Pseudo rough vol-of-vol through Markovian approximation

Henrique Guerreiro *
hguerreiro@iseg.ulisboa.pt

João Guerra †
jguerra@iseg.ulisboa.pt

ISEG - School of Economics and Management, Universidade de Lisboa
REM - Research in Economics and Mathematics, CEMAPRE
Rua do Quelhas 6, 1200-781 Lisboa, Portugal

February 10, 2024

Abstract

We discuss a possible framework for a (pseudo) rough vol-of-vol model through a multi-factor Markovian approximation of the vol-of-vol process. We identify a key martingale condition which may allow to express the VIX in terms of the solution of a certain Riccati ordinary differential equation. We derive this equation and provide sufficient conditions for the existence of solutions. We also provide some partial results regarding the martingale condition. In particular, we verify a local martingale condition.

Keywords — VIX, rough volatility, Markovian approximation, exponentially affine process

---

*Supported by FCT Grant SFRH/BD/147161/2019.

†Partially supported by the project CEMAPRE/REM-UiDB/05069/2020 - financed by FCT/MCTES through national funds.

Contents

1 Introduction

The academic community has shown a great deal of interest in rough volatility modeling since the groundbreaking studies of [Gatheral et al., 2018] and [Bayer et al., 2016]. The log-volatility in rough volatility models behaves as a fractional Brownian motion with a Hurst exponent $H < 1/2$. Because of this, compared to conventional Brownian motion models, the sample trajectories are rougher in the sense that they have a lower Hölder continuity exponent. One of the first rough volatility models proposed, the rBergomi model, adjusts very well to the S&P 500 index and requires a small number of parameters. Empirical evidence of rough volatility is discussed in detail in [Livieri et al., 2018] and [Bennedsen et al., 2021]. The micro-structural basis for rough volatility is provided by [El Euch et al., 2018] and asymptotic results are discussed by [Alòs et al., 2007] and [Fukasawa, 2011].

The rBergomi model's inability to generate anything except flat smiles for the Chicago Board Options Exchange Volatility Index (VIX) options market is one of its main disadvantages since the the empirical smiles associated to VIX are nonflat. This topic can be approached in a variety of ways (see [Alòs et al., 2022]). Similar to the Black-Scholes model, in which volatility is constant and flat S&P 500 smiles are observed, the rBergomi model exhibits flat VIX smiles caused by a constant volatility of volatility (vol-of-vol). Therefore, adding stochastic vol-of-vol to generate nonflat VIX smiles is a natural way to proceed. The Stochastic Volterra models (SVM) introduced in [Horvath et al., 2020], are a natural way to generalize the rBergomi model (see [Bayer et al., 2016]) considering a stochastic vol-of-vol. In these models, the log-variance is a truncated Brownian semi-stationary (TBSS) process (see [Barndorff-Nielsen and Schmiegel, 2009] and [Bennedsen et al., 2017]) with stochastic vol-of-vol. The rBergomi model corresponds to the particular case of a constant vol-of-vol. In [Horvath et al., 2020], SVMs are studied under the assumption that the Brownian motion that drives the volatility is Markovian and independent from the vol-of-vol process. These assumptions allow the authors to derive a semi-closed expression for the CBOE Volatility Index (VIX). Such expression can then be used to simplify the numerical simulation of the VIX, leading to computational feasible option pricing routines.

An extended framework of SVM is studied in [Guerreiro and Guerra, 2022], which admits dependency between the vol-of-vol process and the Brownian motion that drives the volatility. To overcome the computational difficulties posed by this more general framework, a least squares Monte Carlo (LSMC) method is proposed. Typically, LSMC methods are harder to utilize in non-Markovian frameworks, as the dimension of the state variable used in the regression step is infinite. However, exploring the structure of the infinite dimensional state variable, it is possible to tailor the LSMC to this non-Markovian setting. Indeed, it is possible to decompose the state variable in such a way that the regression step is performed on a variable whose dimension depends on the vol-of-vol dynamics. By assuming that the vol-of-vol is Markovian, the dimension of this state variable is finite. In the particular example of a Cox-Ingersoll-Ross (CIR) model for the vol-of-vol considered in [Horvath et al., 2020] and [Guerreiro and Guerra, 2022], this dimension is equal to 1.

As pointed out in [Guerreiro and Guerra, 2022], the SVM framework is able to accommodate even a rough vol-of-vol. A rough vol-of-vol has been suggested for instance in [Da Fonseca and Zhang, 2019]. Indeed, using high-frequency data for major volatility indexes, these authors estimated the volatility of volatility. Their results suggest that its logarithm follows a fractional Brownian motion with Hurst parameter smaller than $1/2$. Therefore, not only the volatility is rough but it seems that empirical data also supports the claim that the vol-of-vol is also rough. Unfortunately, the LSMC approach of [Guerreiro and Guerra, 2022] is not well suited for this framework, as the vol-of-vol is not Markovian. In this paper, we consider the Markovian approximation of the rough CIR process, and formulate a pseudo rough vol-of-vol model. Using such Markovian approximations, we employ ODE techniques (such as the ones applied in [Horvath et al., 2020]) to obtain a semi-closed expression for the VIX.

This paper is organized as follows. In Section 2, we define the SVM and its main properties. Namely, we consider the challenge of VIX simulation and the rough vol-of-vol challenge. Then, in Section 3, we explore the general properties of Markovian approximations. Afterwards, in Section 4, we propose the pseudo rough CIR vol-of-vol model and explore its main properties. In particular, in Section 4.1, we prove the existence and uniqueness of the Markovian approximation, as well as a Feller-type condition in order to ensure that the pseudo rough CIR process remains strictly positive; in Section 4.2 we establish the link between the forward variance curve and a certain process M; in Section 4.3, we provide sufficient conditions for M to be a local martingale; finally, in Section 4.4 we provide some partial results on the true martingale property of M.

2 Stochastic Volterra models

Before introducing stochastic Volterra models, let us present briefly the basic framework for our models. We consider a market with a single risky asset. We assume there are no interest rates or dividends for the sake of simplicity. We model the price of the risky asset as a stochastic process $S = \{S_t\}_{t\geq 0}$, defined in a filtered probability space $(\Omega, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$, where the filtration satisfies the so-called usual conditions. We denote by $v = \{v_t\}_{t\geq 0}$ the variance process, which is an adapted process. Let us now present a definition of SVMs.

