



**HAL**  
open science

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

Eduardo Abi Jaber

► **To cite this version:**

Eduardo Abi Jaber. The characteristic function of Gaussian stochastic volatility models: an analytic expression. Finance and Stochastics, 2022. hal-02946146v3

**HAL Id: hal-02946146**

**<https://hal.science/hal-02946146v3>**

Submitted on 9 May 2022

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

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

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

Eduardo Abi Jaber\*

May 9, 2022

## 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. 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. I would like to thank Shaun Li for interesting discussions.

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:

$$dS_t = S_t \sigma dB_t, \quad S_0 > 0, \quad (1.1)$$

with  $B$  a standard Brownian motion and  $\sigma$  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  $\sigma$  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

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

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 instantaneous variance process  $V_t = \sigma_t^2$  in the Stein–Stein model with  $\theta = 0$  follows a CIR process thanks to Itô’s formula,<sup>1</sup> so that the Stein–Stein model can be recast in the following form

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

---

<sup>1</sup>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)).<sup>2</sup> 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, continuous semi-martingale models driven by a standard Brownian motion fail to reproduce the power-law decay of the at-the-money skew of option prices as shown on Figure 1-(b), see Carr and Wu (2003); Fouque et al. (2003); Lee (2005); Alòs et al. (2007); Bayer et al. (2016); Fukasawa (2011, 2021).

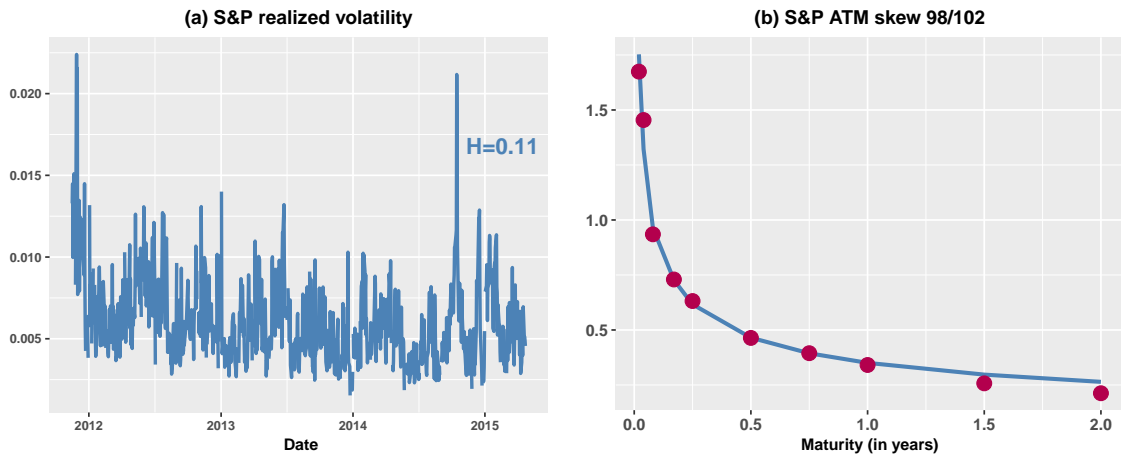


Figure 1: (a) Realized volatility of the S&P downloaded from <https://realized.oxford-man.ox.ac.uk/> with an estimated Hurst index of  $\hat{H} = 0.11$ . (b) Term structure of the at-the-money skew of the implied volatility  $\left. \frac{\partial \sigma_{\text{implicit}}(k, T)}{\partial k} \right|_{k=0}$  for the S&P index on June 20, 2018 (red dots) and a power-law fit  $t \rightarrow 0.35 \times t^{-0.41}$ . Here  $k := \ln(K/S_0)$

<sup>2</sup>The long-term level of the variance  $\nu^2$  in (1.3) can be replaced by a more general coefficient  $\theta \geq 0$ .

stands for the log-moneyness and  $T$  for the time to maturity.

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.<sup>3</sup> 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

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

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 \(2021\)](#), 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,

---

<sup>3</sup>Refer to <https://sites.google.com/site/roughvol/home> for references.

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  $H = 1/2$ .

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, Schöbel–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. (2021) 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:

$$dS_t = S_t X_t dB_t, \quad S_0 > 0, \quad (1.5)$$

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

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,  $K : [0, T]^2 \rightarrow \mathbb{R}$  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  $W$ . In this case, the stochastic integral  $\int_0^\cdot X dB$  cannot be defined in a dynamical way as an Itô integral whenever  $\rho \neq 0$ . We will make sense of (1.5)–(1.6) in a static sense in Section 2.
- 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 and for which the stochastic integral  $\int_0^\cdot X dB$  can be defined in the usual Itô sense, see Section 3. 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 suitable  $u, w \in \mathbb{C}$ , we provide the following analytical expression for the conditional joint Fourier–Laplace transform of the log-price and the integrated variance:

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

with  $\langle f, h \rangle_{L^2} = \int_0^T f(s)h(s)ds$ ,  $\det$  the [Fredholm \(1903\)](#) determinant (see [Appendix A.1](#)),  $g_t$  the adjusted conditional mean given by

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

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

$$\Psi_t = (\text{id} - b\mathbf{K}^*)^{-1} a \left( \text{id} - 2a\tilde{\Sigma}_t \right)^{-1} (\text{id} - b\mathbf{K})^{-1}, \quad t \leq T, \quad (1.9)$$

where  $\mathbf{K}$  denotes the integral operator induced by  $K$ ,  $\mathbf{K}^*$  the adjoint operator,<sup>4</sup>  $\text{id}$  denotes the identity operator, i.e.  $(\text{id}f) = f$  for all  $f \in L^2([0, T], \mathbb{C})$ ,

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

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

$$\tilde{\Sigma}_t = (\text{id} - b\mathbf{K})^{-1} \Sigma_t (\text{id} - b\mathbf{K}^*)^{-1}, \quad (1.11)$$

with  $\Sigma_t$  defined as the integral operator associated with the covariance kernel

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

and finally  $\Phi$  is defined by

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

---

<sup>4</sup>cf. below for detailed notations.

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

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

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  $\Psi_0$  appears naturally after a rearrangement of the terms. The main result is collected in Theorem 2.3.
- **Volterra kernels:** under some  $L^2$ -continuity and boundedness condition, 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. This is the class of kernels which is more suited for financial applications.

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 in the form

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

where  $g_n \in \mathbb{R}^n$  and  $\Phi_0^n, \Psi_0^n \in \mathbb{R}^{n \times n}$  are entirely determined by  $(g_0, K, \nu, \kappa, u, w)$  and  $\det$  is the standard determinant of a matrix, we refer to Section 4.1. 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 4.2.

