of the step size Δ, see Theorems 3.2–3.5. It is worth mentioning that replacing $\{Q_i(t), t\in[0,1]\}$ by its empirical counterpart $\widehat{Q}_i(t)$ introduces an additional error term, $u_i(t) = \log\widehat{Q}_i(t) - \log Q_i(t)$ say, in the AR(1) model (17). That is

$$\begin{aligned}\log\widehat{Q}_i(t) - H(t) =& \varphi\log\widehat{Q}_{i-1}(t) - \varphi H(t) + 2\varepsilon_i + u_i(t) - \varphi u_i(t)\\ =:& \varphi\log\widehat{Q}_{i-1}(t) - \varphi H(t) + \delta_i(t).\end{aligned} \tag{33}$$

By careful inspection, we deduce that the error terms $\delta_i(t)$, $i = 1,2,\ldots,N$, are not necessarily independent. Moreover, the error terms $\delta_i(t)$ encompass three different terms which makes σ_{ε}^2 nonidentifiable. To overcome this issue we establish the *uniform consistency* of $\log\widehat{Q}_i(t)$. Roughly speaking, applying the results of Lemma A.1. , we conclude that the error terms $u_i(t)$ tend to zero *sufficiently fast*, so (33) is a sufficiently good approximation to model (17). Galbraith, Zinde-Walsh and Zhu (2015) confront a similar issue in analysing GARCH models. Differently from the current study, the essence of their problem allows them to interpret the inaccuracy caused by this effect as an error term. Their procedure thus does not require decay of the $u_i(t)$s to zero.

We now turn to the large sample justification of the procedures proposed above. Recall the stochastic volatility model (9)–(10) with the unknown parameter vector $\boldsymbol{\theta}$ given in (16). We study the limiting behavior of the estimates $\check{\boldsymbol{\theta}}$, $\bar{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}$ proposed in (27), (30) and (32) , respectively. Theorem 3.2, in particular, addresses convergence of $\widehat{H}(\cdot) - H(\cdot)$ uniformly in L^1 sense. Convergence of $\check{\varphi}$ and $\check{\sigma}^2_\varepsilon$ as well as their rate of convergence are established in Theorem 3.3. Theorems 3.4 and 3.5 investigate, respectively, $\bar{\varphi}$ and $\bar{\sigma}^2_\varepsilon$ and $\hat{\varphi}$ and $\hat{\sigma}^2_\varepsilon$. The proofs are provided in the online supplement, Kokoszka et al. (2024). The proofs fundamentally rely on the decay of the error induced by replacing the quadratic variation processes by their realized variants. This is quantified in Propositions A.3., A.7. and A.8.. These together with consistency of oracle estimates $\widetilde{H}$ and $(\gamma_{0,N}, \gamma_{1,N})$ addressed in Propositions A.2. and A.6. entail the main results of the current section. Recall that $H(t) = \log \int_0^t \sigma^2(u)du$ and $\widehat{H}(t)$ and $\widehat{G}(t)$ are given by (23) and (24), respectively.

Theorem 3.2. *Assume the stochastic volatility model defined by* (9)–(10) *and conditions 1–6 and recall* (23) *and* (24)*. Then, for any fixed* $0 < \alpha < 1$*,*

$$\mathbb{E} \sup_{t\in[\alpha,1]} \left|\widehat{H}(t) - H(t)\right| = O\left(N^{-\frac{1}{2}} + \Delta^{\frac{1}{2}}\right), \tag{34}$$

$$\sup_{t\in[\alpha,1]} \left|\widehat{G}(t) - G(t)\right| = O_P\left(N^{-\frac{1}{2}} + \Delta^{\frac{1}{2}}\right). \tag{35}$$

Before stating the next results, we need to define the lag–h autocovariances:

$$\kappa_h := \mathbb{E}[\log g_0 \log g_h] = \frac{1}{4}\mathbb{E}\left[\left(\log Q_0(t) - H(t)\right)\left(\log Q_h(t) - H(t)\right)\right], \quad h \in \mathbb{Z}. \tag{36}$$

In the following theorems, no assumptions are needed for the interplay between N and Δ.

Theorem 3.3. *Assume the stochastic volatility model defined by* (9)–(10) *satisfying conditions 1–7 and recall* (28)*. Then,*

$$N^{1/2}\left(\check{\varphi} - \varphi\right) \xrightarrow{\mathcal{L}aw} \mathcal{N}\left(0, \nu\right), \quad \text{as } N \to \infty \text{ and } \Delta \to 0,$$

and

$$N^{1/2}\left(\check{\sigma}^2_\varepsilon - \sigma^2_\varepsilon\right) \xrightarrow{\mathcal{L}aw} \mathcal{N}\left(0, \tau\right), \quad \text{as } N \to \infty \text{ and } \Delta \to 0,$$

where

$$\nu = \left(-\kappa_0^{-2}\kappa_1, \kappa_0^{-1}\right) V \left(-\kappa_0^{-2}\kappa_1, \kappa_0^{-1}\right)^\top \quad \text{and} \quad \tau = \left(1 + \kappa_0^{-2}\kappa_1^2, -\kappa_0^{-1}\right) V \left(1 + \kappa_0^{-2}\kappa_1^2, -\kappa_0^{-1}\right)^\top, \tag{37}$$

and V *is a* 2 *by* 2 *matrix with entries*

$$V_{k,l} = (\eta - 3)\kappa_k\kappa_l + \sum_{h=-\infty}^{\infty} \left(\kappa_h\kappa_{h-k+l} + \kappa_{h-k}\kappa_{h+l}\right), \quad k, l = 0,\ 1. \tag{38}$$

Theorem 3.4. *Assume the stochastic volatility model defined by* (9)–(10) *satisfying conditions 1–7 and recall* (29)*. Then the limiting results of Theorem 3.3 holds true for* $\bar{\varphi}$ *and* $\bar{\sigma}^2$ *as well, i.e. for any*

$0 < \alpha < 1$,

$$N^{1/2}\left(\bar{\varphi} - \varphi\right) \xrightarrow{\mathcal{L}aw} \mathcal{N}\left(0, \nu\right), \quad \text{as } N \to \infty \text{ and } \Delta \to 0,$$

and

$$N^{1/2}\left(\bar{\sigma}_{\varepsilon}^2 - \sigma_{\varepsilon}^2\right) \xrightarrow{\mathcal{L}aw} \mathcal{N}\left(0, \tau\right), \quad \text{as } N \to \infty \text{ and } \Delta \to 0,$$

where ν *and* τ *are defined through* (37)-(38).

