



**HAL**  
open science

# Signature volatility models: pricing and hedging with Fourier

Eduardo Abi Jaber, Louis-Amand Gérard

► **To cite this version:**

Eduardo Abi Jaber, Louis-Amand Gérard. Signature volatility models: pricing and hedging with Fourier. 2024. hal-04435238

**HAL Id: hal-04435238**

**<https://hal.science/hal-04435238v1>**

Preprint submitted on 2 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Signature volatility models: pricing and hedging with Fourier

Eduardo Abi Jaber <sup>\*1</sup> and Louis-Amand Gérard <sup>†2,3</sup>

<sup>1</sup>École Polytechnique, CMAP

<sup>2</sup>Université Paris 1 Panthéon-Sorbonne, CES

<sup>3</sup>Gefip

February 2, 2024

## Abstract

We consider a stochastic volatility model where the dynamics of the volatility are given by a possibly infinite linear combination of the elements of the time extended signature of a Brownian motion. First, we show that the model is remarkably universal, as it includes, but is not limited to, the celebrated Stein-Stein, Bergomi, and Heston models, together with some path-dependent variants. Second, we derive the joint characteristic functional of the log-price and integrated variance provided that some infinite-dimensional extended tensor algebra valued Riccati equation admits a solution. This allows us to price and (quadratically) hedge certain European and path-dependent options using Fourier inversion techniques. We highlight the efficiency and accuracy of these Fourier techniques in a comprehensive numerical study.

**MSC 2020:** 60L10, 91G20, 91G60

**Keywords:** stochastic volatility, path signature, pricing, hedging, calibration, Fourier methods

## 1 Introduction

An important challenge of stochastic volatility modeling is the construction of realistic models that remain tractable for option pricing, risk hedging, trading and calibration purposes. Notable realistic features encompass various aspects, such as inter-temporal and path dependencies, which are inherent phenomena in financial markets. These phenomena have been established empirically at different timescales through extensive research, either in the form of long/short range dependence since Mandelbrot and Van Ness [50] and as documented in [9, 22, 40, 41, 23] or in the form of self-excitation of financial markets [11]. They can also be understood more strategically using the transitory nature of the decisions of market participants and their impact on prices [16, 17].

Incorporating path dependencies and memory effects into the modeling framework gives rise to non-Markovian models, which generally pose computational challenges. Recent developments have identified interesting mathematical classes of models that address this issue. One such class comprises stochastic Volterra models, which provide a flexible framework for capturing certain inter-temporal dependencies while maintaining computational tractability under specific affine and quadratic structures [1, 3, 25, 26, 32]. Another avenue worth exploring lies in the application of signature-based methods. The signature of a path, initially introduced by Chen [20] in 1957, consists of the (infinite) sequence of iterated integrals of a path. It plays a crucial role in the theory of rough paths [37, 46] and is recently gaining considerable momentum in the fields of Machine Learning [21, 35, 51] and Mathematical Finance [10, 13, 14, 18, 19, 27, 28, 31, 48] namely due to its universal

---

<sup>\*</sup>eduardo.abi-jaber@polytechnique.edu. The first author is grateful for the financial support from the Chaires FiME-FDD, Financial Risks, Deep Finance & Statistics and Machine Learning and systematic methods in finance at École Polytechnique.

<sup>†</sup>louis-amand.gerard@etu.univ-paris1.fr. We would like to thank Olivier Guéant for fruitful discussions, and Yuxing Huang for his precious comments.

linearization property: any functional of the path can be approximated by a linear combination of the elements of the signature of the path, provided some regularity. In our present work, we explore the modeling and numerical aspects of these signature-based approaches for stochastic volatility modeling. We uncover their potential in addressing the computational challenges associated with path-dependent models and we highlight their versatility and universality.

We consider a class of stochastic volatility models for a stock price  $S$  in the form

$$dS_t = S_t \Sigma_t \left( \rho dW_t + \sqrt{1 - \rho^2} dW_t^\perp \right),$$

where the stochastic volatility process  $\Sigma$  is of a general path-dependent form

$$f(t, W_{0 \leq s \leq t}), \quad t \leq T,$$

for a certain class of measurable functions  $f$  and  $(W, W^\perp)$  a two-dimensional Brownian motion. The correlation  $\rho \in [-1, 1]$  accounts for the leverage effect. More precisely, we assume that the volatility process  $\Sigma_t$  is a (possibly infinite) linear combination of the elements of the signature process  $\widehat{\mathbb{W}}_t$  of the time extended Brownian motion  $\widehat{W}_t := (t, W_t)$  defined by the infinite sequence of iterated Stratonovich [60] integrals:

$$\widehat{\mathbb{W}}_t = \left( 1, \begin{pmatrix} t \\ W_t \end{pmatrix}, \begin{pmatrix} \frac{t^2}{2!} & \int_0^t s dW_s \\ \int_0^t W_s ds & \frac{W_t^2}{2!} \end{pmatrix}, \begin{pmatrix} \frac{t^3}{3!} & \int_0^t \int_0^s \frac{s^2}{2!} dW_s \\ \int_0^t \int_0^s u dW_u ds & \int_0^t \int_0^s u dW_u \circ dW_s \\ \int_0^t \int_0^s W_u du ds & \int_0^t \int_0^s W_u du \circ dW_s \\ \int_0^t \frac{W_s^2}{2!} ds & \int_0^t \frac{W_s^3}{3!} \end{pmatrix}, \dots \right). \quad (1.1)$$

We call such models signature volatility models. They have been introduced by Arribas et al. [10], with  $\Sigma_t$  a finite linear combination of elements of  $\widehat{\mathbb{W}}_t$ , and their theoretical and empirical properties have been studied further in Cuchiero et al. [27, 28] for pricing and calibration purposes. They allow to naturally embed inter-temporal and path dependencies. The construction is very elementary, yet, from the mathematical perspective, such class of models enjoys a beautiful and powerful universality feature. So far, the universal approximation property of signatures have been invoked to argue that “the framework is universal in the sense that classical models can be approximated arbitrarily well” by signature volatility models, see for instance [27, proposition 2.14]. In contrast to the existing literature on signature volatility models, we allow  $\Sigma_t$  to be an infinite linear combination of the elements of the signature process  $\widehat{\mathbb{W}}_t$ . This introduces intricate theoretical challenges, particularly regarding convergence issues, but provides a more profound comprehension of the elegant universal structure inherent to these models. Our approach draws inspiration from Cuchiero et al. [29] where infinite linear combinations of signature elements are considered in the context of signature stochastic differential equations.

**Universality and flexibility of signature volatility models.** In a first step, by considering infinite linear combinations of signature elements, we go beyond the ‘approximated universality’ and we prove that the class of signature volatility models is universal in the sense of exact representations. We show that many classical and popular Markovian models, and even more advanced not necessarily Markovian models, belong to the class of signature volatility models, using novel exact representations formulas derived in the sequel and in the accompanying paper [6]. This includes:

- (i) **Affine Markovian models:** The models of Stein and Stein [58], Schöbel and Zhu [55] and Heston [42] which became popular because of the explicit knowledge of the characteristic function of the log-price, allowing for fast and accurate pricing and hedging using Fourier inversion techniques.
- (ii) **Non-affine Markovian models:** The models of Bergomi [15] and Hull and White [43], which are more flexible than their affine counterparts but less tractable. In addition, the recently introduced Quintic Ornstein-Uhlenbeck model of Abi Jaber et al. [4], which is able to jointly capture SPX and VIX smiles, also belongs to the class of signature volatility models.

- (iii) **Non-Markovian models:** A large class of models based on stochastic delayed equations and Volterra processes, including the class of polynomial Gaussian volatility models [5], Volterra Stein-Stein model [1], the Volterra [2] and rough [12] Bergomi models.

The above list is far from being exhaustive and we believe that more known and important volatility models can be exactly re-written as a signature stochastic volatility model. In addition, when such exact representation cannot be found for a specific model, building an approximated signature volatility model is possible thanks to the universal approximation property of the signature. The representations are derived in Section 3 and illustrated on numerical examples.

**Tractability of signature volatility models.** In a second step, we develop a generic framework for pricing and quadratic hedging certain vanilla and path-dependent options on the log-price and the integrated variance using Fourier inversion technology. More specifically, we obtain in Theorem 4.1 that for any signature volatility model, the joint characteristic functional of the log-price and the integrated variance is known up to the solution of an infinite dimensional Riccati equation. This result opens the door to fast and accurate Fourier pricing and hedging going beyond the standard affine classes (i) including the whole list of Markovian and non-Markovian models in (ii)-(iii) above, for which pricing and hedging is an intricate task. Our representation formula for the characteristic function can be directly related to the ones that have appeared in Cuchiero et al. [29], and share similarities with the formulas in Friz et al. [38] and Lyons et al. [49].

Using a comprehensive numerical study we highlight the efficiency and the accuracy of the Fourier techniques for pricing and hedging in signature volatility models in Sections 5 and 6. We stress that the numerics are not straightforward, since they involve several approximations and truncations. We use ideas in the spirit of ‘control variate’ with Black-Scholes prices and deltas to stabilize the Fourier inversions and reduce the number of evaluations of the characteristic functional. We also point out that the proposed implementation is both generic and scalable, as it can be applied, or fine-tuned if needed, to any signature volatility model. It only requires as inputs the coefficients (i.e. the parameters) of the signature volatility model, which essentially define the model itself whether it is Markovian/non-Markovian/non-semimartingale, and it generates as outputs option prices and hedging strategies by Fourier methods.

**Outline.** Section 2 introduces the framework of signatures with a focus on infinite linear combinations of signature elements. In Section 3, we define signature volatility models and study their representation properties. Section 4 derives the characteristic functional of the log-price and integrated variance. As applications, pricing and hedging of various European and Asian options under the signature volatility model are implemented in Sections 5 and 6, where we also include calibration examples on simulated and market data. Finally, Appendix A collects some proofs for the representations found in Subsection 3.1.

## 2 A primer on signatures

### 2.1 Tensor algebra

In this section, we setup the framework for dealing with signatures of semimartingales. One can also refer to the first sections in [13, 28, 48].

Let  $d \in \mathbb{N}$  and denote by  $\otimes$  the tensor product over  $\mathbb{R}^d$ , e.g.  $(x \otimes y \otimes z)_{ijk} = x_i y_j z_k$ , for  $i, j, k = 1, \dots, d$ , for  $x, y, z \in \mathbb{R}^d$ . For  $n \geq 1$ , we denote by  $(\mathbb{R}^d)^{\otimes n}$  the space of tensors of order  $n$  and by  $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$ . In the sequel, we will consider mathematical objects, path signatures, that live on the extended tensor algebra space  $T((\mathbb{R}^d))$  over  $\mathbb{R}^d$ , that is the space of (infinite) sequences of tensors defined by

$$T((\mathbb{R}^d)) := \{ \ell = (\ell^n)_{n=0}^{\infty} : \ell^n \in (\mathbb{R}^d)^{\otimes n} \}.$$

Similarly, for  $M \geq 0$ , we define the truncated tensor algebra  $T^M(\mathbb{R}^d)$  as the space of sequences of tensors of order at most  $M$  defined by

$$T^M(\mathbb{R}^d) := \{ \ell \in T((\mathbb{R}^d)) : \ell^n = 0, \text{ for all } n > M \},$$

and the tensor algebra  $T(\mathbb{R}^d)$  as the space of all finite sequences of tensors defined by

$$T(\mathbb{R}^d) := \bigcup_{M \in \mathbb{N}} T^M(\mathbb{R}^d).$$

We clearly have  $T(\mathbb{R}^d) \subset T((\mathbb{R}^d))$ . For  $\ell = (\ell^n)_{n \in \mathbb{N}}$ ,  $\mathbf{p} = (\mathbf{p}^n)_{n \in \mathbb{N}} \in T((\mathbb{R}^d))$  and  $\lambda \in \mathbb{R}$ , we define the following operations:

$$\begin{aligned} \ell + \mathbf{p} &:= (\ell^n + \mathbf{p}^n)_{n \in \mathbb{N}} \\ \ell \otimes \mathbf{p} &:= \left( \sum_{k=0}^n \ell^k \otimes \mathbf{p}^{n-k} \right)_{n \in \mathbb{N}} \\ \lambda \ell &:= (\lambda \ell^n)_{n \in \mathbb{N}}. \end{aligned}$$

These operations induce analogous operations on  $T^M(\mathbb{R}^d)$  and  $T(\mathbb{R}^d)$ .

**Important notations.** Let  $\{e_1, \dots, e_d\} \subset \mathbb{R}^d$  be the canonical basis of  $\mathbb{R}^d$  and  $A_d = \{\mathbf{1}, \mathbf{2}, \dots, \mathbf{d}\}$  be the corresponding alphabet. To ease reading, for  $i \in \{1, \dots, d\}$ , we write  $e_i$  as the blue letter  $\mathbf{i}$  and for  $n \geq 1, i_1, \dots, i_n \in \{1, \dots, d\}$ , we write  $e_{i_1} \otimes \dots \otimes e_{i_n}$  as the concatenation of letters  $\mathbf{i}_1 \cdots \mathbf{i}_n$ , that we call a word of length  $n$ . We note that  $(e_{i_1} \otimes \dots \otimes e_{i_n})_{(i_1, \dots, i_n) \in \{1, \dots, d\}^n}$  is a basis of  $(\mathbb{R}^d)^{\otimes n}$  that can be identified with the set of words of length  $n$  defined by

$$V_n := \{\mathbf{i}_1 \cdots \mathbf{i}_n : \mathbf{i}_k \in A_d \text{ for } k = 1, 2, \dots, n\}. \quad (2.1)$$

Moreover, we denote by  $\emptyset$  the empty word and by  $V_0 = \{\emptyset\}$  which serves as a basis for  $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$ . It follows that  $V := \bigcup_{n \geq 0} V_n$  represents the standard basis of  $T((\mathbb{R}^d))$ . In particular, every  $\ell \in T((\mathbb{R}^d))$ , can be decomposed as

$$\ell = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \mathbf{v}, \quad (2.2)$$

where  $\ell^{\mathbf{v}}$  is the real coefficient of  $\ell$  at coordinate  $\mathbf{v}$ . Representation (2.2) will be frequently used in the paper. We stress again that in the sequel, every blue ‘word’  $\mathbf{v} \in V$  represents an element of the canonical basis of  $T((\mathbb{R}^d))$ , i.e. there exists  $n \geq 0$  such that  $\mathbf{v}$  in the form  $\mathbf{v} = \mathbf{i}_1 \cdots \mathbf{i}_n$ , which represents the element  $e_{i_1} \otimes \dots \otimes e_{i_n}$ . The concatenation  $\ell \mathbf{v}$  of elements  $\ell \in T((\mathbb{R}^d))$  and the word  $\mathbf{v} = \mathbf{i}_1 \cdots \mathbf{i}_n$  means  $\ell \otimes e_{i_1} \otimes \dots \otimes e_{i_n}$ .

In addition to the decomposition (2.2) of elements  $\ell \in T((\mathbb{R}^d))$ , we introduce the projection  $\ell|_{\mathbf{u}} \in T((\mathbb{R}^d))$  as

$$\ell|_{\mathbf{u}} := \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{u}\mathbf{v}} \mathbf{v} \quad (2.3)$$

for all  $\mathbf{u} \in V$ . The projection plays an important role in the space of iterated integrals as it is closely linked to partial differentiation, in contrast with the concatenation that relates to integration. It will be used throughout the paper.

*Remark 2.0.1.* The projection allows us to decompose elements of the extended tensor algebra  $\ell \in T((\mathbb{R}^d))$  as

$$\ell = \ell^{\emptyset} \emptyset + \sum_{\mathbf{i} \in A_d} \ell|_{\mathbf{i}} \mathbf{i}.$$

This decomposition is quite natural as, when iterated, it gives back the decomposition in (2.2). ■

*Example 2.0.1.* Take the alphabet  $A_3 = \{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$  and let  $\ell = 4 \cdot \emptyset + 3 \cdot \mathbf{1} - 1 \cdot \mathbf{12} + 2 \cdot \mathbf{2212}$ , then

- $\ell^{\emptyset} = 4 \cdot \emptyset$ ,
- $\ell|_{\mathbf{1}} = 3 \cdot \emptyset$ ,
- $\ell|_{\mathbf{2}} = -1 \cdot \mathbf{1} + 2 \cdot \mathbf{221}$ ,
- $\ell|_{\mathbf{3}} = 0$ .

■

We now define the bracket between  $\ell \in T(\mathbb{R}^d)$  and  $\mathbf{p} \in T((\mathbb{R}^d))$  by

$$\langle \ell, \mathbf{p} \rangle := \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \mathbf{p}^{\mathbf{v}}. \quad (2.4)$$

Notice that it is actually well defined as  $\ell$  has finitely many non-zero terms. For  $\ell \in T((\mathbb{R}^d))$ , the series in (2.4) involves infinitely many terms and requires special care, this is discussed in Subsection 2.4.

*Example 2.0.2.* Take the alphabet  $A_2 = \{\mathbf{1}, \mathbf{2}\}$  and let  $\mathbf{p} := 1 \cdot \emptyset + 2 \cdot \mathbf{2} - 3 \cdot \mathbf{12} + 4 \cdot \mathbf{111} \in T((\mathbb{R}^d))$ , then

- $\langle \emptyset, \mathbf{p} \rangle = 1,$
- $\langle \mathbf{2} + 2 \cdot \mathbf{12}, \mathbf{p} \rangle = 2 - 6 = -4,$
- $\langle -2 \cdot \emptyset + \mathbf{111}, \mathbf{p} \rangle = -2 + 4 = 2,$
- $\langle \mathbf{p}, \mathbf{p} \rangle = 1 + 4 + 9 + 16 = 20.$

■

We will also consider another operation on the space of words, the shuffle product. The shuffle product plays a crucial role for an integration by parts formula on the space of iterated integrals, see Proposition 2.6 below.

**Definition 2.1** (Shuffle product). *The shuffle product  $\sqcup\sqcup: V \times V \rightarrow T(\mathbb{R}^d)$  is defined inductively for all words  $\mathbf{v}$  and  $\mathbf{w}$  and all letters  $\mathbf{i}$  and  $\mathbf{j}$  in  $A_d$  by*

$$\begin{aligned} (\mathbf{vi}) \sqcup\sqcup (\mathbf{wj}) &= (\mathbf{v} \sqcup\sqcup (\mathbf{wj}))\mathbf{i} + ((\mathbf{vi}) \sqcup\sqcup \mathbf{w})\mathbf{j}, \\ \mathbf{w} \sqcup\sqcup \emptyset &= \emptyset \sqcup\sqcup \mathbf{w} = \mathbf{w}. \end{aligned}$$

*With some abuse of notation, the shuffle product on  $T((\mathbb{R}^d))$  induced by the shuffle product on  $V$  will also be denoted by  $\sqcup\sqcup$ . The shuffle product is clearly commutative. See [53] and [39] for more information on the shuffle product.*

The shuffle product corresponds to all riffle shuffles of two decks of cards together, which keeps the order of each single deck, as illustrated in the following example:

*Example 2.1.1.* We have

- $\mathbf{12} \sqcup\sqcup \mathbf{34} = \mathbf{1234} + \mathbf{1324} + \mathbf{3124} + \mathbf{1342} + \mathbf{3142} + \mathbf{3412},$
- $\mathbf{1} \sqcup\sqcup \mathbf{21} = \mathbf{121} + 2 \cdot \mathbf{211},$
- $(\mathbf{1} + \mathbf{31}) \sqcup\sqcup \mathbf{21} = \mathbf{121} + 2 \cdot \mathbf{211} + \mathbf{2131} + 2 \cdot \mathbf{2311} + \mathbf{3121} + 2 \cdot \mathbf{3211}.$

■

## 2.2 Resolvent and linear equation

For  $n \in \mathbb{N}$  and  $\ell \in T((\mathbb{R}^d))$ , we define the concatenation power of  $\ell$  by

$$\ell^{\otimes n} := \overbrace{\ell \otimes \ell \otimes \dots \otimes \ell}^{n \text{ times}},$$

with the convention that  $\ell^{\otimes 0} = \emptyset$ . For  $\ell \in T((\mathbb{R}^d))$  such that  $\ell^{\emptyset} = 0$ , we define the *resolvent* of  $\ell$  by

$$(\emptyset - \ell)^{-1} := \sum_{n=0}^{\infty} \ell^{\otimes n}. \quad (2.5)$$

The assumption  $\ell^{\emptyset} = 0$  assures that it is well-defined.

The resolvent allows us to solve linear algebraic equations, for instance:

**Proposition 2.2.** Let  $\mathbf{p}, \mathbf{q} \in T((\mathbb{R}^d))$  such that  $\mathbf{q}^\circ = 0$ , then the unique solution  $\ell \in T((\mathbb{R}^d))$  to the linear algebraic equation

$$\ell = \mathbf{p} + \ell \mathbf{q} \quad (2.6)$$

is given by

$$\ell = \mathbf{p}(\circ - \mathbf{q})^{-1},$$

with  $(\circ - \cdot)^{-1}$  as defined in (2.5).