Definition 2.1
Let B and W be two ρ-correlated standard Brownian motions, with $\rho \in (-1, 1)$. Let $H \in (0, 1/2)$ be the Hurst parameter. We say that (S, v) follows a stochastic Volterra model if

$$dS_t = S_t \sqrt{v_t} dB_t. \tag{2.1}$$

The variance process v is given by

$$v_u = A_0(u) \exp(2X_u), \tag{2.2}$$

where A_0 is a deterministic function and X_u is a truncated Brownian semi-stationary process (TBSS) given by

$$X_u = \int_0^u \sqrt{\Gamma_s} K_\alpha(u-s)\, ds. \tag{2.3}$$

The function K_α denotes the fractional kernel $K_\alpha(x) = x^{\alpha-1}$, with $\alpha = H + 1/2$. The vol-of-vol process Γ starts at a given constant $\Gamma_0 = \gamma \geq 0$.

Now we consider the challenge of computing the VIX.

2.1 Computing the VIX

Consider the forward variance curve

$$\xi_t(u) = \mathbb{E}\left[v_u \mid \mathcal{F}_t\right]. \tag{2.4}$$

The VIX is defined from the forward variance curve as follows.

Definition 2.2
The VIX index at time t is given by

$$VIX_t = 100 \times \sqrt{\frac{1}{\Delta} \int_t^{t+\Delta} \xi_t(u) du.} = 100 \times \sqrt{\frac{1}{\Delta} \mathbb{E}\left[\int_t^{t+\Delta} v_u du \mid \mathcal{F}_t\right]}. \tag{2.5}$$

The time horizon Δ is set (by the Chicago Board Options Exchange) to be 30 days.

Fix $t > 0$. We are interested in simulating values of the random variable VIX_t. Suppose we can simulate the Brownian paths W, together with the variance process v. The VIX paths can be easily obtained from $\xi_t(u)$, with u between T and $T+\Delta$. Since $\xi_t(t) = v_t$, we may focus on the case $u > t$.

Following [Guerreiro and Guerra, 2022], we introduce the notation

$$E_{p,q}(u) = \exp\left(2\int_p^q K_\alpha(u-s)\sqrt{\Gamma_s}\,dW_s\right). \tag{2.6}$$

and define the curve

$$h_t(u) = \mathbb{E}\left[E_{t,u}(u) \mid \mathcal{F}_t\right]. \tag{2.7}$$

Proceeding as in [Guerreiro and Guerra, 2022, Proposition 3.1], we decompose the forward variance curve as follows:

$$\xi_t(u) = \frac{\xi_0(u)}{h_0(u)}E_{0,t}(u)h_t(u). \tag{2.8}$$

Now note that apart from h_t, obtaining the other variables should be straightforward. Indeed, the initial forward variance ξ_0 can be implied from the market or considered a constant/piecewise-constant (as a model parameter). The initial curve h_0 can be easily estimated using a simple Monte Carlo simulation. The process $E_{0,t}(u)$ only depends on the paths of the Brownian motion W and the vol-of-vol Γ up to time t. Since $u > t$, there is no singularity in the K_α. Thus the stochastic integral in $E_{0,t}(u)$ can be estimated using a simple approach such as a Riemann sum.

If Γ is assumed Markovian and independent of W, by conditioning on $(\Gamma_s)_{t\le s\le u}$, one obtains

$$h_t(u) = \mathbb{E}\left[\exp(2\int_t^u K_\alpha^2(u-s)\Gamma_s\,ds \mid \Gamma_t\right]. \tag{2.9}$$

By further assuming Γ is an affine process (along with some other regularity conditions), it is shown in [Horvath et al., 2020, Proposition 3] that

$$h_t(u) = \exp[\varphi(u-t) + \psi(u-t)\Gamma_t), \tag{2.10}$$

where (φ, ψ) solves a certain Riccati ODE.

2.2 Rough vol-of-vol

Note that the framework of Definition 2.1 allows in principle for the vol-of-vol process Γ to be a rough process. Working with such model can be motivated by empirical evidence such as the one presented in [Da Fonseca and Zhang, 2019], where the analysis of high-frequency data gives empirical evidence to consider a rough volatility of volatility.

Although it is possible to apply the LSMC in principle to these models, the non-Markovianity of the vol-of-vol will greatly increase the number of predictors for the regression in the LSMC. In this paper, we explore the application of considering a multi-factor Markovian approximation of the vol-of-vol process. The advantage of such approximations is that they can be able to produce processes which possess some of the rough-like features of (truly) rough processes, whilst remaining analytically tractable. The next section contains a brief overview of these Markovian approximation techniques.

3 Markovian approximations

In [Abi Jaber, 2019], the authors apply a multi-factor approximation to the rough Heston model, which functions as a bridge between the classical Heston model (one factor) and the rough Heston model (the limiting case of an infinite number of factors). In [Abi Jaber and El Euch, 2019], the authors apply these techniques to a wider class of rough volatility models. In [Bayer and Breneis, 2023], the authors obtain a super-polynomial rate of convergence by employing a multi-factor Markovian approximation where the choice of nodes is made according to a Gaussian quadrature rule.

Let us briefly consider the intuition behind the approximation discussed in [Alfonsi and Kebaier, 2021]. Suppose we are interested in approximating the solution of the stochastic rough Volterra equation

$$X_t = x_0 + \int_0^t K(t-s)b(X_s)ds + \int_0^t K(t-s)\sigma(X_s)dW_s, \tag{3.1}$$

where $x_0 \in \mathbb{R}$, $b : \mathbb{R} \to \mathbb{R}, \sigma : \mathbb{R} \to \mathbb{R}^+$ are globally Lipschitz functions, W is a standard Brownian motion (sBm), and the kernel $K \in L^2_{loc}(\mathbb{R}^+)$. One approximation strategy would be to replace the potentially problematic kernel K by a more manageable one. Recall that a function f is said to be completely monotone on $\mathbb{R}^+$ if it is smooth and $(-1)^n f^{(n)}(t) \geq 0$ for all n. This definition covers in particular the fractional kernel and any constant positive kernel. Any completely monotone kernel admits a representation of the form

$$K(t) = \int_0^{+\infty} e^{-xt} d\lambda(x)\,,\, t > 0 \tag{3.2}$$

for some measure λ on $\mathbb{R}^+$ (see for instance [Widder, 1941]).