Theorem 3.5. *Assume the stochastic volatility model defined by* (9)–(10) *and conditions 1–7 and recall* (31)*. Then, for any* $0 < \alpha < 1$,

$$N^{\frac{1}{2}}\left(\hat{\varphi} - \varphi\right) \xrightarrow{\mathcal{L}aw} \mathcal{N}\left(0, \pi\right) \quad \text{as } N \to \infty \text{ and } \Delta \to 0, \tag{39}$$

and

$$N^{\frac{1}{2}}\left(\hat{\sigma}_{\varepsilon}^2 - \sigma_{\varepsilon}^2\right) \xrightarrow{\mathcal{L}aw} \mathcal{N}\left(0, \rho\right), \quad \text{as } N \to \infty \text{ and } \Delta \to 0,$$

where

$$\pi = \left(0, \kappa_0^{-1}\right) V \left(0, \kappa_0^{-1}\right)^{\top} \quad \text{and} \quad \rho = \left(1, -\varphi\right) V \left(1, -\varphi\right)^{\top},$$

and matrix V *is given in* (38)*.*

The proofs of all results stated in this section are given in Kokoszka et al. (2024). We explain here briefly where main challenges requiring novel approaches occur. At a heuristic level, the proposed method involves the unobservable quadratic variation processes $Q_i(\cdot)$ which makes it infeasible. Substituting $Q_i(\cdot)$ with discretely observed curves $\widehat{Q}_i(\cdot)$ and studying the decay of the error induced by this approximation locate the problem at the interface between FDA and SDE. Proposition A.7. and Corollaries A.2. and A.3. establish consistency of the empirical autocovariances of the proxy processes $\widehat{Q}_i(\cdot)$. These key results pave the way for applying the delta method to obtain consistency results claimed in Theorems 3.3 and 3.4. Theorem 3.5 is however more demanding and relies on uniform consistency of the empirical autocovariances of the proxy processes $\widehat{Q}_i(\cdot)$ and the reciprocal empirical variance of $\widehat{Q}_i(\cdot)$. These are proved in Proposition A.8. and Lemma A.4.. It is worth mentioning that discarding an arbitrarily narrow interval $[0, \alpha)$ is required to transfer the problem to the level of $\log Q_i(\cdot)$ at which the between curves dependence is formulated. It is an insight that is not obvious from model formulation and is utilized in the proofs of Lemma A.1., Corollary A.1. and Theorem 3.2.

We conclude this section with an almost sure consistency result. Almost sure convergence of the realized quadratic variation does not hold for arbitrary grid points such that $\max |t_k - t_{k-1}| \to 0$. Exercises 1.15 (b) and 1.16 (b) in Mörters and Peres (2010) show that even pointwise almost sure convergence is not guaranteed. However, Theorem 3.6 establishes it for the equidistant design we consider.

Theorem 3.6. *Assume the conditions of Theorem 3.2 and consider the regular design points* $t_k = m^{-1}k = \Delta k$*. Then, for any* $\beta < \frac{1}{2}$*, with probability 1,*

$$\sup_{t \in [\alpha, 1]} \left|\widehat{H}(t) - H(t)\right| = O\left(N^{-\beta} + \Delta^{\beta}\right), \tag{40}$$

$$\sup_{t \in [\alpha, 1]} \left|\widehat{G}(t) - G(t)\right| = O\left(N^{-\beta} + \Delta^{\beta}\right). \tag{41}$$

4. Extension to order p latent autoregression

This section extends the model formulated in Section 3 by replacing the order 1 autoregression in (10) by an AR(p) model. This increases the flexibility of the model. The fundamental properties and estimation approaches remain the same, but the limiting covariance structure has to be worked out carefully. An extension to a more general ARMA structure is more challenging and is not pursued in this paper. It is well known that adding moving average terms, while conceptually simple, often requires theoretically and practically nontrivial modifications. This point is well explained in a high-dimensional context in Wilms et al. (2023) who give numerous relevant references.

For completeness, we begin with model equations, noting that equation (42) is the same as (9):

$$R_i(t) = g_i \int_0^t \sigma(u) dW_i(u), \quad t \in [0,1], \quad i \in \mathbb{Z}, \tag{42}$$

$$\log g_i = \varphi_1 \log g_{i-1} + \varphi_2 \log g_{i-2} + \ldots + \varphi_p \log g_{i-p} + \varepsilon_i, \quad \varepsilon_i \sim \text{ iid } \mathcal{WN}(0, \sigma_\varepsilon^2), \tag{43}$$

the autoregressive polynomial has no zeros in the closed complex unit disk, i.e.

$$1 - \varphi_1 z - \varphi_2 z^2 - \ldots - \varphi_p z^p \neq 0, \ \varphi_p \neq 0, \quad \text{for } |z| \leq 1. \tag{44}$$

It is well-known that (44) is equivalent to the existence of a stationary causal solution to (43), see e.g Theorem 3.1.1. in Brockwell and Davis (1991). Notice again that we could equivalently formulate this model by replacing $\log g_i$ with x_i and g_i with $\exp(x_i)$.

We begin by stating an extension of Theorem 3.1. The proof is analogous, so it is omitted.

Theorem 4.1. *Under* (42)-(43)*, if conditions 1–4 of Section 3 and* (44) *hold, then there exists a unique strictly stationary functional sequence R_i satisfying* (9)–(10)*. Moreover, the scalar g_i is distinguishable from the function σ in* (42)*.*

Turning to estimation, define the vector of autoregressive coefficients:

$$\boldsymbol{\varphi} = \left(\varphi_1, \varphi_2, \ldots, \varphi_p\right)^\top.$$

Recall (36) and define the vector $\boldsymbol{\kappa}_p$ and the matrix $\boldsymbol{\Xi}_p$:

$$\boldsymbol{\kappa}_p = \left(\kappa_1, \kappa_2, \ldots, \kappa_p\right)^\top, \qquad \boldsymbol{\Xi}_p = \left(\kappa_{i-j}\right)_{i,j=1}^p.$$

Define also

$$\boldsymbol{\gamma}_{p,N} = \left(\gamma_{1,N}, \gamma_{2,N}, \ldots, \gamma_{p,N}\right)^\top, \qquad \boldsymbol{\Psi}_{p,N} = \left(\gamma_{i-j,N}\right)_{i,j=1}^p, \tag{45}$$

where $\gamma_{h,N} = \gamma_{-h,N}$ and, for $h = 0, 1, \ldots, p$,

$$\begin{aligned} \gamma_{h,N} =& \frac{1}{4N} \sum_{i=1}^{N-h} \left(\log Q_i(t) - \frac{1}{N} \sum_{i=1}^N \log Q_i(t) \right) \left(\log Q_{i+h}(t) - \frac{1}{N} \sum_{i=1}^N \log Q_i(t) \right) \\ =& \frac{1}{N} \sum_{i=1}^{N-h} \left(\log g_i - \frac{1}{N} \sum_{i=1}^N \log g_i \right) \left(\log g_{i+h} - \frac{1}{N} \sum_{i=1}^N \log g_i \right). \end{aligned} \tag{46}$$

The above definitions allow us to defined the oracle Yule-Walker estimates:

$$\tilde{\boldsymbol{\varphi}} = \boldsymbol{\Psi}_{p,N}^{-1}\boldsymbol{\gamma}_{p,N}, \qquad \tilde{\sigma}_{\varepsilon}^2 = \gamma_{0,N} - \tilde{\boldsymbol{\varphi}}^{\top}\boldsymbol{\gamma}_{p,N}.$$

For each $t \in [0,1]$, the realized counterpart of (45) can be written in the form

$$\boldsymbol{\Gamma}_{p,N}(t) = \big(\Gamma_{1,N}(t), \Gamma_{2,N}(t), \ldots, \Gamma_{p,N}(t)\big)^{\top}, \qquad \boldsymbol{\Sigma}_{p,N}(t) = \big(\Gamma_{i-j,N}(t)\big)_{i,j=1}^{p}, \quad t \in [0,1],$$

where $\Gamma_{h,N}(t) = \Gamma_{-h,N}(t)$ and, for $h = 0, 1, \ldots, p$,

$$\Gamma_{h,N}(t) = \frac{1}{4N}\sum_{i=1}^{N-h}\left(\log\widehat{Q}_i(t) - \frac{1}{N}\sum_{i=1}^{N}\log\widehat{Q}_i(t)\right)\left(\log\widehat{Q}_{i+h}(t) - \frac{1}{N}\sum_{i=1}^{N}\log\widehat{Q}_i(t)\right). \tag{47}$$

In contrast to the sequence $\gamma_{h,N}$, the above empirical autocovariances may depend on t. If we directly plug them in the Yule-Walker equations, we obtain a *family* of estimators, indexed by t, for the constant parameters $\boldsymbol{\varphi}$ and σ_{ε}^2. To overcome this issue, similarly to Section 3, we propose the following three procedures whose asymptotic properties are established in Theorems 4.2–4.5 below.

