The characteristic function of Gaussian stochastic volatility models: an analytic expression

Eduardo Abi Jaber*

September 24, 2020

Abstract

Stochastic volatility models based on Gaussian processes, like fractional Brownian motion, are able to reproduce important stylized facts of financial markets such as rich autocorrelation structures, persistence and roughness of sample paths. This is made possible by virtue of the flexibility introduced in the choice of the covariance function of the Gaussian process. The price to pay is that, in general, such models are no longer Markovian nor semimartingales, which limits their practical use. We derive, in two different ways, an explicit analytic expression for the joint characteristic function of the log-price and its integrated variance in general Gaussian stochastic volatility models. Such analytic expression can be approximated by closed form matrix expressions stemming from Wishart distributions. This opens the door to fast approximation of the joint density and pricing of derivatives on both the stock and its realized variance using Fourier inversion techniques. In the context of rough volatility modeling, our results apply to the (rough) fractional Stein–Stein model and provide the first analytic formulae for option pricing known to date, generalizing that of Stein–Stein, Schöbel–Zhu and a special case of Heston.

Keywords: Gaussian processes, Volterra processes, non-Markovian Stein–Stein/Schöbel-Zhu models, rough volatility.

1 Introduction

In the realm of risk management in mathematical finance, academics and practitioners have been always striving for explicit solutions to option prices and hedging strategies in their models. Undoubtedly, finding explicit expressions to a theoretical problem can be highly satisfying in itself; it also has many practical advantages such as: reducing computational time (compared to brute force Monte-Carlo simulations for instance); achieving a higher precision for option prices and hedging strategies; providing a better understanding of the

*Université Paris 1 Panthéon-Sorbonne, Centre d'Economie de la Sorbonne, 106, Boulevard de l'Hôpital, 75013 Paris, eduardo.abi-jaber@univ-paris1.fr.

role of the parameters of the model and the sensitivities of the prices and strategies with respect to them. As one would expect, explicit expressions usually come at the expense of sacrificing the flexibility and the accuracy of the model. In a nutshell, the aim of the present paper is to show that analytic expressions for option prices can be found in a highly flexible class of non-Markovian stochastic volatility models.

From Black-Scholes to rough volatility

In their seminal paper, Black and Scholes (1973) derived closed form solutions for the prices of European call and put options in the geometric Brownian motion model where the dynamics of the stock price S are given by:

$$dS_t = S_t \sigma dB_t, \quad S_0 > 0, \tag{1.1}$$

with B a standard Brownian motion and σ the constant instantaneous volatility parameter. Although revolutionary, the model remains very simple: it drifts away from the reality of financial markets characterized by non-Gaussian returns, fat tails of stock prices and their volatilities, asymmetric option prices (i.e. the implied volatility smile and skew)... see Cont (2001). Since then a large and growing literature has been developed to refine the Black and Scholes (1973) model. One notable direction is stochastic volatility modeling where the constant volatility σ in (1.1) is replaced by a Markovian stochastic process $(\sigma_t)_{t\geq 0}$. In their celebrated paper, Stein and Stein (1991) modeled $(\sigma_t)_{t\geq 0}$ by a mean-reverting Brownian motion of the form

$$d\sigma_t = \kappa(\theta - \sigma_t)dt + \nu dW_t, \tag{1.2}$$

where W is a standard Brownian motion independent of B. Remarkably, they obtained closed-form expressions for the characteristic function of the log-price, which allowed them to recover the density as well as option prices by Fourier inversion of the characteristic function. Later on the model has been extended by Schöbel and Zhu (1999) to account for the leverage effect, i.e. an arbitrary correlation between W and B. Similar formulas for the characteristic function of the log–price to those of Stein–Stein are derived for the non-zero correlation case.

Prior to the extension by Schöbel and Zhu (1999), Heston (1993) took a slightly different approach to include the leverage effect by introducing a model deeply rooted in the Stein–Stein model. Heston observed that the instantenous variance process $V_t = \sigma_t^2$ in the Stein–Stein model with $\theta = 0$ follows a CIR process thanks to Itô's formula,[1] so that the Stein–Stein model can be recast in the following form

$$\begin{aligned} dS_t &= S_t \sqrt{V_t} dB_t, \\ dV_t &= (\nu^2 - 2\kappa V_t)dt + 2\nu\sqrt{V_t} dW_t, \end{aligned} \tag{1.3}$$

[1] squares of Brownian motion constitute the building blocks of squared Bessel processes, see Revuz and Yor (1999, Chapter XI).

where $B = \rho W + \sqrt{1-\rho^2} W^{\perp}$ with $\rho \in [-1, 1]$ and $W^{\perp}$ a Brownian motion independent of W. Such model remains tractable as it was shown earlier in the context of bond pricing with uncertain inflation by Cox et al. (1985, Equations (51)-(52)).[2] Heston (1993) carried on by deriving closed form expressions for the characteristic function of the log–price, which made his model one of the most, if not the most, popular model among practitioners. As one would expect, the expressions of Heston (1993) and Schöbel and Zhu (1999) share a lot of similarities and they perfectly agree when $\theta = 0$ in (1.2), see Lord and Kahl (2006, equation (44)). Such analytical tractability motivated the development of the theory of finite-dimensional Markovian affine processes, see Duffie et al. (2003).

Unfortunately, Markovian stochastic volatility models, such as the Heston and the Stein–Stein models, are not flexible enough: they generate an auto-correlation structure which is too simplistic compared to empirical observations. Indeed, several empirical studies have documented the persistence in the volatility time series, see Andersen and Bollerslev (1997); Ding et al. (1993). More recently, Gatheral et al. (2018) and Bennedsen et al. (2016) show that the sample paths of the realized volatility are rougher than standard Brownian motion at any realistic time scale as illustrated on Figure 1-(a). From a pricing perspective, Markovian models fail to reproduce the power-law decay of the at-the-money skew of option prices as shown on Figure 1-(b), see Alòs et al. (2007); Bayer et al. (2016); Fukasawa (2011).

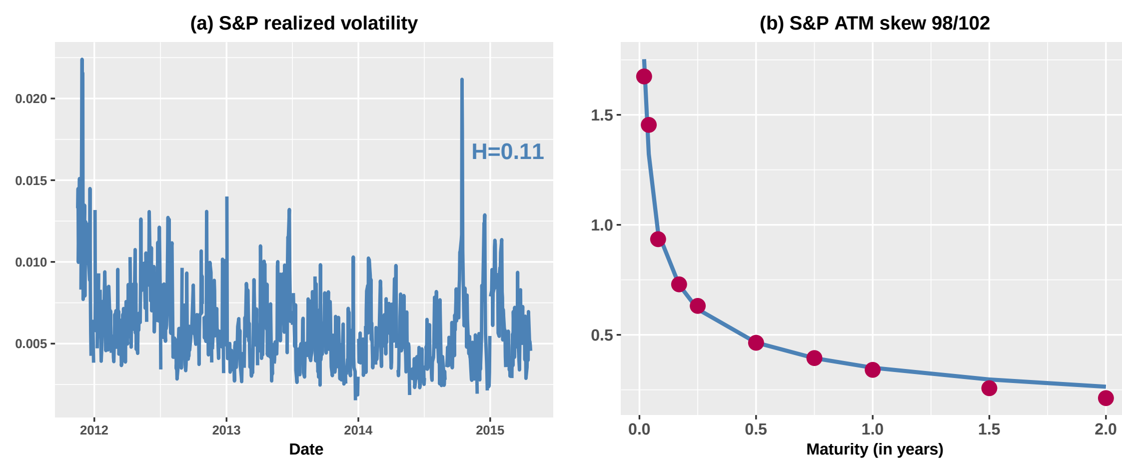


Figure 1: (a) Realized volatility of the S&P with an estimated Hurst index of $\hat{H} = 0.11$. (b) Term structure of the at-the-money skew of the implied volatility $\frac{\partial \sigma_{\text{implicit}}(k,T)}{\partial k}\Big|_{k=0}$ for the S&P index on June 20, 2018 (red dots) and a power-law fit $t \to 0.35 \times t^{-0.41}$. Here $k := \ln(K/S_0)$ stands for the log-moneyness and T for the time to maturity.

[2]The long-term level of the variance ν^2 in (1.3) can be replaced by a more general coefficient $\theta \geq 0$.

These studies have motivated the need to enhance conventional stochastic volatility models with richer auto-correlation structures. This has been initiated in Comte and Renault (1998) by replacing the driving Brownian motion of the volatility process by a fractional Brownian motion W^H:

$$W_t^H = \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} dW_s + \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^0 ((t-s)^{H-1/2} - (-s)^{H-1/2}) dW_s$$

where $H \in (0,1)$ is the Hurst exponent: $H > 1/2$ corresponds to positively correlated returns, $H < 1/2$ to negatively correlated increments and $H = 1/2$ reduces to the case of standard Brownian motion. Sample paths of W^H are locally Hölder continuous of any order strictly less than H, thereby less regular than standard Brownian motion. Initially Comte and Renault (1998) considered the case $H > 1/2$. However, a smaller Hurst index $H \approx 0.1$ allows to match exactly the regularity of the volatility time series and the exponent in the power–law decay of the at-the-money skew measured on the market (Figure 1). Consequently models involving the fractional kernel $t \mapsto t^{H-1/2}$ with $H < 1/2$ have been dubbed "rough volatility models" by Gatheral et al. (2018).

The price to pay is that, in general, such models are no longer Markovian nor semi-martingales, which limits their practical use and make their mathematical analysis quite challenging. This has initiated a thriving branch of research.[3] The need for fast pricing in such non-Markovian models is therefore, more than ever, crucial. One breakthrough in that direction was achieved by El Euch and Rosenbaum (2019) who came up with a rough version of the Heston (1993) model after convolving the dynamics (1.3) with a fractional kernel to get

$$V_t = V_0 + \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} \left((\theta - \kappa V_s) ds + \nu \sqrt{V_s} dW_s \right), \tag{1.4}$$

for $H \in (0, 1/2)$. Remarkably, they show that an analogous formula for the characteristic function of the log price to that of Heston (1993) continue to hold modulo a fractional deterministic Riccati equation. From a theoretical perspective, the rough Heston model falls into the broader class of non-Markovian affine Volterra processes developed in Abi Jaber et al. (2019); Abi Jaber (2019c), and can be recovered as a projection of infinite dimensional Markovian affine processes as illustrated in Abi Jaber and El Euch (2019a); Cuchiero and Teichmann (2020); Gatheral and Keller-Ressel (2019).

Although the rough Heston model can be efficiently implemented (Abi Jaber, 2019b; Abi Jaber and El Euch, 2019b; Callegaro et al., 2018; Gatheral and Radoičić, 2019), no closed-form solution for the fractional deterministic Riccati equation and whence for the characteristic function is known to date, which has to be contrasted with the conventional Heston (1993) model. One possible explanation could be that, unlike the Markovian case, squares of fractional Brownian motion have different dynamics than (1.4), so that the marginals of the process (1.4) are not chi-square distributed, except for the case $H = 1/2$.

[3] Refer to https://sites.google.com/site/roughvol/home for references.

The main objective of the paper is to rely on squares of general Gaussian processes with arbitrary covariance structures by considering the non-Markovian extension of the Stein and Stein (1991) and the Schöbel and Zhu (1999) models. We will show that the underlying Gaussianity makes the problem highly tractable and allows to recover analytic expressions for the joint Fourier–Laplace transform of the log price and the integrated variance in general, which would agree with that of Stein–Stein, Schbel–Zhu and Heston under the Markovian setting. Such models have been already considered several times in the context of non-Markovian and rough volatility literature (Cuchiero and Teichmann, 2019; Gulisashvili et al., 2019; Harms and Stefanovits, 2019; Horvath et al., 2019) but there has been no derivation of the analytic form of the characteristic function. Our methodology takes a step further the recent derivation in Abi Jaber (2019a) for the Laplace transform of the integrated variance and that of Abi Jaber et al. (2020) where the Laplace transform of the forward co-variance curve enters in the context of portfolio optimization.

The Gaussian Stein–Stein model and main results

For $T > 0$, we will consider the following generalized version of the Stein–Stein model:

$$dS_t = S_t X_t dB_t, \quad S_0 > 0, \tag{1.5}$$

$$X_t = g_0(t) + \int_0^T K(t,s)\kappa X_s ds + \int_0^T K(t,s)\nu dW_s, \tag{1.6}$$

with $B = \rho W + \sqrt{1-\rho^2}W^{\perp}$, $\rho \in [-1,1]$, $\kappa, \nu \in \mathbb{R}$, g_0 a suitable deterministic input curve, $K : [0,T]^2 \to \mathbb{R}$ a measurable kernel and $(W, W^{\perp})$ a two-dimensional Brownian motion.