*Proof.* It is easy to verify that  $\ell = \mathbf{q}(\circ - \mathbf{p})^{-1}$  is a solution of (2.6). On the other hand,  $\ell = \ell \mathbf{p}$ , together with the fact that  $\mathbf{p}^\circ \neq 1$ , implies that  $\ell = 0$ , verifying the uniqueness.  $\square$

Interestingly, whenever  $\ell$  is a linear combination of single letters, the resolvent of  $\ell$  is equal to the shuffle exponential  $e^{\sqcup \ell}$  defined by

$$e^{\sqcup \ell} := \sum_{n=0}^{\infty} \frac{\ell^{\sqcup n}}{n!}, \quad (2.7)$$

where

$$\ell^{\sqcup n} := \overbrace{\ell \sqcup \ell \sqcup \dots \sqcup \ell}^{n \text{ times}}, \quad n \geq 1, \quad \ell^{\sqcup 0} = \circ.$$

**Proposition 2.3.** Whenever  $\ell$  is of the form  $\sum_{\mathbf{i} \in A_d} \ell^{\mathbf{i}} \mathbf{i}$ , with  $\ell^{\mathbf{i}} \in \mathbb{R}$ , we have that

$$(\circ - \ell)^{-1} = e^{\sqcup \ell}.$$

In particular, this implies

$$e^{\sqcup \ell} = \circ + e^{\sqcup \ell} \ell = \circ + \ell e^{\sqcup \ell}.$$

*Proof.* Using Lemma 2.4 below, it is easy to see that  $\frac{1}{n!} \ell^{\sqcup n} = \ell^{\otimes n}$  whenever  $\ell = \sum_{\mathbf{i} \in A_d} \ell^{\mathbf{i}} \mathbf{i}$ , hence proving the proposition.  $\square$

**Lemma 2.4.** Let  $\mathbf{p} \in T((\mathbb{R}^d))$  and  $\ell = \sum_{\mathbf{i} \in A_d} \ell^{\mathbf{i}} \mathbf{i}$  with  $\ell^{\mathbf{i}} \in \mathbb{R}$ , then,

$$(\mathbf{p} \ell)^{\sqcup n} = n [\mathbf{p} \sqcup (\mathbf{p} \ell)^{\sqcup n-1}] \ell, \quad n \in \mathbb{N}.$$

*Proof.* By induction, assume that the property holds for some  $n \in \mathbb{N}$ , then

$$\begin{aligned} (\mathbf{p} \ell)^{\sqcup n+1} &= (\mathbf{p} \ell)^{\sqcup n} \sqcup \mathbf{p} \ell \\ &= n [\mathbf{p} \sqcup (\mathbf{p} \ell)^{\sqcup n-1}] \ell \sqcup \mathbf{p} \ell \\ &= n ([\mathbf{p} \sqcup (\mathbf{p} \ell)^{\sqcup n-1}] \ell \sqcup \mathbf{p}) \ell + n ([\mathbf{p} \sqcup (\mathbf{p} \ell)^{\sqcup n-1}] \sqcup \mathbf{p}) \ell \\ &= ((\mathbf{p} \ell)^{\sqcup n} \sqcup \mathbf{p}) \ell + n (\mathbf{p} \sqcup (\mathbf{p} \ell)^{\sqcup n}) \ell \\ &= (n+1) [\mathbf{p} \sqcup (\mathbf{p} \ell)^{\sqcup n}] \ell. \end{aligned}$$

Finally,  $(\mathbf{p} \ell)^{\sqcup 1} = \mathbf{p} \ell = [\mathbf{p} \sqcup \circ] \ell$  concludes the proof.  $\square$

### 2.3 Signatures

We define the (path) signature of a semimartingale process as the sequence of iterated stochastic integrals in the sense of Stratonovich. Throughout the paper, the Itô integral is denoted by  $\int_0^\cdot Y_t dX_t$  and the Stratonovich integral by  $\int_0^\cdot Y_t \circ dX_t$ . If both  $X$  and  $Y$  are semimartingales then, we have the relation  $\int_0^\cdot Y_t \circ dX_t = \int_0^\cdot Y_t dX_t + \frac{1}{2}[X, Y]$ .

**Definition 2.5** (Signature). Fix  $T > 0$ . Let  $(X_t)_{t \geq 0}$  be a continuous semimartingale in  $\mathbb{R}^d$  on some filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . The signature of  $X$  is defined by

$$\begin{aligned} \mathbb{X} : \Omega \times [0, T] &\rightarrow T((\mathbb{R}^d)) \\ (\omega, t) &\mapsto \mathbb{X}_t(\omega) := (1, \mathbb{X}_t^1(\omega), \dots, \mathbb{X}_t^n(\omega), \dots), \end{aligned}$$

where

$$\mathbb{X}_t^n := \int_{0 < u_1 < \dots < u_n < t} \circ dX_{u_1} \otimes \dots \otimes \circ dX_{u_n}$$

takes value in  $(\mathbb{R}^d)^{\otimes n}$ ,  $n \geq 0$ . Similarly, the truncated signature of order  $M \in \mathbb{N}$  is defined by

$$\begin{aligned} \mathbb{X}^{\leq M} : [0, T] &\rightarrow T^M(\mathbb{R}^d) \\ (\omega, t) &\mapsto \mathbb{X}_t^{\leq M}(\omega) := (1, \mathbb{X}_t^1(\omega), \dots, \mathbb{X}_t^M(\omega), 0, \dots, 0, \dots). \end{aligned}$$

The signature plays a similar role to polynomials on path-space. Indeed, in dimension  $d = 1$ , the signature of  $X$  is the sequence of monomials  $(\frac{1}{n!}(X_t - X_0)^n)_{n \in \mathbb{N}}$ . In particular, any finite combination of elements of the signature  $\langle \ell, \mathbb{X}_t \rangle$ , defined in (2.4) for  $\ell \in T^M(\mathbb{R}^d)$ , is a polynomial of degree  $M$  in  $X_t$ .

*Remark 2.5.1.* Explicitly we can write the term  $\mathbb{X}_t^n$  as  $(\mathbb{X}_t^{\mathbf{i}_1 \dots \mathbf{i}_n})_{(\mathbf{i}_1 \dots \mathbf{i}_n) \in V_n}$ . So the definition can be written in an iterated form as

$$\mathbb{X}_t^{\mathbf{i}_1 \dots \mathbf{i}_n} = \int_0^t \mathbb{X}_s^{\mathbf{i}_1 \dots \mathbf{i}_{n-1}} \circ dX_s^{\mathbf{i}_n}.$$

In what will follow, we are exclusively interested in the case  $d = 1$  and  $X_t = \widehat{W}_t := (t, W_t)$  where  $W$  is a 1-dimensional Brownian motion. Its first few signature orders are given by

$$\widehat{\mathbb{W}}_t^0 = 1, \quad \widehat{\mathbb{W}}_t^1 = \begin{pmatrix} t \\ W_t \end{pmatrix}, \quad \widehat{\mathbb{W}}_t^2 = \begin{pmatrix} \frac{t^2}{2!} & \int_0^t s dW_s \\ \int_0^t W_s ds & \frac{W_t^2}{2!} \end{pmatrix}$$

and

$$\widehat{\mathbb{W}}_t^3 = \begin{pmatrix} \frac{t^3}{3!} & \int_0^t \frac{s^2}{2!} dW_s & \\ \int_0^t \int_0^s u dW_u ds & \int_0^t \int_0^s u dW_u \circ dW_s & \\ \int_0^t \int_0^s W_u duds & \int_0^t \int_0^s W_u du \circ dW_s & \\ \int_0^t \frac{W_s^2}{2!} ds & \int_0^t \int_0^s W_u du & \frac{W_t^3}{3!} \end{pmatrix},$$

recall (1.1).

## 2.4 Infinite linear combinations of signature elements

In this section, we recall some results on infinite linear combinations  $\langle \ell, \widehat{\mathbb{W}}_t \rangle$  for certain admissible  $\ell \in T((\mathbb{R}^d))$  for which the infinite series will make sense. Two crucial ingredients for our paper are the shuffle product (Proposition 2.6) and an Itô's formula (Lemma 2.7). We follow the presentation in [6, Section 2], and we refer to [29] for more general results.

We first introduce the space  $\mathcal{A}$  of admissible elements  $\ell$  below, using the associated semi-norm:

$$\|\ell\|_t^{\mathcal{A}} := \sum_{n=0}^{\infty} \left| \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \widehat{\mathbb{W}}_t^{\mathbf{v}} \right|, \quad t \geq 0,$$

recall the definition of  $V_n$  in (2.1) and the decomposition (2.2). Whenever,  $\|\ell\|_t^{\mathcal{A}} < \infty$  a.s., the infinite linear combination

$$\langle \ell, \widehat{\mathbb{W}}_t \rangle = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \widehat{\mathbb{W}}_t^{\mathbf{v}}$$

is well-defined. This leads to the following definition for the admissible set  $\mathcal{A}$ :

$$\mathcal{A} := \{ \ell \in T((\mathbb{R}^2)) : \|\ell\|_t^{\mathcal{A}} < \infty \text{ for all } t \in [0, T] \text{ a.s.} \}.$$

Note that  $T(\mathbb{R}^2) \subset \mathcal{A}$  and that  $\langle \ell, \widehat{\mathbb{W}}_t \rangle$  is an extension of (2.4), as the two bracket operations  $\langle \cdot, \cdot \rangle$  coincide whenever  $\ell \in T(\mathbb{R}^2)$ .



The admissible set  $\mathcal{A}$  has another very interesting property, as it allows us to linearize polynomials on infinite linear combination of the signature, see Proposition 2.6. This is what most of the literature refers to when putting forth the linearization power of the signature.

**Proposition 2.6** (Shuffle property). *If  $\ell_1, \ell_2 \in \mathcal{A}$ , then  $\ell_1 \sqcup \ell_2 \in \mathcal{A}$  and*

$$\langle \ell_1, \widehat{\mathbb{W}}_t \rangle \langle \ell_2, \widehat{\mathbb{W}}_t \rangle = \langle \ell_1 \sqcup \ell_2, \widehat{\mathbb{W}}_t \rangle.$$

*Proof.* This follows from a particular instance of shuffle compatible partitions, see [29, Lemma 4.4]. Alternatively, we refer to [6, Proposition 3.2].  $\square$

*Example 2.6.1.* When  $d = 1$ , for  $\ell_1 \in T^M(\mathbb{R})$  and  $\ell_2 \in T^N(\mathbb{R})$ ,  $\langle \ell_1 \sqcup \ell_2, \widehat{\mathbb{W}}_t \rangle$  can be seen as analogous to the Cauchy product (scaled by  $\frac{1}{n!}$ ) and is thus a polynomial of degree  $M + N$ . Actually, the shuffle product is closely related to the integration by parts in the space of iterated Stratonovich integrals, see [39].  $\blacksquare$

For elements  $\ell \in \mathcal{A}$ , the process  $(\langle \ell, \widehat{\mathbb{W}}_t \rangle)_{t \leq T}$  is well-defined. An important question is to know whether it is a semimartingale and compute its Itô's decomposition. The answer is positive, thanks to an Itô's formula in Lemma 2.7 below, for elements in the set

$$\mathcal{I} := \left\{ \ell \in \mathcal{A} : \text{for all } t \in [0, T], \|\ell\|_t^{\mathcal{I}} < \infty \text{ and } \int_0^T \|\ell\|_t^{\mathcal{I}} dt < \infty \text{ a.s.} \right\},$$

where

$$\|\ell\|_t^{\mathcal{I}} := \|\ell|_{\mathbf{1}}\|_t^{\mathcal{A}} + \|\frac{1}{2}\ell|_{\mathbf{22}}\|_t^{\mathcal{A}} + (\|\ell|_{\mathbf{2}}\|_t^{\mathcal{A}})^2.$$

More generally, we state the result for time dependent linear combinations  $(\langle \ell_t, \widehat{\mathbb{W}}_t \rangle)_{t \leq T}$  with  $\ell : [0, T] \rightarrow \mathcal{A}$  in the set

$$\mathcal{I}' := \left\{ \ell : [0, T] \rightarrow \mathcal{A} : \begin{array}{l} \text{for all } t \in [0, T], \ell_t^{\mathbf{v}} \in C^1([0, T]) \text{ for all } \mathbf{v} \in V, \\ \text{and } \|\dot{\ell}_t\|_t^{\mathcal{A}} < \infty \text{ and } \int_0^T \|\dot{\ell}_t\|_t^{\mathcal{A}} dt < \infty \text{ a.s.} \end{array} \right\},$$

where  $\dot{\ell}_t := \sum_{\mathbf{v} \in V} \frac{d}{dt} \ell_t^{\mathbf{v}} \mathbf{v}$  for all  $t \in [0, T]$ .

**Lemma 2.7** (Itô's decomposition). *Let  $\ell \in \mathcal{I}$ , then*

$$d \langle \ell, \widehat{\mathbb{W}}_t \rangle = \langle \ell|_{\mathbf{1}} + \frac{1}{2} \ell|_{\mathbf{22}}, \widehat{\mathbb{W}}_t \rangle dt + \langle \ell|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \rangle dW_t. \quad (2.8)$$

Let  $\ell \in \mathcal{I}'$ , then

$$d \langle \ell_t, \widehat{\mathbb{W}}_t \rangle = \langle \dot{\ell}_t + \ell_t|_{\mathbf{1}} + \frac{1}{2} \ell_t|_{\mathbf{22}}, \widehat{\mathbb{W}}_t \rangle dt + \langle \ell_t|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \rangle dW_t.$$

*Proof.* The full proof can be found in [6, Section 3], however we will give the reader a sketch of the proof to illustrate the algebraic computations assuming a finite number of non-zero terms, i.e.  $\ell : [0, T] \rightarrow T(\mathbb{R}^2)$ . To obtain the full proof, it suffices then to apply dominated convergence theorems. First we show in (i) that

$$d \langle \ell_t, \widehat{\mathbb{W}}_t \rangle = \langle \dot{\ell}_t + \ell_t|_{\mathbf{1}}, \widehat{\mathbb{W}}_t \rangle dt + \langle \ell_t|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \rangle \circ dW_t,$$

and then in (ii) that

$$\langle \ell_t|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \rangle \circ dW_t = \langle \ell_t|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \rangle dW_t + \frac{1}{2} \langle \ell_t|_{\mathbf{22}}, \widehat{\mathbb{W}}_t \rangle dt.$$

(i) First recall the definition of the bracket (2.4)

$$\langle \ell_t, \widehat{\mathbb{W}}_t \rangle = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell_t^{\mathbf{v}} \widehat{\mathbb{W}}_t^{\mathbf{v}},$$

which allows us to write

$$\begin{aligned}
d\langle \ell_t, \widehat{\mathbb{W}}_t \rangle &= \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \left( \widehat{\mathbb{W}}_t^{\mathbf{v}} d\ell_t^{\mathbf{v}} + \ell_t^{\mathbf{v}} d\widehat{\mathbb{W}}_t^{\mathbf{v}} \right) \\
&= \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \dot{\ell}_t^{\mathbf{v}} \widehat{\mathbb{W}}_t^{\mathbf{v}} dt + \sum_{n=1}^{\infty} \sum_{\mathbf{v} \in V_{n-1}} \left( \ell_t^{\mathbf{v}1} \widehat{\mathbb{W}}_t^{\mathbf{v}} dt + \ell_t^{\mathbf{v}2} \widehat{\mathbb{W}}_t^{\mathbf{v}} \circ dW_t \right) \\
&= \langle \dot{\ell}_t, \widehat{\mathbb{W}}_t \rangle dt + \langle \ell_t|_1, \widehat{\mathbb{W}}_t \rangle dt + \langle \ell_t|_2, \widehat{\mathbb{W}}_t \rangle \circ dW_t.
\end{aligned}$$

The last equality comes from the definition of the projection (2.3).

(ii) We then make the transition between Stratonovich and Itô integrals:

$$\begin{aligned}
\langle \ell_t|_2, \widehat{\mathbb{W}}_t \rangle \circ dW_t &= \langle \ell_t|_2, \widehat{\mathbb{W}}_t \rangle dW_t + \frac{1}{2} d \left[ \langle \ell_t|_2, \widehat{\mathbb{W}}. \rangle, W. \right]_t \\
d \left[ \langle \ell_t|_2, \widehat{\mathbb{W}}. \rangle, W. \right]_t &= d \left[ \int_0^{\cdot} \langle \ell|_{22}, \widehat{\mathbb{W}}_s \rangle \circ dW_s, W. \right]_t \\
&= \langle \ell|_{22}, \widehat{\mathbb{W}}_t \rangle dt.
\end{aligned}$$

□

We note that we clearly have  $T(\mathbb{R}^2) \in \mathcal{I}$ , i.e. showing that finite linear combinations of the signature are always semimartingales. The following example highlights that Lemma 2.7 is indeed an extension of the usual Itô's formula.

*Example 2.7.1.* Fix an analytic function  $f(y) := \sum_{n \geq 0} a_n y^n$  with infinite radius of convergence which we apply to  $W$ . It is clear that

$$f(W_t) = \sum_{n \geq 0} a_n W_t^n = \langle \ell, \widehat{\mathbb{W}}_t \rangle,$$

where

$$\ell = \sum_{n \geq 0} a_n \mathbf{2}^{\sqcup n} = \sum_{n \geq 0} n! a_n \mathbf{2}^{\otimes n}.$$

In particular, the projections read:

$$\begin{aligned}
\ell|_1 &= 0, \\
\ell|_2 &= \sum_{n \geq 0} (n+1) a_{n+1} \mathbf{2}^{\otimes n}, \\
\ell|_{22} &= \sum_{n \geq 0} (n+2) a_{n+2} \mathbf{2}^{\otimes n}.
\end{aligned}$$

It is easy to verify that  $\ell, \ell|_2, \ell|_{22} \in \mathcal{A}$  since  $f$  has infinite radius of convergence. We can thus further derive that

$$f'(W_t) = \langle \ell|_2, \widehat{\mathbb{W}}_t \rangle, \quad f''(W_t) = \langle \ell|_{22}, \widehat{\mathbb{W}}_t \rangle.$$

On the other hand we can see that  $\|\ell|_2\|_t^A = \sum_{n=0}^{\infty} (n+1) |a_{n+1}| |W_t|^n$ , and since  $g(x) := \sum_{n \geq 0} (n+1) |a_{n+1}| x^n$  is also analytic and  $\|\ell|_2\|_t^A$  has continuous sample path almost surely, then

$$\sup_{t \in [0, T]} \|\ell|_2\|_t^A < \infty.$$

With similar arguments we can show that  $\sup_{t \in [0, T]} \|\ell|_{22}\|_t^A < \infty$ . This allows us to verify that  $\ell \in \mathcal{I}$ . An application of Itô's formula (2.8) yields

$$\langle \ell, \widehat{\mathbb{W}}_t \rangle = \ell^\circ + \int_0^t \langle \ell|_{\mathbf{2}}, \widehat{\mathbb{W}}_s \rangle dW_s + \frac{1}{2} \int_0^t \langle \ell|_{\mathbf{22}}, \widehat{\mathbb{W}}_s \rangle ds,$$

which is equivalent to

$$f(W_t) = f(0) + \int_0^t f'(W_s) dW_s + \frac{1}{2} \int_0^t f''(W_s) ds,$$

the standard Itô's formula. ■

### 3 The signature volatility model

Let  $(\Omega, \mathcal{F}, \mathbb{Q})$  be a probability space supporting a two-dimensional Brownian motion  $(W, W^\perp)$ . We denote by  $(\mathcal{F}_t)_{t \geq 0}$  the filtration generated by  $(W, W^\perp)$ . We set

$$B = \rho W + \sqrt{1 - \rho^2} W^\perp, \quad (3.1)$$

for some  $\rho \in [-1, 1]$ . We define the time-augmented process  $\widehat{W}_t = (t, W_t)$  and we consider that the dynamics of the risky asset  $S$ , under the risk neutral probability measure  $\mathbb{Q}$ , are given by a stochastic volatility model where the volatility process  $\Sigma$  is a (possibly infinite) linear combination of the signature of  $\widehat{W}$ :

$$\frac{dS_t}{S_t} = \Sigma_t dB_t, \quad (3.2)$$

$$\Sigma_t = \langle \sigma_t, \widehat{\mathbb{W}}_t \rangle, \quad (3.3)$$

where  $\sigma : [0, T] \rightarrow \mathcal{A}$  corresponds to the parameters of the volatility process and is such that

$$\int_0^T \mathbb{E} [\Sigma_t^2] dt < \infty. \quad (3.4)$$

We recall the definition of the set  $\mathcal{A}$  in (2.4). The condition (3.4) ensures that the stochastic integral

$$\int_0^\cdot \Sigma_s dB_s$$

is well defined as an Itô integral, so that there exists a unique solution to (3.2) given by

$$S_t = S_0 \exp \left( -\frac{1}{2} \int_0^t \Sigma_s^2 ds + \int_0^t \Sigma_s dB_s \right), \quad t \geq 0.$$

Condition (3.4) can be made more explicit by observing that the instantaneous variance  $\Sigma_t^2$  is also linear in the signature, i.e.

$$\Sigma_t^2 = \langle \sigma_t^{\sqcup\sqcup 2}, \widehat{\mathbb{W}}_t \rangle, \quad (3.5)$$

thanks to the shuffle product in Proposition 2.6. We note that the time-independent case  $\sigma_t = \sigma$  for all  $t \in [0, T]$  and some  $\sigma \in \mathcal{A}$ , leads to  $\int_0^T \mathbb{E} [\Sigma_t^2] dt = \langle \sigma^{\sqcup\sqcup 2} \mathbf{1}, \mathbb{E}[\widehat{\mathbb{W}}_T] \rangle$  where the quantity  $\mathbb{E}[\widehat{\mathbb{W}}_T]$  can be computed explicitly using Fawcett's formula [33] extended to time-augmented Brownian motions in [47, Proposition 4.10]:

$$\mathbb{E} [\widehat{\mathbb{W}}_t] = \sum_{n \geq 0} \frac{t^n}{n!} \left( \mathbf{1} + \frac{1}{2} \mathbf{22} \right)^{\otimes n}. \quad (3.6)$$

In practice, we will usually be interested in truncated elements  $\sigma : [0, T] \rightarrow T^M(\mathbb{R}^2)$ , for some  $M \in \mathbb{N}$ , which automatically satisfy the condition (3.4) since in this case  $\sigma$  has only a finite number of non-zero terms and all the integrated moments of the Brownian motion are finite.

Notice that for general  $\sigma : [0, T] \rightarrow \mathcal{A}$  the process  $\Sigma$  is not necessarily Markovian nor a semimartingale. It is a semimartingale if in addition  $\sigma \in \mathcal{I}$  thanks to Itô's formula in Lemma 2.7. For truncated elements  $\sigma : [0, T] \rightarrow T^M(\mathbb{R}^2)$  the process  $\Sigma$  is a semimartingale. So far, truncated elements have been considered in the related literature [10, 28].

In the next subsection, we highlight the flexibility introduced by infinite linear combinations of signature elements in terms of exact representations and provide numerical implementations of their truncated form, i.e.

$$\Sigma_t^{\leq M} := \left\langle \sigma^{\leq M}, \widehat{\mathbb{W}}_t \right\rangle, \quad M \geq 0,$$

where  $\sigma^{\leq M} : [0, T] \rightarrow T^M(\mathbb{R}^2)$  is the truncated form of  $\sigma$  at order  $M$ , i.e. its  $M$  first levels coincide with  $\sigma$  and everything else is set to 0.

### 3.1 Examples of exact representations

We first highlight the flexibility of the signature volatility model (3.2)-(3.3) by showing that it subsumes several known and useful Markovian and non-Markovian models based on Ornstein-Uhlenbeck processes, mean-reverting geometric Brownian motions, square-root processes<sup>1</sup> and processes with path-dependent dynamics including stochastic Volterra processes.