PROCEDURE A: We define

$$\check{\boldsymbol{\theta}} = \left[\widehat{G}(t), t \in [0,1], \quad \check{\boldsymbol{\varphi}}, \quad \check{\sigma}_{\varepsilon}^2\right],$$

where $\widehat{G}(t)$ is the same as (24) and $\check{\varphi}$ and $\check{\sigma}_{\varepsilon}^2$ are obtained by using on the terminal time $t = 1$:

$$\check{\boldsymbol{\varphi}} = \boldsymbol{\Sigma}_{p,N}^{-1}(1)\boldsymbol{\Gamma}_{p,N}(1), \qquad \check{\sigma}_{\varepsilon}^2 = \Gamma_{0,N}(1) - \check{\boldsymbol{\varphi}}^{\top}\boldsymbol{\Gamma}_{p,N}(1). \tag{48}$$

PROCEDURE B: We define

$$\bar{\boldsymbol{\theta}} = \left[\widehat{G}(t), t \in [0,1], \quad \bar{\boldsymbol{\varphi}}, \quad \bar{\sigma}_{\varepsilon}^2\right],$$

where $\widehat{G}(t)$ is the same as (24) and $\bar{\boldsymbol{\varphi}}$ and $\bar{\sigma}_{\varepsilon}^2$ are defined by

$$\bar{\boldsymbol{\varphi}} = \bar{\boldsymbol{\Sigma}}_{p,N}^{-1}\bar{\boldsymbol{\Gamma}}_{p,N}, \qquad \bar{\sigma}_{\varepsilon}^2 = \bar{\Gamma}_{0,N} - \bar{\boldsymbol{\varphi}}^{\top}\bar{\boldsymbol{\Gamma}}_{p,N}, \tag{49}$$

where

$$\bar{\boldsymbol{\Gamma}}_{p,N} = \frac{1}{1-\alpha}\int_{\alpha}^{1}\boldsymbol{\Gamma}_{p,N}(t)dt, \quad \bar{\Gamma}_{0,N} = \frac{1}{1-\alpha}\int_{\alpha}^{1}\Gamma_{0,N}(t)dt, \quad \bar{\boldsymbol{\Sigma}}_{p,N} = \frac{1}{1-\alpha}\int_{\alpha}^{1}\boldsymbol{\Sigma}_{p,N}(t)dt,$$

and where α is any fixed number in the interval $(0,1)$.

PROCEDURE C: We define

$$\hat{\boldsymbol{\theta}} = \left[\widehat{G}(t), t \in [0,1], \quad \hat{\boldsymbol{\varphi}}, \quad \hat{\sigma}_{\varepsilon}^2\right].$$

where $\widehat{G}(t)$ is the same as (24) and $\hat{\boldsymbol{\varphi}}$ and $\hat{\sigma}_{\varepsilon}^2$ are

$$\hat{\boldsymbol{\varphi}} = \frac{1}{1-\alpha}\int_{\alpha}^{1}\boldsymbol{\Sigma}_{p,N}^{-1}(t)\boldsymbol{\Gamma}_{p,N}(t)dt, \qquad \hat{\sigma}_{\varepsilon}^2 = \frac{1}{1-\alpha}\int_{\alpha}^{1}\left(\Gamma_{0,N}(t) - \hat{\boldsymbol{\varphi}}^{\top}\boldsymbol{\Gamma}_{p,N}(t)\right)dt, \tag{50}$$

where α is any fixed number in the interval $(0,1)$.

Theorem 4.2. *Assume the stochastic volatility model defined by* (42)–(43) *and conditions 1–6 except that we replace condition 5 by* (44)*. Define* $\widehat{H}(t)$ *and* $\widehat{G}(t)$ *by* (23) *and* (24)*, respectively. Then, for any fixed* $0 < \alpha < 1$*,*

$$\mathbb{E} \sup_{t\in[\alpha,1]} \left|\widehat{H}(t) - H(t)\right| = O\left(N^{-\frac{1}{2}} + \Delta^{\frac{1}{2}}\right), \tag{51}$$

$$\sup_{t\in[\alpha,1]} \left|\widehat{G}(t) - G(t)\right| = O_P\left(N^{-\frac{1}{2}} + \Delta^{\frac{1}{2}}\right). \tag{52}$$

Before investigating asymptotic properties of procedures A and B. we need to introduce the following notation. Consider the domain $\mathcal{D} \subset \mathbb{R}^p$ defined by

$$\mathcal{D} = \left\{ (x_0, \ldots, x_{p-1}) \mid X = \left[x_{i-j}\right]_{i,j=1}^{p} \text{ is positive definite} \right\},$$

and the functions

$$f : \mathcal{D} \times \mathbb{R} \longrightarrow \mathbb{R}^p,\ f(x_0, \ldots, x_p) = X^{-1}a,$$

$$g : \mathcal{D} \times \mathbb{R} \longrightarrow \mathbb{R},\ g(x_0, \ldots, x_p) = x_0 - \left(X^{-1}a\right)^T a,$$

where $X = [x_{i-j}]_{i,j=1}^{p},\ a = (x_1, \ldots, x_p)^\top$. The above functions $f(\cdot)$ and $g(\cdot)$ are continuously differentiable.

In Theorems 4.3, 4.4 and 4.5, we assume the stochastic volatility model (42)–(43) and conditions 1–7 of Section 3, except that we replace condition 5 by (44), if $p > 1$. We do not assume any interplay between $N \to \infty$ and $\Delta \to 0$. In Theorems 4.4 and 4.5, the same asymptotic distribution holds for any $\alpha \in (0, 1)$.