Under mild assumptions on its covariance function, every Gaussian process can be written in the form (1.6) with $\kappa = 0$, see Sottinen and Viitasaari (2016). Such representation is known as the Fredholm representation. We will be chiefly interested in two classes of kernels K:

- Symmetric kernels, i.e. $K(t,s) = K(s,t)$ for all $s, t \leq T$, for which the integration in (1.6) goes up to time T, meaning that X is not necessarily adapted to the filtration generated by W.

- Volterra kernels, i.e. $K(t,s) = 0$ whenever $s \geq t$, for which integration in (1.6) goes up to time t, which is more in line with standard stochastic volatility modeling. For instance, the conventional mean reverting Stein-Stein model (1.2) can be recovered by setting $g_0(t) = X_0 - \kappa\theta t$, $\kappa \leq 0$ and by considering the Volterra kernel $K(t,s) = \mathbf{1}_{s<t}$. The fractional Brownian motion with a Hurst index $H \in (0,1)$ can be represented using the Volterra kernel

$$K(t,s) = \mathbf{1}_{s<t}\frac{(t-s)^{H-1/2}}{\Gamma(H+\frac{1}{2})}\,{}_2F_1\left(H-\frac{1}{2};\frac{1}{2}-H;H+\frac{1}{2};1-\frac{t}{s}\right),$$

where ${}_2F_1$ is the Gauss hypergeometric function; and the Riemman-Liouville fractional Brownian motion corresponds to the case $K(t,s) = \mathbf{1}_{s<t}(t-s)^{H-1/2}/\Gamma(H+1/2)$.

For $u, w \in \mathbb{C}$ with $\Re(u) \in [0,1]$ and $\Re(w) \le 0$, we provide the following analytical expression for the conditional joint Fourier–Laplace transform of the log-price and the integrated variance:

$$\mathbb{E}\left[\exp\left(u \log \frac{S_T}{S_t} + w \int_t^T X_s^2 ds\right) \Big| \mathcal{F}_t\right] = \frac{\exp\left(\langle g_t, \mathbf{\Psi}_t g_t \rangle_{L^2}\right)}{\det\left(\mathbf{\Phi}_t\right)^{1/2}}, \tag{1.7}$$

with $\langle f, h \rangle_{L^2} = \int_0^T f(s)h(s)ds$, det the Fredholm (1903) determinant (see Simon (1977)), g_t the adjusted conditional mean given by

$$g_t(s) = \mathbf{1}_{t \le s}\mathbb{E}\left[X_s - \int_t^T K(s,r)\kappa X_r dr \Big| \mathcal{F}_t\right]; \tag{1.8}$$

and $\mathbf{\Psi}_t$ a linear operator acting on $L^2([0,T],\mathbb{R})$ defined by

$$\mathbf{\Psi}_t = (\mathrm{id} - b\boldsymbol{K})^{-*}\, a \left(\mathrm{id} - 2a\tilde{\mathbf{\Sigma}}_t\right)^{-1} (\mathrm{id} - b\boldsymbol{K})^{-1}, \quad t \le T, \tag{1.9}$$

where $\boldsymbol{F}^{-*} := (\boldsymbol{F}^{-1})^*$, id denotes the identity operator, i.e. $(\mathrm{id} f) = f$ for all $f \in L^2([0,T],\mathbb{C})$,

$$a = w + \frac{1}{2}(u^2 - u), \quad b = \kappa + \rho\nu u, \tag{1.10}$$

and $\tilde{\mathbf{\Sigma}}_t$ the adjusted covariance integral operator defined by

$$\tilde{\mathbf{\Sigma}}_t = (\mathrm{id} - b\boldsymbol{K})^{-1}\mathbf{\Sigma}_t(\mathrm{id} - b\boldsymbol{K})^{-*}, \tag{1.11}$$

with $\mathbf{\Sigma}_t$ defined as the integral operator associated with the covariance kernel

$$\Sigma_t(s,u) = \nu^2 \int_t^T K(s,z)K(u,z)dz, \quad t \le s, u \le T, \tag{1.12}$$

and finally $\mathbf{\Phi}$ is defined by

$$\mathbf{\Phi}_t = \begin{cases} (\mathrm{id} - b\boldsymbol{K})(\mathrm{id} - 2a\tilde{\mathbf{\Sigma}}_t)(\mathrm{id} - b\boldsymbol{K}) & \text{if } K \text{ is a symmetric kernel} \\ \mathrm{id} - 2a\tilde{\mathbf{\Sigma}}_t & \text{if } K \text{ is a Volterra kernel} \end{cases}.$$

At first glance, the expressions for $\mathbf{\Phi}$ seem to depend on the class of the kernel, but they actually agree. Indeed, for Volterra kernels, i.e. $K(t,s) = 0$ for $s \ge t$, $\det(\mathrm{id} - b\boldsymbol{K}) = \det(\mathrm{id} - b\boldsymbol{K}^*) = 1$ so that using the relation (Simon, 1977, Theorem 3.8) $\det((\mathrm{id} + \boldsymbol{F})(\mathrm{id} + \boldsymbol{G})) = \det(\mathrm{id} + \boldsymbol{F})\det(\mathrm{id} + \boldsymbol{G})$:

$$\det((\mathrm{id} - b\boldsymbol{K})(\mathrm{id} - 2a\tilde{\mathbf{\Sigma}}_t)(\mathrm{id} - b\boldsymbol{K})^*) = \det(\mathrm{id} - 2a\tilde{\mathbf{\Sigma}}_t).$$

As already mentioned, we prove (1.7) for two classes of kernels:

- **Symmetric nonnegative kernels**: we provide an elementary static derivation of (1.7) for $t = 0$ and $\kappa = 0$, based on the spectral decomposition of K which leads to the decomposition of the characteristic function as an infinite product of independent Wishart distributions. The operator $\mathbf{\Psi}_0$ appears naturally after a rearrangement of the terms. The main result is collected in Theorem 2.2.

- **Volterra kernels:** we adopt a dynamical approach to derive the conditional characteristic function (1.7) via Itô's formula on the adjusted conditional mean process $(g_t(s))_{t\leq s}$. The main result is stated in Theorem 3.3.

From the numerical perspective, we will show in Section 4.1 that the expression (1.7) lends itself to approximation by closed form solutions using finite dimensional matrices after a straightforward discretization of the operators. Alternatively, we provide another approximation formula by closed form expressions stemming from Wishart characteristic functions in the form

$$\mathbb{E}\left[\exp\left(u\log\frac{S_T}{S_0}+w\int_0^T X_s^2 ds\right)\right] = \lim_{n\to\infty}\frac{\exp\left(\mu_n^\top w_n\left(I_{2n}-2R_nw_n\right)^{-1}\mu_n\right)}{\det\left(I_{2n}-2R_nw_n\right)^{1/2}},$$

where $\mu_n \in \mathbb{R}^{2n}$ and $w_n, R_n \in \mathbb{R}^{2n\times 2n}$ are entirely determined by $(g_0, K, \nu, \kappa, u, w)$ and det is the standard determinant of a matrix, we refer to Section 4.2. We illustrate the applicability of these formulas on an option pricing and calibration example by Fourier inversion techniques in a (rough) fractional Stein–Stein model in Section 4.3.

Notations

Fix $T > 0$. We let $\mathbb{K}$ denote $\mathbb{R}$ or $\mathbb{C}$. We denote by $\langle\cdot,\cdot\rangle_{L^2}$ the following product

$$\langle f, g\rangle_{L^2} = \int_0^T f(s)^\top g(s)ds, \quad f, g\in L^2\left([0,T],\mathbb{K}\right).$$

We note that $\langle\cdot,\cdot\rangle_{L^2}$ is an inner product on $L^2\left([0,T],\mathbb{R}\right)$, but not on $L^2\left([0,T],\mathbb{C}\right)$. We define $L^2\left([0,T]^2,\mathbb{K}\right)$ to be the space of measurable kernels $K : [0,T]^2 \to \mathbb{K}$ such that

$$\int_0^T\int_0^T |K(t,s)|^2 dtds < \infty.$$

For any $K, L \in L^2\left([0,T]^2,\mathbb{K}\right)$ we define the $\star$-product by

$$(K\star L)(s,u) = \int_0^T K(s,z)L(z,u)dz, \quad (s,u)\in[0,T]^2, \tag{1.13}$$

which is well-defined in $L^2\left([0,T]^2,\mathbb{K}\right)$ due to the Cauchy-Schwarz inequality. For any kernel $K \in L^2\left([0,T]^2,\mathbb{K}\right)$, we denote by $\boldsymbol{K}$ the integral operator induced by the kernel K that is

$$(\boldsymbol{K}g)(s) = \int_0^T K(s,u)g(u)du, \quad g \in L^2\left([0,T],\mathbb{K}\right).$$

$\boldsymbol{K}$ is a linear bounded operator from $L^2\left([0,T],\mathbb{K}\right)$ into itself. If $\boldsymbol{K}$ and $\boldsymbol{L}$ are two integral operators induced by the kernels K and L in $L^2\left([0,T]^2,\mathbb{K}\right)$, then $\boldsymbol{K}\boldsymbol{L}$ is the integral operator induced by the kernel $K \star L$.

When $\mathbb{K}=\mathbb{R}$, we denote by K^* the adjoint kernel of K for $\langle \cdot,\cdot\rangle_{L^2}$, that is

$$K^*(s,u) = K(u,s), \quad (s,u) \in [0,T]^2,$$

and by $\boldsymbol{K}^*$ the corresponding adjoint integral operator.

2 Symmetric kernels: an elementary proof

We provide an elementary derivation of the joint Fourier–Laplace transform in the special case of symmetric kernels with $\kappa = 0$. This will naturally lead to the analytic expression (1.7) in terms of the operator $\boldsymbol{\Psi}$ given in (1.9).

Definition 2.1. *A kernel $K \in L^2\left([0,T]^2,\mathbb{R}\right)$ is symmetric nonnegative if $K = K^*$ and*

$$\int_0^T \int_0^T f(s)^\top K(s,u)f(u)duds \geq 0, \quad \forall f \in L^2\left([0,T],\mathbb{R}\right).$$

In this case, the integral operator $\boldsymbol{K}$ is said to be symmetric nonnegative and $\boldsymbol{K} = \boldsymbol{K}^*$ and $\langle f, \boldsymbol{K}f\rangle_{L^2} \geq 0$. $\boldsymbol{K}$ is said to be symmetric nonpositive, if $(-\boldsymbol{K})$ is symmetric nonnegative.

Throughout this section, we fix $T>0$ and we consider the case of symmetric kernels having the following spectral decomposition

$$K(t,s) = \sum_{n\geq 1} \sqrt{\lambda_n} e_n(t)e_n(s), \quad t,s \leq T, \tag{2.1}$$

where $(e_n)_{n\geq 1}$ is an orthonormal basis of $L^2([0,T],\mathbb{R})$ for the inner product $\langle f,g\rangle_{L^2} = \int_0^T f(s)g(s)ds$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ with $\lambda_n \to 0$, as $n \to \infty$, such that

$$\sum_{n\geq 1} \lambda_n < \infty.$$

Such decomposition is possible whenever the operator $\boldsymbol{K}$ is the (nonnegative symmetric) square-root of a covariance operator $\boldsymbol{C}$ which is generated by a continuous kernel. This is known as Mercer's theorem, see Shorack and Wellner (2009, Theorem 1, p.208) and leads to the so-called Kac–Siegert/Karhunen–Loève representation of the process X, see Kac and Siegert (1947); Karhunen (1946); Loeve (1955). In this case, one can show that any square-integrable Gaussian process X with mean g_0 and covariance $\boldsymbol{C}$ admits the representation (1.6) with $\kappa = 0$ on some filtered probability space supporting a Brownian motion W, see Sottinen and Viitasaari (2016).

We state our main result of the section on the representation of the characteristic function for symmetric kernels.

Theorem 2.2. *Let K be as in (2.1), $g_0 \in L^2([0,T],\mathbb{R})$ and set $\kappa = 0$. Fix $u, w \in \mathbb{C}$ such that $\Re(u) = 0$ and $\Re(w) \leq 0$. Then,*

$$\mathbb{E}\left[\exp\left(u \log \frac{S_T}{S_0} + w \int_0^T X_s^2 ds\right)\right] = \frac{\exp\left(\langle g_0, \boldsymbol{\Psi}_0 g_0\rangle_{L^2}\right)}{\det\left(\boldsymbol{\Phi}_0\right)^{1/2}}, \tag{2.2}$$

with $\boldsymbol{\Psi}_0$ and $\tilde{\boldsymbol{\Sigma}}_0$ respectively given by (1.9) and (1.11), for (a,b) as in (1.10) (with $\kappa = 0$), that is

$$a = w + \frac{1}{2}(u^2 - u), \quad b = \rho\nu u,$$

and $\Phi_0 = (\mathrm{id} - b\boldsymbol{K})(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_0)(\mathrm{id} - b\boldsymbol{K})$.