#### 3.1.1 Models based on the Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck (OU) process  $X$  given by

$$dX_t = \kappa(\theta - X_t)dt + \eta dW_t, \quad X_0 = x \in \mathbb{R}, \quad (3.7)$$

for  $\kappa, \theta, \eta \in \mathbb{R}$ , can be represented as an infinite linear combination of the signature elements of the time-extended Brownian motion, either in a time-independent or time-dependent way as shown in the next Lemma.

**Lemma 3.1.** *The unique solution  $X$  to (3.7) is given by*

$$X_t = \left\langle \ell^{\text{OU}}, \widehat{\mathbb{W}}_t \right\rangle, \quad \ell^{\text{OU}} = (x\mathbf{0} + \kappa\theta\mathbf{1} + \eta\mathbf{2})e^{\llbracket -\kappa \mathbf{1} \rrbracket}, \quad (3.8)$$

such that  $\ell^{\text{OU}} \in \mathcal{A}$ , with  $e^{\llbracket \cdot \rrbracket}$  the exponential shuffle defined in (2.7). Furthermore,  $X$  can also be written in terms of time-dependent coefficients:

$$X_t = \left\langle \tilde{\ell}_t^{\text{OU}}, \widehat{\mathbb{W}}_t \right\rangle, \quad \tilde{\ell}_t^{\text{OU}} = \theta\mathbf{0} + e^{-\kappa t} \left( (x - \theta)\mathbf{0} + \eta e^{\llbracket \kappa \mathbf{1} \rrbracket} \mathbf{2} \right), \quad (3.9)$$

*Proof.* The proof is detailed in Appendix A.1. □

*Example 3.1.1.* To be more explicit, up to order 3, the linear form of the Ornstein-Uhlenbeck process in (3.8) reads

$$\ell^{\text{OU}} = \left( x, \begin{pmatrix} -\kappa(x - \theta) \\ \eta \end{pmatrix}, \begin{pmatrix} \kappa^2(x - \theta) & 0 \\ -\kappa\eta & 0 \end{pmatrix}, \begin{pmatrix} -\kappa^3(x - \theta) & 0 \\ 0 & 0 \\ \kappa^2\eta & 0 \\ 0 & 0 \end{pmatrix}, \dots \right).$$

Although the representations (3.8) and (3.9) are equivalent, from a numerical perspective (3.9) is more advantageous. In the time-independent representation (3.8),  $e^{-\kappa t}$  is approximated by the first terms of its Taylor expansion  $\sum_{n \geq 0} \frac{(-\kappa t)^n}{n!}$ , which becomes numerically unstable for lower orders of truncation in the region  $t > 1/|\kappa|$ , see Figure 1. However, in the time-dependent representation (3.9),  $e^{-\kappa t}$  is exact and even though the approximate solution deteriorates with time, it stays stable and converges towards  $\theta$ , see Figure 2. ■

<sup>1</sup>The theoretical justification for the representation of the square-root process is a work in progress, it is validated numerically in Section 3.1.3 below.

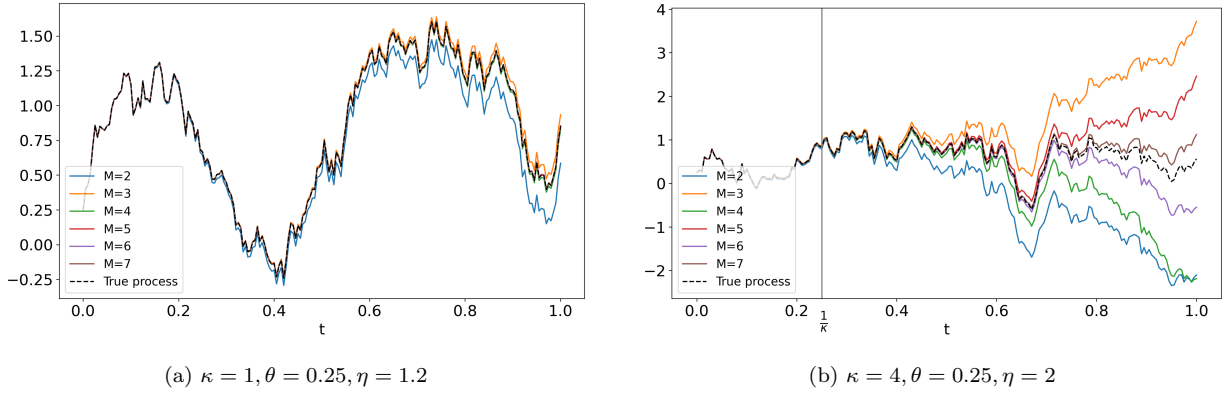


Figure 1: Trajectories of an Ornstein-Uhlenbeck process against their truncated time-independent linear representation (3.8), i.e.  $\langle \ell^{\text{OU}}, \widehat{\mathbb{W}}_t \rangle$ , for several truncation orders  $M$ .

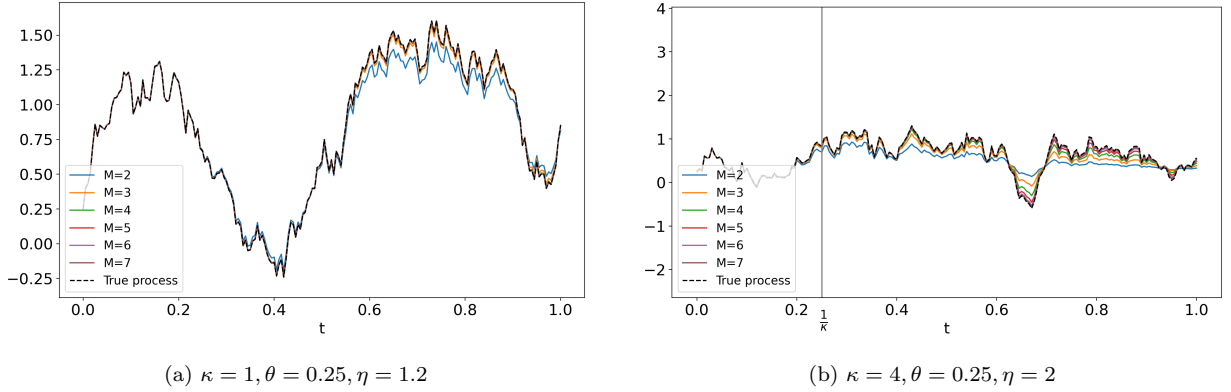


Figure 2: Trajectories of an Ornstein-Uhlenbeck process against their truncated time-dependent linear representation (3.9), i.e.  $\langle \hat{\ell}_t^{\text{OU}}, \widehat{\mathbb{W}}_t \rangle$ , for several truncation orders  $M$ .

If  $|\kappa t| < 1$ , we can see in the left-hand side of Figure 1 and Figure 2 that the truncated linear representations seems to converge quite quickly to the explicit solution of the Ornstein-Uhlenbeck and a truncation order  $M = 4$  is sufficient to get a relatively close fit.

Going back to our signature volatility model, the representations of the Ornstein-Uhlenbeck process in Lemma 3.1 combined with the shuffle property of Proposition 2.6 show that the model (3.2)-(3.3) nests any stochastic volatility model based on an Ornstein-Uhlenbeck process of the form

$$\frac{dS_t}{S_t} = f(t, X_t) dB_t, \quad f(t, x) = \sum_{k \geq 0} \alpha_k(t) x^k,$$

for some coefficients  $\alpha_k : [0, T] \rightarrow \mathbb{R}$  that ensure the absolute convergence of the infinite sum. More precisely, using the representation (3.8) or (3.9) and the shuffle property in Lemma 2.6, we can write

$$f(t, X_t) = \sum_{k \geq 0} \alpha_k(t) \langle \ell^{\text{OU}}, \widehat{\mathbb{W}}_t \rangle^k = \langle \sigma_t, \widehat{\mathbb{W}}_t \rangle \quad \text{with } \sigma_t := \sum_{k \geq 0} \alpha_k(t) (\ell^{\text{OU}})^{\sqcup k}. \quad (3.10)$$

This clearly includes:

- The Stein-Stein model [58] for  $f(t, x) = x$ ,

- The Bergomi model [30, 15] for

$$f(t, x) = \xi_0(t)e^{\eta x} = \xi_0(t) \sum_{k \geq 0} \frac{(\eta x)^k}{k!},$$

for some  $\eta \in \mathbb{R}$  and some deterministic input curve  $\xi_0$ ,

- The Quintic OU model [4] for

$$f(t, x) = \xi_0(t)(\alpha_0 + \alpha_1 x + \alpha_3 x^3 + \alpha_5 x^5),$$

for some  $\alpha_i \geq 0$  and some deterministic input curve  $\xi_0$ , and any other finite polynomial of the Ornstein-Uhlenbeck process.

### 3.1.2 Models based on the mean-reverting geometric Brownian motion

More generally, the mean-reverting geometric Brownian motion (mGBM)  $Y$ , given by

$$dY_t = \kappa(\theta - Y_t)dt + (\eta + \alpha Y_t)dW_t, \quad Y_0 = y \in \mathbb{R}, \quad (3.11)$$

for  $\kappa, \theta, \eta, \alpha \in \mathbb{R}$  can be represented as an infinite linear combination of the signature of the time-extended Brownian motion, either in a time-independent or time-dependent way.

**Lemma 3.2.** *The unique solution  $Y$  to (3.11) is given by*

$$Y_t = \left\langle \ell^{\text{mGBM}}, \widehat{\mathbb{W}}_t \right\rangle, \quad \ell^{\text{mGBM}} = \left( y\theta + \left( \kappa\theta - \frac{\alpha\eta}{2} \right) \mathbf{1} + \eta \mathbf{2} \right) e^{\llbracket \cdot \rrbracket \left( -\left( \kappa + \frac{\alpha^2}{2} \right) \mathbf{1} + \alpha \mathbf{2} \right)}, \quad (3.12)$$

such that  $\ell^{\text{mGBM}} \in \mathcal{A}$ , with  $e^{\llbracket \cdot \rrbracket}$  as defined in (2.7). Equivalently,  $Y$  can also be written in terms of time-dependent coefficients:

$$Y_t = \left\langle \tilde{\ell}_t^{\text{mGBM}}, \widehat{\mathbb{W}}_t \right\rangle, \quad \tilde{\ell}_t^{\text{mGBM}} = \theta\theta + e^{-\lambda t} \left( (\ell^{\text{mGBM}} - \theta\theta) \llbracket \cdot \rrbracket e^{\llbracket \cdot \rrbracket \lambda \mathbf{1}} \right), \quad (3.13)$$

for some  $\lambda \in \mathbb{R}$ .

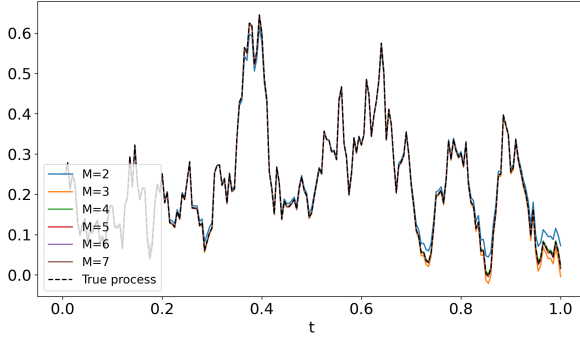
*Proof.* The proof is detailed in Appendix A.2. □

*Example 3.2.1.* Up to order 3, the linear form of a mean-reverting geometric Brownian motion reads

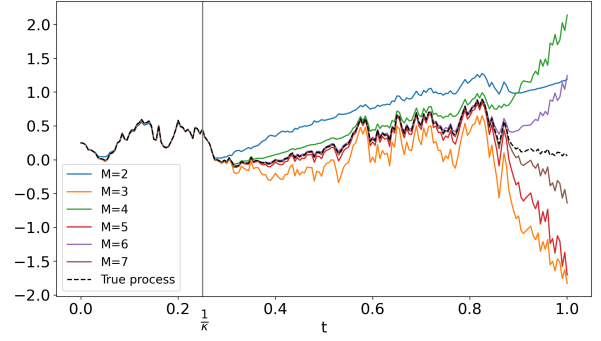
$$\ell^{\text{mGBM}} = \left( y, \begin{pmatrix} \beta y + \gamma \\ \alpha y + \eta \end{pmatrix}, \begin{pmatrix} (\beta y + \gamma)\beta & (\beta y + \gamma)\alpha \\ (\alpha y + \eta)\beta & (\alpha y + \eta)\alpha \end{pmatrix}, \begin{pmatrix} (\beta y + \gamma)\beta^2 & (\beta y + \gamma)\beta\alpha \\ (\beta y + \gamma)\beta\alpha & (\beta y + \gamma)\alpha^2 \\ (\alpha y + \eta)\beta^2 & (\alpha y + \eta)\beta\alpha \\ (\alpha y + \eta)\beta\alpha & (\alpha y + \eta)\alpha^2 \end{pmatrix}, \dots \right),$$

where  $\beta = -\left( \kappa + \frac{\alpha^2}{2} \right)$  and  $\gamma = \kappa\theta - \frac{\alpha\eta}{2}$ . ■

The behaviour and numerical limitations of the linear representations of the mean-reverting geometric Brownian motion are similar to those of the Ornstein-Uhlenbeck, as shown in Figure 3 and Figure 4.

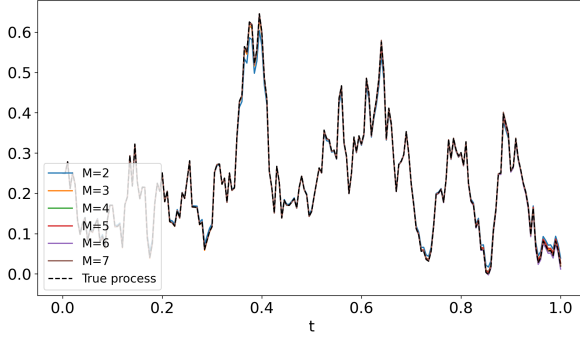


(a)  $\kappa = 1, \theta = 0.25, \eta = 0.5, \alpha = 1.6$

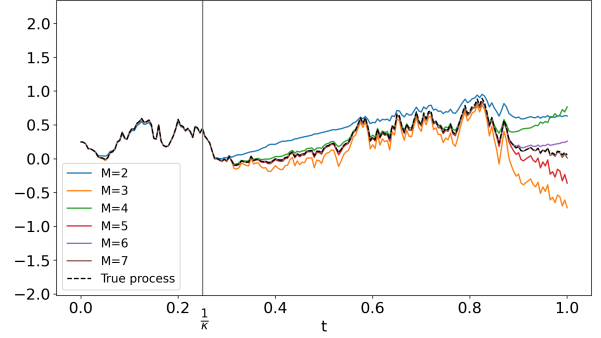


(b)  $\kappa = 4, \theta = 0.25, \eta = 0.5, \alpha = 2$

Figure 3: Trajectories of a mean-reverting geometric Brownian motion against their truncated time-independent linear representation (3.12), i.e.  $\langle \ell^{\text{mGBM}}, \widehat{\mathbb{W}}_t \rangle$ , for several truncation orders  $M$ .



(a)  $\kappa = 1, \theta = 0.25, \eta = 0.5, \alpha = 1.6, \lambda = 0.3$



(b)  $\kappa = 4, \theta = 0.25, \eta = 0.5, \alpha = 2, \lambda = 0.5$

Figure 4: Trajectories of a mean-reverting geometric Brownian motion against their truncated time-dependent linear representation (3.13), i.e.  $\langle \tilde{\ell}_t^{\text{mGBM}}, \widehat{\mathbb{W}}_t \rangle$ , for several truncation orders  $M$ .

*Remark 3.2.1.* For Monte Carlo simulations, it is better to have explicit processes. When  $\alpha \neq 0$ , the mGBM solution to (3.11) can also be formulated explicitly with

$$Y_t := \left( y + \frac{\eta}{\alpha} + \kappa \left( \theta + \frac{\eta}{\alpha} \right) \int_0^t e^{(\kappa + \frac{\alpha^2}{2})s - \alpha W_s} ds \right) e^{-(\kappa + \frac{\alpha^2}{2})t + \alpha W_t} - \frac{\eta}{\alpha}.$$

When  $\alpha = 0$ , the mGBM is an Ornstein-Uhlenbeck process. ■

Going back to the signature volatility model (3.3). The mGBM representation encompasses the following volatility processes

- The volatility in the Hull-White model [43], i.e.

$$\frac{d\sigma_t}{\sigma_t} = \left( \mu - \frac{1}{2}\xi^2 \right) dt + \xi dW_t, \quad \sigma_0 \in \mathbb{R}.$$

- The volatility in Dupire [30], i.e.

$$\frac{d\sigma_t}{\sigma_t} = \frac{1}{2} \left( \frac{\partial \log V_t(0)}{\partial t} - \frac{b^2}{4} \right) dt + \frac{b}{2} dW_t, \quad \sigma_0 \in \mathbb{R}^+,$$

for some deterministic forward variance curve  $V_t(0)$ . Note that the linear representation of  $\sigma$  would be time-dependent.

### 3.1.3 Models based on the square-root process

A square-root or Cox-Ingersoll-Ross [24] (CIR) process  $V$ , driven by

$$dV_t = \kappa(\theta - V_t)dt + \eta\sqrt{V_t}dW_t, \quad V_0 = v > 0,$$

seems to admit a conjectured linear representation

$$V_t = \left\langle \ell^{\text{CIR}}, \widehat{\mathbb{W}}_t \right\rangle = \left( \left\langle \sigma^{\text{CIR}}, \widehat{\mathbb{W}}_t \right\rangle \right)^2, \quad (3.14)$$

where  $\ell^{\text{CIR}} := (\sigma^{\text{CIR}})^{\perp\perp 2}$  with  $\sigma^{\text{CIR}}$  satisfying the non-linear algebraic equation

$$(\sigma^{\text{CIR}})^{\perp\perp 2} = v\mathbf{e} + \left( \left( \kappa\theta - \frac{\eta^2}{4} \right) \mathbf{e} - \kappa(\sigma^{\text{CIR}})^{\perp\perp 2} \right) \mathbf{1} + \eta\sigma^{\text{CIR}}\mathbf{2}.$$

The theoretical convergence, i.e. proving that  $\sigma^{\text{CIR}} \in \mathcal{A}$ , seems intricate to obtain but is ongoing in a separate work. From the numerical perspective, Figure 5 seems to validate our formula under the Feller condition  $2\kappa\theta > \eta$  [34]. Figure 15 below also provides a numerical convergence of prices of our representation (3.14) in the context of the Heston model.

*Example 3.2.2.* Up to order 3, the linear representation of the Cox-Ingersoll-Ross process reads

$$\ell^{\text{CIR}} = \left( v, \begin{pmatrix} \beta \\ \alpha \end{pmatrix}, \begin{pmatrix} -\kappa\beta & \lambda\beta \\ -\kappa\alpha & \lambda\alpha \end{pmatrix}, \begin{pmatrix} \kappa^2\beta & -\left(\kappa + \frac{\beta}{2v}\right)\lambda\beta \\ -\kappa\lambda\beta & 0 \\ \kappa^2\alpha & -\left(\kappa + \frac{\beta}{2v}\right)\lambda\alpha \\ -\kappa\lambda\alpha & 0 \end{pmatrix}, \dots \right),$$

where  $\beta := -\kappa(v - \theta) - \frac{\eta^2}{4}$ ,  $\alpha := \eta\sqrt{v}$  and  $\lambda := \frac{\eta}{2\sqrt{v}}$ . ■

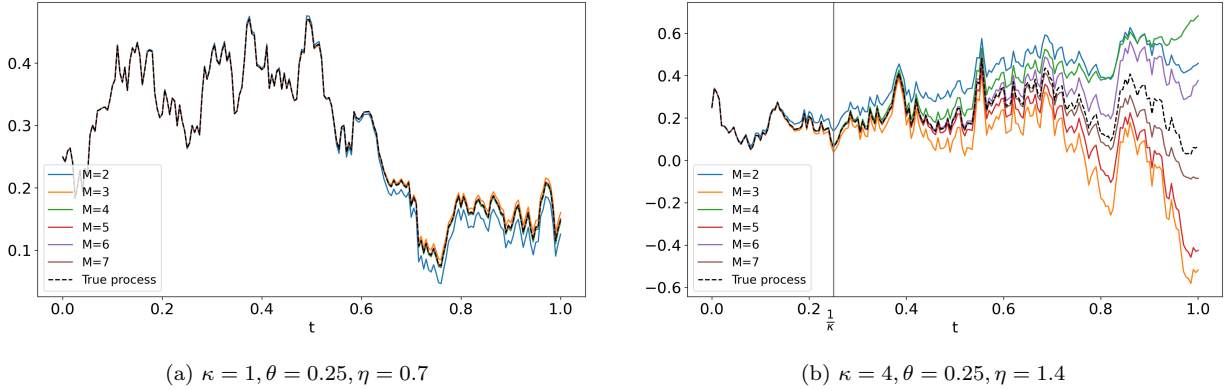


Figure 5: Trajectories of a Cox-Ingersoll-Ross process against their truncated time-independent linear representation (3.14), i.e.  $\left\langle \ell^{\text{CIR}}, \widehat{\mathbb{W}}_t \right\rangle$ , for several truncation orders  $M$ .

Going back to our signature volatility model, the representation (3.14) allows us to include in our framework the volatility of the Heston model [42].

### 3.1.4 Models based on path-dependent processes

As shown in [6], several path-dependent stochastic Volterra processes:

$$Z_t = Z_0 + \int_0^t K(t-s)(a_0 + a_1 Z_s)ds + \int_0^t K(t-s)(b_0 + b_1 Z_s)dW_s,$$



with  $Z_0, a_0, a_1, b_0, b_1 \in \mathbb{R}$ , for certain locally square-integrable kernels  $K$  also admit linear representations in the form  $Z_t = \langle \boldsymbol{\ell}_t, \widehat{\mathbb{W}}_t \rangle$ . This includes non-semimartingale processes such as the Riemann-Liouville fractional Brownian motion

$$W_t^H = \int_0^t (t-s)^{H-1/2} dW_s, \quad H \in (0, 1).$$

For instance, the (time-dependent) representation of  $W^H$  reads

$$W_t^H = \langle \boldsymbol{\ell}_t^{\text{RL}}, \widehat{\mathbb{W}}_t \rangle, \quad \boldsymbol{\ell}_t^{\text{RL}} = t^{H-\frac{1}{2}} \sum_{n=0}^{\infty} \frac{(\frac{1}{2}-H)^{\bar{n}}}{t^n} \mathbf{1}^{\otimes n} \mathbf{2},$$

where  $(\cdot)^{\bar{n}}$  is the rising factorial. Please refer to [6, Section 4] for more details on such representations. Again any volatility process that is an analytic function of such processes falls into the framework of signature volatility models (3.2)-(3.3) thanks to the shuffle property, recall (3.10). This includes for instance the class of Volterra polynomial models [5], in particular Volterra [2] and rough Bergomi models [12].