Now we may approximate the measure λ by a weighted sum of Dirac measures. This results in

$$\tilde{K}(t) = \sum_{i=1}^{n} w_i e^{-c_i t} \tag{3.3}$$

for some weights $w_1, ..., w_n \in \mathbb{R}^+$ and points $c_1, ..., c_n \in \mathbb{R}_0^+$. By replacing K with $\tilde{K}$ in (3.1) we obtain

$$\tilde{X}_t = x_0 + \int_0^t \tilde{K}(t-s) b(X_s) ds + \int_0^t \tilde{K}(t-s)\sigma(X_s) dW_s. \tag{3.4}$$

It happens that the solution to (3.4) can be obtained by solving an n-dimensional (Markovian) SDE, as the following result (adapted from [Alfonsi and Kebaier, 2021]) explains.

Proposition 3.1
Let $n \in \mathbb{N}_1$. Let $x_0 \in \mathbb{R}$ be the initial condition and $w_1, ..., w_n \in \mathbb{R}^+$ be the weights. Suppose $m_1, ..., m_n \in \mathbb{R}$ are such that $\sum_i w_i m_i = x_0$. Then the solution of (3.4) is given by

$$X_t = \sum_{i=1}^{n} w_i X_t^i, \tag{3.5}$$

where $(X_1, ..., X_n)$ solves the n-dimensional SDE

$$\begin{aligned} dX_t^i &= c_i(m_i - X_t^i)dt + b\left(\sum_{j=1}^{n} w_j X_j\right) dt + \sigma\left(\sum_{j=1}^{n} w_j X_j\right) dW_t, \\ X_0^i &= m_i. \end{aligned} \tag{3.6}$$

A key property of this Markovian approximation is the fact that the L^2 error between the random variables X_T and $\tilde{X}_T$ is related to the L^2-error between the kernels K and $\tilde{K}$, as the following result (adapted from [Alfonsi and Kebaier, 2021]) shows.

Proposition 3.2
Let $T > 0$. Suppose the hypothesis of Proposition 3.1 are satisfied. Then there exists a constant depending on $T, |x_0|$, b and σ such that

$$\mathbb{E}\left[|\tilde{X}_T - X_T|^2\right] \leq C \int_0^T |K(t) - \tilde{K}(t)|^2 dt. \tag{3.7}$$

Furthermore, it is possible to obtain a super-polynonimal rate of convergence for the error bound using a Gaussian quadrature rule (see [Bayer and Breneis, 2023, Theorem 2.1] and [Bayer and Breneis, 2023, Theorem 2.14]).

4 A rough CIR vol-of-vol

Let us consider a stochastic Volterra model (as in Definition 2.1) and model the vol-of-vol process Γ as the n-factor Markovian approximation of a rough CIR process.

The so-called rough CIR process satisfies the stochastic Volterra equation

$$X_t = \int_0^t K_\alpha(t-s)\kappa(\mu - X_s)ds + \int_0^t K_\alpha(t-s)\nu\sqrt{X_s}dZ_s, \tag{4.1}$$

where κ, μ, ν are positive constants, Z is a sBm and $\alpha \in (1/2, 1)$.

An n-factor Markovian approximation of X is given by

$$\Gamma_s = \sum_{i=1}^n w_i X_s^i, \tag{4.2}$$

where

$$\begin{aligned} dX^i &= c_i(m_i - X^i)ds + \kappa(\mu - \sum_{j=1}^n w_j X^j)ds + \nu\sqrt{\sum w_j X^j}dZ, \\ X_0^i &= m_i, \end{aligned} \tag{4.3}$$

and

$$\sum_{i=1}^n w_i m_i = \Gamma_0 > 0. \tag{4.4}$$

4.1 Results for non-Lipschitz affine diffusion coefficient

It is important to note that $\sigma(x) = \nu\sqrt{x}$ is not a Lipschitz function. Thus, the results of the previous section do not immediately apply. Moreover, it is not even clear if existence and uniqueness applies to (4.3). Fortunately, as shown in [Abi Jaber, 2019], the Markovian approximation still has desirable properties even in this case. First, we have an existence and uniqueness result.

Proposition 4.1
Let $n \in \mathbb{N}_1$. Suppose

$$\zeta := \sum_{i=1}^n w_i(c_i m_i + \kappa\mu) > 0. \tag{4.5}$$

Then the SDE (4.3) has a unique continuous strong solution such that $\Gamma_t = \sum w_i X_t^i \geq 0$ for all $t \geq 0$ a.s..

Proof. Consider the linear curve

$$g_0(t) = \Gamma_0 + \alpha t. \tag{4.6}$$

Since g_0 is continuous, non-decreasing and such that $g_0(0) \geq 0$, we can use [Abi Jaber, 2019, Theorem A.1] to conclude there exists a unique strong solution to the SDE

$$\begin{aligned} V_t &= g_0(t) + \sum_{i=1}^{n} w_i U_t^i, \\ dU_t^i &= -c_i U_t^i dt - \kappa V_t dt + \nu \sqrt{V_t} dZ_t, \\ U_0^i &= 0. \end{aligned} \tag{4.7}$$

The result follows by noting that solutions to (4.3) and (4.7) are in one-to-one correspondence through the relation

$$X_t^i = m_i + t(c_i m_i + \kappa\mu) + U_t^i. \tag{4.8}$$

□

Remark 1. Note that the condition (4.5) is not very restrictive. Indeed, the only condition on m_i (see Proposition 3.1) is that $\sum w_i m_i = \Gamma_0$. Since w_i and Γ_0 are positive, it is always possible to find m_i that are non-negative.

Since Γ represents a vol-of-vol, we would like to obtain a Feller type of condition so that the process remains strictly positive. We will use the following result of [Duffie and Kan, 1996].

Theorem 4.2
Let $n \in \mathbb{N}_1$. Let $\underset{\sim}{W}$ be an n-dimensional sBm, $a, \Sigma \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^n$. Define the functions

$$v_i(x) = \alpha_i + \beta_i \cdot x, \tag{4.9}$$

where $\alpha_i \in \mathbb{R}$ and $\beta_i \in \mathbb{R}^n$. Moreover, suppose that the following conditions are verified

$$\text{For all } x \text{ such that } v_i(x) = 0, \beta_i^T(ax+b) > \frac{1}{2}\beta_i^T \Sigma\Sigma^T \beta_i \tag{C1}$$

and

$$\text{For all } j, \text{ if } (\beta_i^T \Sigma)_j \neq 0, \text{ then } v_i = v_j. \tag{C2}$$

Then the SDE

$$dX = (aX+b)dt + \Sigma \operatorname{diag}(\sqrt{v_1(X}, ..., \sqrt{v_n(X)})\, d\underset{\sim}{W} \tag{4.10}$$

has a unique continuous strong solution for any initial condition x_0 such that $v_i(x_0) > 0$ for all $i = 1, ..., n$. Moreover, for any $t \geq 0$ and any i, $v_i(X_t) > 0$ a.s..