The rest of the section is dedicated to the proof of Theorem 2.2. The key idea is to rely on the spectral decomposition (2.1) to decompose the characteristic function as an infinite product of independent Wishart distributions. The operators $\tilde{\boldsymbol{\Sigma}}_0$ and $\boldsymbol{\Psi}_0$ will then appear naturally after a rearrangement of the terms.

In the sequel, to ease notations, we drop the subscript L^2 in the product $\langle \cdot, \cdot \rangle_{L^2}$. We will start by computing the joint Fourier–Laplace transform of $\left(\int_0^T X_s^2 ds, \int_0^T X_s dW_s\right)$. To this end, since $g_0 \in L^2([0,T],\mathbb{R})$, we can write $g_0 = \sum_{n\geq 1}\langle g_0, e_n\rangle e_n$. Recalling that $\kappa = 0$, and making use of (2.1) we get that

$$X_t = g_0(t) + \int_0^T K(t,s)\nu dW_s = \sum_{n\geq 1}\left(\langle g_0, e_n\rangle + \sqrt{\lambda_n}\nu\xi_n\right) e_n(t), \tag{2.3}$$

where $\xi_n = \int_0^T e_n(s) dW_s$, for each $n \geq 1$. Since $(e_n)_{n\geq 1}$ is an orthonormal family in L^2, $(\xi_n)_{n\geq 1}$ is a sequence of independent standard Gaussian random variables. Furthermore, the representation (2.3) readily leads to

$$\int_0^T X_s^2 ds = \sum_{n\geq 1}\left(\langle g_0, e_n\rangle + \sqrt{\lambda_n}\nu\xi_n\right)^2, \tag{2.4}$$

$$\int_0^T X_s dW_s = \sum_{n\geq 1}\left(\langle g_0, e_n\rangle + \sqrt{\lambda_n}\nu\xi_n\right)\xi_n. \tag{2.5}$$

Lemma 2.3. *Let K be as in (2.1), $g_0 \in L^2([0,T],\mathbb{R})$, set $\kappa = 0$ and fix $\alpha, \beta \in \mathbb{C}$ such that*

$$\Re(\alpha) \leq 0, \quad \Re(\beta) = 0. \tag{2.6}$$

Then,

$$\mathbb{E}\left[\exp\left(\alpha \int_0^T X_s^2 ds + \beta \int_0^T X_s dW_s\right)\right] = \frac{\exp\left(\left(\alpha + \frac{\beta^2}{2}\right)\sum_{n\geq 1} \frac{\langle g_0, e_n\rangle^2}{1-2\beta\nu\sqrt{\lambda_n}-2\alpha\nu^2\lambda_n}\right)}{\prod_{n\geq 1}\sqrt{1-2\beta\nu\sqrt{\lambda_n}-2\alpha\nu^2\lambda_n}} \tag{2.7}$$

Proof. Define $U_T = \alpha \int_0^T X_s^2 ds + \beta \int_0^T X_s dW_s$. We first observe that (2.6) yields that $|\exp(U_T)| = \exp(\Re(U_T)) \leq 1$, so that $\mathbb{E}[\exp(U_T)]$ is finite. By virtue of the representations (2.4)-(2.5), we have

$$U_T = \sum_{n\geq 1} \alpha\tilde{\xi}_n^2 + \beta\tilde{\xi}_n\xi_n,$$

where $\tilde{\xi}_n = \left(\langle g_0, e_n\rangle + \nu\sqrt{\lambda_n}\xi_n\right)$, for each $n \geq 1$. Setting $Y_n = (\tilde{\xi}_n, \xi_n)^\top$, it follows that $(Y_n)_{n\geq 1}$ are independent such that each Y_n is a two dimensional Gaussian vector with mean μ_n and covariance matrix Σ_n given by

$$\mu_n = \begin{pmatrix} \langle g_0, e_n\rangle \\ 0 \end{pmatrix} \quad \text{and} \quad \Sigma_n = \begin{pmatrix} \nu^2\lambda_n & \nu\sqrt{\lambda_n} \\ \nu\sqrt{\lambda_n} & 1 \end{pmatrix}.$$

Furthermore, we have

$$U_T = \sum_{n\geq 1} Y_n^\top w_n Y_n,$$

with

$$w_n = \begin{pmatrix} \alpha & \frac{\beta}{2} \\ \frac{\beta}{2} & 0 \end{pmatrix}.$$

By successively using the independence of Y_n and the well-known expression for the characteristic function of the Wishart distribution, see for instance Abi Jaber (2019a, Proposition A.1), we get

$$\begin{aligned}
\mathbb{E}[\exp(U_T)] &= \mathbb{E}\left[\exp\left(\sum_{n\geq 1} Y_n^\top w_n Y_n\right)\right] \\
&= \prod_{n\geq 1} \mathbb{E}\left[\exp\left(Y_n^\top w_n Y_n\right)\right] \\
&= \prod_{n\geq 1} \frac{\exp\left(\operatorname{tr}\left(w_n\left(I_2 - 2\Sigma_n w_n\right)^{-1}\mu_n\mu_n^\top\right)\right)}{\det\left(I_2 - 2\Sigma_n w_n\right)^{1/2}}.
\end{aligned}$$

We now compute the right hand side. We have

$$(I_2 - 2\Sigma_n w_n) = \begin{pmatrix} 1 - 2\alpha\nu^2\lambda_n - \beta\nu\sqrt{\lambda_n} & -\beta\nu^2\lambda_n \\ -2\alpha\nu\sqrt{\lambda_n} - \beta & 1 - \beta\nu\sqrt{\lambda_n} \end{pmatrix}$$

so that

$$\det(I_2 - 2\Sigma_n w_n) = 1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n$$

and

$$(I_2 - 2\Sigma_n w_n)^{-1} = \frac{1}{1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n} \begin{pmatrix} 1 - \beta\nu\sqrt{\lambda_n} & \beta\nu^2\lambda_n \\ 2\alpha\nu\sqrt{\lambda_n} + \beta & 1 - 2\alpha\nu^2\lambda_n - \beta\nu\sqrt{\lambda_n} \end{pmatrix}.$$

Straightforward computations lead to the claimed expression (2.7). □

Relying on the spectral decomposition (2.1), we re-express the quantities entering in (2.7) in terms of suitable operators.

Lemma 2.4. *Let K be as in (2.1), set $\kappa = 0$ and fix $\alpha, \beta \in \mathbb{C}$ as in (2.6). Then, the following operator defined by (1.9) with $a = \alpha + \frac{\beta^2}{2}$ and $b = \nu\beta$:*

$$\mathbf{\Psi}_0^{\alpha,\beta} = (\mathrm{id} - b\boldsymbol{K})^{-*}\, a \left(\mathrm{id} - 2\tilde{\mathbf{\Sigma}}_0 a\right)^{-1} (\mathrm{id} - b\boldsymbol{K})^{-1}, \quad t \le T,$$

admits the following decomposition

$$\mathbf{\Psi}_0^{\alpha,\beta} = \sum_{n\ge1} \frac{\alpha + \frac{\beta^2}{2}}{1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n} \langle e_n, \cdot\rangle e_n \tag{2.8}$$

and

$$\det\left(\frac{1}{\alpha + \frac{\beta^2}{2}} \mathbf{\Psi}_0^{\alpha,\beta}\right) = \prod_{n\ge1} \frac{1}{1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n}, \tag{2.9}$$

with the convention that $0/0 = 1$. In particular,

$$\mathbb{E}\left[\exp\left(\alpha\int_0^T X_s^2 ds + \beta\int_0^T X_s dW_s\right)\right] = \det\left(\frac{1}{\alpha + \frac{\beta^2}{2}} \mathbf{\Psi}_0^{\alpha,\beta}\right)^{1/2} \exp\left(\langle g_0, \mathbf{\Psi}_0^{\alpha,\beta} g_0\rangle\right) \tag{2.10}$$

Proof. It follows from (2.1) that

$$(\mathrm{id} - b\boldsymbol{K}) = \sum_{n\ge1} \left(1 - b\sqrt{\lambda_n}\right) \langle e_n, \cdot\rangle e_n.$$

Since $\Re(\beta) = 0$, $\Re(1 - b\sqrt{\lambda_n}) = 1 \neq 0$ fo each $n \geq 1$, so that $(\mathrm{id} - b\boldsymbol{K})$ is invertible with an inverse given by

$$(\mathrm{id} - b\boldsymbol{K})^{-1} = \sum_{n\geq 1} \frac{1}{1 - b\sqrt{\lambda_n}} \langle e_n, \cdot \rangle e_n. \tag{2.11}$$

Similarly, recalling (1.12), (2.1) leads to the representation of $\boldsymbol{\Sigma}_0 = \nu^2 \boldsymbol{K}\boldsymbol{K}^*$:

$$\boldsymbol{\Sigma}_0 = \sum_{n\geq 1} \nu^2 \lambda_n \langle e_n, \cdot \rangle e_n,$$

so that $\tilde{\boldsymbol{\Sigma}}_0$ given by (1.11) reads

$$\tilde{\boldsymbol{\Sigma}}_0 = \sum_{n\geq 1} \frac{\nu^2 \lambda_n}{\left(1 - b\sqrt{\lambda_n}\right)^2} \langle e_n, \cdot \rangle e_n.$$

Whence,

$$\left(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_0\right) = \sum_{n\geq 1} \frac{\left(1 - b\sqrt{\lambda_n}\right)^2 - 2a\nu^2\lambda_n}{\left(1 - b\sqrt{\lambda_n}\right)^2} \langle e_n, \cdot \rangle e_n.$$

Recalling that $a = \alpha + \frac{\beta^2}{2}$ and $b = \nu\beta$, $\left(\left(1 - b\sqrt{\lambda_n}\right)^2 - 2a\nu^2\lambda_n\right) = 1 - 2\nu\beta\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n$. Since $\Re(\alpha) \leq 0$ and $\Re(\beta) = 0$, we have that $\Re(1 - 2\nu\beta\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n) > 0$ so that $\left(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_0\right)$ is invertible with an inverse given by

$$\left(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_0\right)^{-1} = \sum_{n\geq 1} \frac{\left(1 - \nu\beta\sqrt{\lambda_n}\right)^2}{1 - 2\nu\beta\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n} \langle e_n, \cdot \rangle e_n.$$

The representations (2.8)-(2.9) readily follows after composing by $(\mathrm{id} - b\boldsymbol{K})^{-*}a$ from the left, by $(\mathrm{id} - b\boldsymbol{K})^{-*}$ from the right and recalling (2.11). Finally, combining these expressions with (2.7), we obtain (2.10). This ends the proof. $\square$

We can now complete the proof of Theorem 2.2.

Proof of Theorem 2.2. It suffices to prove that

$$\mathbb{E}\left[\exp\left(u \log \frac{S_T}{S_0} + w \int_0^T X_s^2 ds\right)\right] = \mathbb{E}\left[\exp\left(\alpha \int_0^T X_s^2 ds + \beta \int_0^T X_s dW_s\right)\right], \tag{2.12}$$

where

$$\alpha = w + \frac{1}{2}(u^2 - u) - \frac{\rho^2 u^2}{2} \quad \text{and} \quad \beta = \rho u.$$

Indeed, if this the case, then

$$\Re(\alpha) = \Re(w) + \frac{1}{2}(\rho^2 - 1)\Im(u)^2 \leq 0,$$

so that an application of Lemma 2.4 yields the expression (2.2).
It remains to prove (2.12) by means of a projection argument. For this, we recall that $B = \rho W + \sqrt{1-\rho^2} W^{\perp}$ and we write

$$\log S_T = \log S_0 - \frac{1}{2}\int_0^T X_s^2 ds + \rho \int_0^T X_s dW_s + \sqrt{1-\rho^2}\int_0^T X_s dW_s^{\perp}. \qquad (2.13)$$

Denoting by $\mathcal{F}^X$ the filtration generated by $\{X_s : s \leq T\}$, we observe that

$$M_T := \mathbb{E}\left[\exp\left(u\sqrt{1-\rho^2}\int_0^T X_s dW_s^{\perp}\right) \Big| \mathcal{F}^X\right] = \exp\left(\frac{u^2(1-\rho^2)}{2}\int_0^T X_s^2 ds\right) \qquad (2.14)$$

so that, using successively the tower property of the conditional expectation, expression (2.13) and the fact that X and W are $\mathcal{F}^X$-measurable, we obtain

$$\begin{aligned}
\mathbb{E}\left[\exp\left(u \log\frac{S_T}{S_0} + w\int_0^T X_s^2 ds\right)\right] &= \mathbb{E}\left[\mathbb{E}\left[\exp\left(u \log\frac{S_T}{S_0} + w\int_0^T X_s^2 ds\right) \Big| \mathcal{F}^X\right]\right] \\
&= \mathbb{E}\left[\exp\left(\left(w - \frac{u}{2}\right)\int_0^T X_s^2 ds + \rho u \int_0^T X_s dW_s\right) M_T\right]
\end{aligned}$$

leading to (2.12) due to (2.14). This ends the proof. □

3 Volterra kernels: a dynamical approach

In this section, we will consider the class of Volterra kernels of continuous and bounded type in L^2 in the terminology of Gripenberg et al. (1990, Definitions 9.2.1, 9.5.1 and 9.5.2).