As a final example, the delayed equation (DE) process  $U$ , given by

$$dU_t = \left( a + \int_0^t e^{\alpha(t-s)} U_s ds \right) dt + \left( b + \int_0^t e^{\beta(t-s)} U_s ds \right) dW_t, \quad U_0 = u \in \mathbb{R},$$

for some  $a, b, \alpha, \beta \in \mathbb{R}$ , can be represented as a linear combination of the signature of the time-extended Brownian motion with

$$U_t = \langle \boldsymbol{\ell}^{\text{DE}}, \widehat{\mathbb{W}}_t \rangle, \quad \boldsymbol{\ell}^{\text{DE}} = (u\boldsymbol{\phi} + a\mathbf{1} + b\mathbf{2}) (\boldsymbol{\phi} - \mathbf{1} (e^{\llcorner\lrcorner\alpha\mathbf{1}} \mathbf{1} + e^{\llcorner\lrcorner\beta\mathbf{1}} \mathbf{2}))^{-1}. \quad (3.15)$$

The reader can refer to [6, Theorem 4.4] for more details on the linear delayed equation process.

*Example 3.2.3.* Up to order 3, the linear form of a delayed equation process reads

$$\boldsymbol{\ell}^{\text{DE}} = \left( x, \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} z & z \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \alpha z + a & \beta z + a \\ 0 & 0 \\ b & b \\ 0 & 0 \end{pmatrix}, \dots \right),$$

see Figure 6 for a numerical illustration of the signature representation. ■

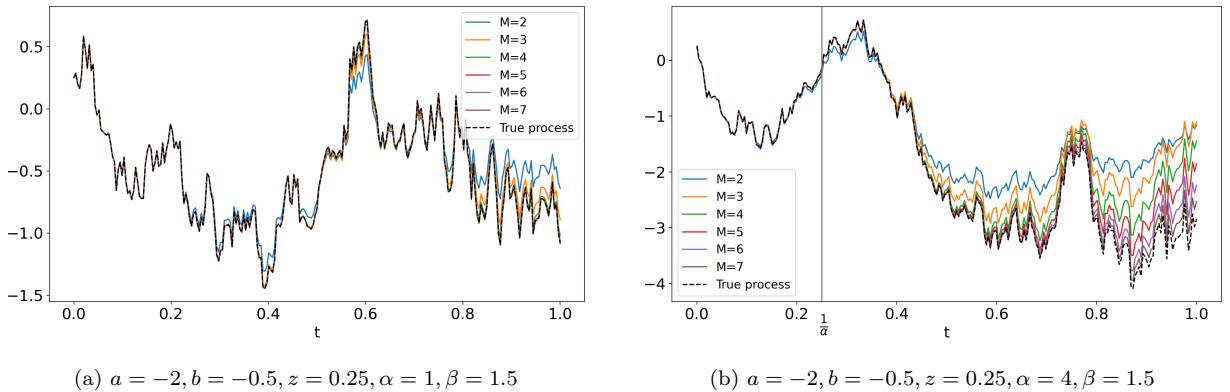


Figure 6: Trajectories of a delayed equation process against their truncated time-independent linear representation (3.15), i.e.  $\langle \boldsymbol{\ell}^{\text{DE}}, \widehat{\mathbb{W}}_t \rangle$ , for several truncation orders  $M$ .

### 3.2 Leverage effect

The leverage effect in the model (3.2)-(3.3) is defined as the instantaneous correlation between the log-price  $S$  and its instantaneous volatility  $|\Sigma|$  defined by

$$\frac{d[\log S, |\Sigma|]_t}{\sqrt{d[\log S]_t} \sqrt{d[|\Sigma|]_t}}.$$

In practice, the leverage effect is negative on equity markets and one usually would like to control its sign via the correlation parameter  $\rho$  between  $B$  and  $W$  in (3.1). Since  $\Sigma$  is not necessarily positive at all time, the leverage effect can flip sign in the model (3.2)-(3.3), which is not realistic. The next Lemma provides a necessary and sufficient condition on the coefficients  $\sigma$  for the leverage effect to be equal to  $\rho$ , at all time.

**Lemma 3.3.** *Assume  $\sigma \in \mathcal{I}'$ , then the leverage effect is given by*

$$\frac{d[\log S, |\Sigma|]_t}{\sqrt{d[\log S]_t} \sqrt{d[|\Sigma|]_t}} = \rho \operatorname{sign} \left( \left\langle \sigma_t|_2, \widehat{W}_t \right\rangle \right),$$

where  $\operatorname{sign}(x) = 1$  if  $x \geq 0$  and  $-1$  otherwise. In particular, it is equal to  $\rho$  for all  $t \geq 0$  if and only if  $\sigma$  satisfies

$$\left\langle \sigma_t|_2, \widehat{W}_t \right\rangle \geq 0, \quad t \geq 0. \quad (3.16)$$

*Proof.* The dynamics of  $\log S$  are in the form

$$d \log S_t = -\frac{1}{2} \Sigma_t^2 dt + \Sigma_t dB_t = -\frac{1}{2} \left\langle \sigma_t^{\llcorner \cup \cup 2}, \widehat{W}_t \right\rangle dt + \left\langle \sigma_t, \widehat{W}_t \right\rangle dB_t. \quad (3.17)$$

The instantaneous volatility of the model is  $|\Sigma|$ . By an application of Itô-Tanaka's formula on  $|\Sigma|$  and of Lemma 2.7, we have

$$\begin{aligned} d|\Sigma|_t &= \operatorname{sign}(\Sigma_t) d\Sigma_t + dL^0(t) \\ &= \operatorname{sign}(\Sigma_t) \left\langle \sigma_t|_2, \widehat{W}_t \right\rangle dW_t + \operatorname{sign}(\Sigma_t) \left\langle \sigma_t|_1 + \frac{1}{2} \sigma_t|_{22} + \dot{\sigma}_t, \widehat{W}_t \right\rangle dt + dL^0(t), \end{aligned}$$

where  $L^0$  is the local time of  $\Sigma$  at 0. It follows that the leverage effect is given by

$$\frac{d[\log S, |\Sigma|]_t}{\sqrt{d[\log S]_t} \sqrt{d[|\Sigma|]_t}} = \rho \frac{\Sigma_t \operatorname{sign}(\Sigma_t) \left\langle \sigma_t|_2, \widehat{W}_t \right\rangle}{|\Sigma_t| \left| \left\langle \sigma_t|_2, \widehat{W}_t \right\rangle \right|} = \rho \operatorname{sign} \left( \left\langle \sigma_t|_2, \widehat{W}_t \right\rangle \right).$$

□

The condition (3.16) can be made explicit in the Stein-Stein, Quintic and Bergomi models as shown in the next example.

*Example 3.3.1.*

- For the Stein-Stein model driven by a Brownian motion,  $\sigma|_2 = \eta \mathbf{e}$ , see (3.8), and  $\left\langle \sigma|_2, \widehat{W}_t \right\rangle = \eta \geq 0$ ,
- For the Quintic model of [4] constructed on a Brownian motion,  $\sigma = \alpha_0 \mathbf{e} + \alpha_1 \mathbf{2} + \alpha_3 \mathbf{222} + \alpha_5 \mathbf{22222}$  so that  $\sigma|_2 = \alpha_1 \mathbf{e} + \alpha_3 \mathbf{22} + \alpha_5 \mathbf{2222}$ , which gives

$$\left\langle \sigma|_2, \widehat{W}_t \right\rangle = \alpha_1 + \frac{\alpha_3}{2} W_t^2 + \frac{\alpha_5}{4!} W_t^4 \geq 0, \quad t \geq 0,$$

as long as  $\alpha_1, \alpha_3, \alpha_5$  are non negative,

- For the Bergomi model, see [30, 15],  $\sigma_t = \xi_0(t) e^{\llcorner \cup \cup \eta^2}$  so that  $\sigma_t|_2 = \eta \xi_0(t) e^{\llcorner \cup \cup \eta^2}$ , which gives

$$\left\langle \sigma_t|_2, \widehat{W}_t \right\rangle = \eta \xi_0(t) e^{\eta W_t} \geq 0,$$

as long as  $\xi_0(t), \eta$  are non negative, for all  $t \leq T$ . ■

### 3.3 Approximated representations

More generally, if exact linear representations are not available for certain processes, approximate representations can be obtained thanks to the universal approximation property of path-signatures:

**Theorem 3.4** (Universal approximation theorem [27]). *For some  $t \in [0, T]$ , let  $K_t$  be a compact subset of  $\{(\widehat{\mathbb{W}}_s^{\leq 2})_{s \in [0, t]}(\omega) : \omega \in \Omega\}$  and consider a continuous map  $F : K_t \rightarrow \mathbb{R}$ . Then for every  $\epsilon > 0$ , there exists some  $M^\epsilon \in \mathbb{N}$  and some  $\mathbf{f}^\epsilon \in T^{M^\epsilon}(\mathbb{R}^2)$  such that*

$$\sup_{W \in K_t} \left| F \left( (\widehat{\mathbb{W}}_s^{\leq 2})_{s \in [0, t]} \right) - \langle \mathbf{f}^\epsilon, \widehat{\mathbb{W}}_t \rangle \right| < \epsilon,$$

almost surely.

In practice, one would perform a linear regression on trajectories of a given process  $X$  against a finite linear combination of the signature elements using the following steps:

**Algorithm 3.5** (Regression against truncated signature). *Assume the spot volatility is of the following form*

$$X_t = F(t, (W_s)_{s \leq t}).$$

1. Generate  $N$  realizations of the Brownian motion  $W$ , denoted by  $W^{(1)}, \dots, W^{(N)}$ ,
2. For each realization  $n = 1, \dots, N$ , compute  $X^{(n)}$  and the truncated signature  $\widehat{\mathbb{W}}^{(n), \leq M}$  up to order  $M$ ,
3. Regress  $(X^{(1)}, \dots, X^{(N)})$  against  $(\widehat{\mathbb{W}}^{(1), \leq M}, \dots, \widehat{\mathbb{W}}^{(N), \leq M})$  with  $L^1$  and  $L^2$  regularization  $\beta_1$  and  $\beta_2$  to learn the coefficients of  $\mathbf{f}$  that minimize

$$\min_{\mathbf{f} \in T^M(\mathbb{R}^d)} \frac{1}{N} \frac{1}{J} \sum_{n=1}^N \sum_{j=1}^J \left| X_{t_j}^{(n)} - \langle \mathbf{f}, \widehat{\mathbb{W}}_{t_j}^{(n), \leq M} \rangle \right|^2 + \sum_{k=0}^M \sum_{\mathbf{v} \in V_k} \left( \beta_1 |\ell^{\mathbf{i}_1 \dots \mathbf{i}_k}| + \beta_2 (\ell^{\mathbf{i}_1 \dots \mathbf{i}_k})^2 \right).$$

Such approach appeared in Cuchiero et al. [27], Lyons et al. [48], Fermanian [36], Arribas et al. [10].

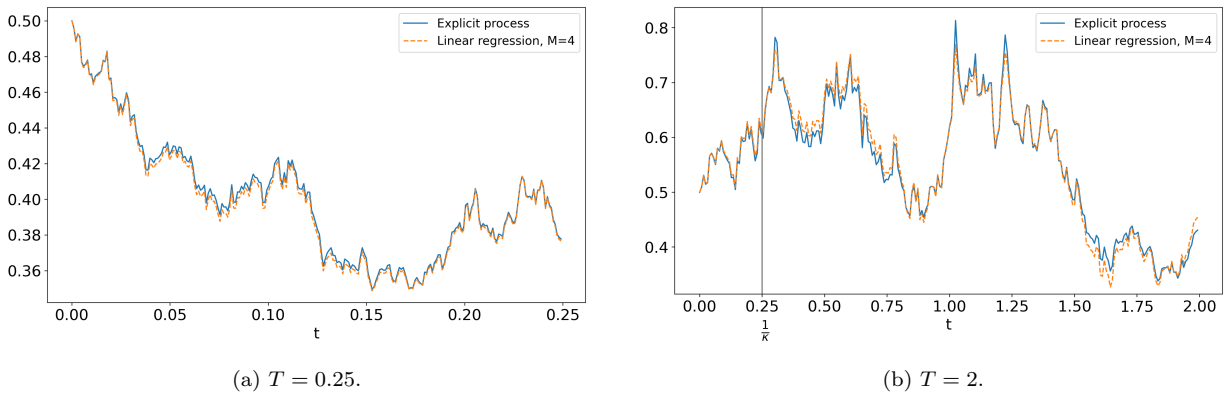


Figure 7: Trajectories of an inverse CIR process against their linear regression.  $\kappa = 4, \theta = 0.5, \eta = 2$

We recall the dynamics of the inverse CIR process, being the variance in the 3/2 model [52], driven by

$$dV_t = \kappa V_t (\theta - V_t) dt + \eta V_t^{\frac{3}{2}} dW_t.$$

We see in Figure 7 that a process without known linear representation can still be approximated quite well through a linear regression.

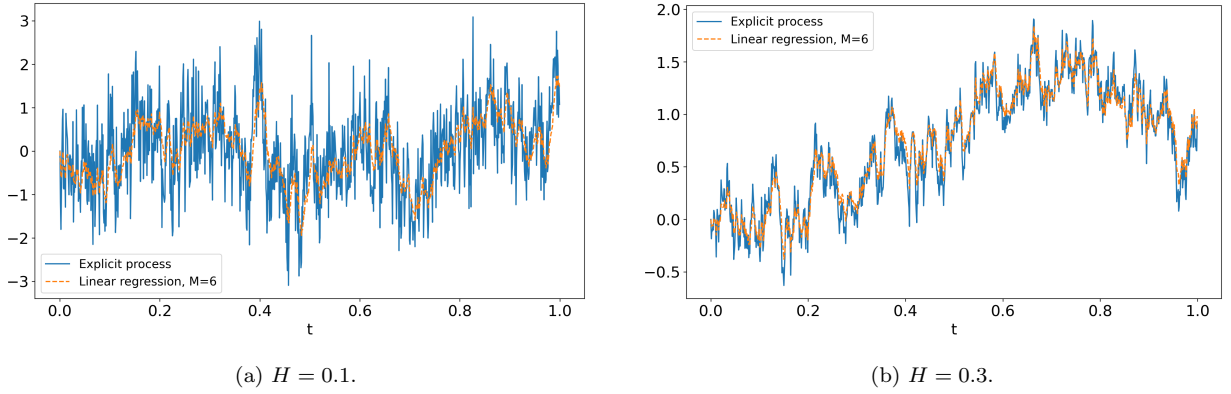


Figure 8: Trajectories of a fractional Brownian motion against their linear regression for  $M = 6$ .

However, as shown on Figure 8, the regression approach does not always work with small truncation levels. For instance, the ability to capture high roughness in a linear functional is not trivial to achieve whilst keeping the natural embedding  $\widehat{W}_t = (t, W_t)$ .

We end this section with a brief comparison of approximate vs exact representations. As seen in the previous subsections, the convergence of the linear representation is quick for short horizons. However, when the horizon gets too large, relatively to the parameters of the represented models, e.g. the mean reversion rate, the truncated representations drastically deteriorates, see Figures 1 to 6. Yet, as seen previously, linear regressions make quite stable representations over their training time and can thus be made over targeted horizons to control the stability. In Figures 9, 10 and 11, we can remark that for short horizons (relative to  $\kappa$ ), the linear regression doesn't fit as well as the linear representation. However for longer horizons, where the linear representation lacks in stability, the linear regression retains hers.

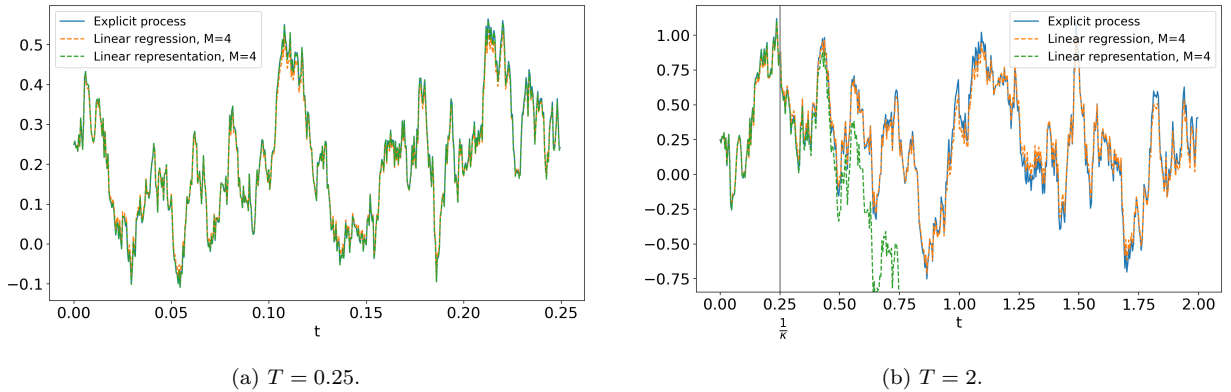


Figure 9: Trajectories of an Ornstein-Uhlenbeck process against their linear representation and linear regression.  $\kappa = 4, \theta = 0.25, \eta = 2$

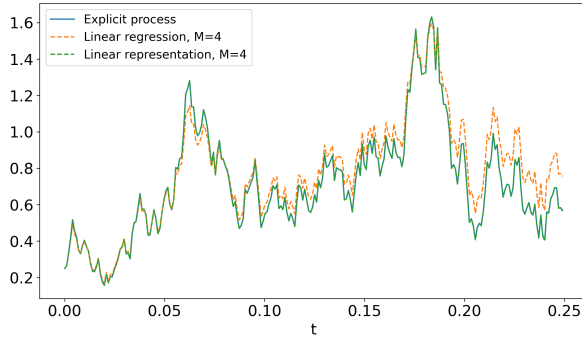
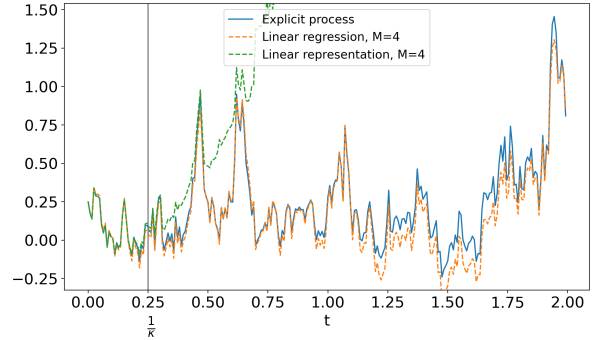
(a)  $T = 0.25$ .(b)  $T = 2$ .

Figure 10: Trajectories of a mean-reverting geometric Brownian motion against their linear representation and linear regression.  $\kappa = 4, \theta = 0.25, \eta = 1, \alpha = 2$

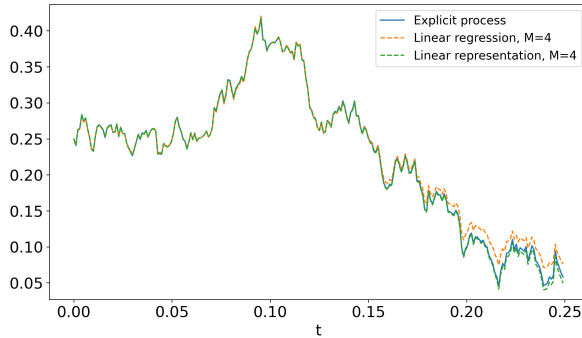
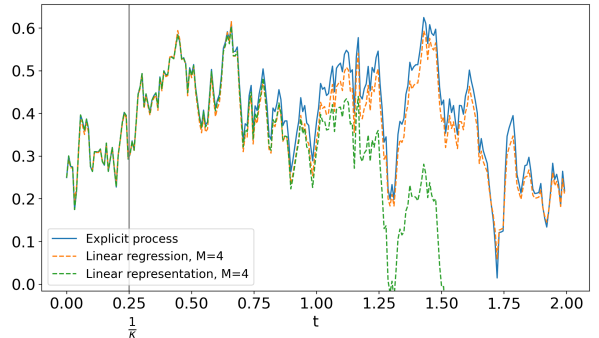
(a)  $T = 0.25$ .(b)  $T = 2$ .

Figure 11: Trajectories of a square-root Cox-Ingersoll-Ross process against their linear representation and linear regression.  $\kappa = 4, \theta = 0.0625, \eta = 1$

However, this stability across the training horizon could easily be interpreted as overfitting. Table 1 below displays the mean squared error between explicit simulations of an Ornstein-Uhlenbeck process and their linear representations (exact truncated and from a regression) where  $\kappa = 4, \theta = 0.25, \eta = 2$  and  $v = 0.25$ . The number of simulations and training horizon have been set to 100 000 and 1 year with 252 time steps respectively, see also Figure 12.

MSE		Test horizon				
		3 months	6 months	1 year	2 years	4 years
$M = 2$	Exact	3.065e-05	5.370e-03	1.969e+00	2.536e+01	2.625e+02
	Regression	6.149e-03	8.564e-03	2.405e-02	6.239e-01	1.189e+01
$M = 4$	Exact	6.541e-08	1.736e-05	1.718e+00	4.243e+02	8.177e+04
	Regression	8.002e-06	3.801e-05	1.079e-04	1.516e+00	1.213e+03
$M = 6$	Exact	5.908e-08	1.167e-06	3.173e-01	1.410e+03	4.872e+06
	Regression	5.735e-06	8.426e-06	1.596e-07	1.066e+00	3.188e+04

Table 1: Mean-squared-error between simulations of an Ornstein-Uhlenbeck process against its linear representation and linear regression when trained over only 1 year.  $\kappa = 4, \theta = 0.25, \eta = 2$ .