## 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

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

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

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



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

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

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

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

$\mathbf{K}$  is a linear bounded operator from  $L^2([0, T], \mathbb{K})$  into itself. If  $\mathbf{K}$  and  $\mathbf{L}$  are two integral operators induced by the kernels  $K$  and  $L$  in  $L^2([0, T]^2, \mathbb{K})$ , then  $\mathbf{K}\mathbf{L}$  is also an 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  $\mathbf{K}^*$  the corresponding adjoint integral operator.

The square-root of a complex number  $\sqrt{z}$  is defined through its main branch, i.e.  $\sqrt{z} = |z|e^{i\arg(z)/2}$  with  $z = |z|e^{i\arg(z)}$  such that  $\arg(z) \in (-\pi, \pi]$ .

## 2 Symmetric kernels: an elementary static approach

We provide an elementary static derivation of the joint Fourier–Laplace transform in the special case of symmetric kernels with  $\kappa = 0$ . We stress that, although the case of symmetric kernels is not of interest for practical applications, it naturally leads through direct computations to the analytic expression (1.7) in terms of the operator  $\Psi$  given in (1.9). Later on, in Section 3, such expressions are shown to hold in the more practical case of Volterra kernels using a dynamical approach.

**Definition 2.1.** *A linear operator  $\mathbf{K}$  from  $L^2([0, T], \mathbb{R})$  into itself is symmetric nonnegative if  $\mathbf{K} = \mathbf{K}^*$  and  $\langle f, \mathbf{K}f \rangle_{L^2} \geq 0$ , for all  $f \in L^2([0, T], \mathbb{R})$ . Whenever  $\mathbf{K}$  is an integral operator induced by some kernel  $K \in L^2([0, T]^2, \mathbb{R})$ , we will say that  $K$  is symmetric nonnegative. In this case, it follows that  $K = K^*$  a.e. and*

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

$\mathbf{K}$  is said to be symmetric nonpositive, if  $(-\mathbf{K})$  is symmetric nonnegative.

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

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

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  $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$  with  $\lambda_n \rightarrow 0$ , as  $n \rightarrow \infty$ , such that

$$\sum_{n \geq 1} \lambda_n < \infty. \quad (2.2)$$

Such decomposition is possible whenever the operator  $\mathbf{K}$  is the (nonnegative symmetric) square-root of a covariance operator  $\mathbf{C}$  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  $\mathbf{C}$  admits the representation (1.6) with  $\kappa = 0$  on some filtered probability space supporting a Brownian motion  $W$ , see [Sottinen and Viitasaari \(2016, Theorem 12\)](#).

We start by making precise how one should understand (1.5)–(1.6) in the case of symmetric kernels and  $\kappa = 0$ . We rewrite (1.5) in the equivalent form

$$\log S_t = \log S_0 - \frac{1}{2} \int_0^t X_s^2 ds + \rho \int_0^t X_s dW_s + \sqrt{1 - \rho^2} \int_0^t X_s dW_s^\perp. \quad (2.3)$$

We fix  $T > 0$ ,  $g_0 \in L^2([0, T], \mathbb{R})$  and a complete probability space  $(\Omega, \mathcal{F}, \mathbb{Q})$  supporting a two dimensional Brownian motion  $(W, W^\perp)$  and, for each  $t \leq T$ , we set

$$X_t = g_0(t) + \int_0^t K(t, s) \nu dW_s.$$

We note that (2.1)–(2.2) imply that  $K \in L^2([0, T]^2, \mathbb{R})$  so that the stochastic integral  $\int_0^t K(t, s) \nu dW_s$  is well-defined as an Itô integral for almost every  $t \leq T$  and  $X$  has sample paths in  $L^2([0, T], \mathbb{R})$  almost surely. Setting  $\mathcal{F}_t = \mathcal{F}_t^X \vee \mathcal{F}_t^{W^\perp}$  where  $(\mathcal{F}_t^Y)_{t \geq 0}$  stands for the filtration generated by the process  $Y$ , we have that  $W^\perp$  is still a Brownian motion with respect to  $(\mathcal{F}_t)_{t \geq 0}$  and, up to a modification,  $X$  is progressively measurable<sup>5</sup> with respect to  $(\mathcal{F}_t^X)_{t \geq 0}$  (and whence w.r.t. the enlarged filtration  $(\mathcal{F}_t)_{t \geq 0}$ ) so that

$$\int_0^\cdot X_s dW_s^\perp$$

---

<sup>5</sup>Every jointly measurable and adapted process admits a progressively measurable modification, see [Ondreját and Seidler \(2013\)](#).

is well defined as an Itô integral with respect to  $(\mathcal{F}_t)_{t \geq 0}$ . If  $\rho = 0$ , (2.3) is therefore well-defined in the classical way. However, for  $\rho \neq 0$ , since  $X$  is not necessarily adapted to the filtration generated by  $W$  (and vice versa),  $W$  is no longer necessarily a Brownian motion with respect to the extended filtration  $\mathcal{F}^X \vee \mathcal{F}^W$ , and one cannot make sense of the stochastic integral  $\int_0^\cdot X dW$  in the usual dynamical sense. We provide a static interpretation of (2.3) valid only at the terminal time  $T$ . To this end, since  $g_0 \in L^2([0, T], \mathbb{R})$ , we can write  $g_0 = \sum_{n \geq 1} \langle g_0, e_n \rangle e_n$ . Making use of (2.1), we first observe that, an application of Fubini's theorem (Veraar, 2012, Theorem 2.2), justified by the fact that

$$\int_0^T \sum_{n \geq 1} \mathbb{E} \left[ \int_0^T |\sqrt{\lambda_n} e_n(t) e_n(s)|^2 ds \right] dt = \sum_{n \geq 1} \lambda_n \int_0^T e_n(t)^2 dt \leq \sum_{n \geq 1} \lambda_n < \infty,$$

yields that

$$X_t = g_0(t) + \int_0^T K(t, s) \nu dW_s = \sum_{n \geq 1} \left( \langle g_0, e_n \rangle + \sqrt{\lambda_n} \nu \xi_n \right) e_n(t), \quad dt \otimes \mathbb{Q} - a.e. \quad (2.4)$$

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 that are  $\mathcal{F}_T^W$  measurable. We set

$$N_T = \sum_{n \geq 1} \left( \langle g_0, e_n \rangle + \sqrt{\lambda_n} \nu \xi_n \right) \xi_n. \quad (2.5)$$

**Remark 2.2.** We note that  $N_T$  plays the role of  $\int_0^T X_s dW_s$ , since a formal interchange leads to

$$\begin{aligned} N_T &= \sum_{n \geq 1} \left( \langle g_0, e_n \rangle + \sqrt{\lambda_n} \nu \xi_n \right) \int_0^T e_n(s) dW_s \\ & (= ) \int_0^T \sum_{n \geq 1} \left( \langle g_0, e_n \rangle + \sqrt{\lambda_n} \nu \xi_n \right) e_n(s) dW_s \\ & (= ) \int_0^T X_s dW_s. \end{aligned}$$

Obviously, since  $\xi_n$  are not adapted the integral  $\int_0^\cdot \xi_n e_n(s) dW_s$  cannot be defined in the non-anticipative sense.