Definition 3.1. *A kernel $K : [0,T]^2 \to \mathbb{R}$ is a Volterra kernel of continuous and bounded type in L^2 if $K(t,s) = 0$ whenever $s \geq t$ and*

$$\sup_{t\in[0,T]}\int_0^T |K(t,s)|^2 ds < \infty, \quad \lim_{h\to 0}\int_0^T |K(u+h,s) - K(u,s)|^2 ds = 0, \quad u \leq T. \qquad (3.1)$$

The following kernels are of continuous and bounded type in L^2.

Example 3.2. (i) Any convolution kernel of the form $K(t,s) = k(t-s)\mathbf{1}_{s<t}$ with $k \in L^2([0,T],\mathbb{R})$.

(ii) For $H \in (0,1)$,

$$K(t,s) = \mathbf{1}_{s<t} \frac{(t-s)^{H-1/2}}{\Gamma(H+\frac{1}{2})} \, {}_2F_1\left(H-\frac{1}{2}; \frac{1}{2}-H; H+\frac{1}{2}; 1-\frac{t}{s}\right),$$

where ${}_2F_1$ is the Gauss hypergeometric function. Such kernel enters in the Volterra representation (1.6) of the the fractional Brownian motion whose covariance function is $\Sigma_0(s,u) = \frac{1}{2}(s^{2H} + u^{2H} - |s-u|^{2H})$, see Decreusefond and Ustunel (1999).

(iii) Continuous kernels K on $[0,T]^2$. This is the case for instance for the Brownian Bridge W^{T_1} conditioned to be equal to $W_0^{T_1}$ at a time T_1: for all $T < T_1$, W^{T_1} admits the Volterra representation (1.6) on $[0,T]$ with the continuous kernel $K(t,s) = \mathbf{1}_{s<t}(T_1 - t)/(T_1 - s)$, for all $s, t \leq T$.

(iv) If K_1 an K_2 satisfy (3.1) then so does $K_1 \star K_2$ by an application of Cauchy-Schwarz inequality.

Throughout this section, we fix a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\leq T}, \mathbb{Q})$ supporting a two dimensional Brownian motion $(W, W^\perp)$ and we set $B = \rho W + \sqrt{1-\rho^2} W^\perp$. For any Volterra kernel K of continuous and bounded type in L^2, and any $g_0 \in L^2([0,T], \mathbb{R})$, there exists a progressively measurable $\mathbb{R} \times \mathbb{R}_+$-valued strong solution (X, S) to (1.5)-(1.6) such that

$$\sup_{t\leq T} \mathbb{E}\left[|X_t|^p\right] < \infty, \quad p \geq 1, \tag{3.2}$$

we refer to Theorem A.1 below for the proof.

We now state our main result on the representation of the Fourier–Laplace transform for Volterra kernels under the following additional assumption on the kernel:

$$\sup_{t\leq T} \int_0^T |K(s,t)|^2 ds < \infty. \tag{3.3}$$

Theorem 3.3. *Let $g_0 \in L^2([0,T], \mathbb{R})$ and K be a Volterra kernel as in Definition 3.1 satisfying (3.3). Fix $u, w \in \mathbb{C}$, such that $0 \leq \Re(u) \leq 1$ and $\Re(w) \leq 0$. Then,*

$$\mathbb{E}\left[\exp\left(u \log \frac{S_T}{S_t} + w \int_t^T X_s^2 ds\right) \Big| \mathcal{F}_t\right] = \frac{\exp\left(\langle g_t, \mathbf{\Psi}_t g_t \rangle_{L^2}\right)}{\det\left(\mathrm{id} - 2a\tilde{\mathbf{\Sigma}}_t\right)^{1/2}}, \tag{3.4}$$

for all $t \leq T$, with $\mathbf{\Psi}_t$ given by (1.9) for (a,b) as in (1.10).

Proof. We refer to Appendix B. □

Finally, for $K(t,s) = \mathbf{1}_{s<t}$ and an input curve of the form

$$g_0(t) = X_0 + \theta t, \quad t \geq 0, \tag{3.5}$$

for some $X_0, \theta \in \mathbb{R}$, one recovers from Theorem 3.3 the well-known closed form expressions of Stein and Stein (1991) and Schöbel and Zhu (1999), and that of Heston (1993) when $\theta = 0$.

Corollary 3.4. *Assume that $K(t,s) = \mathbf{1}_{s<t}$ and that g_0 is of the form (3.5), then, the expression (3.4) reduces to*

$$\mathbb{E}\left[\exp\left(u \log \frac{S_T}{S_t} + w \int_t^T X_s^2 ds\right) \Big| \mathcal{F}_t\right] = \exp\left(A(t) + B(t)X_t + C(t)X_t^2\right) \tag{3.6}$$

where A, B, C solve the following system of (Backward) Riccati equations

$$\begin{aligned}
\dot{A} &= -\theta B - \frac{1}{2}\nu^2 B^2 - \nu^2 C, & A(T) &= 0, \\
\dot{B} &= -2\theta C - (\kappa + \rho\nu u + 2\nu C)B, & B(T) &= 0, \\
\dot{C} &= -2\nu^2 C^2 - 2(\kappa + \rho\nu u)C - w - \frac{1}{2}(u^2 - u), & C(T) &= 0.
\end{aligned}$$

In particular, (A, B, C) can be computed in closed form as in Lord and Kahl (2006, Equations (43)-(44)-(45)).

Sketch of proof. The characteristic function is given by (3.4). Assume that $K(t,s) = \mathbf{1}_{s<t}$ and g_0 is as in (3.5). Then,

$$X_s = X_t + (s-t)\theta + \int_t^s \kappa X_u du + \int_t^s \nu dW_u, \quad s \geq t,$$

so that taking conditional expectation yields

$$g_t(s) = \mathbf{1}_{t \leq s}\left(X_t + (s-t)\theta\right).$$

It follows that

$$\langle g_t, \mathbf{\Psi}_t g_t \rangle_{L^2} = \tilde{A}(t) + B(t)X_t + C(t)X_t^2$$

with

$$\tilde{A}(t) = \theta^2 \langle \mathbf{1}_{t\leq \cdot}(\cdot - t), \mathbf{\Psi}_t \mathbf{1}_{t\leq \cdot}(\cdot - t)\rangle_{L^2}, \; B(t) = 2\theta \langle \mathbf{1}_{t\leq \cdot}(\cdot - t), \mathbf{\Psi}_t \mathbf{1}_{t\leq \cdot}\rangle_{L^2}, \; C(t) = \langle \mathbf{1}_{t\leq \cdot}, \mathbf{\Psi}_t \mathbf{1}_{t\leq \cdot}\rangle.$$

Combined with (B.5) and (B.6) below, we obtain (3.6) with A such that $A_T = 0$ and

$$\dot{A}(t) = \dot{\tilde{A}}(t) + \mathrm{Tr}(\mathbf{\Psi}_t \dot{\mathbf{\Sigma}}_t)$$

with

$$\dot{\Sigma}_t(s,u) = -\nu^2 \mathbf{1}_{t \leq s \wedge u}.$$

Using the operator Riccati equation satisfied by $t \mapsto \mathbf{\Psi}_t$, see Lemma B.1 below, and straightforward computations as in Abi Jaber et al. (2020, Corollary 5.14) lead to the claimed system of Riccati equations for (A, B, C). □

4 Numerical illustration

In this section, we make use of the analytic expression for the characteristic function in (1.7) to price options. We first present two approximations of the formula (1.7) using closed form expressions. The first one is obtained from a natural discretization of the operators; the second one is more informative and establishes the link with finite-dimensional Wishart distributions.

4.1 A straightforward approximation by closed form expressions

For simplicity, we consider the case $t = 0$ and K a Volterra kernel, i.e. $K(t,s) = 0$ if $s \geq t$. The expression (1.7) lends itself to approximation by closed form solutions by a simple discretization of the operator $\boldsymbol{\Psi}_0$ given by (1.9) *la* Fredholm (1903). Fix $n \in \mathbb{N}$ and let $t_i = iT/n$, $i = 0, 1, \ldots, n$ be a partition of $[0, T]$. Discretizing the $\star$-product given in (1.13) yields the following approximation for $\boldsymbol{\Psi}_0$ by the $n \times n$ matrix:

$$\Psi_0^n = a\,(I_n - bK^n)^{-\top}\,(I_n - 2\frac{aT}{n}\tilde{\Sigma}^n)^{-1}\,(I_n - bK^n)^{-1},$$

where I_n is the $n \times n$ identity matrix, K^n is the lower triangular matrix with components

$$K_{ij}^n = \int_{t_{j-1}}^{t_j} K(t_{i-1}, s)ds, \quad j \leq i-1,$$

and

$$\tilde{\Sigma}^n = (I_n - bK^n)^{-1}\,\Sigma^n\,(I_n - bK^n)^{-\top}$$

with Σ^n the $n \times n$ discretized covariance matrix, recall (1.12), given by

$$\Sigma_{ij}^n = \nu^2 \int_0^T K(t_i, s)K(t_j, s)ds, \quad i, j \leq n. \tag{4.1}$$

Defining the n-dimensional vector $g_n = (g_0(t_1), \ldots, g_0(t_n))^\top$, the discretization of the inner product $\langle \cdot, \cdot \rangle_{L^2}$ leads to the approximation

$$\mathbb{E}\left[\exp\left(u \log S_T + w \int_0^T X_s^2 ds\right)\right] \approx \frac{\exp\left(u \log S_0 + \frac{T}{n} g_n^\top \Psi_0^n g_n\right)}{\det(\Phi_0^n)^{1/2}}$$

with $\Phi_0^n = \left(I_n - 2a\frac{T}{n}\tilde{\Sigma}^n\right)$.

Remark 4.1. *Depending on the smoothness of the kernel, other quadrature rules might be more efficient for the choice of the discretization of the operator and the approximation of the Fredholm determinant based on the so-called Nyström method, see for instance Bornemann (2009, 2010); Corlay (2010); Kang et al. (2003).*

4.2 An alternative approximation coming from Wishart distributions

We now provide an alternative representation stemming from Wishart distributions. Without loss of generality, we restrict to the case $\kappa = 0$.[4] The key idea is to rely on the projection argument already used in the proof of Theorem 2.2 to write

$$\mathbb{E}\left[\exp\left(u \log \frac{S_T}{S_0} + w \int_0^T X_s^2 ds\right)\right] = \mathbb{E}\left[\exp\left(\alpha \int_0^T X_s^2 ds + \beta \int_0^T X_s dW_s\right)\right], \quad (4.2)$$

with

$$\alpha = w + \frac{1}{2}(u^2 - u) - \frac{\rho^2 u^2}{2} \quad \text{and} \quad \beta = \rho u.$$

We now approximate the right hand in (4.2) by Wishart distributions in the following way. Let $n \in \mathbb{N}$, $t_i = iT/n$, $i = 0, \ldots, n$ and consider the Euler discretization of the quantities $(\int_0^T X_s^2 ds, \int_0^T X_s dW_s)$:

$$\alpha \int_0^T X_s^2 ds + \beta \int_0^T X_s dW_s \approx \sum_{i=1}^n \frac{\alpha T}{n} X_{t_i}^2 + \beta X_{t_i} Y_i,$$

with

$$X_i = X_{t_{i-1}} \quad \text{and} \quad Y_i = \int_{t_{i-1}}^{t_i} dW_s, \quad i = 1, \ldots, n.$$

Define the $2n$-dimensional vector $Z_n = (X_1, \ldots, X_n, Y_1, \ldots Y_n)^\top$ and the $2n \times 2n$ dimensional matrix

$$w_n = \begin{pmatrix} \frac{\alpha T}{n} I_n & \frac{\beta}{2} I_n \\ \frac{\beta}{2} I_n & 0_{\mathbb{R}^{n\times n}} \end{pmatrix}.$$

Then, Z_n is a $2n$-dimensional Gaussian vector with mean the $2n$-dimensional vector $\mu_n = (g_0(t_0), \ldots, g_0(t_{n-1}), 0, \ldots, 0)^\top$ and $2n \times 2n$ covariance matrix

$$R_n = \begin{pmatrix} \Sigma^X & \Sigma^{XY} \\ (\Sigma^{XY})^\top & \frac{T}{n} I_n \end{pmatrix}$$

with Σ^{XY} and Σ^X the $n \times n$-matrices with components

$$\Sigma_{ij}^X = \nu^2 \int_0^T K(t_{i-1}, s) K(t_{j-1}, s) ds, \quad \Sigma_{ij}^{XY} = \mathbf{1}_{j \le i-1} \nu \int_{t_{j-1}}^{t_j} K(t_{i-1}, s) ds. \quad (4.3)$$

[4] If $\kappa \neq 0$, then making use of the resolvent kernel R_T^κ of κK on $[0, T]$ (we refer to Appendix A), we reduce to the case $\kappa = 0$ as illustrated on (A.3) by working on the kernel $(K + R_T^\kappa) = \frac{1}{\kappa} R_T^\kappa$ instead of K and considering $g_0^\kappa = (\mathrm{id} + \boldsymbol{R}_T^\kappa) g_0$ instead of g_0.