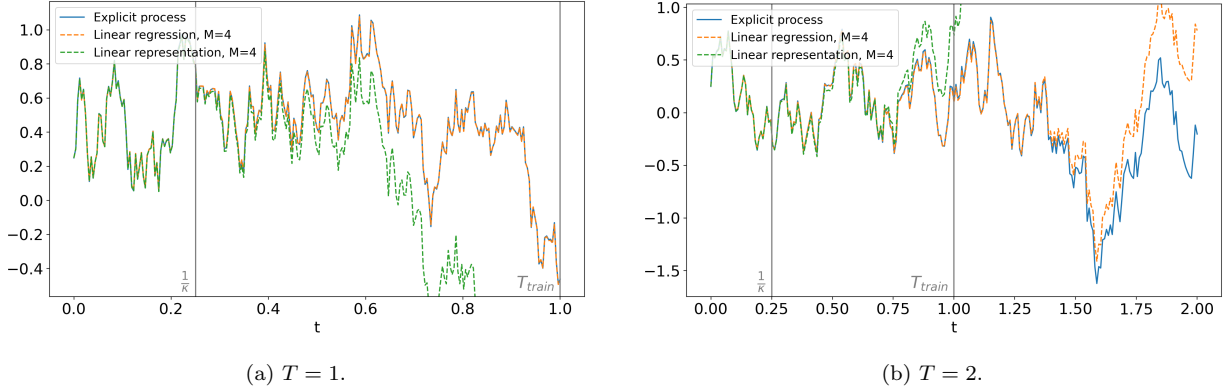


Figure 12: Trajectories of an Ornstein-Uhlenbeck process against their linear representation and linear regression when trained over only 1 year.  $\kappa = 4, \theta = 0.25, \eta = 2$ .

## 4 The joint characteristic functional

The following theorem provides the joint conditional characteristic functional of the log-price  $\log S$  and the integrated variance  $\bar{V} := \int_0^T \Sigma_s^2 ds$  in the model (3.2)–(3.3) in terms of a solution to an infinite-dimensional system of time-dependent  $T((\mathbb{R}^2))$ -valued Riccati differential equations.

**Theorem 4.1.** *Let  $f, g : [0, T] \rightarrow \mathbb{C}$  be measurable and bounded functions and  $\sigma : [0, T] \rightarrow \mathcal{A}$ . Assume that there exists  $\psi \in \mathcal{I}'$ , solution to the following infinite-dimensional system of time-dependent Riccati equations*

$$\begin{aligned} -\dot{\psi}_t &= \frac{1}{2}(\psi_t|_2)^{\llcorner\llcorner 2} + \rho f(t)(\sigma_t \llcorner\llcorner \psi_t|_2) + \frac{1}{2}\psi_t|_{22} + \psi_t|_1 + \left( \frac{f(t)^2 - f(t)}{2} + g(t) \right) \sigma_t^{\llcorner\llcorner 2}, \\ \psi_T &= 0. \end{aligned} \quad (4.1)$$

Define the processes

$$U_t = \langle \psi_t, \widehat{\mathbb{W}}_t \rangle + \int_0^t f(s) d \log S_s + \int_0^t g(s) d \bar{V}_s, \quad (4.2)$$

and

$$M_t = e^{U_t}. \quad (4.3)$$

Then  $M$  is a local martingale. If in addition it is a true martingale, then the following expression holds for the joint characteristic functional of  $(\log S, \bar{V})$ :

$$\mathbb{E} \left[ \exp \left( \int_t^T f(s) d \log S_s + \int_t^T g(s) d \bar{V}_s \right) \middle| \mathcal{F}_t \right] = \exp \left( \langle \psi_t, \widehat{\mathbb{W}}_t \rangle \right), \quad t \leq T. \quad (4.4)$$

*Proof.* To show that  $M$  is a local martingale we show that its part in  $dt$  vanishes using Itô's formula. The dynamics of  $M$  read

$$dM_t = M_t \left( dU_t + \frac{1}{2} d[U]_t \right). \quad (4.5)$$

We start by deriving the dynamics of  $U$  in (4.2). By using Lemma 2.7, since  $\psi \in \mathcal{I}'$  by assumption, we have that  $(\langle \psi_t, \widehat{\mathbb{W}}_t \rangle)_{t \geq 0}$  is a semimartingale with dynamics

$$d \langle \psi_t, \widehat{\mathbb{W}}_t \rangle = \langle \dot{\psi}_t + \psi_t|_1 + \frac{1}{2} \psi_t|_{22}, \widehat{\mathbb{W}}_t \rangle dt + \langle \psi_t|_2, \widehat{\mathbb{W}}_t \rangle dW_t.$$

Combining the previous identity with  $d\bar{V}_t = \langle \sigma_t^{\sqcup\sqcup 2}, \widehat{\mathbb{W}}_t \rangle dt$ , recall (3.5), and the dynamics of  $\log S$  in (3.17), we obtain that

$$\begin{aligned} dU_t &= d \langle \psi_t, \widehat{\mathbb{W}}_t \rangle + f(t) d \log S_t + g(t) d\bar{V}_t \\ &= \left\langle \dot{\psi}_t + \psi_t|_{\mathbf{1}} + \frac{1}{2} \psi_t|_{\mathbf{22}}, \widehat{\mathbb{W}}_t \right\rangle dt + \left\langle \psi_t|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle dW_t \\ &\quad + (g(t) - \frac{1}{2} f(t)) \left\langle \sigma_t^{\sqcup\sqcup 2}, \widehat{\mathbb{W}}_t \right\rangle dt + f(t) \left\langle \sigma_t, \widehat{\mathbb{W}}_t \right\rangle dB_t \\ &= \left\langle \dot{\psi}_t + \psi_t|_{\mathbf{1}} + \frac{1}{2} \psi_t|_{\mathbf{22}} + (g(t) - \frac{1}{2} f(t)) \sigma_t^{\sqcup\sqcup 2}, \widehat{\mathbb{W}}_t \right\rangle dt \\ &\quad + \left\langle \psi_t|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle dW_t + f(t) \left\langle \sigma_t, \widehat{\mathbb{W}}_t \right\rangle dB_t. \end{aligned}$$

Using the shuffle product of Proposition 2.6 and the fact that  $B$  and  $W$  are correlated, recall (3.1), we get that the quadratic variation of  $U$  is

$$d[U]_t = \left\langle (\psi_t|_{\mathbf{2}})^{\sqcup\sqcup 2} + 2\rho f(t)(\psi_t|_{\mathbf{2}} \sqcup\sqcup \sigma_t) + f(t)^2 \sigma_t^{\sqcup\sqcup 2}, \widehat{\mathbb{W}}_t \right\rangle dt.$$

This yields that the drift of  $dM_t/M_t$  in (4.5) is given by

$$\begin{aligned} &\left\langle \dot{\psi}_t + \psi_t|_{\mathbf{1}} + \frac{1}{2} \psi_t|_{\mathbf{22}} + (g(t) - \frac{1}{2} f(t)) \sigma_t^{\sqcup\sqcup 2} \right. \\ &\quad \left. + \frac{1}{2} [(\psi_t|_{\mathbf{2}})^{\sqcup\sqcup 2} + 2\rho f(t)(\psi_t|_{\mathbf{2}} \sqcup\sqcup \sigma_t) + f(t)^2 \sigma_t^{\sqcup\sqcup 2}], \widehat{\mathbb{W}}_t \right\rangle, \end{aligned}$$

which is equal to 0 from the Riccati equations (4.1). This shows that  $M$  is a local martingale. By assumption  $M$  is even a true martingale. After observing that the terminal value of  $M$ , is given by

$$M_T = \exp \left( \int_0^T f(s) d \log S_s + \int_0^T g(s) d\bar{V}_s \right),$$

recall that  $\psi_T = 0$ , we obtain

$$\mathbb{E} \left[ \exp \left( \int_0^T f(s) d \log S_s + \int_0^T g(s) d\bar{V}_s \right) \middle| \mathcal{F}_t \right] = \mathbb{E} [M_T | \mathcal{F}_t] = M_t = \exp(U_t),$$

which yields (4.4).  $\square$

Theorem 4.1 is a verification result to obtain the exponentially affine representation of the joint characteristic functional (4.4). It disentangles the algebraic affine structure in infinite dimension. It can be related to [29, Theorem 4.24] if one considers the signature of the three dimensional process  $((t, W_t, B_t))_{t \geq 0}$  there. It relies on the two crucial assumptions that a well-behaved  $\mathcal{I}$ -valued solution  $\psi$  exists to the Riccati equation (4.1), together with the true martingality of  $M$ . These assumptions seem very intricate to prove, even in one dimensional settings, partial results in these directions for  $T(\mathbb{R})$ -valued Riccati equations can be found in [29, Section 6] and [7]. We note that no semimartingality assumption for  $\Sigma_t = \langle \sigma_t, \widehat{\mathbb{W}}_t \rangle$  is required in Theorem 4.1.

In the following two sections, we will validate the representation (4.4) numerically and we will highlight the application of Theorem 4.1 to the pricing and the hedging of several contingent claims by Fourier inversion techniques. The general functions  $f$  and  $g$  allow for enough flexibility to cover a broad set of contingent claims. For instance:

1. Certain vanilla options that depend on the values  $(S_T, \bar{V}_T)$ , like European call and put options and volatility swaps, are recovered using constant  $f$  and  $g$ .
2. Geometric Asian options on the average  $\frac{1}{T} \int_0^T \log S_s ds$  can be recovered by setting  $f(s) := iu \frac{T-s}{T}$ , since

$$\begin{aligned} \int_t^T f(s) d \log S_s &= iu \int_t^T \frac{T-s}{T} d \log S_s \\ &= iu \frac{1}{T} \int_t^T \log S_s ds - iu \frac{T-t}{T} \log S_t. \end{aligned}$$

## 5 Pricing by Fourier methods

In this section, we show how Theorem 4.1 can be applied to price European and Asian call and put options as well as  $q$ -Volatility swaps using Fourier inversion techniques in our signature volatility model (3.2)-(3.3). All of our numerical results validate the exponentially affine representation (4.4). We also provide a calibration to market volatility surface for the S&P 500.

For the numerical implementation, we consider a truncated version of  $S$  from (3.2)-(3.3), denoted by  $S^{\leq M}$  and defined by

$$dS_t^{\leq M} = S_t^{\leq M} \left\langle \sigma_t^{\leq M}, \widehat{\mathbb{W}}_t \right\rangle dB_t,$$

where  $\sigma_t^{\leq M} \in T^M(\mathbb{R}^2)$  has its first  $M$  levels coincide with  $\sigma_t$  and is 0 elsewhere. Finally, in order to ease notations in the sequel we assume  $S_0 = 1$  (recall that the short rate here is assumed to be 0).

### 5.1 European options

Let us consider a European call option on  $S$  with maturity  $T > 0$  and strike  $K > 0$ . Its price at time  $t \leq T$  is given by  $C_t(T, K) = \mathbb{E}[(S_T - K)^+ | \mathcal{F}_t]$ . From Lewis [45], one can price this option using the Fourier inversion formula

$$C_t(S_t; T, K) = S_t - \frac{K}{\pi} \int_0^\infty \Re \left[ e^{i(u - \frac{i}{2})k_t} \phi_t \left( u - \frac{i}{2} \right) \right] \frac{du}{\left( u^2 + \frac{1}{4} \right)},$$

where  $\phi$  is the conditional characteristic function  $\phi_t(u) = \mathbb{E} \left[ e^{iu \log \frac{S_T}{S_t}} | \mathcal{F}_t \right]$  and  $k_t = \log \frac{S_t}{K}$ .

Moreover, we add a ‘control variate’ to quicken the convergence, as in Andersen and Andreasen [8]. Given  $\sigma_{\text{BS}} < \infty$ , one has

$$C_t(S_t; T, K) = C_t^{\text{BS}}(S_t; T, K) - \frac{K}{\pi} \int_0^\infty \Re \left[ e^{i(u - \frac{i}{2})k_t} \left( \phi_t \left( u - \frac{i}{2} \right) - \phi_t^{\text{BS}} \left( u - \frac{i}{2} \right) \right) \right] \frac{du}{\left( u^2 + \frac{1}{4} \right)}, \quad (5.1)$$

where

$$C_t^{\text{BS}}(S_t; K, T) = \mathcal{N}(d_1)S_t - \mathcal{N}(d_2)K,$$

with

$$d_1 = \frac{1}{\sigma_{\text{BS}} \sqrt{T-t}} \left( \log \frac{S_t}{K} + \frac{\sigma_{\text{BS}}^2}{2} (T-t) \right), \quad d_2 = d_1 - \sigma_{\text{BS}} \sqrt{T-t}$$

and

$$\phi_t^{\text{BS}}(u) = \exp \left[ -\frac{\sigma_{\text{BS}}^2}{2} (u^2 + iu) (T-t) \right]. \quad (5.2)$$

This Black-Scholes control variate can also be applied to other products and, as will be seen in Section 6, to Fourier hedging. The quantity  $\sigma_{\text{BS}}$  can be determined to ensure approximate moment matching for instance, and one can use the characteristic function  $\phi_t$  to approximate (by finite differences) the second order cumulant of the distribution of the log price  $\log S$ .

Numerically, for the discretization of the Fourier integral, our numerical experiments show that Gauss-Laguerre quadrature outperform other quadrature rules in our class of models, which is in line with the empirical findings that appeared in [54]. In particular, the higher the maturity is, the lower the degree of the quadrature needs to be for the same set of parameters. In addition, the use of a control variate makes computations much quicker yielding fewer calls of the characteristic function to make the quadrature converge, see Figure 13. Moreover, we can see that the added value of the control variate, in terms of speed of convergence, grows with the time horizon. See [54, Section 6] for more details on numerical refinements.

For a given signature volatility model (3.2)-(3.3), i.e. a set of parameters  $\sigma : [0, T] \rightarrow \mathcal{A}$ , an application of Theorem 4.1 with  $f(t) = iu$  and  $g = 0$  provides the characteristic function  $\phi_t(u)$  modulo the solution  $\psi$  to the



$T((\mathbb{R}^2))$ -valued Riccati equation (4.1), which allows us to compute the price of call options in the signature volatility model using (5.1) together with a suitable truncation and discretization of the Riccati equation (4.1). The truncation rule for  $\psi$  is not trivial and requires a little bit of care, because the shuffle products cannot be exact, as each step of the discretized ODE would double the truncation order of  $\psi$ . For the numerical implementation, we decided to fix the order  $\tilde{M}$  of  $\psi$  for each step and hence only have a shuffle product projected on  $T^{\tilde{M}}(\mathbb{R}^2)$ , i.e.  $\widetilde{\square} : (T^{\tilde{M}}(\mathbb{R}^2))^2 \rightarrow T^{\tilde{M}}(\mathbb{R}^2)$ . Obviously, choosing  $\tilde{M}$  lower than  $2M$ , where  $M$  is the truncation order of  $\sigma$ , also induces an approximation of shuffle product in  $\sigma^{\llbracket \cdot \rrbracket 2}$ , which greatly deteriorated the quality of the convergence, if not prevented it altogether. We thus fixed  $\tilde{M} = 2M$  throughout our experiments. Said differently for a given signature volatility  $\sigma \in T^M(\mathbb{R}^2)$  the truncated solution  $\psi$  to the Riccati equation (4.1) is an element of  $T^{2M}(\mathbb{R}^2)$ . Regarding the numerical discretization and in order for the Riccati to converge in a realistic amount of time, we use Runge-Kutta to the 4th order to solve the ODE, which computes 4 times as many points as the Euler direct algorithm, but converges more than 4 times as fast. We also compute the characteristic function both JIT and in parallel so that it is drastically faster.

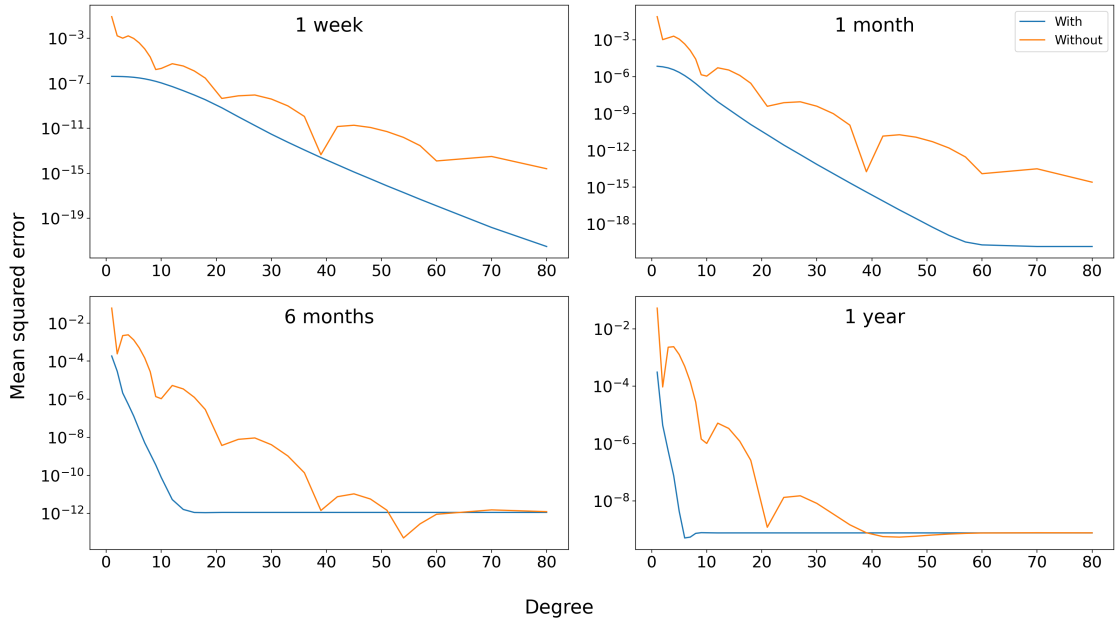


Figure 13: Gauss-Laguerre quadrature mean-squared-error of put pricing with and without control variate, under Stein-Stein model.  $\kappa = 1, \theta = 0.25, \eta = 1.2$  and  $\rho = -0.5$ .

In Figures 14 and 15, we compare our Signature volatility pricing using the linear representations of the Ornstein-Uhlenbeck (3.8) and Cox-Ingersoll-Ross (3.14) representations, truncated at order  $M = 4$ , to the explicit Fourier pricing of the Stein-Stein [58] and Heston [42] models. Lewis' approach together with Black-Scholes control variate, see (5.1), was used. This serves as a numerical validation for Theorem 4.1 as well as for our conjectured representation for the square-root process (3.14).

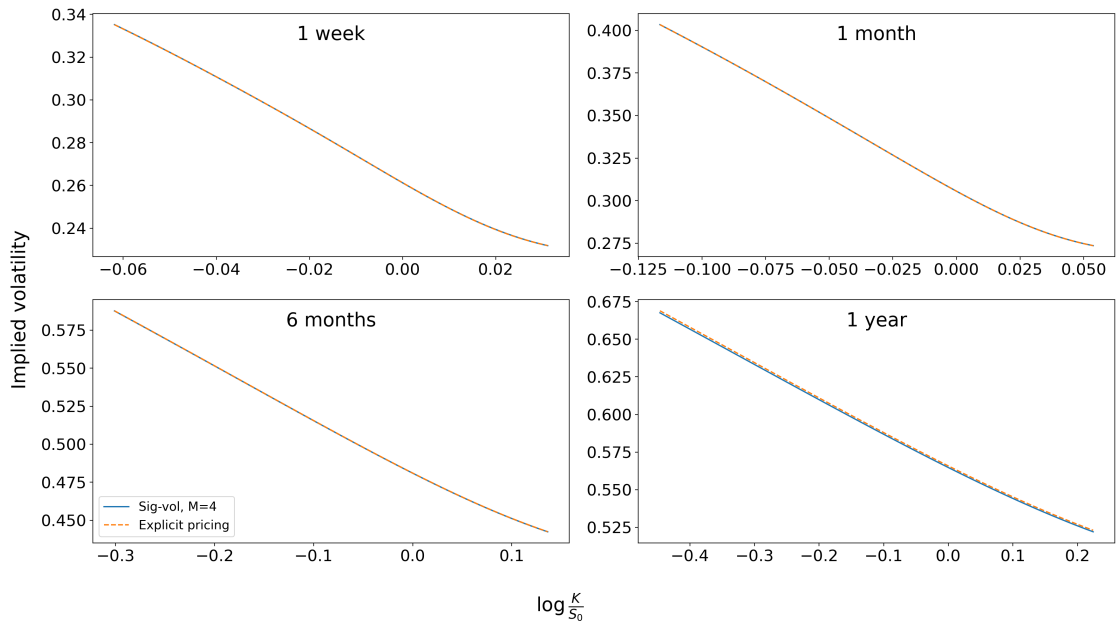


Figure 14: European call Fourier pricing of Stein-Stein model vs Signature Ornstein-Uhlenbeck volatility model (3.8).  $\kappa = 1, \theta = 0.25, \eta = 1.2$  and  $\rho = -0.5$ .

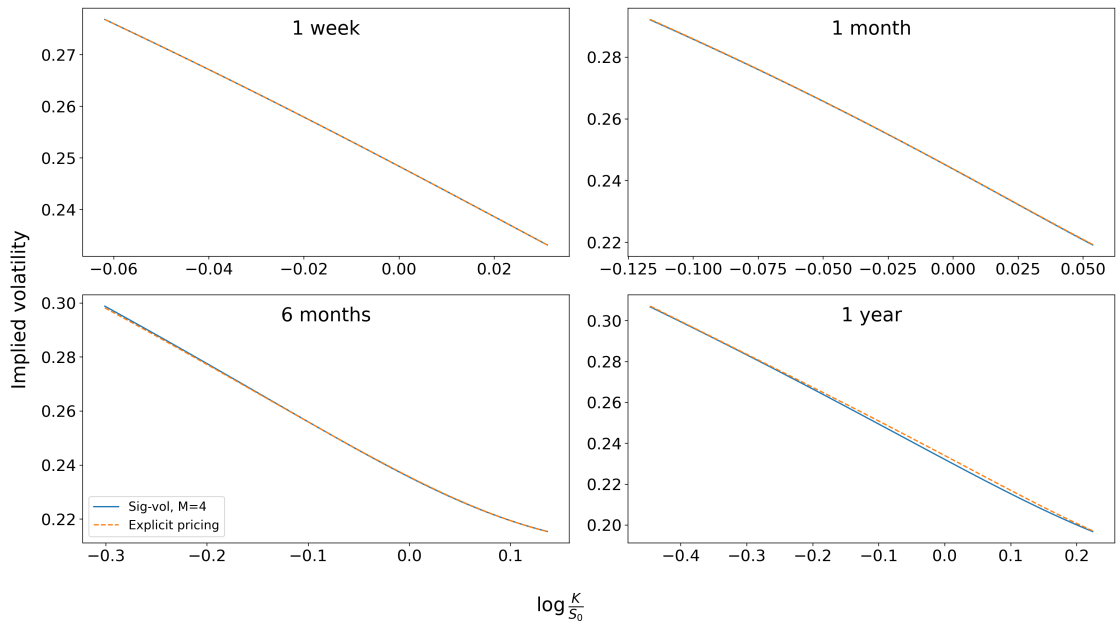


Figure 15: European call Fourier pricing of Heston model vs Signature square-root Cox-Ingersoll-Ross volatility model (3.14).  $\kappa = 4, \theta = 0.0625, \eta = 0.7$  and  $\rho = -0.7$ .