Finally, we take as definition for the log-price at the terminal time  $T$ :

$$\log S_T = \log S_0 - \frac{1}{2} \int_0^T X_s^2 ds + \rho N_T + \sqrt{1 - \rho^2} \int_0^T X_s dW_s^\perp, \quad S_0 > 0, \quad (2.6)$$

which is an  $\mathcal{F}_T^W \vee \mathcal{F}_T$ -measurable random variable.

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

**Theorem 2.3.** 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,

$$\mathbb{E} \left[ \exp \left( u \log \frac{S_T}{S_0} + w \int_0^T X_s^2 ds \right) \right] = \frac{\exp(\langle g_0, \Psi_0 g_0 \rangle_{L^2})}{\det(\Phi_0)^{1/2}}, \quad (2.7)$$

with  $\Psi_0$  and  $\tilde{\Sigma}_0$  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  $\Phi_0 = (\text{id} - b\mathbf{K})(\text{id} - 2a\tilde{\Sigma}_0)(\text{id} - b\mathbf{K})$ .

The rest of the section is dedicated to the proof of Theorem 2.3. 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  $\tilde{\Sigma}_0$  and  $\Psi_0$  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  $\left( \int_0^T X_s^2 ds, N_T \right)$ . Furthermore, the representation (2.4) readily leads to

$$\int_0^T X_s^2 ds = \sum_{n \geq 1} \left( \langle g_0, e_n \rangle + \sqrt{\lambda_n} \nu \xi_n \right)^2. \quad (2.8)$$

**Lemma 2.4.** 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

$$\Re(\alpha) \leq 0, \quad \Re(\beta) = 0. \quad (2.9)$$

Then,

$$\mathbb{E} \left[ \exp \left( \alpha \int_0^T X_s^2 ds + \beta N_T \right) \right] = \frac{\exp \left( \left( \alpha + \frac{\beta^2}{2} \right) \sum_{n \geq 1} \frac{\langle g_0, e_n \rangle^2}{1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n} \right)}{\prod_{n \geq 1} \sqrt{1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n}}. \quad (2.10)$$

*Proof.* Define  $U_T = \alpha \int_0^T X_s^2 ds + \beta N_T$ . We first observe that (2.9) 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 (2.5) and (2.8), we have

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

where  $\tilde{\xi}_n = (\langle g_0, e_n \rangle + \nu\sqrt{\lambda_n}\xi_n)$ , 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  $\mu_n$  and covariance matrix  $\Sigma_n$  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

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

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 [\(2.10\)](#).  $\square$

Relying on the spectral decomposition [\(2.1\)](#), we re-express the quantities entering in [\(2.10\)](#) in terms of suitable operators.

**Lemma 2.5.** Let  $K$  be as in (2.1), set  $\kappa = 0$  and fix  $\alpha, \beta \in \mathbb{C}$  as in (2.9). Then, the following operator defined by (1.9) with  $a = \alpha + \frac{\beta^2}{2}$  and  $b = \nu\beta$ :

$$\Psi_0^{\alpha, \beta} = (\text{id} - b\mathbf{K}^*)^{-1} a \left( \text{id} - 2\tilde{\Sigma}_0 a \right)^{-1} (\text{id} - b\mathbf{K})^{-1}, \quad t \leq T,$$

admits the following decomposition

$$\Psi_0^{\alpha, \beta} = \sum_{n \geq 1} \frac{\alpha + \frac{\beta^2}{2}}{1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n} \langle e_n, \cdot \rangle e_n \quad (2.11)$$

and

$$\det \left( \frac{1}{\alpha + \frac{\beta^2}{2}} \Psi_0^{\alpha, \beta} \right) = \prod_{n \geq 1} \frac{1}{1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n}, \quad (2.12)$$

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

$$\mathbb{E} \left[ \exp \left( \alpha \int_0^T X_s^2 ds + \beta N_T \right) \right] = \det \left( \frac{1}{\alpha + \frac{\beta^2}{2}} \Psi_0^{\alpha, \beta} \right)^{1/2} \exp \left( \langle g_0, \Psi_0^{\alpha, \beta} g_0 \rangle \right). \quad (2.13)$$

*Proof.* Throughout the proof, we will make use of the following rule for computing the decomposition of a product of operators in terms of the orthonormal basis  $(e_n)_{n \geq 1}$ : for  $\mathbf{K}$  and  $\mathbf{L}$  in the form

$$\mathbf{K} = \sum_{n \geq 1} a_n \langle e_n, \cdot \rangle e_n, \quad \mathbf{L} = \sum_{n \geq 1} b_n \langle e_n, \cdot \rangle e_n,$$

the composition is given by

$$\mathbf{K}\mathbf{L} = \sum_{n \geq 1} a_n \langle e_n, \sum_{m \geq 1} b_m e_m \langle e_n, \cdot \rangle \rangle e_n = \sum_{n \geq 1} a_n b_n \langle e_n, \cdot \rangle e_n.$$

It follows from (2.1) that

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

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

$$(\text{id} - b\mathbf{K})^{-1} = \sum_{n \geq 1} \frac{1}{1 - b\sqrt{\lambda_n}} \langle e_n, \cdot \rangle e_n. \quad (2.14)$$

Similarly, recalling (1.12), (2.1) leads to the representation of  $\Sigma_0 = \nu^2 \mathbf{K} \mathbf{K}^*$ :

$$\Sigma_0 = \sum_{n \geq 1} \nu^2 \lambda_n \langle e_n, \cdot \rangle e_n,$$

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

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

Whence,

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

Recalling that  $a = \alpha + \frac{\beta^2}{2}$  and  $b = \nu\beta$ ,  $\left( (1 - b\sqrt{\lambda_n})^2 - 2a\nu^2 \lambda_n \right) = 1 - 2\nu\beta\sqrt{\lambda_n} - 2\alpha\nu^2 \lambda_n$ . 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  $\left( \text{id} - 2a\tilde{\Sigma}_0 \right)$  is invertible with an inverse given by

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

The representations (2.11)-(2.12) readily follows after composing by  $(\text{id} - b\mathbf{K}^*)^{-1}a$  from the left, by  $(\text{id} - b\mathbf{K})^{-1}$  from the right and recalling (2.14). Finally, combining these expressions with (2.10), we obtain (2.13). This ends the proof.  $\square$

We can now complete the proof of Theorem 2.3.