Furthermore, we have

$$\sum_{i=1}^{n} \frac{\alpha T}{n} X_i^2 + \beta X_i Y_i = Z_n^\top w_n Z_n,$$

so that, the well-known expression for the characteristic function of the Wishart distribution, see for instance Abi Jaber (2019a, Proposition A.1), yields

$$\mathbb{E}\left[\exp\left(Z_n^\top w_n Z_n\right)\right] = \frac{\exp\left(\mu_n^\top w_n \left(I_{2n} - 2R_n w_n\right)^{-1} \mu_n\right)}{\det\left(I_{2n} - 2R_n w_n\right)^{1/2}}. \tag{4.4}$$

By dominated convergence in (4.2) we obtain

$$\mathbb{E}\left[\exp\left(u \log \frac{S_T}{S_0} + w \int_0^T X_s^2 ds\right)\right] = \lim_{n\to\infty} \frac{\exp\left(\mu_n^\top w_n \left(I_{2n} - 2R_n w_n\right)^{-1} \mu_n\right)}{\det\left(I_{2n} - 2R_n w_n\right)^{1/2}}.$$

For practical implementation, one can take advantage of the block matrix structure of $(R_n, w_n, \mu_n \mu_n^\top, w_n)$ to reduce to an expression involving $n \times n$ matrices instead of $2n \times 2n$. The well-known inversion formula for 2×2 block matrices using Schur complements combined with straightforward computations reduce expression (4.4) to

$$\mathbb{E}\left[\exp\left(Z_n^\top w_n Z_n\right)\right] = \frac{\exp\left(g_n^\top \left(\alpha \frac{T}{n} \tilde{A}_n + \frac{\beta}{2} \tilde{C}_n\right) g_n\right)}{\det\left(A_n D_n - B_n C_n\right)^{1/2}},$$

where $g_n = (g_0(t_0), \ldots, g_0(t_{n-1}))^\top$ and $A_n, B_n, C_n, D_n, \tilde{A}_n, \tilde{C}_n$ are the $n \times n$ matrices defined by

$$I_{2n} - 2R_n w_n = \begin{pmatrix} I_n - 2\alpha \frac{T}{n} \Sigma^X - \beta \Sigma^{XY} & -\beta \Sigma^X \\ -2\alpha \frac{T}{n} (\Sigma^{XY})^\top - \beta \frac{T}{n} I_n & I_n - \beta (\Sigma^{XY})^\top \end{pmatrix} =: \begin{pmatrix} A_n & B_n \\ C_n & D_n \end{pmatrix}$$

and

$$\tilde{A}_n := (A_n - B_n D_n^{-1} C_n)^{-1}, \quad \tilde{C}_n := -D_n^{-1} C_n \tilde{A}_n.$$

4.3 Option pricing in the fractional Stein–Stein model

In this section, we illustrate the applicability of our results on the following fractional Stein–Stein model based on the Riemann–Liouville fractional Brownian motion with the Volterra convolution kernel $K(t,s) = \mathbf{1}_{s<t}(t-s)^{H-1/2}/\Gamma(H+1/2)$:

$$dS_t = S_t X_t dB_t, \quad S_0 > 0,$$
$$X_t = g_0(t) + \frac{\nu}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} dW_s,$$

with $B = \rho W + \sqrt{1-\rho^2} W^{\perp}$, for $\rho \in [-1,1]$, $\nu \in \mathbb{R}$ and a Hurst index $H \in (0,1)$. For illustration purposes we will consider that the input curve g_0, which can be used in general to fit at-the-money curves observed in the market, has the following parametric form

$$g_0(t) = X_0 + \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} \theta ds = X_0 + \theta \frac{t^{H+1/2}}{\Gamma(H+1/2)(H+1/2)}, \quad t \geq 0.$$

Remark 4.2. *It would have also been possible to take instead of the fractional Riemman–Liouville Brownian motion the true fractional Brownian motion by considering*

$$X_t = g_0(t) + \frac{\nu}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} {}_2F_1\left(H-1/2, 1/2-H; H+1/2, 1-\frac{t}{s}\right) dW_s,$$

where ${}_2F_1$ is the Gaussian hypergeometric function.

Taking $H < 1/2$ allows one to reproduce the stylized facts observed in the market as in Figure 1. Indeed, the simulated sample paths of the instantaneous variance process X^2 with $H = 0.1$ in Figure 2 has the same regularity as the realized variance of the S&P in Figure 1-(a). In the case $H < 1/2$, we refer to the model as the rough Stein–Stein model.

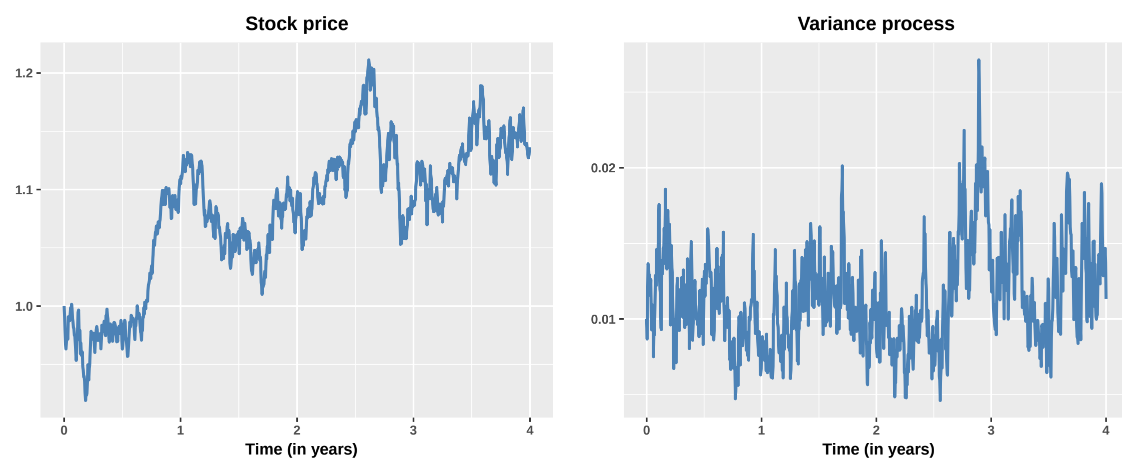


Figure 2: One simulated sample path of the stock price S and the instantaneous variance process X^2 in the rough Stein–Stein model with parameters: $X_0 = 0, 1$, $\theta = 0.01$, $\nu = 0.02$, $\rho = -0.7$ and $H = 0.1$.

We now move to pricing. The expression (1.7) for the joint characteristic function allows one to recover the joint density $p_T(x,y)$ of $(\log S_T, \int_0^T X_s^2 ds)$ by Fourier inversion:

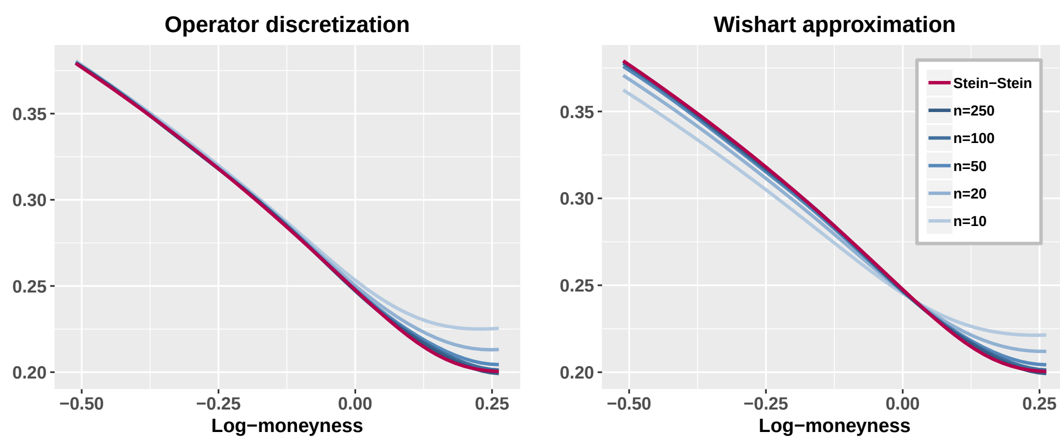
$$p_T(x,y) = \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{-i(z_1 x + z_2 y)} \mathbb{E}\left[\exp\left(i z_1 \log S_T + i z_2 \int_0^T X_s^2 ds\right)\right] dz_1 dz_2,$$

but also to price derivatives on the stock price and the integrated variance by Fourier inversion techniques, see Carr and Madan (1999); Fang and Oosterlee (2009); Lewis (2001) among many others. In the sequel we will make use of the cosine method of Fang and Oosterlee (2009) to price European call options on the stock S combined with our approximation formulae of Sections 4.1 and 4.2. We refer to Abi Jaber (2019a, Section 4) for examples of pricing options on the integrated variance. We start by observing that the covariance function of X is given in the following closed form

$$\begin{aligned}\Sigma_0(s,u) &= \frac{\nu^2}{\Gamma(H+1/2)}\int_0^{s\wedge u}(s-z)^{H-1/2}(u-z)^{H-1/2}dz\\ &= \frac{\nu^2}{\Gamma(\alpha)\Gamma(1+\alpha)}\frac{s^\alpha}{u^{1-\alpha}}\ {}_2F_1\left(1,1-\alpha;1+\alpha;\frac{s}{u}\right)\end{aligned}$$

where $\alpha = H+1/2$ and ${}_2F_1$ is the Gaussian hypergeometric function, see for instance Malyarenko (2012, page 71).[5] It follows that the matrices (4.1) and (4.3) can be computed in closed form.

As a sanity check, for $H = 0.5$ we visualize on Figure 3 the convergence of both approximation methods on the implied volatility. The benchmark is computed via the cosine method with the closed form expressions for the characteristic function of the conventional Stein–Stein model, see Lord and Kahl (2006). Both methods converge. After some numerical experiments, we found that the approximation method based on Wishart distributions is slightly more stable for small values of H. We will therefore use it in the sequel with $n = 50$. Other discretization rules might turn out more efficient and would require less points to achieve the same accuracy, which makes the implementation much faster, recall Remark 4.1. The main challenge for applying such methods is the singularity of the kernel at $s = t$ when $H < 1/2$ and is left for future research.



[5] Note that in the case of Remark 4.2, the expression for the covariance function simplifies to $\Sigma_0(s,u) = \frac{\nu^2}{2}(s^{2H}+u^{2H}-|s-u|^{2H})$.

Figure 3: Convergence of the implied volatility slice for $T = 1$ year of the operator discretization of Section 4.1 (left) and the approximation via Wishart distributions of Section 4.2 (right) for $H = 0.5$ towards that of the conventional Stein–Stein model (red). The parameters are $X_0 = 0.2$, $\theta = 0$, $\nu = 0.3$ and $\rho = -0.7$.

Going back to real market data, we calibrate the fractional Stein–Stein model to the at-the-money skew of Figure 1-(b). Keeping the parameters $X_0 = 0.44$, $\theta = 0.3$ fixed, the calibrated parameters are given by

$$\hat{\nu} = 0.5231458, \quad \hat{\rho} = -0.9436174 \quad \text{and} \quad \hat{H} = 0.2234273. \tag{4.5}$$

This power-law behaviour of the at-the-money skew observed on the market is perfectly captured by the fractional Stein–Stein model as illustrated on Figure 4 with only three parameters. Again $\hat{H} < 0.5$ indicates that the rough regime of the fractional Stein–Stein model is coherent with the observations on the market.