Theorem 4.3. *For the estimators defined by* (48) *(Procedure A),*

$$N^{1/2}\left(\check{\boldsymbol{\varphi}} - \boldsymbol{\varphi}\right) \xrightarrow{\mathcal{L}aw} \mathcal{N}\left(0, \boldsymbol{\nu}_p\right), \quad \text{as } N \to \infty \text{ and } \Delta \to 0,$$

and

$$N^{1/2}\left(\check{\sigma}_{\varepsilon}^2 - \sigma_{\varepsilon}^2\right) \xrightarrow{\mathcal{L}aw} \mathcal{N}\left(0, \tau_p\right), \quad \text{as } N \to \infty \text{ and } \Delta \to 0,$$

where

$$\boldsymbol{\nu}_p = \left(\nabla f\left(\kappa_0, \boldsymbol{\kappa}_p\right)\right)^\top V \nabla f\left(\kappa_0, \boldsymbol{\kappa}_p\right), \quad \tau_p = \left(\nabla g\left(\kappa_0, \boldsymbol{\kappa}_p\right)\right)^\top V \nabla g\left(\kappa_0, \boldsymbol{\kappa}_p\right) \tag{53}$$

and V *is a* $(p+1) \times (p+1)$ *matrix with entries*

$$V_{k,l} = (\eta - 3)\kappa_k \kappa_l + \sum_{h=-\infty}^{\infty} \left(\kappa_h \kappa_{h-k+l} + \kappa_{h-k}\kappa_{h+l}\right), \quad k, l = 0, \ldots, p. \tag{54}$$

Theorem 4.4. *The estimators defined by* (49) *(Procedure B) have the same asymptotic distribution as the estimators* (48) *(Procedure A), i.e.*

$$N^{1/2}\left(\bar{\boldsymbol{\varphi}} - \boldsymbol{\varphi}\right) \xrightarrow{\mathcal{L}aw} \mathcal{N}\left(0, \boldsymbol{\nu}_p\right), \quad \text{as } N \to \infty \text{ and } \Delta \to 0,$$

and

$$N^{1/2}\left(\bar{\sigma}_{\varepsilon}^{2}-\sigma_{\varepsilon}^{2}\right)\xrightarrow{\mathcal{L}aw}\mathcal{N}\left(0,\tau_{p}\right),\quad \text{as } N\to\infty \text{ and } \Delta\to 0,$$

where $\boldsymbol{\nu}_p$ *and* τ_p *are defined through* (53)-(54).

Theorem 4.5. *For the estimators* (50) *(procedure C),*

$$N^{\frac{1}{2}}\left(\hat{\boldsymbol{\varphi}}-\boldsymbol{\varphi}\right)\xrightarrow{\mathcal{L}aw}\mathcal{N}\left(0,\boldsymbol{\pi}_{p}\right)\quad \text{as } N\to\infty \text{ and } \Delta\to 0, \tag{55}$$

and

$$N^{\frac{1}{2}}\left(\hat{\sigma}_{\varepsilon}^{2}-\sigma_{\varepsilon}^{2}\right)\xrightarrow{\mathcal{L}aw}\mathcal{N}\left(0,\rho_{p}\right),\quad \text{as } N\to\infty \text{ and } \Delta\to 0,$$

where

$$\boldsymbol{\pi}_{p}=\boldsymbol{\Xi}_{p}^{-1}WVW^{\top}\boldsymbol{\Xi}_{p}^{-1},\quad \rho_{p}=\left(1,-\boldsymbol{\varphi}^{\top}\right)V\begin{pmatrix}1\\-\boldsymbol{\varphi}\end{pmatrix},$$

with W *a* $p\times(p+1)$ *matrix in the form:*

$$W=\begin{pmatrix}0&1&0&0&\dots&0\\0&0&1&0&\dots&0\\0&0&0&1&\dots&0\\\vdots&\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&0&\dots&1\end{pmatrix}, \tag{56}$$

and matrix V *is given in* (54).

REMARK 4.1. A salient feature of stock returns is the leverage effect, which is generally understood as an association between negative returns and increased volatility. In the scalar SV framework, it is modeled by introducing correlation between the noise sequences in the state and observation equations, see Harvey and Shephard (1996). Our approach could potentially be extended to model leverage by introducing suitably formulated dependence between $\{W_i\}$ and $\{g_i\}$. Within such an extended framework, one could potentially develop tests of independence of sequences $\{W_i\}$ and $\{g_i\}$.

REMARK 4.2. To facilitate practical application of the model studied in this section, methods of selecting the order p are desirable. For scalar models, most broadly used techniques are based on information criteria and minimization of prediction errors. The problem is more complex even in linear *functional* settings, see Kokoszka and Reimherr (2013). It can be hoped that suitable techniques can be developed in the context of this paper because parameter estimators enter into all order selection criteria.

REMARK 4.3. An important application of scalar stochastic volatility models is volatility forecasting. This issue will require a separate detailed study in the context of the model of this section. In particular direct and iterative methods would need to be compared via asymptotic and finite sample prediction errors. To explain the potential of our approach, we focus on one step ahead prediction. Denoting by $\varphi_k^{\star}$ any of the derived estimates, we can predict $x_{N+1}=\log g_{N+1}$ by $\hat{x}_{N+1}=\varphi_1^{\star}x_N+\varphi_2^{\star}x_{N-1}+\cdots+\varphi_p^{\star}x_{N+p-1}$. We can then predict $R_{N+1}^2(t)$ by $\widehat{R}_{N+1}^2(t)=e^{2\hat{x}_{N+1}}\widehat{G}(t)$.

5. Empirical analysis

We first present in Section 5.1 the estimation of our model on intraday price curves, suitably transformed to stationarity, of a large number of U.S. stocks. Such an analysis will explain the meaning of the model elements introduced in previous sections and will suggest relevant parameter settings for a simulation study presented in Section 5.2.

5.1. Application to U.S. stocks

By way of introduction, we begin with the analysis of curves derived from price data of Apple Inc. (Permno: 14593). The sample period is from Jan 3, 2016 to Dec 31, 2021, corresponding to $N = 4021$ trading days. In each trading day i, we have the opening price $P_i(t_0)$ and the following 78 of 5-min intraday price observations $P_i(t_k), k = 1, ..., 78$, the last trading prices in every 5-min time interval. A different time resolution could be used, but the five minute resolution allows us to analyze most stocks traded in U.S. because not all of them are traded as frequently as Apple. Additionally, the intraday price data at 5-min frequency provides a good balance between informative signals and effects of market microstructure errors, see e.g. Barndorff-Nielsen and Shephard (2002). Thus, the step size is $\Delta = 1/78$ and the design points are $t_k = k \times \Delta, k = 0, 1, ..., m$, where $m = 78$.

We calculate the cumulative intraday return (CIDR) curves as

$$R_i(t_k) = \log(P_i(t_k)) - \log(P_i(t_0)), \qquad k = 0, 1, ..., m,\ i = 1, 2, ..., N.$$

By definition, the CIDR curves always start from zero, i.e. $R_i(t_0) = 0$, and are scale invariant. This generally leads to a stationary sequence of curves, as investigated in Horváth, Kokoszka and Rice (2014). We calculate the realized version of the quadratic variation as

$$\widehat{Q}_i(t) = \sum_{k=1}^{K} |R_i(t_k) - R_i(t_{k-1})|^2 \mathbb{I}\{t_k \le t\}, \qquad t \in [0, 1], \qquad i = 1, 2, ..., N.$$

Figure 1 shows the Apple intraday price $P_i(t_0)$ (upper panel), the CIDRs $R_i(t_k)$ (middle panel) and the realized quadratic variation $\widehat{Q}_i(t)$ (lower panel). Note that we can have $\widehat{Q}_i(t) = 0$ for a few 5-min long intervals at the beginning of the trading day. This is because the price can remain the same as the opening price in the absence of any trades for some time after the opening. In such cases, $\log \widehat{Q}_i(t)$ is not computable. This means that the truncation at $\alpha > 0$ required by our theory is often practically needed. In the definitions below we assume that α is such that $\log \widehat{Q}_i(t)$ can be computed for $t > \alpha$. For the stocks we consider, $\alpha = \Delta = 1/78$ is generally sufficient. With this caveat, we can calculate $\widehat{G}(t)$ given by (24). We can also compute numerically the derivative $\hat{\sigma}_u^2(t) = d\widehat{G}(t)/dt$. Figure 2 shows both curves.