*Proof of Theorem 2.3.* It suffices to prove that

$$\mathbb{E} \left[ \exp \left( u \log \frac{S_T}{S_0} + w \int_0^T X_s^2 ds \right) \right] = \mathbb{E} \left[ \exp \left( \alpha \int_0^T X_s^2 ds + \beta N_T \right) \right], \quad (2.15)$$

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 2.5 yields the expression (2.7).

It remains to prove (2.15) by means of a projection argument. Conditional on  $(\mathcal{F}_t^X \vee \mathcal{F}_t^W)_{t \leq T}$ , by independence of  $X$  and  $W^\perp$ , the random variable  $\int_0^T X_s dW_s^\perp$  is centered gaussian with variance  $\int_0^T X_s^2 ds$  so that

$$\begin{aligned} M_T &:= \mathbb{E} \left[ \exp \left( u \sqrt{1 - \rho^2} \int_0^T X_s dW_s^\perp \right) \mid (\mathcal{F}_t^X \vee \mathcal{F}_t^W)_{t \leq T} \right] \\ &= \exp \left( \frac{u^2(1 - \rho^2)}{2} \int_0^T X_s^2 ds \right). \end{aligned} \quad (2.16)$$

A successive application of the tower property of the conditional expectation on the expression (2.6) yields that

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

leading to (2.15) due to (2.16). This ends the proof.  $\square$

### 3 Volterra kernels: a dynamical approach

In this section, we treat the class of Volterra kernels which are practically relevant in mathematical finance. 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  $K : [0, T]^2 \rightarrow \mathbb{R}$  is a Volterra kernel of continuous and bounded type in  $L^2$  if  $K(t, s) = 0$  whenever  $s \geq t$  and*

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

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  $k \in L^2([0, T], \mathbb{R})$ . Indeed,

$$\sup_{t \leq T} \int_0^T |K(t, s)|^2 ds = \sup_{t \leq T} \int_0^t |k(t - s)|^2 ds \leq \int_0^T |k(s)|^2 ds < \infty$$

yielding the first part of (3.1). The second part follows from the  $L^2$ -continuity of  $k$ , see (Brezis, 2010, Lemma 4.3).



(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 fractional Brownian motion whose covariance function is  $\Sigma_0(s, u) = (K \star K^*)(s, u) = \frac{1}{2}(s^{2H} + u^{2H} - |s - u|^{2H})$ , see [Decreusefond and Ustunel \(1999\)](#). In this case,

$$\sup_{t \leq T} \int_0^T |K(t, s)|^2 ds = \sup_{t \leq T} \Sigma_0(t, t) \leq T^{2H}$$

and by developing the square

$$\int_0^T |K(u+h, s) - K(u, s)|^2 ds = \Sigma_0(u+h, u+h) - 2\Sigma_0(u+h, u) + \Sigma_0(u, u)$$

which goes to 0 as  $h \rightarrow 0$ .

- (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  $K_1$  and  $K_2$  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

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

we refer to Theorem A.3 below for the proof. It follows in particular from (3.2) that  $\int_0^T X_s^2 ds < \infty$  almost surely, so that  $X$  has sample paths in  $L^2([0, T], \mathbb{R})$ .

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,

$$\mathbb{E} \left[ \exp \left( u \log \frac{S_T}{S_t} + w \int_t^T X_s^2 ds \right) \middle| \mathcal{F}_t \right] = \exp(\phi_t + \langle g_t, \Psi_t g_t \rangle_{L^2}), \quad (3.4)$$

for all  $t \leq T$ , with  $\Psi_t$  given by (1.9) for  $(a, b)$  as in (1.10) and

$$\phi_t = - \int_t^T \text{Tr}(\Psi_t \dot{\Sigma}_t) dt, \quad (3.5)$$

where  $\dot{\Sigma}_t$  is the strong derivative<sup>6</sup> of  $t \mapsto \Sigma_t$  induced by the kernel

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

and  $\text{Tr}$  is the trace operator, see Appendix A.1.

*Proof.* We refer to Appendix B. □

The following remark establishes the link between  $\phi$  and the Fredholm determinant.

**Remark 3.4.** Assume  $u, w$  are real. We recall the definition

$$\Phi_t = \text{id} - 2\tilde{\Sigma}_t a, \quad t \leq T,$$

and that  $\tilde{\Sigma}_t$  is an integral operator of trace class with continuous kernel by virtue of Lemma A.5 below so that the determinant  $\det(\Phi_t)$  is well defined and non-zero by the invertibility of  $(\text{id} - 2\tilde{\Sigma}_t a)$ , see Lemma A.6 and Simon (1977, Theorem 3.9). We set

$$\phi_t = \log(\det(\Phi_t)^{-1/2}) = -\frac{1}{2} \log(\det(\Phi_t)). \quad (3.6)$$

Differentiation using the logarithmic derivative of the Fredholm's determinant (see (Gohberg and Krein, 1978, Chap IV, p.158 (1.3))) and (1.11) yields

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

Finally, using (1.9) and the identity  $\text{Tr}(\mathbf{F}\mathbf{G}) = \text{Tr}(\mathbf{G}\mathbf{F})$ , we obtain

$$\dot{\phi}_t = \text{Tr}(\Psi_t \dot{\Sigma}_t). \quad (3.7)$$

When  $u, w$  are complex numbers, the definition of (3.6) requires the use of several branches of the complex logarithm. For numerical implementation, to prevent complex discontinuities, one should either use (3.6) with multiple branches or stick with the discretization of expression (3.5). We refer to section 4.1 for the numerical implementation.

---

<sup>6</sup>See Lemma B.1 below.