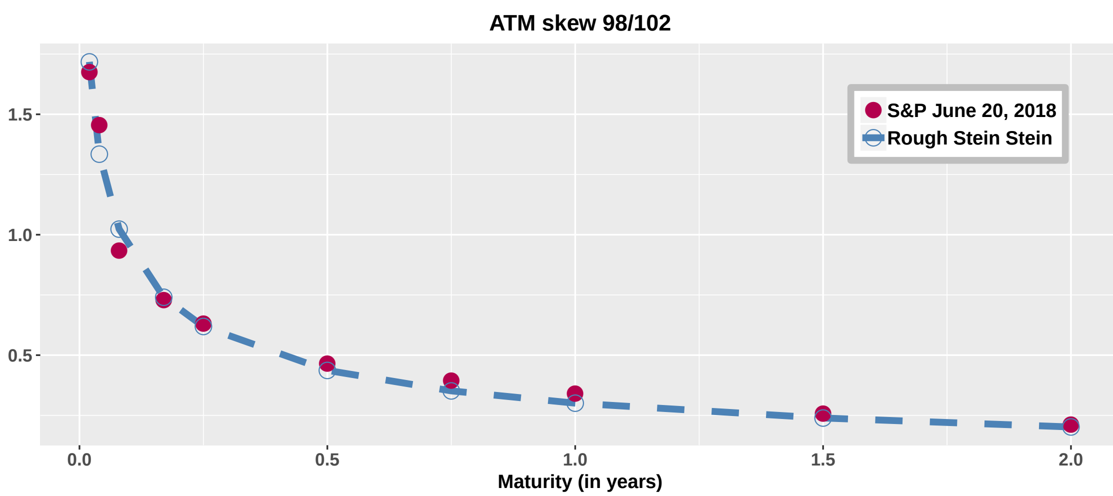


Figure 4: Term structure of the at-the-money skew for the S&P index on June 20, 2018 (red dots) and for the rough Stein–Stein model with calibrated parameters (4.5) (blue circles with dashed line).

A Resolvents

For a kernel $K \in L^2([0,T]^2, \mathbb{K})$, we define its resolvent $R_T \in L^2([0,T]^2, \mathbb{K})$ by the unique solution to

$$R_T = K + K \star R_T, \qquad K \star R_T = R_T \star K. \tag{A.1}$$

In terms of integral operators, this translates into

$$\boldsymbol{R}_T = \boldsymbol{K} + \boldsymbol{K}\boldsymbol{R}_T, \quad \boldsymbol{K}\boldsymbol{R}_T = \boldsymbol{R}_T\boldsymbol{K}.$$

In particular, if K admits a resolvent, $(\mathrm{id} - \boldsymbol{K})$ is invertible and

$$(\mathrm{id} - \boldsymbol{K})^{-1} = \mathrm{id} + \boldsymbol{R}_T. \tag{A.2}$$

Whenever $K \in L^2([0,T]^2, \mathbb{K})$, $\boldsymbol{K}$ is a Hilbert–Schmidt operator on $L^2([0,T], \mathbb{K})$ into itself. Whence, it has a countable spectrum $\mathrm{sp}(\boldsymbol{K})$. If $1 \notin \mathrm{sp}(\boldsymbol{K})$, then $(\mathrm{id} - \boldsymbol{K})$ is invertible and $\boldsymbol{K}$ admits a resolvent. We refer to Smithies (1958) for more details.

Using the resolvent we can provide the explicit solution to the system (1.5)–(1.6).

Theorem A.1. *Fix $T > 0$, $g_0 \in L^2([0,T], \mathbb{R})$ and a kernel K as in Definition 3.1. Then, there exists a progressively measurable strong solution (X, S) to (1.5)–(1.6) on $[0,T]$ such that (3.2) holds.*

Proof. Since K is a Volterra kernel of continuous and bounded type, it follows from Gripenberg et al. (1990, Lemma 9.3.3, Theorem 9.5.5(i)) that κK admits a resolvent R_T^κ which is again a Volterra kernel of continuous and bounded type. A straightforward application of the resolvent equation (A.1) together with stochastic Fubini's theorem, yields that the solution for (1.6) is given in the following closed form

$$X_t = g_0(t) + \int_0^t R_T^\kappa(t,s) g_0(s) ds + \int_0^t (K(t,s) + R_T^\kappa(t,s)) \nu dW_s, \tag{A.3}$$

The existence of S readily follows from that of X and is given by the stochastic exponential:

$$S_t = S_0 \exp\left(-\frac{1}{2} \int_0^t X_s^2 ds + \int_0^t X_s dB_s \right).$$

Finally, (3.2) follows from the fact that $\sup_{s \leq T} \int_0^T |R_T^\kappa(s,u)|^2 du < \infty$ combined with the Burkholder-Davis-Gundy inequality. $\square$

We now justify in the three following lemmas that the quantities $(\mathrm{id} - b\boldsymbol{K})$ and $\left(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_t\right)$ appearing in the definition of $t \mapsto \boldsymbol{\Psi}_t$ in (1.9) are invertible so that $\boldsymbol{\Psi}_t$ is well-defined for any kernel K as in Definition 3.1.

Lemma A.2. *Let K satisfy (3.1) and $L \in L^2([0,T]^2, \mathbb{R})$. Then, $K \star L$ satisfies (3.1). Furthemore, if L satisfies (3.1), then, $(s,u) \mapsto (K \star L^*)(s,u)$ is continuous.*

Proof. An application of the Cauchy-Schwarz inequality yields the first part. The second part follows along the same lines as in the proof of Abi Jaber (2019a, Lemma 3.2). $\square$

Lemma A.3. *Fix $b \in \mathbb{C}$ and a kernel K as in Definition 3.1. Then, $(\mathrm{id} - b\boldsymbol{K})$ is invertible. Furthermore, for all $t \le T$, $\tilde{\boldsymbol{\Sigma}}_t$ given by (1.11) can be re-written in the form*

$$\tilde{\boldsymbol{\Sigma}}_t = (\mathrm{id} - b\boldsymbol{K}_t)^{-1}\boldsymbol{\Sigma}_t(\mathrm{id} - b\boldsymbol{K}_t)^{-*} \tag{A.4}$$

where $\boldsymbol{K}_t$ is the integral operator induced by the kernel $K_t(s,u) = K(s,u)\mathbf{1}_{u\ge t}$, for $s, u \le T$.

Proof. Since K is a Volterra kernel of continuous and bounded type, it follows from Gripenberg et al. (1990, Lemma 9.3.3, Theorem 9.5.5(i)) (see also Smithies (1958, Theorem 2.7.1)) that bK admits a resolvent R_T^b which is again a Volterra kernel of continuous and bounded type. Whence, (A.2) yields that $(\mathrm{id} - b\boldsymbol{K})$ is invertible with an inverse given by $(\mathrm{id} + R_T^b)$. To prove (A.4), we fix $t \le T$ and we observe that since $\Sigma_t(s,u) = 0$ whenever $s \wedge u \le t$, we have

$$(R_T^b \star \Sigma_t)(s,u) = \int_t^T R_T^b(s,u)\Sigma_t(z,u)dz = (R_{t,T}^b \star \Sigma_t(z,u))(s,u),$$

where we defined the kernel $R_{t,T}^b(s,u) = R_T^b(s,u)\mathbf{1}_{u\ge t}$. Similarly, $\Sigma_t \star (R_T^b)^* = \Sigma_t \star (R_{t,T}^b)^*$. Using the resolvent equation (A.1) of R_T^b, it readily follows that $R_{t,T}^b$ is the resolvent of bK_t so that $(\mathrm{id} - b\boldsymbol{K_t})^{-1} = (\mathrm{id} + \boldsymbol{R}_{t,T}^b)$. Combining all of the above leads to

$$\begin{aligned}
\tilde{\boldsymbol{\Sigma}}_t &= (\mathrm{id} - b\boldsymbol{K})^{-1}\boldsymbol{\Sigma}_t(\mathrm{id} - b\boldsymbol{K})^{-*} \\
&= (\mathrm{id} + \boldsymbol{R}_T^b)\boldsymbol{\Sigma}_t(\mathrm{id} + \boldsymbol{R}_T^b)^* \\
&= \mathrm{id} + \boldsymbol{R}_T^b\boldsymbol{\Sigma}_t + \boldsymbol{\Sigma}_t(\boldsymbol{R}_T^b)^* + \boldsymbol{R}_T^b\boldsymbol{\Sigma}_t(\boldsymbol{R}_T^b)^* \\
&= \mathrm{id} + \boldsymbol{R}_{t,T}^b\boldsymbol{\Sigma}_t + \boldsymbol{\Sigma}_t(\boldsymbol{R}_{t,T}^b)^* + \boldsymbol{R}_{t,T}^b\boldsymbol{\Sigma}_t(\boldsymbol{R}_{t,T}^b)^* \\
&= (\mathrm{id} + \boldsymbol{R}_{t,T}^b)\boldsymbol{\Sigma}_t(\mathrm{id} + \boldsymbol{R}_{t,T}^b)^* \\
&= (\mathrm{id} - b\boldsymbol{K}_t)^{-1}\boldsymbol{\Sigma}_t(\mathrm{id} + b\boldsymbol{K}_t)^{-*},
\end{aligned}$$

which ends the proof. □

Lemma A.4. *Fix $a, b \in \mathbb{C}$ such that $\Re(a) \le 0$ and $\Re(a) + \frac{\Im(b)^2}{2\nu^2} \le 0$. Let $t \le T$ and K be a kernel as in Definition 3.1. Then, $(\mathrm{id} - 2\tilde{\boldsymbol{\Sigma}}_t a)$ is invertible and $\boldsymbol{\Psi}_t$ given by (1.9) is well-defined. Furthermore, if $\Im(a) = \Im(b) = 0$ then, $\boldsymbol{\Psi}_t$ is a symmetric nonpositive operator in the sense of Definition 2.1.*

Proof. • Using Lemma A.3, we write

$$(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_t) = (\mathrm{id} - b\boldsymbol{K}_t)^{-1}\boldsymbol{A}_t(\mathrm{id} - b\boldsymbol{K}_t)^{-*}$$

with

$$\begin{aligned}
\boldsymbol{A}_t &= (\mathrm{id} - b\boldsymbol{K}_t)(\mathrm{id} - b\boldsymbol{K}_t^*) - 2a\boldsymbol{\Sigma}_t \\
&= \mathrm{id} - b\boldsymbol{K}_t - b\boldsymbol{K}_t^* + b^2\boldsymbol{K}_t\boldsymbol{K}_t^* - 2a\boldsymbol{\Sigma}_t.
\end{aligned}$$

It suffices to prove that $\boldsymbol{A}_t$ is invertible, that is $0 \notin \mathrm{sp}(\boldsymbol{A}_t)$. Taking real parts and observing that $\Sigma_t = \nu^2 \boldsymbol{K}_t \boldsymbol{K}_t^*$ yields

$$\begin{aligned}
\Re(\boldsymbol{A}_t) &= \mathrm{id} - \Re(b)\boldsymbol{K}_t - \Re(b)\boldsymbol{K}_t^* + \Re(b)^2 \boldsymbol{K}_t \boldsymbol{K}_t^* - \Im(b)^2 \boldsymbol{K}_t \boldsymbol{K}_t^* - 2\Re(a)\boldsymbol{\Sigma}_t \\
&= (\mathrm{id} - \Re(b)\boldsymbol{K}_t)(\mathrm{id} - \Re(b)\boldsymbol{K}_t)^* - \left(2\Re(a) + \frac{\Im(b)^2}{\nu^2}\right)\boldsymbol{\Sigma}_t \\
&= \mathbf{I} + \mathbf{II}
\end{aligned}$$

The operator $\mathbf{I}$ is symmetric nonnegative and invertible so that $\mathrm{sp}(\mathbf{I}) \in (0, \infty)$. Furthermore, since $\left(2\Re(a) + \frac{\Im(b)^2}{\nu^2}\right) \leq 0$ by assumption and $\boldsymbol{\Sigma}_t$ is symmetric nonnegative we have $\mathrm{sp}(\mathbf{II}) \in [0, \infty)$. It follows that $\mathrm{sp}(\Re(\boldsymbol{A}_t)) \in (0, \infty)$, showing that $0 \notin \mathrm{sp}(\boldsymbol{A}_t)$ and that $\boldsymbol{A}_t$ is invertible. Whence, $(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_t)$ is invertible. Combined with Lemma A.3, we obtain that $\boldsymbol{\Psi}_t$ is well-defined.