The following proposition is an easy consequence of the above theorem.

Proposition 4.3
Suppose there exists j such that $\zeta > \frac{1}{2}\nu^2 w_j^2$. Then, for all $t \geq 0$, $\Gamma_s > 0$ a.s.

Proof. Note that (4.3) is in the form of Theorem 4.2. The matrix Σ and functions v_i are given by

$$\Sigma_{ij} = \begin{cases} \nu & , i = j \\ 0 & , i \neq j \end{cases}. \tag{4.11}$$

and

$$v_i(x) = \begin{cases} \sum_{i=1}^n w_i x_i & , i = j \\ 1 & , i \geq j \end{cases}. \tag{4.12}$$

For (C1), it is vacuously true for $i \neq j$. Now note that

$$\beta_j^T(ax + b) = \sum w_i(c_i m_i + \kappa\mu - c_i x_i - \kappa \sum w_i x_i), \tag{4.13}$$

which is equal to

$$\sum w_i(c_i m_i + \kappa\mu) = \zeta \tag{4.14}$$

if $0 = v_j(x) = \sum w_i x_i$. Condition (C1) follows by noting that

$$\beta_i^T \Sigma \Sigma^T \beta_i = \nu^2 w_j^2. \tag{4.15}$$

Condition (C2) is trivially satisfied for $i \neq j$ since $\beta_i = 0, i \neq j$. For $i = j$, we obtain the vector $\beta_j^T \Sigma$, which is equal to $(0, ...0, w_j, 0..., 0)$. Thus, $(\beta_j^T \Sigma)_k = 0$ implies $k = j = i$, as required. Finally, $v_i > 0$ by construction for $i \neq j$ and Γ_0 ensures $v_j(X_0) > 0$. □

Recall that results like Proposition 3.2 and [Bayer and Breneis, 2023, Theorems 2.1, 2.14] relate the approximation error of the rough kernel with error measures between the target rough process and its Markovian approximation. These kinds of results are also available for the rough CIR process. Indeed, see [Abi Jaber, 2019, Theorem 4.1], [Abi Jaber, 2019, Proposition 4.3], [Bayer and Breneis, 2023, Corollary 3.7] and [Bayer and Breneis, 2023, Corollary 3.8].

4.2 The forward variance curve

Recall the conditional expectation h_t defined in (2.7) and its connection to the VIX described in (2.8). Suppose that Z is independent of W so that Γ is independent of W. Then, we have

$$h_t(u) = \mathbb{E}\left[\mathbb{E}\left[\exp\left(2\int_t^u \sqrt{\Gamma_s}K(u-s)dW_s\right)\Bigg|\mathcal{F}_t \vee (\Gamma_s)_{t\le s\le u}\right]\Bigg|\mathcal{F}_t\right] \tag{4.16}$$

$$= \mathbb{E}\left[\exp\left(2\int_t^u \Gamma_s K^2(u-s)\,ds\right)\Bigg|\mathcal{F}_t\right]. \tag{4.17}$$

By virtue of $(X_1, ..., X_n)$ being Markovian, the above conditional expectation has to be a function of $(X^1, ..., X^n)$. We wish to investigate if it is actually exponentially affine. To this end, we generalize the approach of [Horvath et al., 2020, Section 4]. Fix $u > t$. We will build a process of the form

$$M_s = \exp\left(2\int_t^s \Gamma_r K^2(u-r)\,dr\right)\exp\left(\varphi(u-s) + \sum_{i=1}^n w_i\psi_i(u-s)X_s^i\right) \tag{4.18}$$

for $t \le s \le u$. In Section 4.4, we will see that if M is a true martingale, we have an exponentially affine formula for h_t. The functions $(\varphi, \psi_1, ..., \psi_n)$ will be constructed so that M is a local martingale. To achieve this, we apply Itô's formula and obtain a certain n-dimensional ODE that makes the drift term vanish.

4.3 Local martingale property of M

Consider the function

$$f(s,x) = f(s,y,x_1,...,x_n) = \exp\left(y + \varphi(u-s) + \sum_{i=1}^n w_i\psi_i(u-s)x_i\right), \tag{4.19}$$

and note that

$$M_s = f(s, Y_s, X_s^1, ..., X_s^n), \tag{4.20}$$

where

$$Y_s = 2\int_t^s \Gamma_r K^2(u-r)\,dr. \tag{4.21}$$

We will apply Itô's formula to M. Let us compute the partial derivatives of f:

$$\frac{\partial f}{\partial s}(s,y,x) = f(s,y,x)\left(-\varphi'(u-s) - \sum_{i=1}^n w_i\psi_i'(u-s)x_i\right) \tag{4.22}$$

and

$$\frac{\partial f}{\partial x_i}(s,y,x) = f(s,y,x)w_i\psi_i(u-s). \tag{4.23}$$

Hence

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(s,y,x) = f(s,y,x)w_iw_j\psi_i(u-s)\psi_j(u-s). \tag{4.24}$$

Moreover

$$\frac{\partial f}{\partial y} = f(s, y, x). \tag{4.25}$$

By Itô's formula, with $\underset{\sim}{X} = (X_1, ..., X_n)$,

$$\begin{aligned} df(s, Y, \underset{\sim}{X})/f(s, Y, \underset{\sim}{X}) = &\left(-\varphi'(u-s) - \sum_{i=1}^{n} w_j \psi_j'(u-s) x_i\right) ds \\ &+ dY \\ &+ \sum_{i=1}^{n} w_i \psi_i(u-s) dX^i \\ &+ \frac{1}{2} \sum_{i,j \le N} w_i w_j \psi_i(u-s) \psi_j(u-s) d\langle X^i, X^j \rangle. \end{aligned}$$