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

$$g_0(t) = X_0 + \theta t, \quad t \geq 0, \quad (3.8)$$

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 3.5.** *Assume that  $K(t, s) = \mathbf{1}_{s < t}$  and that  $g_0$  is of the form (3.8), then, the expression (3.4) reduces to*

$$\mathbb{E} \left[ \exp \left( u \log \frac{S_T}{S_t} + w \int_t^T X_s^2 ds \right) \middle| \mathcal{F}_t \right] = \exp (A(t) + B(t)X_t + C(t)X_t^2) \quad (3.9)$$

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

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

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 (3.8). 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

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

It follows that

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

with

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

Combined with (3.6) and (B.7) below, we obtain (3.9) with  $A$  such that  $A_T = 0$  and

$$\dot{A}(t) = \dot{\tilde{A}}(t) + \text{Tr}(\Psi_t \dot{\Sigma}_t)$$

with  $\text{Tr}$  the trace of an operator (see Appendix A.1 below) and

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

Using the operator Riccati equation satisfied by  $t \mapsto \Psi_t$ , see Lemma B.1 below, and straightforward computations as in Abi Jaber et al. (2021, Corollary 5.14) lead to the claimed system of Riccati equations for  $(A, B, C)$ .  $\square$

## 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 an approximation of the formula (1.7) using closed form expressions obtained from a natural discretization of the operators. Throughout this section, we consider the case  $t = 0$  and we fix a Volterra kernel  $K$ , i.e.  $K(t, s) = 0$  if  $s \geq t$ , as in Section 3.

### 4.1 A straightforward approximation by closed form expressions

The expression (1.7) lends itself to approximation by closed form solutions by a simple discretization of the operator  $\Psi_0$  given by (1.9) à la Fredholm (1903). Fix  $n \in \mathbb{N}$  and let  $t_i = iT/n$ ,  $i = 0, 1, \dots, n$  be a partition of  $[0, T]$ . Discretizing the  $\star$ -product given in (1.13) yields the following approximation for  $\Psi_0$  by the  $n \times n$  matrix:

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

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

$$K_{ij}^n = \mathbf{1}_{j \leq i-1} \int_{t_{j-1}}^{t_j} K(t_{i-1}, s) ds, \quad 1 \leq i, j \leq n, \quad (4.1)$$

and

$$\tilde{\Sigma}^n = (I_n - bK^n)^{-1} \Sigma^n \left( I_n - b(K^n)^\top \right)^{-1}$$

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

$$\Sigma_{ij}^n = \nu^2 \int_0^T K(t_{i-1}, s) K(t_{j-1}, s) ds, \quad 1 \leq i, j \leq n. \quad (4.2)$$

Defining the  $n$ -dimensional vector  $g_n = (g_0(t_0), \dots, g_0(t_{n-1}))^\top$ , the discretization of the inner product  $\langle \cdot, \cdot \rangle_{L^2}$  leads to the approximation

$$\mathbb{E} \left[ \exp \left( u \log S_T + w \int_0^T X_s^2 ds \right) \right] \approx \frac{\exp \left( u \log S_0 + \frac{T}{n} g_n^\top \Psi_0^n g_n \right)}{\det(\Phi_0^n)^{1/2}} \quad (4.3)$$

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

**Remark 4.1.** Recalling Remark 3.4, one needs to be careful with the numerical implementation of the square root of the determinant that appears in equation (4.3) to avoid complex discontinuities, either by switching the sign of the determinant each time it crosses the axis of negative real numbers or by discretizing (3.5) which would require the computation of  $\Psi_t$  for several values of  $t$  but has the advantage of being analytic on the whole domain. We refer to Mayerhofer (2019) for more details for finite-dimensional Wishart distributions.

**Remark 4.2.** 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).

**Remark 4.3.** For the case  $u = 0$  and  $\kappa = 0$ , the previous approximation formulas agree with the ones derived in (Abi Jaber, 2019a, Section 2.3) where a numerical illustration for the integrated squared fractional Brownian motion is provided.

## 4.2 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  $K(t, s) = \mathbf{1}_{s < t}(t - s)^{H-1/2}/\Gamma(H + 1/2)$ :

$$dS_t = S_t X_t dB_t, \quad S_0 > 0,$$

$$X_t = g_0(t) + \frac{\kappa}{\Gamma(H + 1/2)} \int_0^t (t - s)^{H-1/2} X_s ds + \frac{\nu}{\Gamma(H + 1/2)} \int_0^t (t - s)^{H-1/2} dW_s,$$

with  $B = \rho W + \sqrt{1 - \rho^2} W^\perp$ , for  $\rho \in [-1, 1]$ ,  $\kappa, \nu \in \mathbb{R}$  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<sup>7</sup>

$$g_0(t) = X_0 + \frac{1}{\Gamma(H + 1/2)} \int_0^t (t - s)^{H-1/2} \theta ds = X_0 + \theta \frac{t^{H+1/2}}{\Gamma(H + 1/2)(H + 1/2)}. \quad (4.4)$$

**Remark 4.4.** It would have also been possible to take instead of the fractional Riemann–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.

---

<sup>7</sup>In conventional Markovian stochastic volatility models, the input curve  $g_0$  is usually in the parametric form (4.4). However, if one is interested in a practical implementation, then more general forms of  $g_0$  (non-parametric) would allow more flexibility (by making  $\theta$  time dependent for instance). The advantage is that  $g_0$  can be estimated from the market to match certain term structures today (e.g. term structure of forward variance, etc...). For illustration purposes here, and since a comparison with the standard Stein–Stein model is given, we restrict to such parametric forms of  $g_0$ .

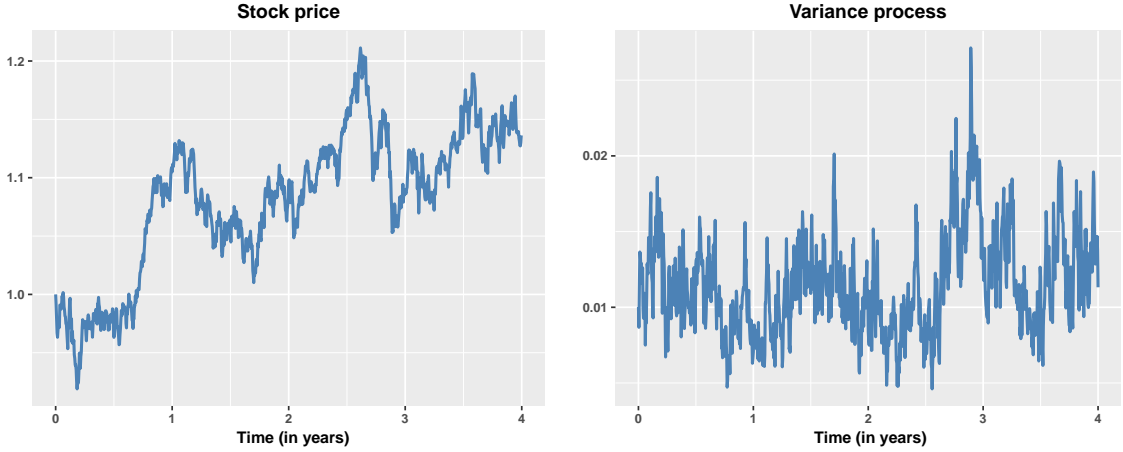


Figure 2: One simulated sample path of the stock price  $S$  and the instantaneous variance process  $X^2$  in the rough Stein–Stein model with parameters:  $X_0 = 0, 1$ ,  $\kappa = 0$ ,  $\theta = 0.01$ ,  $\nu = 0.02$ ,  $\rho = -0.7$  and  $H = 0.1$ .