• Assume that $\Im(a) = \Im(b) = 0$. $\tilde{\boldsymbol{\Sigma}}_t$ defined as in (1.11) is clearly a symmetric nonnegative operator with a continuous kernel on $[0,T]^2$, recall Lemma A.2, an application of Mercer's theorem (Shorack and Wellner, 2009, Theorem 1, p.208) yields the existence of an orthonormal basis $(e_n)_{n\geq 1}$ of $L^2([0,T], \mathbb{R})$ and nonnegative eigenvalues $(\lambda_n)_{n\geq 1}$ such that

$$\tilde{\boldsymbol{\Sigma}}_t = \sum_{n\geq 1} \lambda_n \langle e_n, \cdot \rangle_{L^2} e_n.$$

Whence,

$$\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_t = \sum_{n\geq 1} (1 - 2a\lambda_n) \langle e_n, \cdot \rangle_{L^2} e_n.$$

Since $a \leq 0$, $(1 - 2a\lambda_n) \geq 1 > 0$, for each $n \geq 1$, so that the inverse of $(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_t)$ is a symmetric nonnegative operator given by

$$\left(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_t\right)^{-1} = \sum_{n\geq 1} \frac{1}{1 - 2a\lambda_n} \langle e_n, \cdot \rangle_{L^2} e_n.$$

Finally, $\boldsymbol{\Psi}_t$ is clearly symmetric and for any $f \in L^2([0,T], \mathbb{R})$

$$\langle f, \boldsymbol{\Psi}_t f \rangle_{L^2} = a \langle \tilde{f}, \left(\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_t\right)^{-1} \tilde{f} \rangle_{L^2} \geq 0,$$

with $\tilde{f} = (\mathrm{id} - b\boldsymbol{K})^{-1} f$. This shows that $\boldsymbol{\Psi}_t$ is nonpositive. □

B Proof of Theorem 3.3

This section is dedicated to the proof of Theorem 3.3. We fix $T > 0$, a Volterra kernel K as in Definition 3.1 satisfying (3.3) and $u, w \in \mathbb{C}$, such that $0 \leq \Re(u) \leq 1$ and $\Re(w) \leq 0$. It follows that a, b defined by (1.10) satisfy

$$\Re(a) = \Re(w) + \frac{1}{2}(\Re(u)^2 - \Re(u)) - \frac{1}{2}\Im(u)^2 \leq 0$$

and

$$\Re(a) + \frac{\Im(b)^2}{2\nu^2} = \Re(w) + \frac{1}{2}(\Re(u)^2 - \Re(u)) + \frac{1}{2}(\rho^2 - 1)\Im(u)^2 \leq 0.$$

so that an application of Lemma A.4 yields that $\boldsymbol{\Psi}_t$ is well-defined.

We now collect from Abi Jaber et al. (2020, Lemma 5.8) further properties of $t \mapsto \boldsymbol{\Psi}_t$. In particular, its link with an operator Riccati equation. We recall that $t \mapsto \boldsymbol{\Psi}_t$ is said to be strongly differentiable at time $t \geq 0$, if there exists a bounded linear operator $\dot{\boldsymbol{\Psi}}_t$ from $L^2([0,T],\mathbb{C})$ into itself such that

$$\lim_{h\to 0} \frac{1}{h} \|\boldsymbol{\Psi}_{t+h} - \boldsymbol{\Psi}_t - h\dot{\boldsymbol{\Psi}}_t\|_{\text{op}} = 0, \quad \text{where } \|\boldsymbol{G}\|_{\text{op}} = \sup_{f\in L^2([0,T],\mathbb{C})} \frac{\|\mathbf{G}f\|_{L^2}}{\|f\|_{L^2}}.$$

Lemma B.1. *Fix a kernel K as in Definition 3.1 satisfying (3.3). Then, for each $t \leq T$, $\boldsymbol{\Psi}_t$ given by (1.9) is a bounded linear operator from $L^2([0,T],\mathbb{R})$ into itself. Furthermore,*

(i) *$(-a\text{id} + \boldsymbol{\Psi}_t)$ is an integral operator induced by a symmetric kernel $\psi_t(s,u)$ such that*

$$\sup_{t\leq T} \int_{[0,T]^2} |\psi_t(s,u)|^2 ds du < \infty.$$

(ii) *For any $f \in L^2([0,T],\mathbb{R})$,*

$$(\boldsymbol{\Psi}_t f 1_t)(t) = (a\text{id} + b\boldsymbol{K}^*\boldsymbol{\Psi}_t)(f 1_t)(t),$$

where $1_t : s \mapsto \mathbf{1}_{t\leq s}$.

(iii) *$t \mapsto \boldsymbol{\Psi}_t$ is strongly differentiable and satisfies the operator Riccati equation*

$$\begin{aligned} \dot{\boldsymbol{\Psi}}_t &= 2\boldsymbol{\Psi}_t\dot{\boldsymbol{\Sigma}}_t\boldsymbol{\Psi}_t, \qquad t \in [0,T] \\ \boldsymbol{\Psi}_T &= a\,(\text{id} - b\boldsymbol{K})^{-*}\,(\text{id} - b\boldsymbol{K})^{-1} \end{aligned} \tag{B.1}$$

where $\dot{\boldsymbol{\Sigma}}_t$ is the strong derivative of $t \mapsto \boldsymbol{\Sigma}_t$ induced by the kernel

$$\dot{\Sigma}_t(s,u) = -\nu^2 K(s,t)K(u,t), \quad a.e.$$

Proof. The proof follows from a straighforward adaptation of the proof of Abi Jaber et al. (2020, Lemma 5.8). □

Using the previous lemma and observing that the adjusted conditional mean given in (1.8) has the following dynamics

$$g_t(s) = \mathbf{1}_{t\leq s}\left(g_0(s) + \int_0^t K(s,u)\kappa X_u du + \int_0^t K(s,u)\nu dW_u\right) \tag{B.2}$$

we derive in the next lemma the dynamics of $t \mapsto \langle g_t, \boldsymbol{\Psi}_t g_t\rangle_{L^2}$.

Lemma B.2. *The dynamics of* $t \mapsto \langle g_t, \mathbf{\Psi}_t g_t \rangle_{L^2}$ *are given by*

$$d\langle g_t, \mathbf{\Psi}_t g_t\rangle_{L^2} = \Big(\langle g_t, \dot{\mathbf{\Psi}}_t g_t\rangle_{L^2} - aX_t^2 - 2u\rho\nu X_t \left(\boldsymbol{K}^*\mathbf{\Psi}_t\right)(g_t)(t) - \operatorname{Tr}\left(\mathbf{\Psi}_t \dot{\mathbf{\Sigma}}_t\right)\Big) dt + 2\nu \left(\left(\boldsymbol{K}^*\mathbf{\Psi}_t\right) g_t\right)(t) dW_t. \tag{B.3}$$

Proof. We first note that

$$\langle g_t, \mathbf{\Psi}_t g_t\rangle_{L^2} = \int_0^T g_t(s)(\mathbf{\Psi}_t g_t)(s) ds,$$

and compute the dynamics of $t \mapsto g_t(s)(\mathbf{\Psi}_t g_t)(s)$. For fixed $s \le T$, it follows from (B.2) and the fact that $g_t(t) = X_t$, that

$$dg_t(s) = -\delta_{t=s} X_t dt + K(s,t)\kappa X_t dt + K(s,t)\nu dW_t.$$

We recall that

$$(\mathbf{\Psi}_t f)(t) = af(t) + \int_0^T \psi_t(s,u) f(u) du, \tag{B.4}$$

see Lemma B.1–(i). Together with Lemma B.1-(iii), we deduce that $t \mapsto (\mathbf{\Psi}_t g_t)(s)$ is a semimartingale on $[0,s)$ with the following dynamics

$$\begin{aligned} d(\mathbf{\Psi}_t g_t)(s) &= (\dot{\mathbf{\Psi}}_t g_t)(s) dt + (\mathbf{\Psi}_t dg_t)(s) \\ &= (\dot{\mathbf{\Psi}}_t g_t)(s) dt - X_t \psi_t(s,t) dt + X_t(\mathbf{\Psi}_t K(\cdot,t)\kappa)(s) dt + (\mathbf{\Psi}_t K(\cdot,t)\nu)(s) dW_t. \end{aligned}$$

Here, we used the fact that $\mathrm{id}\delta_t = 0$: indeed, for every $f \in L^2([0,T],\mathbb{R})$ we have $(\mathrm{id}\delta_{t=\cdot})(f) = (f(\cdot)\delta_{t=\cdot}) = 0_{L^2}$. Moreover, the quadratic covariation between $t \mapsto g_t(s)$ and $t \mapsto (\mathbf{\Psi}_t g_t)(s)$ is given by

$$\begin{aligned} d\left[g_\cdot(s), (\mathbf{\Psi}_\cdot g_\cdot)(s)\right]_t &= a\nu^2 K(s,t)^2 dt + \nu^2 \int_0^T \psi_t(s,u) K(u,t) K(s,t) du dt \\ &= -a\dot{\Sigma}_t(s,s) dt - \int_0^T \psi_t(s,u)\dot{\Sigma}_t(u,s) du dt \\ &= -\left(\mathbf{\Psi}_t \dot{\Sigma}_t(\cdot,s)\right)(s). \end{aligned}$$

Whence, combining the previous three identities, we get

$$\begin{aligned} d\left(g_t(s)(\mathbf{\Psi}_t g_t)(s)\right) &= dg_t(s)(\mathbf{\Psi}_t g_t)(s) + g_t(s) d(\mathbf{\Psi}_t g_t)(s) + d\left[g_\cdot(s), (\mathbf{\Psi}_\cdot g_\cdot)(s)\right]_t \\ &= -\delta_{t=s} X_t (\mathbf{\Psi}_t g_t)(s) dt + X_t \kappa K(s,t)(\mathbf{\Psi}_t g_t)(s) dt \\ &\quad + g_t(s)(\dot{\mathbf{\Psi}}_t g_t)(s) dt - g_t(s)\psi_t(s,t) X_t dt + g_t(s) X_t (\mathbf{\Psi}_t K(\cdot,t)\kappa)(s) dt \\ &\quad - \left(\mathbf{\Psi}_t \dot{\Sigma}_t(\cdot,s)\right)(s) dt \\ &\quad + \left(\nu K(s,t)(\mathbf{\Psi}_t g_t)(s) + g_t(s)(\mathbf{\Psi}_t K(\cdot,t)\nu)(s)\right) dW_t \\ &= \Big[\mathbf{I}(s) + \mathbf{II}(s) + \mathbf{III}(s) + \mathbf{IV}(s) + \mathbf{V}(s) + \mathbf{VI}(s)\Big] dt \\ &\quad + \left(\mathbf{VII}(s) + \mathbf{VIII}(s)\right) dW_t. \end{aligned}$$

We now integrate in s. First, using (B.4) and Lemma B.1–(ii) successively we get that

$$\begin{aligned}\int_0^T \big[\mathbf{I}(s) + \mathbf{IV}(s)\big] ds &= -X_t(\boldsymbol{\Psi}_t g_t)(t) - X_t \int_0^T \psi_t(t,u) g_t(u) du \\ &= aX_t^2 - 2X_t(\boldsymbol{\Psi}_t g_t)(t) \\ &= -aX_t^2 - 2b((\boldsymbol{K}^*\boldsymbol{\Psi}_t)g_t)(t)\end{aligned}$$

On the other hand, since, $\boldsymbol{\Psi}^* = \boldsymbol{\Psi}$, we have

$$\int_0^T \big[\mathbf{II}(s) + \mathbf{V}(s)\big] ds = 2\kappa X_t \Big((\boldsymbol{K}^*\boldsymbol{\Psi}_t) g_t \Big)(t).$$

Therefore, summing the above and recalling that $b = \kappa + u\rho\nu$ yield

$$\int_0^T \big[\mathbf{I}(s) + \mathbf{IV}(s) + \mathbf{II}(s) + \mathbf{V}(s)\big] ds = -aX_t^2 - 2u\rho\nu X_t \Big((\boldsymbol{K}^*\boldsymbol{\Psi}_t) g_t \Big)(t).$$

Finally, observing that

$$\int_0^T \mathbf{III}(s) ds = \langle g_t, \dot{\boldsymbol{\Psi}}_t g_t \rangle_{L^2}, \quad \int_0^T \mathbf{VI}(s) ds = \operatorname{Tr}\Big(\boldsymbol{\Psi}_t \dot{\boldsymbol{\Sigma}}_t \Big),$$

$$\int_0^T \big[\mathbf{VII}(s) + \mathbf{VIII}(s)\big] ds = 2\nu \left(\boldsymbol{K}^*\boldsymbol{\Psi}_t \right) (g_t)(t) dW_t,$$

we obtain the claimed dynamics (B.3). □

We now recall the definition

$$\Phi_t = \mathrm{id} - 2\tilde{\boldsymbol{\Sigma}}_t a, \quad t \le T,$$

so that

$$\det(\Phi_t)^{-1/2} = \det\left(\mathrm{id} - 2\tilde{\boldsymbol{\Sigma}}_t a \right)^{-1/2},$$

which is well-defined by virtue of the invertibility of $(\mathrm{id} - 2\tilde{\boldsymbol{\Sigma}}_t a)$, see Simon (1977, Theorem 3.9). Furthermore, Lidskii's theorem, see Simon (1977), ensures that $\det(\mathrm{id} + \boldsymbol{F}) = \exp\left(\operatorname{Tr}(\log(\mathrm{id} + \boldsymbol{F}))\right)$. Hence,

$$\det(\Phi_t)^{-1/2} = \exp\left(\phi_t \right), \tag{B.5}$$

with

$$\phi_t = -\frac{1}{2} \operatorname{Tr}\left(\log\left(\mathrm{id} - 2\tilde{\boldsymbol{\Sigma}}_t a \right) \right).$$

Differentiation using (1.11) yields

$$\dot{\phi}_t = \mathrm{Tr}\left(a\left(\mathrm{id} - 2\tilde{\boldsymbol{\Sigma}}_t a\right)^{-1}\dot{\tilde{\boldsymbol{\Sigma}}}_{\boldsymbol{t}}\right) = \mathrm{Tr}\left(a\left(\mathrm{id} - 2\tilde{\boldsymbol{\Sigma}}_t a\right)^{-1}(\mathrm{id} - b\boldsymbol{K})^{-1}\dot{\boldsymbol{\Sigma}}_{\boldsymbol{t}}(\mathrm{id} - b\boldsymbol{K})^{-*}\right).$$

Finally, using (1.9) and the identity $\mathrm{Tr}(\boldsymbol{FG}) = \mathrm{Tr}(\boldsymbol{GF})$, we obtain

$$\dot{\phi}_t = \mathrm{Tr}(\boldsymbol{\Psi}_t\dot{\boldsymbol{\Sigma}}_{\boldsymbol{t}}). \tag{B.6}$$

We can now complete the proof of Theorem 3.3.