Table 1. Estimation results for Apple with $\alpha = \Delta$

	Proc. A	Proc. B	Proc. C
φ	0.537	0.512	0.535
σ_ε^2	0.262	0.281	0.276

After computing the covariances $\Gamma_{0,N}(t)$ and $\Gamma_{1,N}(t)$, given, respectively, by (25) and (26), we can compute all estimators in Procedures A, B and C introduced in Section 3. Table 1 shows the

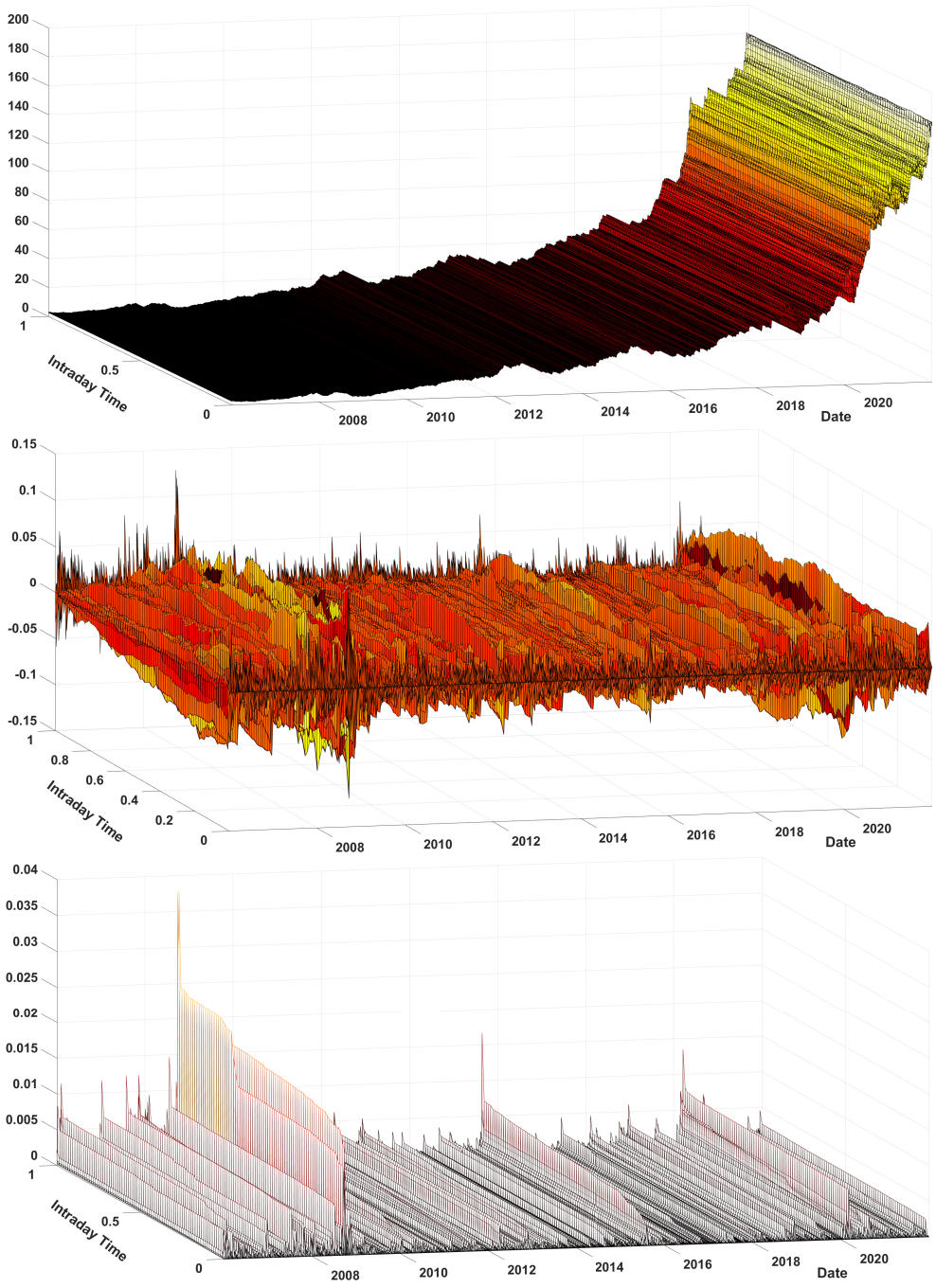

Figure 1. Time series of functional objects derived from intraday Apple prices. Upper Panel: Intraday Prices $P_i(t_k)$; Middle Panel: the CIDRs $R_i(t_k)$; Lower Panel: Realized Quadratic Variation $\bar{Q}_i(t)$.

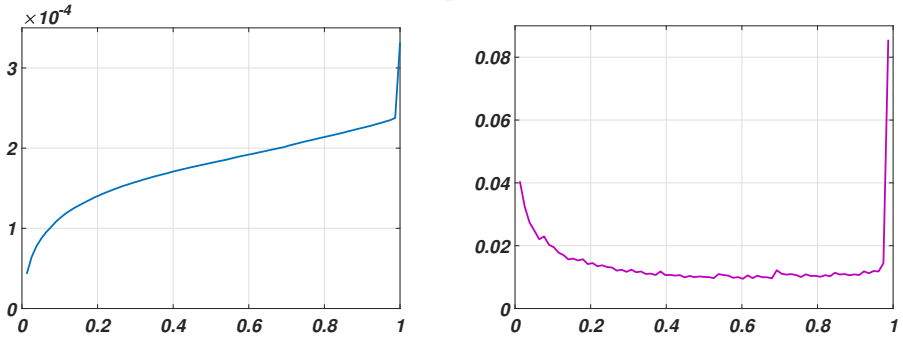

Figure 2. Apple volatility curves: Left: the cumulative volatility $\widehat{G}$; Right: pointwise volatility $\hat{\sigma}_u$.

estimation results of the three procedures with $\alpha = \Delta$. We see that they yield similar estimates. This is encouraging because it indicates that they all could be close to the true values of these parameters. This will be investigated in Section 5.2, but before we do it, we need to get a more comprehensive picture of possible parameter ranges. For this purpose, we repeat the same analysis for 7293 stocks in the U.S. stock markets. The original dataset includes all U.S. stocks from 2006 to 2021. To ensure data quality, the intraday price data is cleaned based on the rules explained in Section C of Kokoszka et al. (2024). The summary statistics of the estimators are presented in Table 2. We see that the three procedure produce estimates in similar ranges.

Table 2. Summary statistics of the three estimators based on 7293 stocks

	φ			σ_ε^2		
	Proc. A	Proc. B	Proc. C	Proc. A	Proc. B	Proc. C
Mean	0.557	0.504	0.529	0.234	0.292	0.287
SD	0.123	0.121	0.117	0.060	0.072	0.072
Skewness	0.119	0.257	0.158	4.453	3.340	3.434
Kurtosis	2.857	2.967	2.967	39.421	22.813	23.662
Min	0.119	0.093	0.112	0.059	0.133	0.131
Q. 25%	0.470	0.419	0.447	0.205	0.250	0.246
Median	0.551	0.495	0.522	0.227	0.278	0.272
Q. 75%	0.642	0.584	0.608	0.252	0.314	0.308
Max	0.961	0.941	0.943	0.955	1.020	1.019

We conclude this section with an investigation of the impact of the truncation parameter $\alpha \in (0, 1)$. This parameter is absent in Procedure A, but we included it for comparison. We repeated the comprehensive analysis with $\alpha \in \{5\Delta, 20\Delta, 40\Delta\}$. Since $\Delta = 1/78$, we have $\alpha = \{0.0641, 0.2564, 0.5128\}$. Table 3 presents selected, most informative, summary statistics based on different values of α. It shows that larger α leads to larger estimates of φ and smaller estimates of σ, but all those estimates are similar.