In Figure 16 we compare our signature volatility pricing using the linear mean-reverting geometric Brownian motion representation (3.12) truncated at order  $M = 4$  to Monte Carlo simulations, see Remark 3.2.1. Lewis' approach together with Black-Scholes control variate was also used.

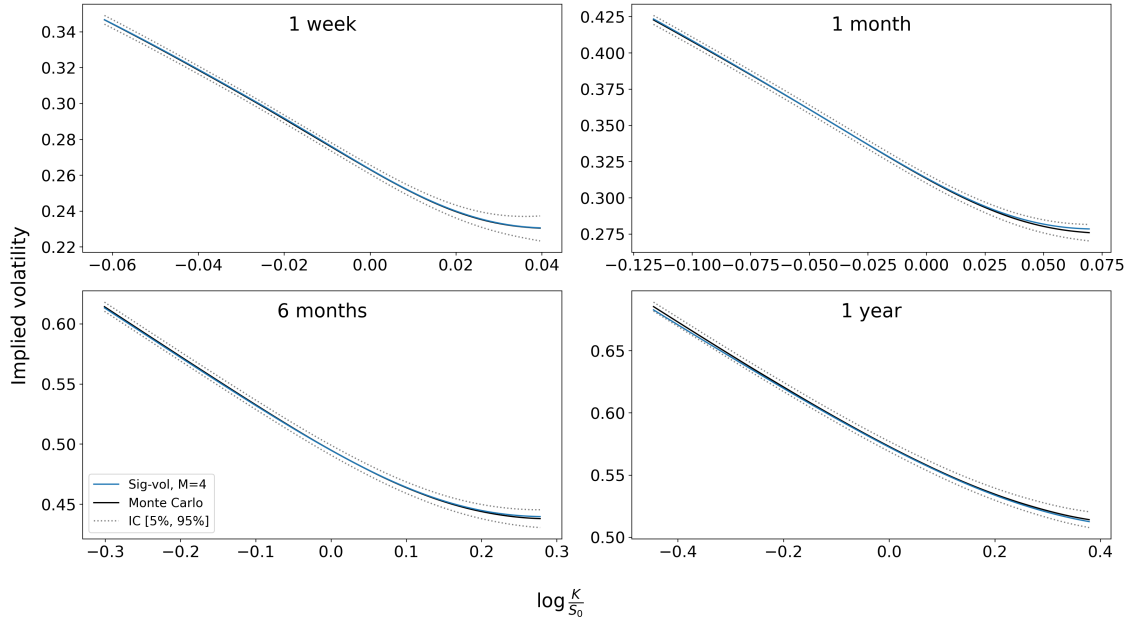


Figure 16: European call Monte Carlo pricing of mGBM volatility model vs Fourier pricing of Signature mean-reverting geometric Brownian motion volatility model (3.12).  $\kappa = 1, \theta = 0.25, \eta = 1.2, \alpha = 0.6$  and  $\rho = -0.5711$ .

In Figure 17 we work on a linear functional drawn at random under the leverage effect condition (LER), see (3.16). In this example, the coefficients  $\sigma^{\text{LER}}$  are drawn such that

1.  $\sigma^{\text{LER}\phi} = x$ ,
2. the coefficients that must remain positive have been drawn from a  $\mathcal{U}_{[0,0.5]}$ ,
3. the unconstrained coefficients have been drawn from a  $\mathcal{U}_{[-0.5,0.5]}$ .

Bellow is the draw used in Figure 17:

$$\sigma^{\text{LER}} = \left( 0.25, \begin{pmatrix} -0.0644715 \\ 0.36712438 \end{pmatrix}, \begin{pmatrix} 0.05250405 & 0 \\ -0.02796828 & 0 \end{pmatrix}, \begin{pmatrix} -0.3698671 & 0 \\ 0.05678161 & 0 \\ 0.3307244 & 0 \\ 0.2119773 & 0.2379793 \end{pmatrix}, \mathbf{0} \right). \quad (5.3)$$

Then, using  $\sigma^{\text{LER}}$ , we compute Monte Carlo simulations, i.e. we simulate Brownian motions, augment them with time and compute their signature. We then compute the linear combination of  $\sigma^{\text{LER}}$  against each simulated signature. This further validates Theorem 4.4 and highlights the efficiency of our numerical implementation using Gauss-Laguerre quadrature, a Black-Scholes control variate and the fourth-order Runge-Kutta scheme for the truncated tensor algebra valued Riccati equation.

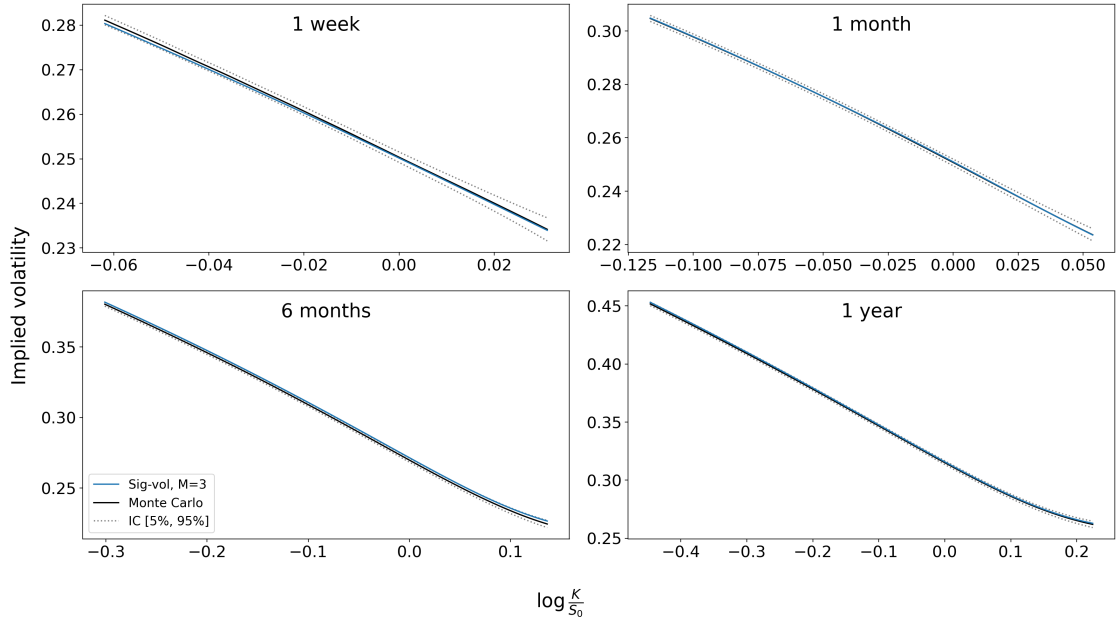


Figure 17: European put Monte Carlo pricing of LER linear volatility model vs Fourier pricing of Signature LER volatility model.  $\sigma^{\text{LER}}$  defined in (5.3).

## 5.2 Asian options

Let us consider a geometric Asian call option with price  $\bar{C}_t(T, K) = \mathbb{E}[(\bar{S}_T - K)^+ | \mathcal{F}_t]$ , where  $\bar{S}_t = \exp\left(\frac{1}{T} \int_0^t \log S_s ds\right)$ . Similarly to (5.1), we obtain the Fourier representation

$$\bar{C}_t(\bar{S}_t; T, K) = \bar{C}_t^{\text{BS}}(\bar{S}_t; T, K) - \frac{K}{\pi} \int_0^\infty \Re \left[ e^{i(u - \frac{i}{2})\bar{k}_t} \left( \bar{\phi}_t(u - \frac{i}{2}) - \bar{\phi}_t^{\text{BS}}(u - \frac{i}{2}) \right) \right] \frac{du}{(u^2 + \frac{1}{4})},$$

where  $\bar{\phi}$  is the characteristic function  $\bar{\phi}_t(u) = \mathbb{E}\left[e^{iu \log \frac{\bar{S}_T}{\bar{S}_t} | \mathcal{F}_t}\right]$  and  $\bar{k}_t = \log \frac{\bar{S}_t}{K}$  with

$$\bar{\phi}_t^{\text{BS}} = \exp\left(-\frac{\sigma_{\text{BS}}^2}{2} \left(u^2 \frac{(T-t)^3}{3T^2} + iu \frac{(T-t)^2}{2T}\right)\right), \quad (5.4)$$

and where

$$\bar{C}_t^{\text{BS}}(\bar{S}_t; T, K) := \mathcal{N}(\bar{d}_1) \bar{S}_t \exp\left(\frac{T-t}{T} \log S_t - \frac{\sigma_{\text{BS}}^2}{2} \left(\frac{(T-t)^2}{2T} - \frac{(T-t)^3}{3T^2}\right)\right) + \mathcal{N}(\bar{d}_2) K, \quad (5.5)$$

and

$$\bar{d}_2 := \frac{1}{\sigma_{\text{BS}} \sqrt{\frac{(T-t)^3}{3T^2}}} \left( \log \frac{\bar{S}_t}{K} + \frac{T-t}{T} \log S_t - \frac{\sigma_{\text{BS}}^2}{2} \frac{(T-t)^2}{2T} \right), \quad \bar{d}_1 := \bar{d}_2 - \sigma_{\text{BS}} \sqrt{\frac{(T-t)^3}{3T^2}}.$$

In Figure 18, we compare our signature volatility Fourier pricing for Asian options using the truncated, at order  $M = 4$ , representation of the linear Ornstein-Uhlenbeck (3.8) to Monte Carlo simulations. Lewis' approach together with Black-Scholes control variate was also used.

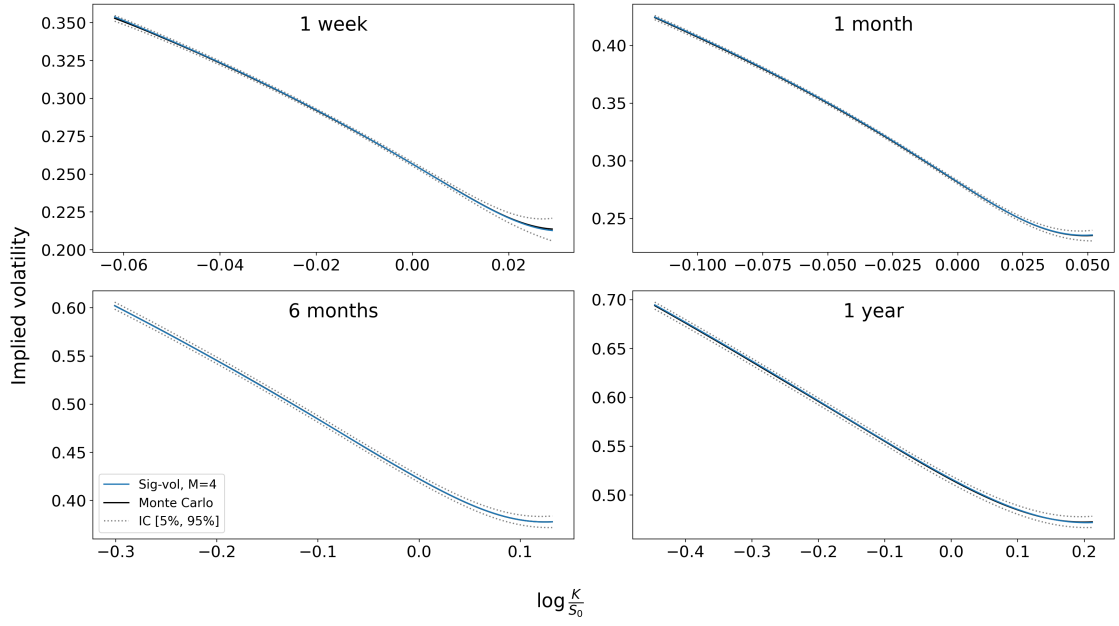


Figure 18: Asian put Monte Carlo pricing of Stein-Stein model vs Signature Ornstein-Uhlenbeck volatility model.  $\kappa = 1, \theta = 0.25, \eta = 1.2$  and  $\rho = -0.9$ .

### 5.3 $q$ -Volatility swaps

The payoff of a  $q$ -volatility swap is  $R_T^q - K^q$  where  $R_T^q = \left(\frac{1}{T}\bar{V}_T\right)^q$  is the realized  $q$ -volatility and  $K^q$  the  $q$ -volatility strike. The aim in pricing  $q$ -volatility swaps is to find the fair strike price, i.e.  $K^q = \mathbb{E}[R_T^q]$ . This is made possible by Laplace inversion:

**Theorem 5.1** (Schürger [56]). *Let  $X \geq 0$  be a random variable, then*

$$\mathbb{E}[X^q] = \frac{q}{\Gamma(1-q)} \int_0^\infty \frac{1 - \mathbb{E}[e^{-uX}]}{u^{q+1}} du, \quad q \in (0, 1).$$

It is straightforward to see that setting  $f = 0$  and  $g(t) = -\frac{u}{T}$  in (4.2) gives  $\tilde{M}_0(u) = \mathbb{E}[e^{-\frac{u}{T}\bar{V}_T}]$ , which allows us to compute analytically  $q$ -volatility swaps in the framework of the signature volatility models with

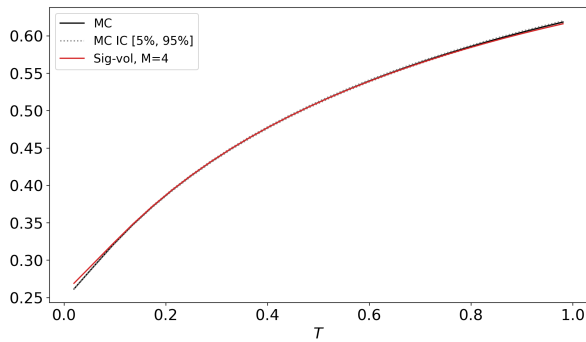
$$\mathbb{E}\left[\left(\frac{1}{T}\bar{V}_T\right)^q\right] = \frac{q}{\Gamma(1-q)} \int_0^\infty \frac{1 - \tilde{M}_0(u)}{u^{q+1}} du.$$

Specifically, the fair strike of the volatility swap, i.e.  $q = \frac{1}{2}$ , is thus of the form

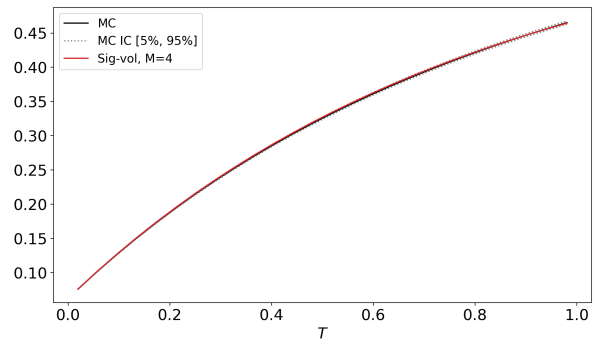
$$\mathbb{E}\left[\sqrt{\frac{1}{T}\bar{V}_T}\right] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - \tilde{M}_0(u)}{u^{3/2}} du.$$

Moreover, the fair strike of the variance swap, i.e.  $q = 1$ , can be written in closed form thanks to Fawcett's formula (3.6) for time-independent representations, i.e.

$$K_T^1 = \frac{1}{T} \left\langle \sigma^{\perp \perp 2} \mathbf{1}, \mathbb{E}[\widehat{\mathbb{W}}_T] \right\rangle.$$



(a)  $q = \frac{1}{2}$  (volatility swap).



(b)  $q = 1$  (variance swap).

Figure 19: Strikes of  $q$ -volatility swaps as a function of the maturity, for a Signature Ornstein-Uhlenbeck volatility model.  $\kappa = 1, \theta = 0.25, \eta = 1.2$  and  $\rho = -0.7$ .

## 5.4 Calibration examples

Another way to learn the dynamics  $\sigma$  of the signature volatility is to calibrate it against implied volatility surfaces. In this section, we consider  $\sigma \in T^3(\mathbb{R}^2)$ . We recall that the number of non-zero terms in an object of  $T^M(\mathbb{R}^d)$ , is  $\frac{d^{M+1}-1}{d-1}$ . In our case,  $d = 2$  and we use  $M = 3$  as it is sufficient to make relatively good fits, i.e. we calibrate 15 parameters. For Figures 20, 21 and 22, we minimized the MSE between a given implied volatility and the one of the signature model, using the differential evolution global minimizer [59].

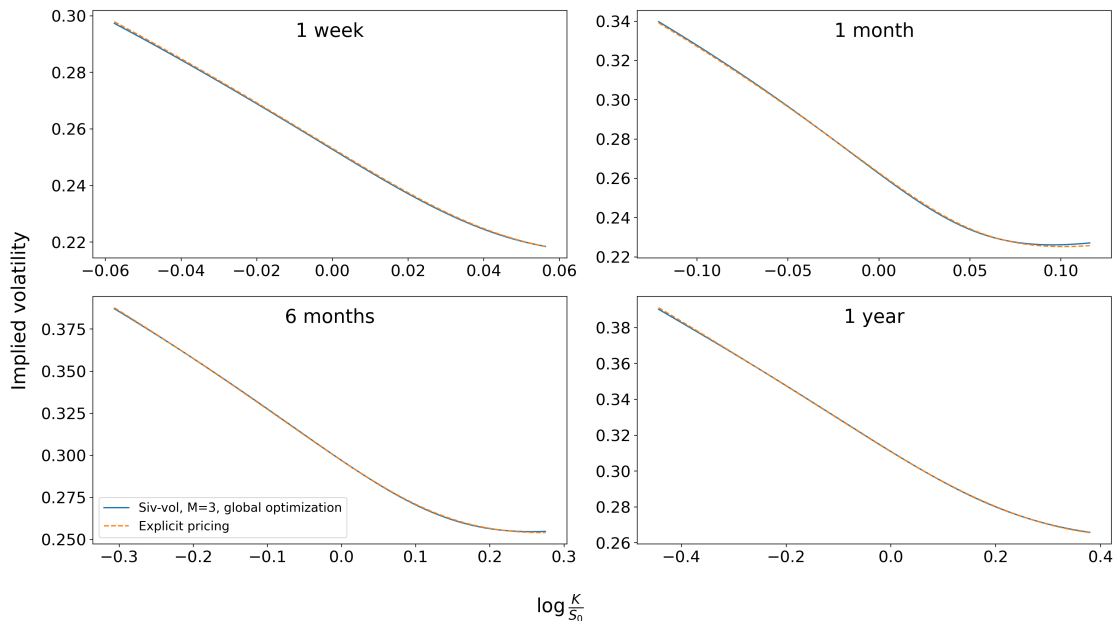


Figure 20: Implied volatility of time-independent Signature model calibrated against Heston implied volatility surface.  $V_0 = 0.25^2, \kappa = 4, \theta = 0.0625, \eta = 0.7$  and  $\rho = -0.6$ .

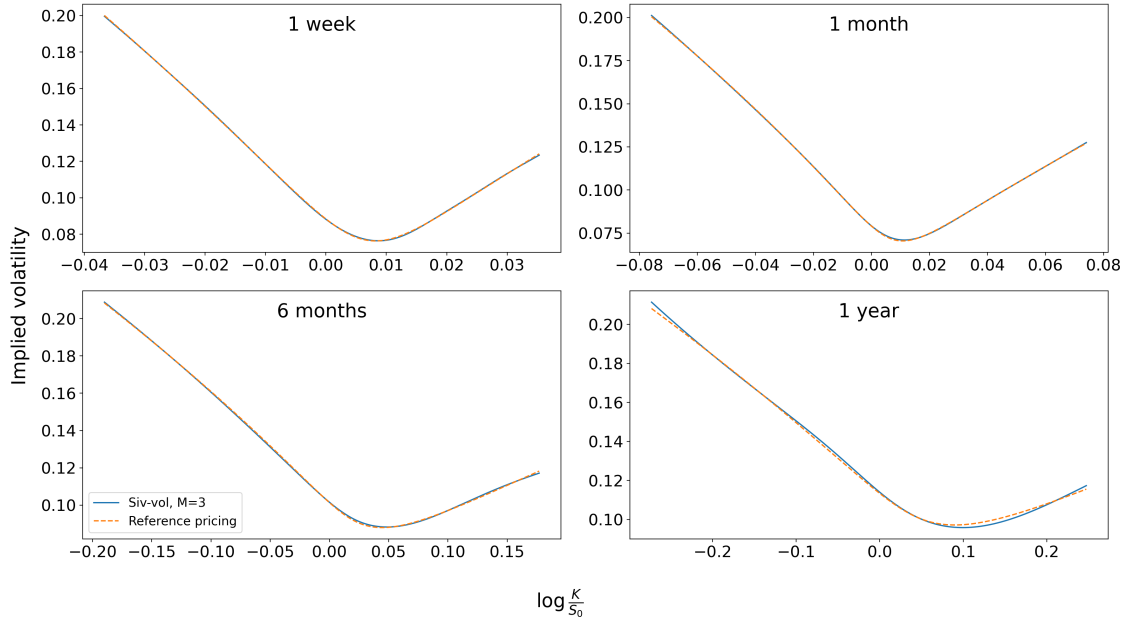


Figure 21: Implied volatility of time-dependent Signature model calibrated against Quintic implied volatility surface.  $\xi_0(t) \equiv 0.25^2$ ,  $H = -0.1$ ,  $\epsilon = \frac{1}{52}$ ,  $\alpha = \{0.01, 1, 0, 0.214, 0, 0.227\}$  and  $\rho = -0.65$ .

Finally, we show how our signature volatility model is well adapted to produce a good fit with daily SPX implied volatility surface data purchased from the CBOE website <https://datashop.cboe.com/>.