Proof of Theorem 3.3. Since $\boldsymbol{\Psi}_t$ and ϕ_t given by (1.9) and (B.6) are clearly analytic in (a, b), see for instance Smithies (1958, Corollary on p.31), it suffices to prove that (3.4) holds for all $0 \le u \le 1$ and $w \le 0$ to obtain the claimed expression by analytic continuation of the characteristic function on $\{(u, w) \in \mathbb{C}^2 : 0 \le \Re(u) \le 1 \text{ and } \Re(v) \le 0\}$. Fix $u \in [0, 1]$, $w \in \mathbb{R}_-$. Set

$$U_t = u \log S_t + w \int_0^t X_s^2 ds + \phi_t + \langle g_t, \boldsymbol{\Psi}_t g_t\rangle_{L^2}, \tag{B.7}$$

and $M_t = \exp(U_t)$. It suffices to prove that M is a martingale. Indeed, if this is the case, then observing that the terminal value of M is

$$M_T = u \log S_T + w \int_0^T X_s^2 ds$$

and writing the martingale property $\mathbb{E}[M_T|\mathcal{F}_t] = M_t$, for $t \le T$, yields (3.4).
Step 1. We prove that M is a local martingale by expliciting its dynamics. We first observe that

$$dM_t = M_t\big(dU_t + \frac{1}{2}d\langle U\rangle_t\big). \tag{B.8}$$

Using (1.5), we have

$$d\log S_t = -\frac{1}{2}X_t^2 dt + \rho X_t dW_t + \sqrt{1-\rho^2}X_t dW_t^\perp.$$

Combined with the dynamics (B.3) and the fact that $a = w + \frac{1}{2}(u^2 - u)$, we get that

$$\begin{aligned} dU_t = &\left(\langle g_t, \dot{\boldsymbol{\Psi}}_t g_t\rangle_{L^2} - \frac{u^2}{2}X_t^2 - 2u\rho\nu X_t\left(\boldsymbol{K}^*\boldsymbol{\Psi}_t\right)(g_t)(t) + \dot{\phi}_t - \mathrm{Tr}\left(\boldsymbol{\Psi}_t\dot{\boldsymbol{\Sigma}}_t\right)\right)dt \\ &+ \left(\rho u X_t + 2\nu\left(\boldsymbol{K}^*\boldsymbol{\Psi}_t\right)(g_t)(t)\right)dW_t + u\sqrt{1-\rho^2}X_t dW_t^\perp, \end{aligned}$$

so that

$$d\langle U\rangle_t = \left(u^2 X_t^2 + 4\rho u\nu X_t\left(\boldsymbol{K}^*\boldsymbol{\Psi}_t\right)(g_t)(t) + 4\nu^2\left(\left(\boldsymbol{K}^*\boldsymbol{\Psi}_t\right)(g_t)(t)\right)^2\right)dt.$$

Observing that

$$4\nu^2 \left(\left(\boldsymbol{K}^* \boldsymbol{\Psi}_t \right) (g_t)(t) \right)^2 = -4\langle g_t, \boldsymbol{\Psi}_t \dot{\boldsymbol{\Sigma}}_t \boldsymbol{\Psi}_t g_t \rangle_{L^2},$$

we get that the drift part in (B.8) is given by

$$M_t \left(\langle g_t, \dot{\boldsymbol{\Psi}}_t - 2\boldsymbol{\Psi}_t \dot{\boldsymbol{\Sigma}}_t \boldsymbol{\Psi}_t g_t \rangle_{L^2} + \dot{\phi}_t - \operatorname{Tr}\left(\boldsymbol{\Psi}_t \dot{\boldsymbol{\Sigma}}_t \right) \right) = 0,$$

by virtue of the Riccati equations (B.1) and (B.6). This shows that M is a local martingale.
Step 2. It remains to argue that the local martingale M is a true martingale. To this end, we fix $t \leq T$. An application of the second part of Lemma A.4 yields that $\boldsymbol{\Psi}_t$ is a symmetric nonpositive operator so that, recall (B.6),

$$\langle g_t, \boldsymbol{\Psi}_t g_t \rangle_{L^2} \leq 0 \quad \text{and} \quad \phi_t = -\int_t^T \operatorname{Tr}(\boldsymbol{\Psi}_s \dot{\boldsymbol{\Sigma}}_s) ds \leq 0.$$

Whence, since $w \leq 0$ and $0 \leq u \leq 1$, it follows from (B.7) that

$$\begin{aligned} U_t &\leq u \log S_t \\ &= u \log S_0 - \frac{u}{2} \int_0^t X_s^2 ds + u \int_0^t X_s dB_s \\ &\leq u \log S_0 - \frac{u^2}{2} \int_0^t X_s^2 ds + u \int_0^t X_s dB_s \end{aligned}$$

Therefore,

$$|M_t| = \exp(U_t) \leq \exp(u \log S_t) \leq N_t$$

with $N_t = S_0^u \exp\left(-\frac{u^2}{2} \int_0^t X_s^2 ds + u \int_0^t X_s dB_s \right)$ which can be shown to be a true martingale by a similar argument to that used in Abi Jaber et al. (2019, Lemma 7.3). Finally, we have showed that the local martingale M is bounded by a martingale, which gives that M is also a true martingale. The proof is complete. □

References

Abi Jaber, E. (2019a). The Laplace transform of the integrated Volterra Wishart process. *arXiv preprint arXiv:1911.07719*.

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

Abi Jaber, E. (2019c). Weak existence and uniqueness for affine stochastic Volterra equations with L1-kernels. *arXiv preprint arXiv:1912.07445*.

Abi Jaber, E. and El Euch, O. (2019a). Markovian structure of the Volterra heston model. *Statistics & Probability Letters*, 149:63–72.

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

Abi Jaber, E., Larsson, M., Pulido, S., et al. (2019). Affine Volterra processes. *The Annals of Applied Probability*, 29(5):3155–3200.

Abi Jaber, E., Miller, E., and Pham, H. (2020). Markowitz portfolio selection for multivariate affine and quadratic Volterra models. *arXiv preprint arXiv:2006.13539*.

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.

Andersen, T. G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of empirical finance*, 4(2-3):115–158.

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

Bennedsen, M., Lunde, A., and Pakkanen, M. S. (2016). Decoupling the short-and long-term behavior of stochastic volatility. *arXiv preprint arXiv:1610.00332*.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654.

Bornemann, F. (2009). On the numerical evaluation of distributions in random matrix theory: a review. *arXiv preprint arXiv:0904.1581*.

Bornemann, F. (2010). On the numerical evaluation of Fredholm determinants. *Mathematics of Computation*, 79(270):871–915.

Callegaro, G., Grasselli, M., and Pages, G. (2018). Rough but not so tough: fast hybrid schemes for fractional Riccati equations. *arXiv preprint arXiv:1805.12587*.

Carr, P. and Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of computational finance*, 2(4):61–73.

Comte, F. and Renault, E. (1998). Long memory in continuous-time stochastic volatility models. *Mathematical finance*, 8(4):291–323.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues.

Corlay, S. (2010). The Nyström method for functional quantization with an application to the fractional Brownian motion. *arXiv preprint arXiv:1009.1241*.

Cox, J. C., Ingersoll Jr, J. E., and Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica: Journal of the Econometric Society*, pages 363–384.

Cuchiero, C. and Teichmann, J. (2019). Markovian lifts of positive semidefinite affine Volterra type processes. *Decisions in Economics and Finance*, 42(2):407–448.

Cuchiero, C. and Teichmann, J. (2020). Generalized Feller processes and markovian lifts of stochastic Volterra processes: the affine case. *Journal of Evolution Equations*, pages 1–48.

Decreusefond, L. and Ustunel, A. S. (1999). Stochastic analysis of the fractional Brownian motion. *Potential analysis*, 10(2):177–214.

Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1):83–106.

Duffie, D., Filipović, D., and Schachermayer, W. (2003). Affine processes and applications in finance. *Ann. Appl. Probab.*, 13(3):984–1053.

El Euch, O. and Rosenbaum, M. (2019). The characteristic function of rough Heston models. *Mathematical Finance*, 29(1):3–38.

Fang, F. and Oosterlee, C. W. (2009). A novel pricing method for european options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, 31(2):826–848.

Fredholm, I. (1903). Sur une classe déquations fonctionnelles. *Acta mathematica*, 27(1):365–390.

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

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

Gatheral, J. and Keller-Ressel, M. (2019). Affine forward variance models. *Finance and Stochastics*, 23(3):501–533.

Gatheral, J. and Radoičić, R. (2019). Rational approximation of the rough Heston solution. *International Journal of Theoretical and Applied Finance*, 22(03):1950010.

Gripenberg, G., Londen, S.-O., and Staffans, O. (1990). *Volterra integral and functional equations*, volume 34 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge.

Gulisashvili, A., Viens, F., and Zhang, X. (2019). Extreme-strike asymptotics for general Gaussian stochastic volatility models. *Annals of Finance*, 15(1):59–101.

Harms, P. and Stefanovits, D. (2019). Affine representations of fractional processes with applications in mathematical finance. *Stochastic Processes and their Applications*, 129(4):1185 – 1228.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343.

Horvath, B., Jacquier, A., and Lacombe, C. (2019). Asymptotic behaviour of randomised fractional volatility models. *Journal of Applied Probability*, 56(2):496–523.

Kac, M. and Siegert, A. J. (1947). On the theory of noise in radio receivers with square law detectors. *Journal of Applied Physics*, 18(4):383–397.

Kang, S.-Y., Koltracht, I., and Rawitscher, G. (2003). Nyström-Clenshaw-Curtis quadrature for integral equations with discontinuous kernels. *Mathematics of computation*, 72(242):729–756.

Karhunen, K. (1946). Zur spektraltheorie stochastischer prozesse. *Ann. Acad. Sci. Fennicae, AI*, 34.

Lewis, A. L. (2001). A simple option formula for general jump-diffusion and other exponential lévy processes. *Available at SSRN 282110.*

Loeve, M. (1955). Probability theory: foundations, random sequences.

Lord, R. and Kahl, C. (2006). Why the rotation count algorithm works.

Malyarenko, A. (2012). *Invariant random fields on spaces with a group action.* Springer Science & Business Media.

Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion*, volume 293. Springer-Verlag, Berlin, third edition.

Schöbel, R. and Zhu, J. (1999). Stochastic volatility with an Ornstein–Uhlenbeck process: an extension. *Review of Finance*, 3(1):23–46.

Shorack, G. R. and Wellner, J. A. (2009). *Empirical processes with applications to statistics.* SIAM.

Simon, B. (1977). Notes on infinite determinants of Hilbert space operators. *Advances in Mathematics*, 24(3):244–273.

Smithies, F. (1958). Integral equations.

Sottinen, T. and Viitasaari, L. (2016). Stochastic analysis of Gaussian processes via Fredholm representation. *International journal of stochastic analysis*, 2016.

Stein, E. M. and Stein, J. C. (1991). Stock price distributions with stochastic volatility: an analytic approach. *The review of financial studies*, 4(4):727–752.