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:

$$p_T(x, y) = \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{-i(z_1 x + z_2 y)} \mathbb{E} \left[ \exp \left( iz_1 \log S_T + iz_2 \int_0^T X_s^2 ds \right) \right] dz_1 dz_2,$$

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. We start by observing that the kernel  $\Sigma_0$  is given in the following closed form

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

where  $\alpha = H + 1/2$  and  ${}_2F_1$  is the Gaussian hypergeometric function, see for instance Malyarenko (2012, page 71).<sup>8</sup> Fix  $n \in \mathbb{N}$  and a given partition  $0 = t_0 < t_1 < \dots < t_n = T$ .

<sup>8</sup>Note that in the case of Remark 4.4, the expression for the covariance function simplifies to  $\Sigma_0(s, u) = \frac{\nu^2}{2} (s^{2H} + u^{2H} - |s - u|^{2H})$ .

It follows that the  $n \times n$  matrices (4.1)-(4.2) can be computed in closed form:

$$K_{ij}^n = \mathbf{1}_{j \leq i-1} \frac{1}{\Gamma(1+\alpha)} [(t_{i-1} - t_{j-1})^\alpha - (t_{i-1} - t_j)^\alpha], \quad 1 \leq i, j \leq n,$$

$$\Sigma_{ij}^n = \frac{\nu^2}{\Gamma(\alpha)\Gamma(1+\alpha)} \frac{t_{i-1}^\alpha}{t_{j-1}^{1-\alpha}} {}_2F_1\left(1, 1-\alpha; 1+\alpha; \frac{t_{i-1}}{t_{j-1}}\right), \quad \Sigma_{ji}^n = \Sigma_{ij}^n, \quad 1 \leq i \leq j \leq n,$$

with the convention that  $0/0 = 0$ . We note that  $K^n$  is lower triangular with zeros on the diagonal and that the symmetric matrix  $\Sigma^n$  has zeros on its first row and first column. The final ingredient to compute (4.3) is the vector  $g_n$  whose elements are given by:

$$g_n^i = g_0(t_{i-1}) = X_0 + \theta \frac{t_{i-1}^\alpha}{\Gamma(1+\alpha)}, \quad 1 \leq i \leq n.$$

As a sanity check, we visualize on Figure 3 the convergence of the approximation methods on the implied volatility for  $H = 0.2$  and  $H = 0.5$  with the uniform partition  $t_i = iT/n$ . The benchmark is computed for  $H = 0.5$  via the cosine method with the closed form expressions for the characteristic function of the conventional Stein–Stein model, see Lord and Kahl (2006); and for  $H = 0.2$  using Monte Carlo simulation. The smaller the maturity the faster the convergence. Other discretization rules might turn out to be more efficient and would require less points to achieve the same accuracy, which makes the implementation even faster, recall Remark 4.2. 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.

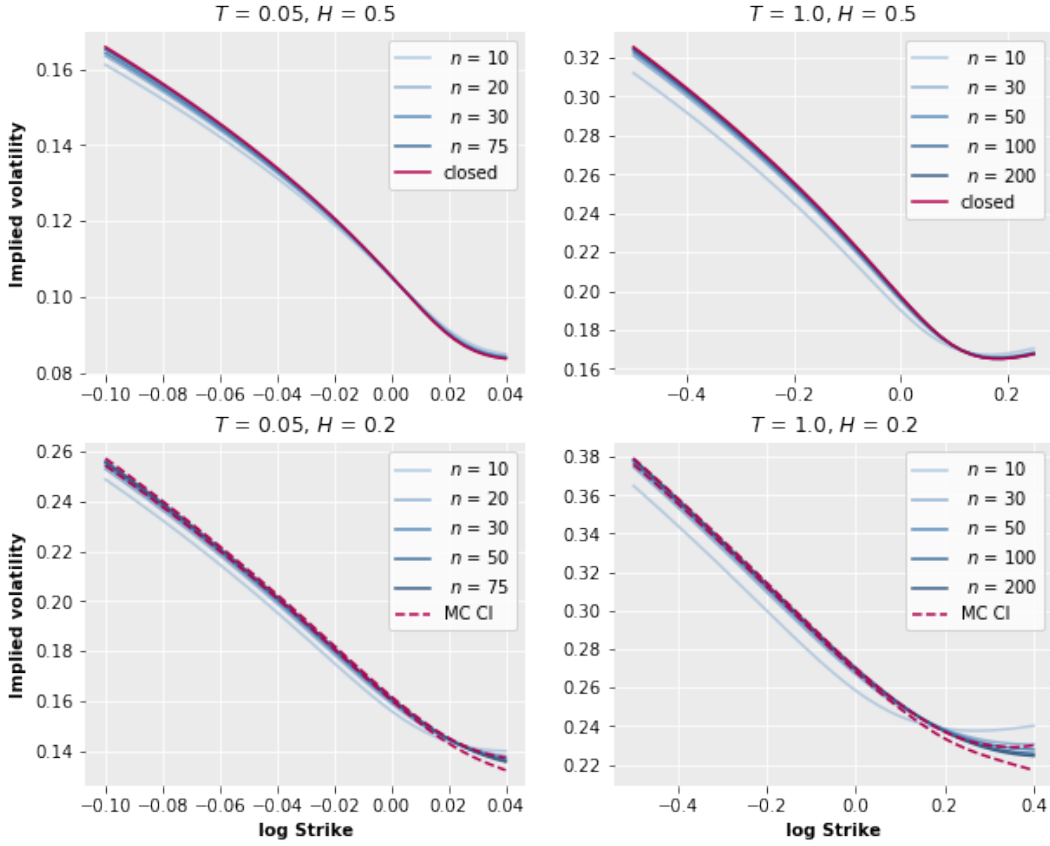


Figure 3: Convergence of the implied volatility slices for short ( $T = 0.05$  year) and long maturities ( $T = 1$  year) of the operator discretization of Section 4.1 towards: (i) the explicit solution of the conventional Stein–Stein model ( $H = 0.5$  upper graphs); (ii) the 95% Monte-Carlo confidence intervals ( $H = 0.2$  lower graphs). The parameters are  $X_0 = \theta = 0.1$ ,  $\kappa = 0$ ,  $\nu = 0.25$  and  $\rho = -0.7$ .