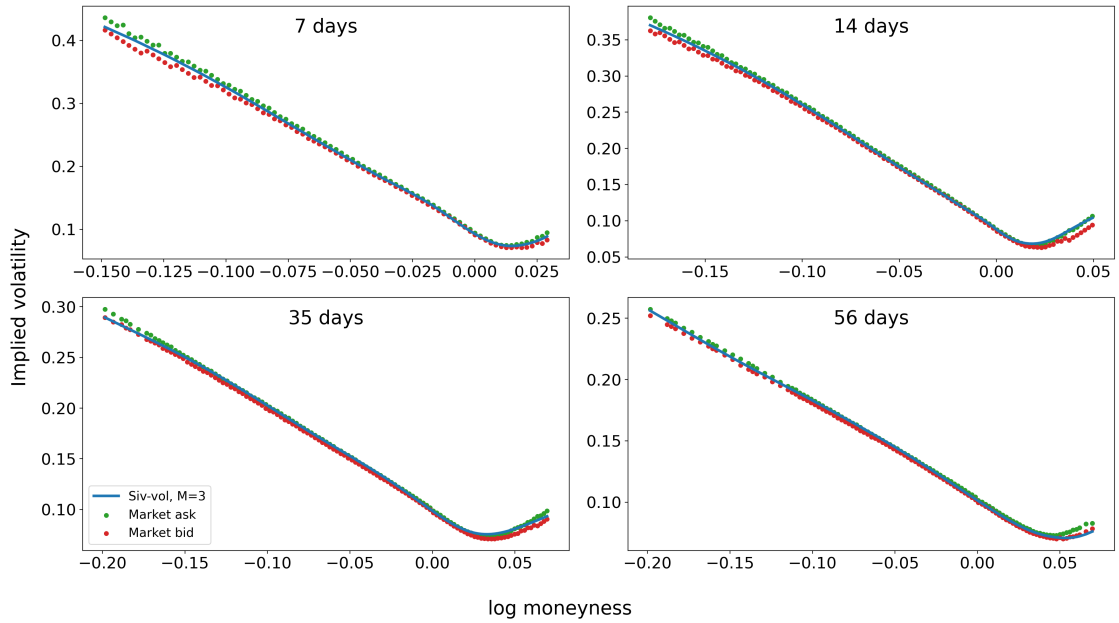


Figure 22: Implied volatility of time-dependent Signature model calibrated against SPX data 2017-05-19.

## 6 Quadratic hedging by Fourier methods

The signature volatility model (3.2)-(3.3) generates an incomplete market in general (unless the correlation  $\rho$  between the two Brownian motions  $B$  and  $W$  is  $\pm 1$ ). Therefore contingent claims on the stock  $S$  cannot be perfectly hedged. We will consider quadratic hedging methods instead, we refer to [57] for a detailed overview of quadratic hedging approaches. We will show that in our setup, quadratic hedging remains highly tractable

in signature volatility models using Fourier techniques on the conditional characteristic function (4.4). We do this in two steps:

1. We first solve in Section 6.1 the quadratic hedging problem for a generic contingent claim  $\xi$ , in a general stochastic volatility model driven by the two dimensional Brownian motion  $(W, W^\perp)$ . The hedging strategy depends on the quantities that appear in the martingale representation theorem of  $\mathbb{E}[\xi|\mathcal{F}_t]$ . We provide a concise proof in this setting.
2. We then show in Section 6.2 that in the setting of a signature volatility model and for contingent claims that admit Fourier representation, such as European and Asian call and put options, the hedging strategy can be recovered from the conditional characteristic function (4.4).

## 6.1 A generic solution

Let  $\xi$  be an  $\mathcal{F}_T$ -measurable non-negative random variable such that  $\mathbb{E}[\xi^2] < \infty$  that we are looking to hedge using a self-financing portfolio. We recall that  $(\mathcal{F}_t)_{t \geq 0}$  is the filtration generated by  $(W, W^\perp)$ . A self-financing hedging portfolio  $X$  consists of an initial wealth  $X_0 \in \mathbb{R}$  and a progressively measurable strategy  $(\vartheta_u)_{u \leq T}$  of the amount of shares invested in asset  $S$  given in (3.2) at time  $u \leq T$ . It has the following dynamics

$$X_t^\alpha = X_0 + \int_0^t \vartheta_u dS_u = X_0 + \int_0^t \alpha_u dB_u,$$

with  $\alpha_u := \vartheta_u S_u \Sigma_u$ . The set of admissible hedging strategies  $\alpha$  is defined by

$$\mathcal{H} = \left\{ \alpha \text{ progressively measurable such that } \int_0^T \mathbb{E}[\alpha_s^2] ds < \infty \right\}.$$

We stress that in this section we do not impose specific dynamics for the stochastic volatility  $\Sigma$ , i.e. (3.3),  $\Sigma$  is only assumed to be adapted to the Brownian motion  $(W, W^\perp)$ .

A quadratic hedging strategy aims at minimizing the following objective function

$$J(X_0, \alpha) = \mathbb{E} \left[ (X_T^\alpha - \xi)^2 \right] \quad (6.1)$$

over  $X_0 \in \mathbb{R}$  and  $\alpha \in \mathcal{H}$ .

The next theorem provides a solution of the quadratic hedging problem using the martingale representation theorem. Note that  $(\mathbb{E}[\xi|\mathcal{F}_t])_{t \leq T}$  is a square integrable martingale with terminal value  $\xi$  at  $T$ . An application of the martingale representation theorem [44, Theorem 4.15] yields the existence of two progressively measurable and square integrable processes  $Z$  and  $Z^\perp$  such that

$$\mathbb{E}[\xi|\mathcal{F}_t] = \xi - \int_t^T Z_s dW_s - \int_t^T Z_s^\perp dW_s^\perp.$$

**Theorem 6.1.** *The value of the quadratic hedging problem is given by*

$$\inf_{X_0 \in \mathbb{R}, \alpha \in \mathcal{H}} J(X_0, \alpha) = \mathbb{E} \left[ \int_0^T (Z_t^2 + (Z_t^\perp)^2) dt - \int_0^T \left( \rho Z_t + \sqrt{1 - \rho^2} Z_t^\perp \right)^2 dt \right], \quad (6.2)$$

where the optimum is attained for  $(X_0^*, \alpha^*)$  given by

$$X_0^* = \mathbb{E}[\xi] \quad \text{and} \quad \alpha_t^* = \rho Z_t + \sqrt{1 - \rho^2} Z_t^\perp, \quad t \leq T. \quad (6.3)$$

*Proof.* We start by expanding the square inside the objective function  $J$  in (6.1):

$$(X_T^\alpha - \xi)^2 = (X_T^\alpha)^2 - 2\xi X_T^\alpha + \xi^2,$$

and we write

$$\begin{aligned} (X_T^\alpha)^2 &= X_0^2 + 2X_0 \int_0^T \alpha_u dB_u + \left( \int_0^T \alpha_u dB_u \right)^2 \\ \xi X_T^\alpha &= \mathbb{E}[\xi]X_0 + \int_0^T \mathbb{E}[\xi|\mathcal{F}_s] dX_s^\alpha + \int_0^T X_s^\alpha d\mathbb{E}[\xi|\mathcal{F}_s] + \int_0^T \alpha_s \left( \rho Z_s + \sqrt{1 - \rho^2} Z_s^\perp \right) ds \end{aligned}$$



so that by using the fact that stochastic integrals are centered and by Itô's isometry we obtain that

$$\begin{aligned} J(X_0, \alpha) &= \mathbb{E} \left[ (X_0 - \mathbb{E}[\xi])^2 + (\xi^2 - \mathbb{E}[\xi]^2) + \int_0^T \left( \alpha_s^2 - 2\alpha_s \left( \rho Z_s + \sqrt{1 - \rho^2} Z_s^\perp \right) \right) ds \right] \\ &= \mathbb{E} \left[ (X_0 - \mathbb{E}[\xi])^2 + \int_0^T (Z_s^2 + (Z_s^\perp)^2) ds + \int_0^T (\alpha_s - \alpha_s^*)^2 ds - \int_0^T \left( \rho Z_s + \sqrt{1 - \rho^2} Z_s^\perp \right)^2 ds \right]. \end{aligned}$$

It immediately follows that the minimum is clearly attained for  $X_0^*$  and  $\alpha^*$  given by (6.3) and the claimed expression for the value function (6.2) follows.  $\square$

## 6.2 Fourier implementation in the signature volatility model

We now illustrate how the optimal hedging strategy  $X_0^*$  and  $\alpha^*$  given in (6.3) can be recovered numerically from the knowledge of the conditional characteristic function (4.4) in the specific case of a signature volatility model, i.e. when  $\Sigma$  is of the form (3.3) and for contingent claims that admit a Fourier representation.

### 6.2.1 European call option

In this section, we consider a European call options. In order to implement the quadratic hedging, the idea is to re-express the Fourier inversion formula (5.1) in terms of the process  $M_t(u)$  in (4.3) with  $f(t) = iu$  and  $g = 0$  and apply Itô. Since in this case

$$M_t(u) = \phi_t(u) e^{iu \log S_t},$$

the representation (5.1) directly leads to

$$C_t(S_t; T, K) = C_t^{\text{BS}}(S_t; T, K) - \frac{K}{\pi} \int_0^\infty \Re \left[ e^{i(u - \frac{i}{2}) \log K} \left( M_t \left( u - \frac{i}{2} \right) - M_t^{\text{BS}} \left( u - \frac{i}{2} \right) \right) \right] \frac{du}{\left( u^2 + \frac{1}{4} \right)},$$

where

$$M_t^{\text{BS}}(u) = \phi_t^{\text{BS}}(u) e^{iu \log S_t},$$

and  $\phi_t^{\text{BS}}$  as defined in (5.2). Setting  $w(u) := \frac{K}{\pi} \frac{e^{-i(u - \frac{i}{2}) \log K}}{u^2 + \frac{1}{4}}$  and  $\tilde{u} := (u - \frac{i}{2})$ , an application of Itô's formula yields that

$$\begin{aligned} dC_t &= dC_t^{\text{BS}} + \int_0^\infty \Re \left[ w(u) \left( dM_t(\tilde{u}) - dM_t^{\text{BS}}(\tilde{u}) \right) \right] du \\ &= \Delta_t^{\text{BS}} dS_t + \left( \Theta_t^{\text{BS}} + \frac{1}{2} \Gamma_t^{\text{BS}}(S_t \Sigma_t)^2 \right) dt \\ &\quad + \int_0^\infty \Re \left[ w(u) \left( M_t(\tilde{u}) \left( dU_t + \frac{1}{2} d[U]_t \right) \right. \right. \\ &\quad \left. \left. - M_t^{\text{BS}}(\tilde{u}) \left( d \log M_t^{\text{BS}}(\tilde{u}) + \frac{1}{2} d[\log M^{\text{BS}}(\tilde{u})]_t \right) \right) \right] du, \end{aligned}$$

where  $\Delta_t^{\text{BS}} = \frac{\partial}{\partial S_t} C_t^{\text{BS}}$ ,  $\Theta_t^{\text{BS}} = \frac{\partial}{\partial t} C_t^{\text{BS}}$  and  $\Gamma_t^{\text{BS}} = \frac{\partial^2}{\partial S_t^2} C_t^{\text{BS}}$ . Furthermore, using equalities between Black-Scholes Greeks, one can show that

$$\begin{aligned} dC_t &= \Delta_t^{\text{BS}} S_t \Sigma_t dB_t + \frac{\Theta_t^{\text{BS}}}{\sigma_{\text{BS}}^2} (\sigma_{\text{BS}}^2 - \Sigma_t^2) dt \\ &\quad + \int_0^\infty \Re \left[ w(u) \left( M_t(\tilde{u}) \left( \left\langle \psi_t(\tilde{u}) \right|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle dW_t + i\tilde{u} \Sigma_t dB_t \right) \right. \\ &\quad \left. - M_t^{\text{BS}}(\tilde{u}) \left[ \frac{1}{2} (\tilde{u}^2 + i\tilde{u}) (\sigma_{\text{BS}}^2 - \Sigma_t^2) dt + i\tilde{u} \Sigma_t dB_t \right] \right) \right] du \\ &= \Delta_t^{\text{BS}} S_t \Sigma_t dB_t + \\ &\quad + \int_0^\infty \Re \left[ w(u) \left( M_t(\tilde{u}) \left( \left\langle \psi_t(\tilde{u}) \right|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle dW_t + i\tilde{u} \Sigma_t dB_t \right) \right. \\ &\quad \left. - M_t^{\text{BS}}(\tilde{u}) i\tilde{u} \Sigma_t dB_t \right) \right] du. \end{aligned}$$

The final equality comes from the fact that  $\Theta_t^{\text{BS}}$  can also be written as a Fourier integral. This gives option price dynamics of the following form

$$dC_t = Z_t dW_t + Z_t^\perp dW_t^\perp,$$

where  $Z$  and  $Z^\perp$  are defined as follows

$$\begin{aligned} Z_t &= \Sigma_t S_t \Delta_t^{\text{BS}} \rho + \int_0^\infty \Re [\zeta_t(\tilde{u}) w(u)] du \\ Z_t^\perp &= \Sigma_t S_t \Delta_t^{\text{BS}} \sqrt{1 - \rho^2} + \int_0^\infty \Re [\zeta_t^\perp(\tilde{u}) w(u)] du, \end{aligned}$$

with

$$\begin{aligned} \zeta_t(u) &:= iu \Sigma_t (M_t(u) - M_t^{\text{BS}}(u)) \rho + M_t(u) \left\langle \boldsymbol{\psi}_t(u) |_{\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle \\ \zeta_t^\perp(u) &:= iu \Sigma_t (M_t(u) - M_t^{\text{BS}}(u)) \sqrt{1 - \rho^2}. \end{aligned}$$

This allows us to solve the quadratic hedging problem in (6.3) numerically for European call options in the framework of signature volatility models. Moreover, applying the put-call parity allows us to easily extend it to European put options.

In Figure (23), we simulate price trajectories under the Stein-Stein model [58] and compare the performance of the explicit hedging strategy to the Fourier hedging of the signature Ornstein-Uhlenbeck volatility model for a European put option with multiple strikes and two horizons.

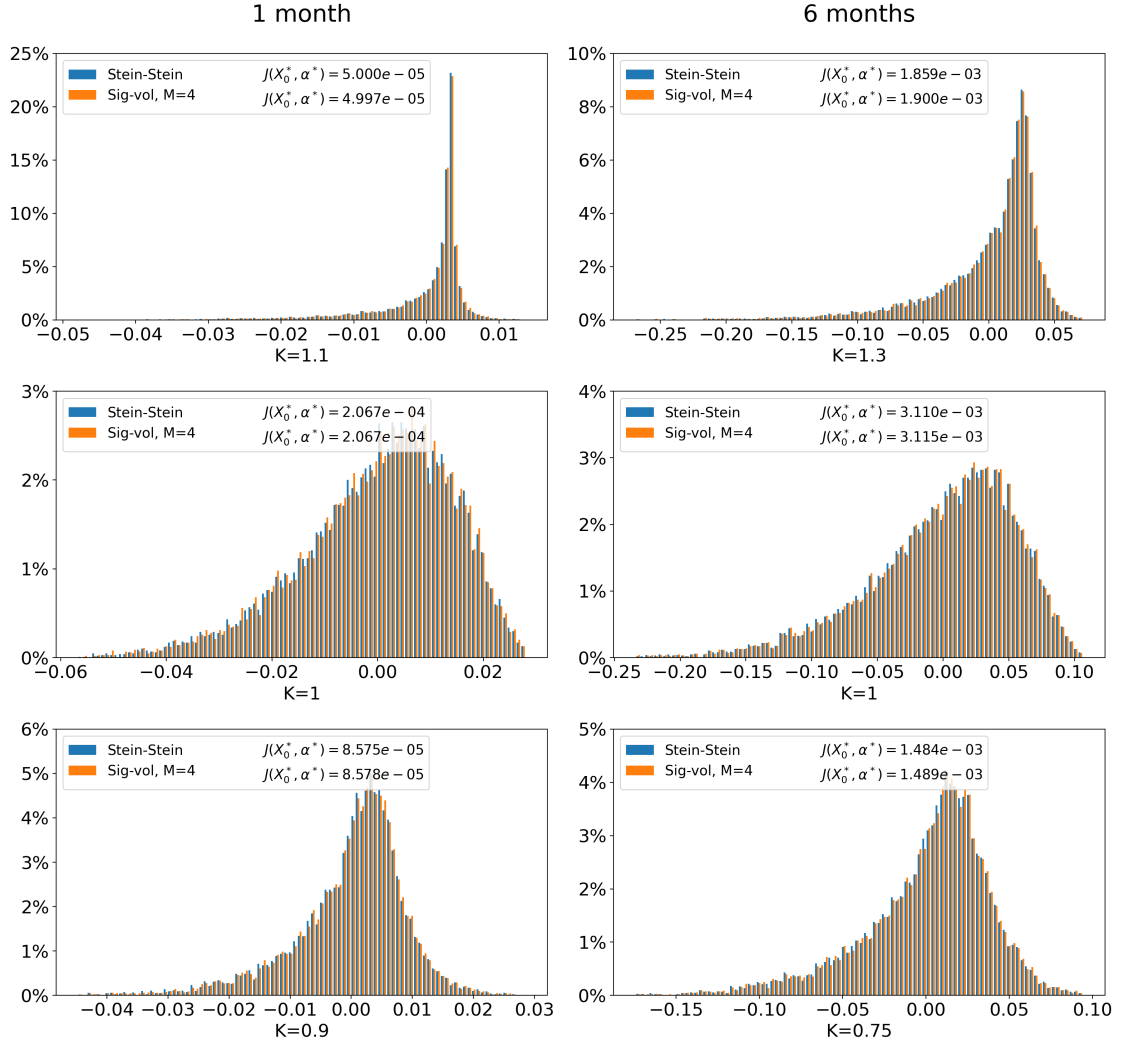


Figure 23: P&L of Stein-Stein model (blue) vs signature Ornstein-Uhlenbeck volatility model (orange) quadratic hedging strategies.  $\kappa = 1, \theta = 0.25, \eta = 1.2$  and  $\rho = -0.6$ .

Remark that both strategies mostly coincide. It illustrates that both the method in Section 6.2 works well within our framework and that the truncated signature representation of the volatility process is a good approximation when horizons are short enough, relative to the stiffness of the model's parameters.

### 6.2.2 Asian call option

In the same spirit as for the European call option, we will express the Fourier inversion formula (5.1) in terms of the process  $\bar{M}_t(u)$  in (4.4) with  $f(s) := iu \frac{T-s}{T}$  and  $g = 0$  and apply Itô. Using the same notations as in Section 5.2, for all  $t \leq T$

$$\begin{aligned} \exp\left(\langle \psi_t, \widehat{\mathbb{W}}_t \rangle\right) &= \mathbb{E} \left[ \exp \left( \int_t^T f(s) d \log S_s \right) \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[ \exp \left( iu \log \frac{\bar{S}_T}{\bar{S}_t} - iu \frac{T-t}{T} \log S_t \right) \middle| \mathcal{F}_t \right], \end{aligned}$$

so that, recall (4.2),

$$\begin{aligned}\bar{\phi}_t(u) &= \exp\left(\langle \boldsymbol{\psi}_t, \widehat{\mathbb{W}}_t \rangle + iu \frac{T-t}{T} \log S_t\right) \\ &= \bar{M}_t(u) \exp\left(-\int_0^t f(s) d\log S_s + iu \frac{T-t}{T} \log S_t\right) \\ &= \bar{M}_t(u) e^{-iu \log \bar{S}_t}.\end{aligned}$$

Thus one can now write  $\bar{C}$  as a Fourier integral on  $\bar{M}$

$$\bar{C}_t(\bar{S}_t; T, K) = \bar{S}_t - \frac{K}{\pi} \int_0^\infty \Re \left[ e^{-i(u-\frac{i}{2}) \log K} \bar{M}_t\left(u - \frac{i}{2}\right) \right] \frac{du}{\left(u^2 + \frac{1}{4}\right)},$$

together with its Black-Scholes control variate version

$$\bar{C}_t(\bar{S}_t; T, K) = \bar{C}_t^{\text{BS}}(\bar{S}_t; T, K) - \frac{K}{\pi} \int_0^\infty \Re \left[ e^{i(u-\frac{i}{2}) \log K} \left( \bar{M}_t\left(u - \frac{i}{2}\right) - \bar{M}_t^{\text{BS}}\left(u - \frac{i}{2}\right) \right) \right] \frac{du}{\left(u^2 + \frac{1}{4}\right)}$$

with

$$\bar{M}_t^{\text{BS}}(u) := \bar{\phi}_t^{\text{BS}}(u) e^{iu \log \bar{S}_t},$$

where  $\bar{\phi}^{\text{BS}}$  is defined in (5.4) and  $\bar{C}^{\text{BS}}$  in (5.5).

Finally,  $w(u)$  and  $\tilde{u}$  are defined as in Subsection 6.2.1 and with very similar computations, one can get  $\bar{Z}$  and  $\bar{Z}^\perp$  as follows

$$\begin{aligned}\bar{Z}_t &= \Sigma_t S_t \bar{\Delta}_t^{\text{BS}} \rho + \int_0^\infty \Re \left[ \bar{\zeta}_t(\tilde{u}) w(u) \right] du \\ \bar{Z}_t^\perp &= \Sigma_t S_t \bar{\Delta}_t^{\text{BS}} \sqrt{1-\rho^2} + \int_0^\infty \Re \left[ \bar{\zeta}_t^\perp(\tilde{u}) w(u) \right] du,\end{aligned}$$

with

$$\begin{aligned}\bar{\zeta}_t(u) &:= \frac{T-t}{T} iu \Sigma_t \left( \bar{M}_t(u) - \bar{M}_t^{\text{BS}}(u) \right) \rho + \bar{M}_t(u) \left\langle \boldsymbol{\psi}_t(u) |_{\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle \\ \bar{\zeta}_t^\perp(u) &:= \frac{T-t}{T} iu \Sigma_t \left( \bar{M}_t(u) - \bar{M}_t^{\text{BS}}(u) \right) \sqrt{1-\rho^2},\end{aligned}$$

where  $\bar{\Delta}_t^{\text{BS}} = \frac{\partial}{\partial S_t} \bar{C}_t^{\text{BS}}$ .

This allows us to solve the quadratic hedging problem in (6.3) numerically for Asian call and put options in the framework of sig-volatility models.

In Figure 24, we simulate price trajectories under mean-reverting geometric Brownian motion volatility model and compare the performance of the Black-Scholes hedging strategy, simply as a point of reference with  $\sigma_{\text{BS}} = \theta$ , to the Fourier hedging of the signature mean-reverting geometric Brownian motion volatility model for an Asian put option.