Table 3. Estimation results for different α.

	φ			σ_ε^2		
	Proc. A	Proc. B	Proc. C	Proc. A	Proc. B	Proc. C
$\alpha = 5\Delta$						
Mean	0.557	0.532	0.542	0.234	0.263	0.261
SD	0.123	0.120	0.119	0.060	0.071	0.071
Q. 25%	0.470	0.447	0.459	0.205	0.224	0.222
Median	0.551	0.526	0.537	0.227	0.249	0.247
Q. 75%	0.642	0.613	0.622	0.252	0.283	0.281
$\alpha = 20\Delta$						
Mean	0.557	0.565	0.567	0.234	0.232	0.231
SD	0.123	0.119	0.119	0.060	0.067	0.067
Q. 25%	0.470	0.482	0.485	0.205	0.196	0.196
Median	0.551	0.561	0.564	0.227	0.219	0.218
Q. 75%	0.642	0.647	0.649	0.252	0.248	0.248
$\alpha = 40\Delta$						
Mean	0.557	0.586	0.587	0.234	0.213	0.213
SD	0.123	0.118	0.118	0.060	0.065	0.065
Q. 25%	0.470	0.505	0.505	0.205	0.180	0.180
Median	0.551	0.584	0.585	0.227	0.201	0.201
Q. 75%	0.642	0.668	0.669	0.252	0.228	0.228

5.2. A simulation study

The purpose of this section is to obtain more detailed insights into the finite sample performance of the proposed estimators and to compare them. Functional time series are generated according to (2), but with $i = 1, \ldots, N$. We use $\epsilon_i \sim i.i.d.\ \mathcal{N}(0, \sigma_\varepsilon^2)$. Based on the results of Section 5.1, we use $\varphi = 0.55$ and $\sigma_\varepsilon^2 = 0.25$.

We use four intraday volatility functions:

- Flat: $\sigma(u) = 0.2$, the same intraday volatility throughout the day.
- Slope: $\sigma(u) = 0.1 + 0.2u$, intraday volatility increases in a linear manner.
- Sine: $\sigma(u) = 0.1\sin(2\pi u) + 0.2$, higher volatility in the morning, but lower volatility in the afternoon.
- U-shape: $\sigma(u) = (u - 0.5)^2 + 0.1145299$. This choice is most relevant as it reflects the stylized fact that the volatility is typically highest at the beginning and the end of a trading day.

The coefficients in the four σ functions are set to have a similar level of average daily volatility. Figure 3 in Kokoszka et al. (2024) displays the four choices of σ function and their corresponding (theoretical) H and G functions. The H and G functions under the U-shape σ function exhibit a generally similar pattern to those shown in Figure 1, with a minor difference at the final 5-min interval of the trading day.

To simulate the integral $\int_0^t \sigma(u)dW(u)$, we use the time change formula explained in Section 2, cf. Corollary 2.1, i.e. we set

$$\int_0^t \sigma(u)dW(u) = W\left(\int_0^t \sigma^2(u)du\right). \tag{57}$$

Corresponding to the analysis in Section 5.1, the continuous time t in $[0,1]$ is discretized as $[t_0, t_1, ..., t_K]$, where $t_k = k\Delta$, $k = 1, ..., K$. The stepsize is chosen to be $\Delta = 1/78$ which corresponds to 5-min frequency. Using (57), we generate $\int_0^{t_k} \sigma(u)dW(u) = \sum_{s=1}^k d(t_k)$, where $d(t_k) \sim \mathcal{N}(0, G(t_k) - G(t_{k-1}))$ are independent random variables.

We consider sample sizes $N = 100, 500, 1000, 2000$, with the larger sizes being most relevant (we used over 4,000 trading days in Section 5.1).

To compare the different procedures, we use the following evaluation metrics. For the estimators of φ and σ_ε^2, we calculate the empirical bias (B) and the empirical mean root squared error (RMSE). For Procedure A, they are defined as

$$\mathrm{B}(\breve{\varphi}) = \frac{1}{R}\sum_{r=1}^R \breve{\varphi}_r - \varphi, \qquad \mathrm{RMSE}(\breve{\varphi}) = \left\{\frac{1}{R}\sum_{r=1}^R (\breve{\varphi}_r - \varphi)^2\right\}^{1/2},$$

and

$$\mathrm{B}(\breve{\sigma}_\varepsilon^2) = \frac{1}{R}\sum_{r=1}^R \breve{\sigma}_{\varepsilon,r}^2 - \sigma_\varepsilon^2, \qquad \mathrm{RMSE}(\breve{\sigma}_\varepsilon^2) = \left\{\frac{1}{R}\sum_{r=1}^R (\breve{\sigma}_{\varepsilon,r}^2 - \sigma_\varepsilon^2)^2\right\}^{1/2},$$

where the subscript r denotes the r-th simulation repetition. For Procedures B and C, the above metrics are defined analogously.

As for evaluating the estimation of $G(t)$, we know $\sigma(u)$ since we simulate the data, and thus we can compute the exact value of $G(t) = \int_0^t \sigma^2(u)du$. Then we can compute the functional empirical bias (fB) and the functional empirical root mean squared error (fRMSE):

$$\mathrm{fB}(\widehat{G}) = \left(\frac{1}{R}\sum_{r=1}^R \int_0^1 (\widehat{G}_r(t) - G(t))dt\right), \quad \mathrm{fRMSE}(\widehat{G}) = \left(\frac{1}{R}\sum_{r=1}^R \int_0^1 (\widehat{G}_r(t) - G(t))^2 dt\right)^{1/2}.$$

The above two measures are very close to zero even for $N = 100$, and decrease with N even further. We therefore report only the functional relative error

$$\mathrm{fRE}(\widehat{G}) = \left(\frac{1}{R}\sum_{r=1}^R \frac{\int_0^1 (\widehat{G}_r(t) - G(t))^2 dt}{\int_0^1 G^2(t)dt}\right)^{1/2}.$$

Table 4 shows the estimation error of φ and σ_ε^2 under the four different σ functions with $\alpha = \Delta$. Generally, the consistency of the three procedures for estimating φ and σ_ε^2 is well supported by the simulation results because the RMSE decreases with the increase in sample size. Comparing between the three procedures, Procedure A outperforms the other two procedures in term of lowest bias and RMSE. When considering the four shapes of the σ functions, there is no substantial distinction in the results of the estimation of φ and σ_ε^2, although the U-shape has marginally higher bias and RMSE than the others. Lastly, our estimation procedures have some minor bias (underestimate φ and overestimate σ_ε^2), which is consistent with the Yule-Walker estimator of the scalar AR(1) process, as noted by Shaman and Stine (1988). We provide a bias-corrected version of our three procedures in E of Kokoszka