Now by (4.3), for any $1 \le i, j \le n$ we have,

$$d\langle X^i, X^j \rangle = \nu^2 \sum_{k=1}^{n} w_k X^k ds. \tag{4.26}$$

Also

$$dY_s = 2\Gamma_s K^2(u-s) ds. \tag{4.27}$$

Thus

$$\begin{aligned} df(s, Y, X)/f(s, Y, X) = &\left(-\varphi'(u-s) - \sum_{i=1}^{n} w_i \psi_i'(u-s) X^i\right) ds \\ &+ 2(\sum w_j X^j) K^2(u-s) ds \\ &+ \sum_{i=1}^{n} w_i \psi_i(u-s) c_i (m_i - X^i) ds \\ &+ \sum_{i=1}^{n} w_i \psi_i(u-s) \kappa(\mu - \sum_{j=1}^{n} w_j X_s^j) ds \\ &+ \frac{1}{2} \sum_{i,j \le N} w_i w_j \psi_i(u-s) \psi_j(u-s) \nu^2 \sum w_j X_s^j ds \\ &+ \sum_{i=1}^{n} w_i \psi_i(u-s) \nu \sqrt{\sum w_j X_s^j} dW. \end{aligned} \tag{4.28}$$

Now introduce the quadratic function

$$R(z_1, ..., z_n) = -\kappa \sum_{i=1}^{n} w_i z_i + \frac{1}{2} \nu^2 \sum_{1 \le i,j \le n} w_i w_j z_i z_j \tag{4.29}$$

and the linear function

$$F(z_1, ..., z_n) = \sum_{i=1}^{n} w_i z_i (c_i m_i + \kappa\mu). \tag{4.30}$$

We can re-arrange (4.28) and use (4.2) to obtain

$$\frac{dM_s}{M_s} = A_s ds + \left[\nu \sum_{i=1}^{n} w_i \psi_i (u-s) \sqrt{\sum_{j=1}^{n} w_j X_s^j} \right] dW_s, \tag{4.31}$$

where

$$\begin{aligned} A_s = &\sum_{i=1}^{n} w_i X^i \left(-\psi_i'(u-s) - c_i \psi_i(u-s) \right) & (4.32)\\ &+ \Gamma_s([R(\psi_1, ..., \psi_n)](u-s) + 2K^2(u-s)) & (4.33)\\ &+ [F(\psi_1, ..., \psi_n)](u-s) - \varphi'(u-s). & (4.34) \end{aligned}$$

Thus, to ensure $A \equiv 0$ we first solve the n-dimensional ODE

$$\begin{aligned} \psi_i' &= -c_i \psi_i + R(\psi_1, ..., \psi_n) + 2K^2 \\ \psi_i(0) &= 0 \end{aligned}, i = 1, .., n, \tag{4.35}$$

and then we find φ by solving

$$\begin{aligned} \varphi' &= F(\psi_1, ... \psi_n), \\ \varphi(0) &= 0. \end{aligned} . \tag{4.36}$$

To avoid the singularity of K at $t = 0$, we can modify ODE (4.35) by considering the auxiliary functions

$$y_i(t) = \psi_i(t) - 2G(t), \tag{4.37}$$

where

$$G(t) = \int_0^t K^2(s)\, ds. \tag{4.38}$$

Then ODE (4.35) becomes

$$\begin{aligned} y_i' &= -c_i(y_i + 2G) + R((y_1 + 2G, ..., y_n + 2G)) = -c_i \psi_i + R(\psi_1, ..., \psi_n), \\ y_i(0) &= 0. \end{aligned} . \tag{4.39}$$

Thus, if y solves ODE (4.39) and φ solves ODE (4.36), we know that M is a local martingale. That is, we obtain the following result.

Proposition 4.4
Let $0 < t \leq u$ be fixed time points. Let $(X_1, ..., X_n)$ solve SDE (4.3) and let Γ be defined as in (4.2). Suppose $(y_1, ..., y_n)$ solves ODE (4.39). Define the functions ψ_i by $\psi_i = y_i + 2G$ and suppose φ solves ODE (4.36). Define the process M by (4.18) using these functions $(\varphi, \psi_1, ..., \psi_n)$. Then, M is a local martingale in $[t, u]$.

Remark 2. Technically, the random variables and functions mentioned in the previous proposition depend on the choice of t and u, but we omit sub and superscripts to alleviate the notation.

Now we present sufficient conditions for the ODEs (4.39) and (4.36) to have a solution. The result and proof are based on [Horvath et al., 2020, Proposition 3] but adapted to the n-dimensional case.

Proposition 4.5
Define the functions R_i as

$$R_i(x_1, ...x_n) = -c_i x_i + R(x_1, ...x_n). \tag{4.40}$$

Let $\Lambda > 0$ be a fixed time horizon. Assume there exists a vector $(a_1, .., a_n)$ with non-negative entries such that for all $i = 1, ..., n$ we have

$$2G(\Lambda) + \Lambda \max_Q R_i \leq a_i, \tag{4.41}$$

where Q is the 2^n vertices set

$$Q = \{x \in \mathbb{R}^n \mid x_i \in \{0, a_i\}, i = 1, ..., n\}\,. \tag{4.42}$$

Then (4.39) and (4.36) have solutions on $[0, \Lambda]$. Moreover, we have the bounds

$$0 \leq \psi_i := y_i + 2G \leq a_i \tag{4.43}$$

and

$$0 \leq \varphi \leq \Lambda \max_Q F. \tag{4.44}$$

Proof. First note that using the function R_i, equation (4.39) can be simply written as

$$y_i' = R_i(y_1 + 2G, ..., y_n + 2G). \tag{4.45}$$

We will start by considering a truncated version of the equation to obtain a solution. Then we will use the bound in the assumption to prove this solution solves the original equation. To simplify notation, let us define the truncation function as

$$T_a(x) = \operatorname{sign}(x) \min(|x|, a). \tag{4.46}$$

The truncated equation is defined as

$$y_i' = R_i(T_{a_1}(y_1 + 2G), ..., T_{a_n}(y_n + 2G)). \tag{4.47}$$

The right-hand side is Lipschitz in y and continuous in t on $[0, \Lambda]$. Thus, by the Picard-Lindelöf theorem, it has a unique solution y^a on $[0, \Lambda]$. Now we prove that this solution is actually a solution to (4.39) on $[0, \Lambda]$.