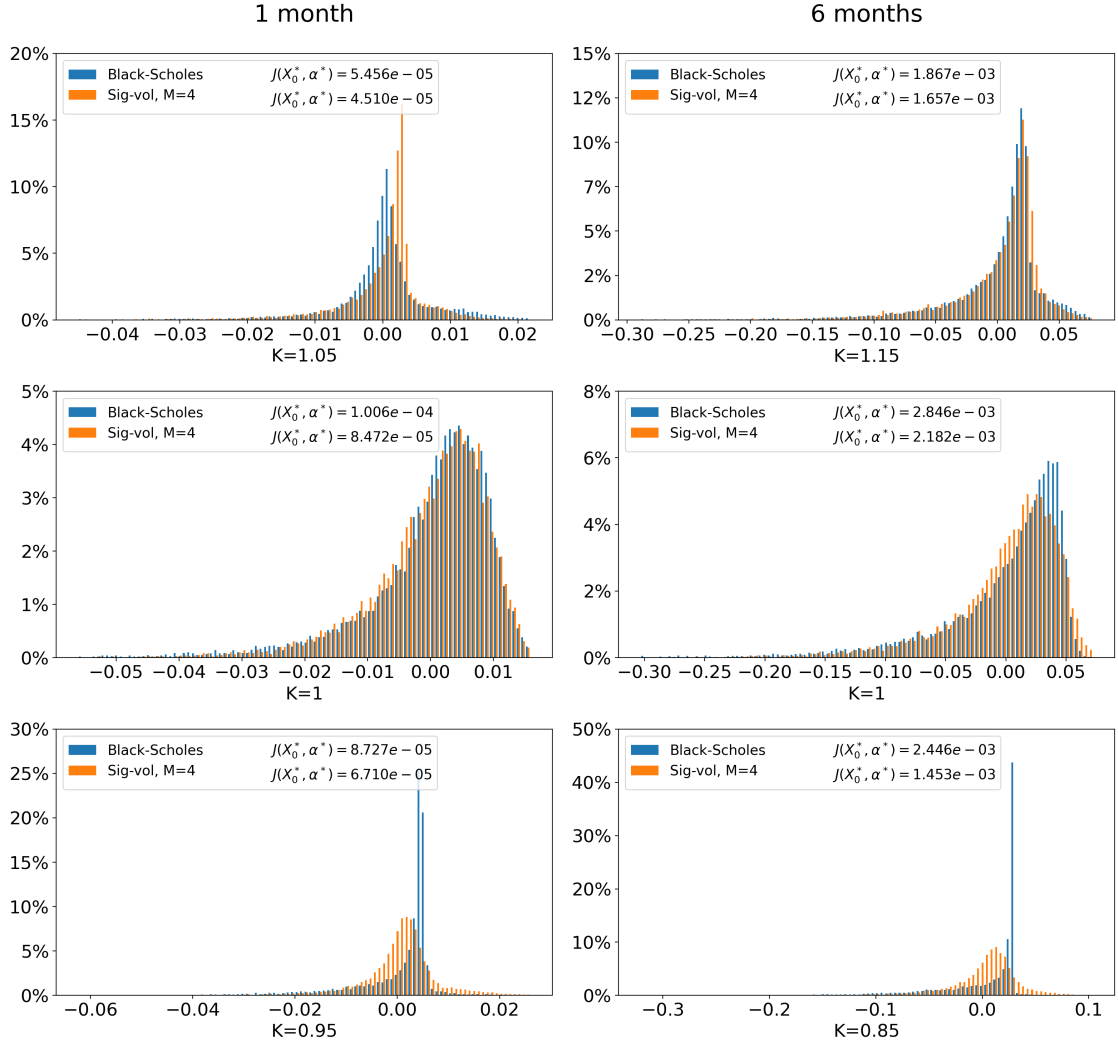


Figure 24: P&L of mGBM volatility model (blue) vs signature mGBM volatility model (orange) quadratic hedging strategies.  $\kappa = 1, \theta = 0.25, \eta = 1.2, \alpha = 0.6$  and  $\rho = -0.6$ .

Remark that the signature volatility systematically outperforms, in terms of minimized squared P&L, the naive Black-Scholes quadratic hedging strategy by 10 to 25%. This suggests that our framework might specifically be relevant for path-dependent options and more complex volatility dynamics where hedging strategies are not known explicitly or tractable.

## A Proofs

### A.1 Proof of Lemma 3.1

**Lemma A.1.**  $\sigma^{\text{OU}} = (x\boldsymbol{\varphi} + \kappa\theta\mathbf{1} + \eta\mathbf{2})e^{\mathbb{L}\mathbb{I} - \kappa\mathbf{1}}$  solves the equation

$$\sigma^{\text{OU}} = (x\boldsymbol{\varphi} + \kappa\theta\mathbf{1} + \eta\mathbf{2}) - \kappa\sigma^{\text{OU}}\mathbf{1}. \quad (\text{A.1})$$

*Proof.* Straightforward by applying Proposition 2.3.  $\square$

Using [6, Theorem 4.2], one has

$$\sigma^{\text{OU}} = (x\boldsymbol{\varphi} + \kappa\theta\mathbf{1} + \eta\mathbf{2})e^{\mathbb{L}\mathbb{I} - \kappa\mathbf{1}} \in \mathcal{A}.$$

We can therefore define the process

$$X_t = \langle \sigma^{\text{OU}}, \widehat{\mathbb{W}}_t \rangle.$$

Now, showing

$$\begin{aligned} (\sigma^{\text{OU}})^{\boldsymbol{\varphi}} &= y\boldsymbol{\varphi}, \\ \sigma^{\text{OU}}|_{\mathbf{1}} &= \kappa\theta\boldsymbol{\varphi} - \kappa\sigma^{\text{OU}}, \\ \sigma^{\text{OU}}|_{\mathbf{2}} &= \eta\boldsymbol{\varphi}, \\ \sigma^{\text{OU}}|_{\mathbf{22}} &= 0. \end{aligned}$$

makes it clear that  $\|\sigma^{\text{OU}}\|_t^{\mathcal{I}} < \infty$  a.s.. Moreover, remarking that  $\sigma^{\text{OU}}|_{\mathbf{1}\mathbf{1}} = \sigma^{\text{OU}} - x\boldsymbol{\varphi} - \eta\mathbf{2}$  by using (A.1), allows us to write

$$\int_0^t \langle \sigma^{\text{OU}}|_{\mathbf{1}}, \widehat{\mathbb{W}}_s \rangle ds = \langle \sigma^{\text{OU}}|_{\mathbf{1}\mathbf{1}}, \widehat{\mathbb{W}}_t \rangle < \infty,$$

and thus have  $\int_0^t \|\sigma^{\text{OU}}\|_s^{\mathcal{I}} ds < \infty$  a.s. and  $\sigma^{\text{OU}} \in \mathcal{I}$ .

We are now ready to apply Proposition 2.7

$$\begin{aligned} dY_t &= d \langle \sigma^{\text{OU}}, \widehat{\mathbb{W}}_t \rangle \\ &= \langle \sigma^{\text{OU}}|_{\mathbf{1}} + \frac{1}{2}\sigma^{\text{OU}}|_{\mathbf{22}}, \widehat{\mathbb{W}}_t \rangle dt + \langle \sigma^{\text{OU}}|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \rangle dW_t \\ &= \langle \kappa\theta\boldsymbol{\varphi} - \kappa\sigma^{\text{OU}}, \widehat{\mathbb{W}}_t \rangle dt + \langle \eta\boldsymbol{\varphi}, \widehat{\mathbb{W}}_t \rangle dW_t \\ &= \left( \kappa\theta - \kappa \langle \sigma^{\text{OU}}, \widehat{\mathbb{W}}_t \rangle \right) dt + \eta dW_t \\ &= \kappa(\theta - Y_t) dt + \eta dW_t. \end{aligned}$$

By uniqueness of the solution of the Ornstein-Uhlenbeck, the representation (3.8) follows.

### A.2 Proof of Lemma 3.2

**Lemma A.2.**  $\sigma^{\text{mGBM}} = (y\boldsymbol{\varphi} + (\kappa\theta - \frac{\alpha\eta}{2})\mathbf{1} + \eta\mathbf{2})e^{\mathbb{L}\mathbb{I} - (\kappa + \frac{\alpha^2}{2})\mathbf{1} + \alpha\mathbf{2}}$  solves the equation

$$\sigma^{\text{mGBM}} = \left( x\boldsymbol{\varphi} + \left( \kappa\theta - \frac{\alpha\eta}{2} \right) \mathbf{1} + \eta\mathbf{2} \right) + \sigma^{\text{mGBM}} \left( - \left( \kappa + \frac{\alpha^2}{2} \right) \mathbf{1} + \alpha\mathbf{2} \right).$$

*Proof.* Straightforward by applying Proposition 2.3.  $\square$

Using [6, Theorem 4.2], one has

$$\sigma^{\text{mGBM}} = \left( y\boldsymbol{\varphi} + \left( \kappa\theta - \frac{\alpha\eta}{2} \right) \mathbf{1} + \eta\mathbf{2} \right) e^{\mathbb{L}\mathbb{I} - (\kappa + \frac{\alpha^2}{2})\mathbf{1} + \alpha\mathbf{2}} \in \mathcal{A}.$$

We can therefore define the process

$$Y_t = \langle \sigma^{\text{mGBM}}, \widehat{\mathbb{W}}_t \rangle.$$

Moreover, showing

$$\begin{aligned}
(\boldsymbol{\sigma}^{\text{mGBM}})^\phi &= y\phi, \\
\boldsymbol{\sigma}^{\text{mGBM}}|_{\mathbf{1}} &= \left(\kappa\theta - \frac{\alpha\eta}{2}\right)\phi - \left(\kappa + \frac{\alpha^2}{2}\right)\boldsymbol{\sigma}^{\text{mGBM}}, \\
\boldsymbol{\sigma}^{\text{mGBM}}|_{\mathbf{2}} &= \eta\phi + \alpha\boldsymbol{\sigma}^{\text{mGBM}}, \\
\boldsymbol{\sigma}^{\text{mGBM}}|_{\mathbf{22}} &= \alpha\eta\phi + \alpha^2\boldsymbol{\sigma}^{\text{mGBM}},
\end{aligned}$$

makes it clear that  $\|\boldsymbol{\sigma}^{\text{mGBM}}\|_t^Z < \infty$ . As it requires much more care to prove  $\int_0^t \|\boldsymbol{\sigma}^{\text{mGBM}}\|_s^Z ds < \infty$ , it is left for the interested reader to refer to [6, Section 7] for a detailed proof.

We can thus finally use Proposition 2.7

$$\begin{aligned}
dY_t &= d\left\langle \boldsymbol{\sigma}^{\text{mGBM}}, \widehat{\mathbb{W}}_t \right\rangle \\
&= \left\langle \boldsymbol{\sigma}^{\text{mGBM}}|_{\mathbf{1}} + \frac{1}{2}\boldsymbol{\sigma}^{\text{mGBM}}|_{\mathbf{22}}, \widehat{\mathbb{W}}_t \right\rangle dt + \left\langle \boldsymbol{\sigma}^{\text{mGBM}}|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle dW_t. \\
&= \left\langle \left(\kappa\theta - \frac{\alpha\eta}{2}\right)\phi - \left(\kappa + \frac{\alpha^2}{2}\right)\boldsymbol{\sigma}^{\text{mGBM}} + \frac{1}{2}(\alpha\eta\phi + \alpha^2\boldsymbol{\sigma}^{\text{mGBM}}), \widehat{\mathbb{W}}_t \right\rangle dt \\
&\quad + \left\langle \eta\phi + \alpha\boldsymbol{\sigma}^{\text{mGBM}}, \widehat{\mathbb{W}}_t \right\rangle dW_t. \\
&= \left(\kappa\theta - \kappa\left\langle \boldsymbol{\sigma}^{\text{mGBM}}, \widehat{\mathbb{W}}_t \right\rangle\right) dt + \left(\eta + \alpha\left\langle \boldsymbol{\sigma}^{\text{mGBM}}, \widehat{\mathbb{W}}_t \right\rangle\right) dW_t \\
&= \kappa(\theta - Y_t) dt + (\eta + \alpha Y_t) dW_t
\end{aligned}$$

By uniqueness of the solution of the mean-reverting geometric Brownian motion, the representation (3.12) follows.

## References

- [1] Eduardo Abi Jaber. The characteristic function of Gaussian stochastic volatility models: an analytic expression. *Finance and Stochastics*, 26(4):733–769, 2022.
- [2] Eduardo Abi Jaber and Shaun Xiaoyuan Li. Volatility models in practice: Rough, path-dependent or markovian? *arXiv preprint arXiv:2401.03345*, 2024.
- [3] Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine Volterra processes. *The Annals of Applied Probability*, 29(5):3155–3200, 2019.
- [4] Eduardo Abi Jaber, Camille Illand, and Shaun Li. The quintic Ornstein-Uhlenbeck volatility model that jointly calibrates spx & vix smiles. *arXiv preprint arXiv:2212.10917*, 2022.
- [5] Eduardo Abi Jaber, Camille Illand, and Shaun Xiaoyuan Li. Joint SPX–VIX calibration with Gaussian polynomial volatility models: deep pricing with quantization hints. *arXiv preprint arXiv:2212.08297*, 2022.
- [6] Eduardo Abi Jaber, Louis-Amand Gérard, and Yuxing Huang. Linear path-dependent processes from signatures. *Working paper*, 2023.
- [7] Eduardo Abi Jaber, Shaun Li, and Xuyang Lin. Characteristic function in polynomial Ornstein-Uhlenbeck volatility models. *Working paper*, 2024.
- [8] Leif Andersen and Jesper Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4(3):231–262, 2000. doi: 10.1023/A:1011354913068. URL <https://doi.org/10.1023/A:1011354913068>.
- [9] Torben G Andersen and Tim Bollerslev. Intraday periodicity and volatility persistence in financial markets. *Journal of empirical finance*, 4(2-3):115–158, 1997.

- [10] Imanol Perez Arribas, Cristopher Salvi, and Lukasz Szpruch. Sig-SDEs model for quantitative finance. In *Proceedings of the First ACM International Conference on AI in Finance*, pages 1–8, 2020.
- [11] Emmanuel Bacry, Iacopo Mastromatteo, and Jean-François Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(01):1550005, 2015.
- [12] Christian Bayer, Peter Friz, and Jim Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904, 2016.
- [13] Christian Bayer, Paul P Hager, Sebastian Riedel, and John Schoenmakers. Optimal stopping with signatures. *The Annals of Applied Probability*, 33(1):238–273, 2023.
- [14] Christian Bayer, Luca Pelizzari, and John Schoenmakers. Primal and dual optimal stopping with signatures. *arXiv preprint arXiv:2312.03444*, 2023.
- [15] Lorenzo Bergomi. Smile dynamics II. *Derivatives*, 2005.
- [16] Jean-Philippe Bouchaud, Yuval Gefen, Marc Potters, and Matthieu Wyart. Fluctuations and response in financial markets: the subtle nature of random price changes. *Quantitative finance*, 4(2):176, 2003.
- [17] Jean-Philippe Bouchaud, J Doyne Farmer, and Fabrizio Lillo. How markets slowly digest changes in supply and demand. In *Handbook of financial markets: dynamics and evolution*, pages 57–160. Elsevier, 2009.
- [18] Hans Buehler, Blanka Horvath, Terry Lyons, Imanol Perez Arribas, and Ben Wood. Generating financial markets with signatures. *Available at SSRN 3657366*, 2020.
- [19] Alvaro Cartea, Imanol Perez Arribas, and Leandro Sanchez-Betancourt. Optimal execution of foreign securities: A double-execution problem, 2020. URL <https://ssrn.com/abstract=3562251>.
- [20] Kuo-Tsai Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Annals of Mathematics*, 65(1):163–178, 1957.
- [21] Ilya Chevyrev and Andrey Kormilitzin. A primer on the signature method in machine learning. *arXiv preprint arXiv:1603.03788*, 2016.
- [22] Fabienne Comte and Eric Renault. Long memory in continuous-time stochastic volatility models. *Mathematical finance*, 8(4):291–323, 1998.
- [23] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001. doi: 10.1080/713665670. URL <https://doi.org/10.1080/713665670>.
- [24] John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985. ISSN 00129682, 14680262. URL <http://www.jstor.org/stable/1911242>.
- [25] Christa Cuchiero and Josef Teichmann. Markovian lifts of positive semidefinite affine Volterra-type processes. *Decisions in Economics and Finance*, 42:407–448, 2019.
- [26] Christa Cuchiero and Josef Teichmann. Generalized feller processes and markovian lifts of stochastic Volterra processes: the affine case. *Journal of evolution equations*, 20(4):1301–1348, 2020.
- [27] Christa Cuchiero, Guido Gazzani, and Sara Svaluto-Ferro. Signature-based models: theory and calibration. *arXiv preprint arXiv:2207.13136*, 2022.
- [28] Christa Cuchiero, Guido Gazzani, Janka Möller, and Sara Svaluto-Ferro. Joint calibration to SPX and VIX options with signature-based models. *arXiv preprint arXiv:2301.13235*, 2023.
- [29] Christa Cuchiero, Sara Svaluto-Ferro, and Josef Teichmann. Signature SDEs from an affine and polynomial perspective. *arXiv preprint arXiv:2302.01362*, 2023.
- [30] Bruno Dupire. Arbitrage pricing with stochastic volatility, 1993. URL <https://api.semanticscholar.org/CorpusID:4870064>.



- [31] Bruno Dupire and Valentin Tissot-Daguette. Functional expansions. *arXiv preprint arXiv:2212.13628*, 2022.
- [32] Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. *Mathematical Finance*, 29(1):3–38, 2019.
- [33] Thomas Fawcett. *Problems in stochastic analysis. Connections between rough paths and noncommutative harmonic analysis*. PhD thesis, University of Oxford, 2003.
- [34] William Feller. Two singular diffusion problems. *Annals of Mathematics*, 54(1):173–182, 1951. ISSN 0003486X. URL <http://www.jstor.org/stable/1969318>.
- [35] Adeline Fermanian. Embedding and learning with signatures. *Computational Statistics & Data Analysis*, 157:107148, 2021.
- [36] Adeline Fermanian. Functional linear regression with truncated signatures. *Journal of Multivariate Analysis*, 192:105031, 2022. ISSN 0047-259X. doi: <https://doi.org/10.1016/j.jmva.2022.105031>.
- [37] Peter K. Friz and Nicolas B. Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010. doi: 10.1017/CBO9780511845079.
- [38] Peter K Friz, Jim Gatheral, and Radoš Radoičić. Forests, cumulants, martingales. *The Annals of Probability*, 50(4):1418–1445, 2022.
- [39] J. G. Gaines. The algebra of iterated stochastic integrals. *Stochastics and Stochastic Reports*, 49(3-4): 169–179, 1994. doi: 10.1080/17442509408833918. URL <https://doi.org/10.1080/17442509408833918>.
- [40] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative finance*, 18(6):933–949, 2018.
- [41] Julien Guyon and Jordan Lekeufack. Volatility is (mostly) path-dependent. *Volatility Is (Mostly) Path-Dependent (July 27, 2022)*, 2022.
- [42] Steven Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.
- [43] John Hull and Alan White. The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2):281–300, 1987. doi: <https://doi.org/10.1111/j.1540-6261.1987.tb02568.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1987.tb02568.x>.
- [44] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer New York, NY, 1988. URL <https://doi.org/10.1007/978-1-4684-0302-2>.
- [45] Alan Lewis. A simple option formula for general jump-diffusion and other exponential Levy processes, 2001. URL <https://EconPapers.repec.org/RePEc:vsv:svpubs:explevy>.
- [46] Terry Lyons. Rough paths, signatures and the modelling of functions on streams. *arXiv preprint arXiv:1405.4537*, 2014.
- [47] Terry Lyons and Nicolas Victoir. Cubature on Wiener space. *Proceedings: Mathematical, Physical and Engineering Sciences*, 460(2041):169–198, 2004. ISSN 13645021. URL <http://www.jstor.org/stable/4143098>.
- [48] Terry Lyons, Sina Nejad, and Imanol Perez Arribas. Non-parametric pricing and hedging of exotic derivatives. *Applied Mathematical Finance*, 27(6):457–494, 2020. doi: 10.1080/1350486X.2021.1891555. URL <https://doi.org/10.1080/1350486X.2021.1891555>.
- [49] Terry Lyons, Hao Ni, and Jiajie Tao. A pde approach for solving the characteristic function of the generalised signature process. *arXiv preprint arXiv:2401.02393*, 2024.
- [50] Benoit B Mandelbrot and John W Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437, 1968.

- [51] Imanol Perez Arribas, Guy M Goodwin, John R Geddes, Terry Lyons, and Kate EA Saunders. A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder. *Translational psychiatry*, 8(1):274, 2018.
- [52] Eckhard Platen. A non-linear stochastic volatility model. *Financial Mathematics Research Report No. FMRR 005-97*, Center for Financial Mathematics, Australian National University, Canberra, 1997.
- [53] Rimhak Ree. Lie elements and an algebra associated with shuffles. *Annals of Mathematics*, 68(2):210–220, 1958. ISSN 0003486X. URL <http://www.jstor.org/stable/1970243>.
- [54] Martin Schmelzle. Option pricing formulae using Fourier transform: Theory and application, 04 2010.
- [55] Rainer Schöbel and Jianwei Zhu. Stochastic volatility with an Ornstein–Uhlenbeck process: an extension. *Review of Finance*, 3(1):23–46, 1999.
- [56] Klaus Schürger. *Laplace Transforms and Suprema of Stochastic Processes*, pages 285–294. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002. ISBN 978-3-662-04790-3. doi: 10.1007/978-3-662-04790-3\_15.
- [57] Martin Schweizer. A guided tour through quadratic hedging approaches. SFB 373 Discussion Papers 1999,96, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, 1999. URL <https://EconPapers.repec.org/RePEc:zbw:sfb373:199996>.
- [58] Elias M. Stein and Jeremy C. Stein. Stock price distributions with stochastic volatility: An analytic approach. *The Review of Financial Studies*, 4(4):727–752, 1991. ISSN 08939454, 14657368.
- [59] Rainer Storn and Kenneth Price. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359, Dec 1997. doi: 10.1023/A:1008202821328. URL <https://doi.org/10.1023/A:1008202821328>.
- [60] R. L. Stratonovich. A new representation for stochastic integrals and equations. *SIAM Journal on Control*, 4(2):362–371, 1966. doi: 10.1137/0304028. URL <https://doi.org/10.1137/0304028>.