Table 4. Estimation error for the estimation of φ and σ_ε^2 with $\alpha = \Delta$

		φ			σ_ε^2		
		Proc. A	Proc. B	Proc. C	Proc. A	Proc. B	Proc. C
Flat							
B	$N = 100$	-0.040	-0.083	-0.069	0.006	0.045	0.043
	$N = 500$	-0.016	-0.060	-0.046	0.008	0.048	0.045
	$N = 1000$	-0.014	-0.057	-0.043	0.008	0.049	0.046
	$N = 2000$	-0.011	-0.055	-0.041	0.008	0.049	0.046
RMSE	$N = 100$	0.095	0.120	0.110	0.038	0.060	0.058
	$N = 500$	0.042	0.071	0.060	0.018	0.051	0.049
	$N = 1000$	0.030	0.063	0.051	0.014	0.050	0.047
	$N = 2000$	0.022	0.058	0.045	0.012	0.050	0.047
Slope							
B	$N = 100$	-0.045	-0.087	-0.073	0.008	0.047	0.045
	$N = 500$	-0.019	-0.062	-0.048	0.010	0.050	0.047
	$N = 1000$	-0.015	-0.058	-0.044	0.011	0.051	0.048
	$N = 2000$	-0.014	-0.057	-0.043	0.011	0.051	0.048
RMSE	$N = 100$	0.097	0.121	0.112	0.038	0.060	0.058
	$N = 500$	0.042	0.072	0.060	0.019	0.053	0.050
	$N = 1000$	0.031	0.064	0.051	0.015	0.052	0.049
	$N = 2000$	0.024	0.060	0.047	0.014	0.051	0.049
Sine							
B	$N = 100$	-0.046	-0.087	-0.073	0.008	0.046	0.044
	$N = 500$	-0.020	-0.061	-0.047	0.011	0.049	0.047
	$N = 1000$	-0.017	-0.058	-0.044	0.012	0.051	0.048
	$N = 2000$	-0.016	-0.057	-0.043	0.012	0.051	0.048
RMSE	$N = 100$	0.098	0.122	0.112	0.038	0.060	0.058
	$N = 500$	0.044	0.072	0.061	0.020	0.052	0.050
	$N = 1000$	0.032	0.064	0.052	0.017	0.052	0.050
	$N = 2000$	0.025	0.060	0.047	0.014	0.051	0.049
U-Shape							
B	$N = 100$	-0.047	-0.091	-0.077	0.010	0.051	0.048
	$N = 500$	-0.021	-0.065	-0.052	0.012	0.054	0.051
	$N = 1000$	-0.018	-0.062	-0.048	0.013	0.054	0.052
	$N = 2000$	-0.017	-0.061	-0.047	0.013	0.054	0.052
RMSE	$N = 100$	0.098	0.125	0.116	0.038	0.064	0.062
	$N = 500$	0.044	0.076	0.065	0.021	0.057	0.054
	$N = 1000$	0.032	0.068	0.055	0.017	0.056	0.053
	$N = 2000$	0.025	0.064	0.051	0.015	0.055	0.053

et al. (2024). They improve the accuracy in small sample sizes, but do not make much difference for large N.

We have examined the estimation of the function G. The fRE decreases with the larger sample size, from approximately 22% for a sample of 100 to about 5% for a sample of 2000. This provides evidence supporting the consistency of our estimation of G. There are no obvious differences quality of estimation between the four shapes of σ functions.

The choice of α affects the estimation of ε and σ_ε in Procedures B and C, but not much. Table 6 in Kokoszka et al. (2024) shows the bias and the RMSE of estimation error of φ and σ_ε^2 under the U-shaped $\sigma(u)$ with $\alpha \in \{5\Delta, 20\Delta, 40\Delta\}$. The bias and the RMSE exhibit a marginal reduction when a larger value of α is used. Lastly, even setting $\alpha = 40\Delta$, Procedures B and C still still underperform Procedure A.

6. Summary and further work

We have proposed a model for a sequence of curves of normalized intraday asset prices. The curves are functional analogs of point-to-point daily returns. The model extends the extensively used stochastic volatility paradigm to the setting of functional time series. It includes day-to-day dynamics that quantify the dependence between the daily curves as well as a diffusion model for the evolution of the curves within each day. We have developed estimation methodology for this model and justified it both theoretically and via a data application and an informative simulation study.

The formulation of the model framework and effective estimation methodology opens up paths toward further developments. In particular, order selection, the selection of p in Section 4, may be practically relevant. Related to this, goodness-of-fit tests that consider the suitability of any model in the model family we introduced is another important problem. The ability of the model to predict future curves should be explored. Our theorems show that the estimators attain the optimal $N^{-1/2}$ rate, but their finite sample (and even asymptotic) standard errors are difficult to estimate. Block bootstrap or subsampling approaches could be explored. Another direction that builds on the theoretical optimality of our estimators, is their aggregation, e.g. Lavancier and Rochet (2016). One could explore if optimal weights on the three estimators could be derived that would result in an estimator superior to the one used in Procedure A. Beyond that Diffusions driven by the Wiener process could be replaced by diffusions with a jump component, and the advantages of such an extension could be explored. All these tasks will require new theoretical derivations and extensive numerical studies. It is hoped that this paper will motivate research in these and other similar directions, including more extensive applications.

Acknowledgments

We are grateful for Dr. Jian Chen's help on collecting and cleaning the U.S. stock data. We thank two referees and the Associate Editor for insightful comments that helped us put this work in a broader perspective and substantively improve the presentation.

Funding

This research was partially supported by the United States NSF grant DMS–2123761.

Supplementary Material

The online supplement, Kokoszka et al. (2024), contains the proofs of the results of Sections 3 and 4, as well as information on data processing and additional tables and graphs referred to above.

References

AÏT-SAHALIA, Y. and JACOD, J. (2014). *High-Frequency Financial Econometrics*, Illustrated edition ed. Princeton University Press, Princeton. https://doi.org/10.1515/9781400850327

ASAI, M., MCALEER, M. and YU, J. (2006). Multivariate stochastic volatility: a review. *Econometric Rev.* **25** 145–175. https://doi.org/10.1080/07474930600713564

AUE, A., HORVÁTH, L. and PELLATT, D. (2017). Functional generalized autoregressive conditional heteroskedasticity. *J. Time Series Anal.* **38** 3–21. https://doi.org/10.1111/jtsa.12192

BACHELIER, L. (1900). Théorie de la spéculation. *Annales Scientifiques de l'Ecole Normale Supérieure* **III-17** 21–86. Translated in *The Random Character of Stock Market Prices*, Ed. P.H. Cootner, pp. 17-78, Cambridge, MIT Press, 1964.