We start by noting that

$$\psi_i^a := y_i^a + 2G \geq 0, i = 1, ...n. \tag{4.48}$$

To arrive at a contradiction, let us suppose ψ_i^a is negative at some time in $[0, \Lambda]$. That is

$$A := \{s \geq 0 \mid \psi_i^a(s) < 0\} \neq \varnothing. \tag{4.49}$$

Let $\tau = \inf A$. Since ψ_i^a is continuous and $\psi_i(0) = 0$, it follows that $\psi_i^a(\tau) = 0$. Now note that, since $R(0, ..., 0) = 0$, we have,

$$\frac{d}{dt}\psi_i^a(\tau) = R_i(\psi_1^a(\tau), ..., \psi_n^a(\tau)) + 2G'(\tau) = R(0, ..., 0) + 2K^2(\tau) > 0, \tag{4.50}$$

where we consider also the case $K^2(0) = +\infty$. But then $\psi_i^a > 0$ in $(\tau, \tau+\varepsilon]$ for some $\varepsilon > 0$. Hence $\tau + \varepsilon > \tau = \inf A$ is a lower bound of A, which is a contradiction.

Now note that the Hessian matrix of R_i is

$$H_{k,j}^i = \nu^2 w_k w_j. \tag{4.51}$$

Moreover, for any vector $x \in \mathbb{R}^n$,

$$x^T H^i x = \sum_{1 \leq k,j \leq n} w_k w_j x_i x_j = \left(\sum_{k=1}^n w_k x_k\right)^2 \geq 0. \tag{4.52}$$

Thus, H^i is positive semi-definite and hence R_i is convex. Since the domain

$$D = [0, a_1] \times [0, a_2] \times ... \times [0, a_n]$$

is the convex hull of the vertices set Q, the maximum of the convex function R_i in D is attained at one of the vertices in Q. Thus,

$$\max_D R_i = \max_Q R_i. \tag{4.53}$$

By (4.48), for any $s \in [0, \Lambda]$,

$$[T_{a_1}(y_1^a + 2G), ..., T_{a_n}(y_n^a + 2G)](s) \in D. \tag{4.54}$$

Hence

$$R_i(T_{a_1}(y_1^a + 2G), ..., T_{a_n}(y_n^a + 2G)) \leq \max_Q R_i. \tag{4.55}$$

Since $y_i^a(0) = 0$, it easily follows that

$$\max_{[0,\Lambda]} y_i^a(t) \leq \max_{[0,\Lambda]} (y_i^a)' \tag{4.56}$$

$$= \Lambda \max_{[0,\Lambda]} R_i(T_{a_1}(y_1^a + 2G), ..., T_{a_n}(y_n^a + 2G)) \tag{4.57}$$

$$\leq \Lambda \max_Q R_i. \tag{4.58}$$

This together with assumption (4.41) implies that

$$y_i^a + 2G \leq \Lambda \max_Q R_i + 2G(\Lambda) \leq a_i. \tag{4.59}$$

Thus, the function y_i^a actually solves the original equation (4.39) on $[0, \Lambda]$ and (4.43) is verified.

Finally, since F is linear, (4.36) has a unique solution on $[0, \Lambda]$. Since $F(0, ..., 0) = 0$, by a similar argument to what we made above, $\varphi \geq 0$ in $[0, \Lambda]$. For the upper bound, we simply note that

$$\varphi \leq \Lambda \max_{[0,\Lambda]} F(\psi_1, ..., \psi_n). \tag{4.60}$$

Again, using (4.43) and applying a similar argument as we did for R_i, we arrive at the conclusion the maximum of F is attained at the set of vertices Q and thus (4.44) follows.

□

4.4 True martingale property of M

Proposition 4.4 gives us conditions for M to be a local martingale. But if M is actually a *true martingale*, we have an exponentially affine formula for h_t. Such formula allows for efficient VIX pricing (recall Section 2.1).

Proposition 4.6
Suppose M is a true martingale. Then

$$h_t(u) = \exp\left(\varphi(u-t) + \sum_{i=1}^n w_i \psi_i(u-t) X_t^i\right). \tag{4.61}$$

Proof. Note that since for $i = 1, ..., n$ we have $\phi(0) = \psi_i(0) = 0, i = 1, ..., n$. It follows that

$$h_t(u) = \mathbb{E}\left[M_u \mid \mathcal{F}_t\right]. \tag{4.62}$$

Then, since M is a true martingale,

$$\mathbb{E}\left[M_u \mid \mathcal{F}_t\right] = M_t = \exp\left(\varphi(u-t) + \sum_{i=1}^{n} w_i \psi_i(u-t) X_t^i\right). \tag{4.63}$$

□

The process M is in fact the stochastic exponential of a local martingale. Indeed consider the process

$$Y_s = \int_t^s \Psi_u(r) d\Gamma_r^c, \tag{4.64}$$

where

$$d\Gamma_s^c = \nu \sqrt{\Gamma_s} dZ_s \tag{4.65}$$

and

$$\Psi_u(s) = \sum_{i=1}^{n} w_i \psi_i(u-s). \tag{4.66}$$

Since

$$\frac{dM_s}{M_s} = \Psi_u(s) d\Gamma_s^c = dY_s, \tag{4.67}$$

we have

$$\mathcal{E}(Y) = M. \tag{4.68}$$

In [Horvath et al., 2020], where $n = 1$, the authors apply Theorem 3.1 of [Kallsen and Muhle-Karbe, 2010] to the pair (Y, Γ) to prove that $M = \mathcal{E}(Y)$ is a true martingale. The analogous way of proceeding for the general case $n \geq 1$ would be to apply the same theorem to the process $(Y, X_1, ..., X_n)$. Note that Γ is not in general an affine process when $n > 1$ but the n-dimensional process $(X_1, ..., X_n)$ is. Unfortunately, the strong admissibility assumption of [Kallsen and Muhle-Karbe, 2010, Theorem 3.1] fails.

Since we are working with processes with no jumps, we provide a simplified definition of strong admissibility ([Kallsen and Muhle-Karbe, 2010, Definition 2.4]) for this case.

Definition 4.1
Let $d \in \mathbb{N}_1$. Lévy-Khintchine triplets $(\beta_j, \gamma_j, \phi_j)$ are called strongly admissible in the case $\phi_j \equiv 0$ if there exists $m \in \mathbb{N}_0, m \leq d$ such that for any $t \in \mathbb{R}^+$

1.
$$\beta_j^k(t) \geq 0 \text{ for } 0 \leq j \leq m, 1 \leq k \leq m, k \neq j; \tag{4.69}$$

2.
$$\gamma_j^{k,l} = 0 \text{ for } 0 \leq j \leq m, 1 \leq k, l \leq m \text{ unless } k = l = j; \tag{4.70}$$

3.
$$\beta_j^k(t) = 0 \text{ for } j \geq m+1, 1 \leq k \leq m; \tag{4.71}$$

4.
$$\gamma_j(t) = 0 \text{ for } j \geq m+1; \tag{4.72}$$

Moreover,

5. β_j, γ_j are continuous in $\mathbb{R}^+$ for $0 \leq j \leq d$.