Going back to real market data, we calibrate the fractional Stein–Stein model to

- (i) the at-the-money skew of Figure 1-(b). Keeping the parameters  $X_0 = 0.44$ ,  $\theta = 0.3$ ,  $\kappa = 0$  fixed, the calibrated parameters are given by

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

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.



(ii) the implied volatility surface of the S&P across several maturities for in Figure 5.

Both calibration lead to  $\hat{H} < 0.5$  indicating that the rough regime of the fractional Stein–Stein model is coherent with the observations on the market.

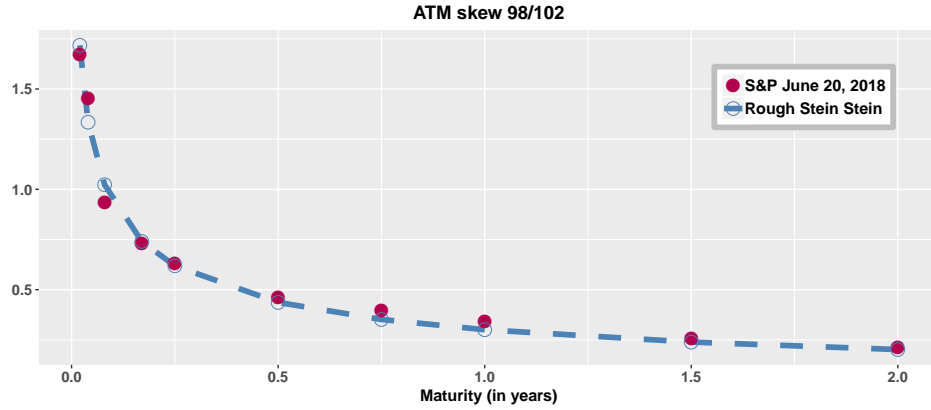


Figure 4: Term structure of the at-the-money skew for the S&P index on June 20, 2018 (red dots) and for the rough Stein–Stein model with calibrated parameters (4.5) (blue circles with dashed line).

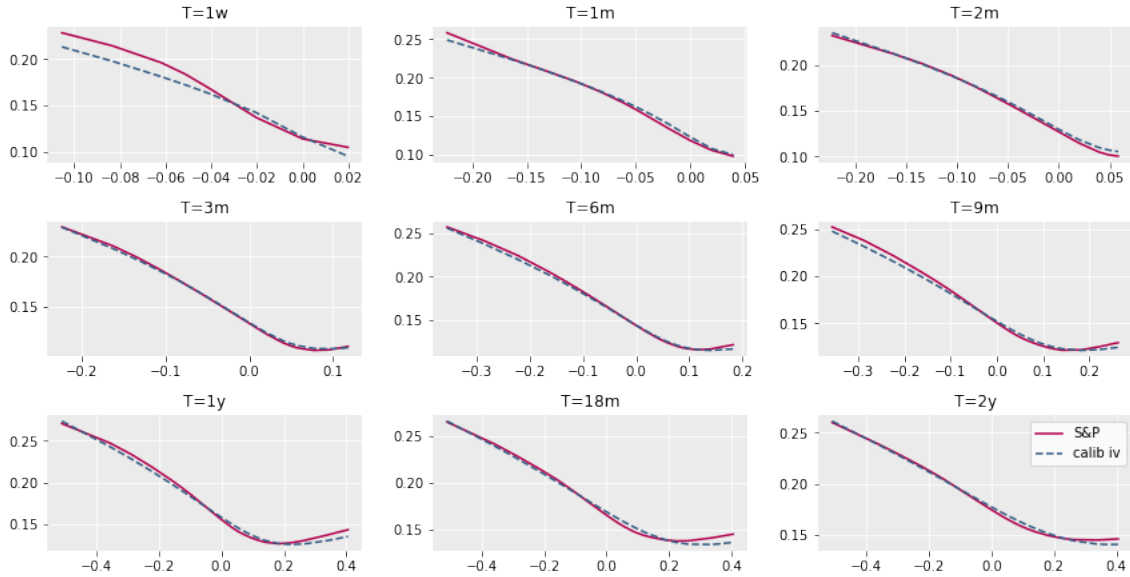


Figure 5: The implied volatility surface of the S&P index (red) and the calibrated fractional

Stein–Stein model (blue) with parameters:  $\hat{X}_0 = 0.113$ ,  $\hat{\theta} = -0.044$ ,  $\hat{\kappa} = -8.9e - 5$ ,  $\hat{\nu} = 0.176$ ,  $\hat{\rho} = -0.704$ , and  $\hat{H} = 0.279$ .

## A Trace, determinants and resolvents

### A.1 Trace and determinants

In this section we recall classical results on operator theory in Hilbert spaces regarding mainly their trace and their determinant. For further details we refer to [Gohberg and Krein \(1978\)](#); [Gohberg et al. \(2012\)](#); [Simon \(1977, 2005\)](#); [Smithies \(1958\)](#), and also [Bornemann \(2009, Section 2 and 3\)](#). Let  $\mathbf{A}$  be a linear compact operator acting on  $L^2([0, T], \mathbb{C})$ . Then, the operator  $\mathbf{A}$  has a countable spectrum<sup>9</sup> denoted by  $\text{sp}(\mathbf{A}) = (\lambda_n(\mathbf{A}))_{n \leq N(\mathbf{A})}$ , where  $N(\mathbf{A})$  is either a finite integer or infinity. Whenever  $\mathbf{A}$  is a linear operator induced by a kernel  $A \in L^2([0, T]^2, \mathbb{C})$ ,  $\mathbf{A}$  is a Hilbert–Schmidt operator on  $L^2([0, T], \mathbb{C})$  into itself and is in particular compact.