BARNDORFF-NIELSEN, O. and SHEPHARD, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. Roy. Statist. Soc. Ser. B* **64** 253–280. https://doi.org/10.1111/1467-9868.00336

BAUWENS, L., LAURENT, S. and ROMBOUTS, J. V. K. (2006). Multivariate GARCH models: a survey. *J. Appl. Econometrics* **21** 79–109. https://doi.org/10.1002/jae.842

BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

BOLLERSLEV, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Rev. Econ. Stat.* **72** 498–505. https://doi.org/10.2307/2109358

BOLLERSLEV, T., ENGLE, R. F. and WOOLDRIDGE, J. M. (1988). A capital asset pricing model with time-varying covariances. *J. Political Econ.* **96** 116–131. https://www.jstor.org/stable/1830713

BOSQ, D. (2000). *Linear Processes in Function Spaces*. Springer. https://doi.org/10.1007/978-1-4612-1154-9

BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*. Springer. https://doi.org/10.1007/978-3-319-29854-2_4

CEROVECKI, C., FRANCQ, C., HÖRMANN, S. and ZAKOÏAN, J. (2019). Functional GARCH models: the quasi-likelihood approach and its applications. *J. Econometrics* **209** 353–375. https://doi.org/10.1016/j.jeconom.2019.01.006

CHONG, C., HOFFMANN, M., LIU, Y., ROSENBAUM, M. and SZYMANSKI, G. (2022). Statistical inference for rough volatility: Minimax Theory. *arXiv:2210.01214 [math.ST]*.

CHRISTENSEN, K., HOUNYO, U. and PODOLSKIJ, M. (2018). Is the diurnal pattern sufficient to explain intraday variation in volatility? a nonparametric assessment. *J. Econometrics* **205** 336–362. https://doi.org/10.1016/j.jeconom.2018.03.016

DANIELSSON, J. (1998). Multivariate stochastic volatility models: estimation and a comparison with VGARCH models. *J. Empir. Finance.* **5** 155–173. https://doi.org/10.1016/S0927-5398(97)00016-9

ENGLE, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. https://doi.org/10.2307/1912773

ENGLE, R. (2002). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econom. Statist.* **20** 339–350. https://doi.org/10.1198/073500102288618487

ENGLE, R. F. and KRONER, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econom. Theory* **11** 122–150. https://doi.org/10.1017/S0266466600009063

GALBRAITH, J. W., ZINDE-WALSH, V. and ZHU, J. (2015). GARCH model estimation using estimated quadratic variation. *Econometric Rev.* **34**. https://doi.org/10.1080/07474938.2014.956629

LE GALL, J.-F. (2016). *Brownian Motion, Martingales, and Stochastic Calculus*. Springer. https://doi.org/10.1007/978-3-319-31089-3

HALL, P., MÜLLER, H. G. and WANG, J. L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Ann. Statist.* **34** 1493–1517. https://doi.org/10.1214/009053606000000272

HARVEY, A., RUIZ, E. and SHEPHARD, N. (1994). Multivariate stochastic variance models. *Rev. Econ. Stud.* **61** 247–264. https://doi.org/10.2307/2297980

HARVEY, A. and SHEPHARD, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *J. Bus. Econom. Statist.* **14** 429–434. https://doi.org/10.2307/1392251

HÖRMANN, S., HORVÁTH, L. and REEDER, R. (2013). A functional version of the ARCH model. *Econom. Theory* **29** 267–288. https://doi.org/DOI:10.1017/S0266466612000345

HORVÁTH, L., KOKOSZKA, P. and RICE, G. (2014). Testing stationarity of functional time series. *J. Econometrics* **179** 66–82. http://dx.doi.org/10.1016/j.jeconom.2013.11.002

JANG, P. A., JAUCH, M. and MATTESON, D. S. (2022). Functional stochastic volatility in financial option surfaces. *Data Scien. Scien.* **1** 6–19. https://doi.org/10.1080/26941899.2022.2152764

KARATZAS, I. and SHREVE, S. (1991). *Brownian Motion and Stochastic Calculus*, 2nd edition ed. Springer, New York. https://doi.org/10.1007/978-1-4612-0949-2

KOKOSZKA, P. and REIMHERR, M. (2013). Determining the order of the functional autoregressive model. *J. Time Series Anal.* **24** 116–129. https://doi.org/10.1111/j.1467-9892.2012.00816.x

KOKOSZKA, P., MOHAMMADI JOUZDANI, N., WANG, H. and WANG, S. (2024). Supplement to "Functional diffusion driven stochastic volatility model". *Bernoulli* Online supplement to this paper.

KUENZER, T. (2024). Estimation of functional ARMA models. *Bernoulli* **30** 117–142. https://doi.org/10.3150/23-BEJ1591

LAVANCIER, F. and ROCHET, P. (2016). A general procedure to combine estimators. *Comput. Statist. Data Anal.* **94** 175–192. https://doi.org/10.1016/j.csda.2015.08.001

LI, Y. and HSING, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Ann. Statist.* **38** 3321–3351. https://doi.org/DOI:10.1214/10-AOS813

MÖRTERS, P. and PERES, Y. (2010). *Brownian Motion*. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511750489

MÜLLER, H.-G., SEN, R. and STADTMÜLLER, U. (2011). Functional data analysis for volatility. *J. Econometrics* **165** 233–245. https://doi.org/10.1016/j.jeconom.2011.08.002

ØKSENDAL, B. (2003). *Stochastic Differential Equations*. *Universitext*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-14394-6

PAPARODITIS, E. and SHANG, H. (2023). Bootstrap prediction bands for functional time series. *J. Amer. Statist. Assoc.* **00** 1–15. Forthcoming. https://doi.org/10.1080/01621459.2023.2231056

ROSENBAUM, M. (2008). Estimation of the volatility persistence in a discretely observed diffusion model. *Stochastic Process. Appl.* **118** 1434–1462. https://doi.org/10.1016/j.spa.2007.09.004

SHAMAN, P. and STINE, R. (1988). The bias of autoregressive coefficient estimators. *J. Amer. Statist. Assoc.* **83** 842–848. https://doi.org/10.2307/2289315

SHEPHARD, N. (1996). Statistical Aspects of ARCH and Stochastic Volatility. In *Time Series Models* Chapman and Hall/CRC.

SILVENNOINEN, A. and TERÄSVIRTA, T. (2009). Multivariate GARCH Models. In *Handbook of Financial Time Series* (T. Mikosch, J.-P. Kreiß, R. A. Davis and T. G. Andersen, eds.) 201–229. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-71297-8_9

TAYLOR, S. J. (1982). Financial returns modelled by the product of two stochastic processes: a study of daily sugar prices, 1961-79. *Time series analysis: theory and practice* **1**.

TAYLOR, S. J. (1994). Modeling stochastic volatility: a review and comparative study. *Math. Finance* **4** 183–204. https://doi.org/10.1111/j.1467-9965.1994.tb00057.x

TAYLOR, S. J. (2008). *Modelling Financial Time Series*. World Scientific. https://doi.org/10.1142/6578

VISSER, M. (2011). GARCH Parameter Estimation Using High-Frequency Data. *J. Fin. Econom.* **9** 162–197. https://doi.org/10.1093/jjfinec/nbq017

VRONTOS, I. D., DELLAPORTAS, P. and POLITIS, D. N. (2003). A full-factor multivariate GARCH model. *Econom. J.* **6** 312–334. https://doi.org/10.1111/1368-423X.t01-1-00111

WILMS, I., BASU, S., BIEN, J. and MATTESON, D. (2023). Sparse identification and estimation of large-scale vector autoregressive moving averages. *J. Amer. Statist. Assoc.* **118** 571–582. https://doi.org/10.1080/01621459.2021.1942013

YAO, F., MÜLLER, H.-G. and WANG, J.-L. (2005). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590. https://doi.org/10.1198/016214504000001745

YU, J. and MEYER, R. (2006). Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Rev.* **25** 361–384. https://doi.org/10.1080/07474930600713465