Lemma 4.7
Let $t \leq u$. Let Y be as in (4.64). The process $U = (X_1, ..., X_n, Y)$ is a $\mathbb{R}^{n+1}$-valued time-inhomogeneous semi-martingale with affine characteristics relative to Lévy-Khintchine triplets $(\beta_j(\cdot), \gamma_j(\cdot), \phi_j(\cdot)), 1 \leq j \leq n+1$.

The triplets are given as follows. First

$$\beta_0^k = \begin{cases} c_k m_k + \kappa\mu & , 1 \leq k \leq n \\ 0 & , k = n+1 \end{cases}. \tag{4.73}$$

Then, for $1 \leq j \leq n$,

$$\beta_j^k = \begin{cases} -\kappa w_j & , 1 \leq k \leq n, k \neq j \\ -\kappa w_j - c_k & , k = j \\ 0 & , k = n+1 \end{cases}. \tag{4.74}$$

and

$$\beta_{n+1} \equiv 0. \tag{4.75}$$

Also

$$\phi_j \equiv 0. \tag{4.76}$$

Finally

$$\gamma_0 \equiv \gamma_{n+1} \equiv 0, \tag{4.77}$$

$$\gamma_j^k(s) = \begin{cases} \widehat{\gamma}_j(s) & , 1 \leq k \leq n \\ \Psi_u(s)\widehat{\gamma}_j(s) & , k = n+1 \end{cases}, \tag{4.78}$$

where $\gamma_j^k(s)$ denotes the k-th row of $\gamma_j(s)$ and $\widehat{\gamma}_j$ is the $\mathbb{R}^{1\times d}$ valued function

$$\widehat{\gamma}_j(s) = \nu^2 w_j[1, ..., 1, \Psi_u(s)]. \tag{4.79}$$

Furthermore, these triplets are strongly admissible if and only if $n = 1$.

Proof. Let $d = n + 1$. By construction, the process U is a $\mathbb{R}^d$-valued affine semimartingale. Let us compute the Lévy-Khintchine triplets. Consider the function

$$p(x) = p(x_1, ..., x_n) = \nu\sqrt{\sum_{j=1}^{n} x_j}. \tag{4.80}$$

Note that

$$dX^i = D^i(\underset{\sim}{X})dt + p(\underset{\sim}{X})dZ, i = 1, ..., n. \tag{4.81}$$

where $D^i(\underset{\sim}{X})$ is the drift coefficient in (4.3). Moreover,

$$dY_s = \Psi_u(s)p(\underset{\sim}{X})dZ_s. \tag{4.82}$$

Consider now the matrix valued function $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ given by

$$\sigma(x, y, s) = \begin{cases} \hat{\sigma}(x, y) & , 1 \leq i \leq n \\ \hat{\sigma}(x, y)\Psi_u(s) & , i = n + 1 \end{cases}, \tag{4.83}$$

where $\sigma_i(x, y, s)$ denotes the i-th row of $\sigma(x, y, s)$ and the $\hat{\sigma}$ is the $\mathbb{R}^{1\times d}$ valued function

$$\hat{\sigma}(x, y) = [0, ..., 0, p(x)]. \tag{4.84}$$

Then

$$dU_s = R(U_s)ds + \sigma(U, s)d\underset{\sim}{B}_s, \tag{4.85}$$

where $\underset{\sim}{B} = (B^1, ..., B^d)$ is a d-dimensional sBm with $B^d = Z$ and

$$R_i(x, y) = \begin{cases} D^i(x, y) & , 1 \leq i \leq n \\ 0 & , i = n + 1 \end{cases} = \beta_0^i + \sum_{j=1}^{n} x_j\beta_j^i + y\beta_{n+1}^i, \tag{4.86}$$

with β_j as in (4.73), (4.74) and (4.75). Consider

$$a(x, y, s) := \sigma(x, y, s)^T\sigma(x, y, s). \tag{4.87}$$

Then

$$a_i(x, y, s) = \begin{cases} \hat{a}(x, y)\Psi_u(s) & , i = 1 \\ \hat{a}(x, y) & , i \geq 2 \end{cases}, \tag{4.88}$$

where $a_i(x, y, s)$ denotes the i-th row of $a(x, y, s)$ and $\hat{a}$ is the $\mathbb{R}^{1\times d}$ valued function

$$\hat{a}(x, y, s) = p^2(x)[\Psi_u(s), 1, ..., 1]. \tag{4.89}$$

So

$$a(x, y, s) = \gamma_0(s) + \sum_{j=1}^{n} x_j \gamma_j(s) + \gamma_{n+1}(s) y, \tag{4.90}$$

with γ_j as in (4.77) and (4.78).

Now we verify the admissibility conditions.

Note that the matrices γ_j for $1 \leq j \leq n$ do not have any non-zero entries. Thus, the only choice of m that allows to satisfy (4.72) is $m = n$. So let us choose $m = n$. Since again the matrices γ_j for $1 \leq j \leq n$ do not have any non-zero entries, condition (4.70) is satisfied if and only if it is the empty condition. This occurs only for $n = 1$. Thus, the triplets are not admissible for $n > 1$. To see they are admissible for $n = 1$, we only need to verify the other three conditions. Condition (4.69) reduces to $\beta_0^1 \geq 0$, i.e. $c_1 m_1 + k\mu \geq 0$, which is satisfied since $m_1 = X_0^1 = \Gamma_0 / w_1 \geq 0$ and $c_1, \kappa, \mu \geq 0$. Condition (4.71) is satisfied by virtue of (4.75). Finally, (4.1) is satisfied as the functions β_j are constant and the functions γ_j are continuous. □

Remark 3. If we try to use the process $(Y, X_1, ..., X_n)$ instead, we will obtain $\gamma_1 = 0$ but all the entries of matrices $\gamma_2, ..., \gamma_{n+1}$ will be non-zero. Thus, condition (4.72) will fail unless $m = d$. But if we choose $m = d$, condition (4.70) will fail, as for instance $\gamma_{n+1}^{1,1}$ will not be zero. Similar effects happen for different reshufflings.

Under the no jump assumption $\phi_j \equiv 0$, all conditions in [Kallsen and Muhle-Karbe, 2010, Theorem 3.1] are trivially satisfied except for the third one. We state the simplified version of this theorem under the no jump assumption.

Theorem 4.8
Let $d \in \mathbb{N}_1$ and X be an $\mathbb{R}^d$-valued semimartingale with affine characteristics relative to strongly admissible Lévy-Khintchine triplets $(\beta_j, \gamma_j, \phi_j)$, with $\phi_j \equiv 0$. Suppose that for some $1 \leq i \leq d$ and $T \in \mathbb{R}^+$ the following condition holds

$$\beta_j^i(t) = 0 \; \forall j = 0, ..., d, t \in [0, T]. \tag{4.91}$$

Then the stopped process $\mathcal{E}(X^i)^T$ is a (true) martingale.

Remark 4. Condition (4.91) is satisfied for the process $(X_1, ..., X_n, Y)$ with $i = n+1$ as a consequence of Eqs. (4.73) to (4.75).

5 Conclusions

The Markovian approximation of non-Markovian Volterra processes is a powerful tool which allows the treatment of otherwise too complex processes. It allows to tackle models which are in general analytically untractable and often also numerically too costly. There are multiple techniques to construct such approximations which have been recently presented in the literature.

One application of Markovian approximations is to allow for a process mimicking rough vol-of-vol. We study this approximation for the rough CIR case. The Markovian approximation is written as a weighted sum of mean reverting processes driven by a common Brownian motion. With VIX pricing in mind, we explore the problem of finding an expression for the conditional expectation of the exponential stochastic Volterra process. To do this, we build a process M and a Riccati ODE. If the process M is a true martingale, the aforementioned conditional expectation can be written in terms of the solution of the Riccati ODE. This in turn allows for efficient VIX simulations and pricing. We provide sufficient conditions for the existence of solutions for the Riccati ODE, and prove that M is a local martingale.

The question of whether M is a true martingale is left for future research. We still offer a brief treatment of it in this paper. We explore an approach which is known to work for the non-rough CIR process. This approach involves a strong admissibility on the Lévy-Khintchine triplets of a certain associated process. We arrive at the conclusion that in general the Lévy-Khintchine triplets are not strongly admissible, except when the Markovian approximation is made of a single component.

References

[Abi Jaber, 2019] Abi Jaber, E. (2019). Lifting the Heston model. *Quantitative Finance*, 19(12):1995–2013.

[Abi Jaber and El Euch, 2019] Abi Jaber, E. and El Euch, O. (2019). Multifactor approximation of rough volatility models. *SIAM Journal on Financial Mathematics*, 10(2):309–349.

[Alfonsi and Kebaier, 2021] Alfonsi, A. and Kebaier, A. (2021). Approximation of stochastic Volterra equations with kernels of completely monotone type. *arXiv preprint arXiv:2102.13505*.

[Alòs et al., 2022] Alòs, E., García-Lorite, D., and Gonzalez, A. M. (2022). On smile properties of volatility derivatives: Understanding the VIX skew. *SIAM Journal on Financial Mathematics*, 13(1):32–69.

[Alòs et al., 2007] Alòs, E., León, J. A., and Vives, J. (2007). On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, 11(4):571–589.

[Barndorff-Nielsen and Schmiegel, 2009] Barndorff-Nielsen, O. E. and Schmiegel, J. (2009). Brownian semistationary processes and volatility/intermittency. In Albrecher, H., Runggaldier, W. J., and Schachermayer, W., editors, *Advanced Financial Modelling*, volume 8 of *Radon Series on Applied and Computational Mathematics*, pages 1–26. De Gruyter.

[Bayer and Breneis, 2023] Bayer, C. and Breneis, S. (2023). Markovian approximations of stochastic Volterra equations with the fractional kernel. *Quantitative Finance*, 23(1):53–70.

[Bayer et al., 2016] Bayer, C., Friz, P., and Gatheral, J. (2016). Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904.

[Bennedsen et al., 2021] Bennedsen, M., Lunde, A., , and Pakkanen, M. S. (2021). Decoupling the short- and long-term behavior of stochastic volatility. *Journal of Financial Econometrics*, 20(5):961–1006.

[Bennedsen et al., 2017] Bennedsen, M., Lunde, A., and Pakkanen, M. S. (2017). Hybrid scheme for Brownian semistationary processes. *Finance and Stochastics*, 21(4):931–965.

[Da Fonseca and Zhang, 2019] Da Fonseca, J. and Zhang, W. (2019). Volatility of volatility is (also) rough. *Journal of Futures Markets*, 39(5):600–611.

[Duffie and Kan, 1996] Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406.

[El Euch et al., 2018] El Euch, O., Fukasawa, M., and Rosenbaum, M. (2018). The microstructural foundations of leverage effect and rough volatility. *Finance and Stochastics*, 22(2):241–280.

[Fukasawa, 2011] Fukasawa, M. (2011). Asymptotic analysis for stochastic volatility: martingale expansion. *Finance and Stochastics*, 15(4):635–654.

[Gatheral et al., 2018] Gatheral, J., Jaisson, T., and Rosenbaum, M. (2018). Volatility is rough. *Quantitative Finance*, 18(6):933–949.

[Guerreiro and Guerra, 2022] Guerreiro, H. and Guerra, J. (2022). Least squares Monte Carlo methods in stochastic Volterra rough volatility models. *Journal of Computational Finance*, 26(3):73–101.

[Horvath et al., 2020] Horvath, B., Jacquier, A., and Tankov, P. (2020). Volatility options in rough volatility models. *SIAM Journal on Financial Mathematics*, 11(2):437–469.

[Kallsen and Muhle-Karbe, 2010] Kallsen, J. and Muhle-Karbe, J. (2010). Exponentially affine martingales, affine measure changes and exponential moments of affine processes. *Stochastic Processes and their Applications*, 120(2):163–181.

[Livieri et al., 2018] Livieri, G., Mouti, S., Pallavicini, A., and Rosenbaum, M. (2018). Rough volatility: Evidence from option prices. *IISE Transactions*, 50(9):767–776.

[Widder, 1941] Widder, D. V. (1941). *The Laplace Transform*, volume 6 of *Princeton Mathematical Series*. Princeton University Press, Princeton, N.J.