The trace and the determinant are two important functionals on the space of compact operators. Such quantities are defined for operators of trace class. A compact operator  $\mathbf{A}$  is said to be of trace class if the quantity

$$\text{Tr } \mathbf{A} = \sum_{n \geq 1} \langle \mathbf{A}v_n, v_n \rangle \tag{A.1}$$

is finite for a given orthonormal basis  $(v_n)_{n \geq 1}$ . It can be shown that the quantity on the right hand side of (A.1) is independent of the choice of the orthonormal basis and will be called the trace of the operator  $\mathbf{A}$ . Furthermore, Lidskii’s theorem ([Simon, 2005](#), Theorem 3.7) ensures that

$$\text{Tr } \mathbf{A} = \sum_{n=1}^{N(\mathbf{A})} \lambda_n(\mathbf{A}).$$

**Remark A.1.** *The product of two Hilbert-Schmidt operators  $\mathbf{K}$  and  $\mathbf{L}$  is of trace class. If in addition, both  $\mathbf{K}$  and  $\mathbf{L}$  are integral operators on  $L^2([0, T])$  induced by  $K$  and  $L$ , then*

$$\text{Tr}(\mathbf{KL}) = \int_0^T (K \star L)(s, s) ds, \tag{A.2}$$

see [Brislaw \(1988, Proposition 3\)](#).

Furthermore, the equivalence

$$\prod_{n \geq 1} (1 + |\lambda_n|) < \infty \Leftrightarrow \sum_{n \geq 1} |\lambda_n| < \infty,$$

---

<sup>9</sup>We recall that the spectrum  $\text{sp}(\mathbf{A})$  is defined as the set of points  $\lambda \in \mathbb{C}$  for which there does not exist a bounded inverse operator  $(\lambda \text{id} - \mathbf{A})^{-1}$ .

allows one to define a determinant functional for a trace class operator  $\mathbf{A}$  by

$$\det(\text{id} + z\mathbf{A}) = \prod_{n=1}^{N(\mathbf{A})} (1 + z\lambda_n(\mathbf{A})),$$

for all  $z \in \mathbb{C}$ . If in addition  $\mathbf{A}$  is an integral operator induced by a continuous kernel  $A$ , then one can show that

$$\det(\text{id} + z\mathbf{A}) = \sum_{n \geq 0} \frac{z^n}{n!} \int_0^T \dots \int_0^T \det [(A(s_i, s_j))_{1 \leq i, j \leq n}] ds_1 \dots ds_n. \quad (\text{A.3})$$

The determinant (A.3) is named after [Fredholm \(1903\)](#) who defined it for the first time for integral operators with continuous kernels.

## A.2 Resolvents

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

$$R_T = K + K \star R_T, \quad K \star R_T = R_T \star K. \quad (\text{A.4})$$

In terms of integral operators, this translates into

$$\mathbf{R}_T = \mathbf{K} + \mathbf{K}\mathbf{R}_T, \quad \mathbf{K}\mathbf{R}_T = \mathbf{R}_T\mathbf{K}. \quad (\text{A.5})$$

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

$$(\text{id} - \mathbf{K})^{-1} = \text{id} + \mathbf{R}_T. \quad (\text{A.6})$$

**Lemma A.2.** *Any  $K$  as in Definition 3.1 admits a resolvent kernel  $R_T$  which is again a Volterra kernel and satisfies (3.1).*

*Proof.* It follows from (3.1) that  $K$  is a Volterra kernel of continuous and bounded type in the terminology of [Gripenberg et al. \(1990, Definitions 9.5.1 and 9.5.2\)](#). But since, we are considering kernels on the compact set  $[0, T]$ , then every kernel of bounded and continuous type is of bounded and uniformly continuous type, see [Gripenberg et al. \(1990, p.243, paragraph 1\)](#). An application of [Gripenberg et al. \(1990, Theorem 9.5.5-\(ii\)\)](#), yields that  $K$  admits a resolvent kernel  $R_T$  which is again a Volterra kernel of bounded and continuous type. In particular,

$$|R_T|_{L^1} := \sup_{t \leq T} \int_0^T |R_T(t, s)| ds < \infty.$$

It remains to prove that  $R_T$  inherits condition (3.1) from  $K$  using the resolvent equation (A.4). We first show that

$$\int_0^T \int_0^T |R_T(t, s)|^2 dt ds < \infty. \quad (\text{A.7})$$

An application of Jensen's inequality on the normalized measure  $|R_T(t, z)| dz / \int_0^T |R_T(t, z')| dz'$  yields

$$\int_0^T \int_0^T |(R_T \star K)(t, s)|^2 dt ds \leq |R_T|_{L^1} \sup_{r \leq T} \int_0^T |K(r, s)|^2 ds \int_0^T \int_0^T |R_T(t, z)| dt dz < \infty.$$

Combined with the resolvent equation (A.4) and the first part of (3.1), we obtain (A.7). Using (A.7) and the Cauchy-Schwarz inequality we now get

$$\int_0^T |(K \star R_T)(t, u)|^2 du \leq \sup_{t' \leq T} \int_0^T |K(t', z)|^2 dz \int_0^T \int_0^T |R_T(u, z)|^2 du dz < \infty, \quad t \leq T,$$

which combined with (3.1) and (A.7) gives

$$\sup_{t \leq T} \int_0^T |R_T(t, s)|^2 ds < \infty,$$

which shows that  $R_T$  satisfies the first condition in (3.1). Finally, another application of the Cauchy-Schwarz inequality, for all  $t, h \geq 0$ , shows that

$$\int_0^T |(K \star R_T)(t+h, s) - (K \star R_T)(t, s)|^2 ds \leq \int_0^T \int_0^T |R_T(u', s)|^2 du' ds \int_0^T |K(t+h, u) - K(t, u)|^2 du$$

where the left hand side goes to 0 as  $h \rightarrow 0$  from the second part of (3.1). Combined with the resolvent equation (A.4), we can deduce that

$$\lim_{h \rightarrow 0} \int_0^T |R_T(t+h, u) - R_T(t, u)|^2 du = 0,$$

which yields the second condition in (3.1) for  $R_T$ .  $\square$

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

**Theorem A.3.** *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 unique progressively measurable strong solution  $(X, S)$  to (1.5)–(1.6) on  $[0, T]$  given by*

$$X_t = g_0(t) + \int_0^t R_T^\kappa(t, s) g_0(s) ds + \frac{1}{\kappa} \int_0^t R_T^\kappa(t, s) \nu dW_s, \quad (\text{A.8})$$

$$S_t = S_0 \exp \left( -\frac{1}{2} \int_0^t X_s^2 ds + \int_0^t X_s dB_s \right), \quad (\text{A.9})$$

where  $R_T^\kappa$  is the resolvent kernel of  $\kappa K$  with the convention that  $R_T^\kappa/\kappa = K$  when  $\kappa = 0$ . In particular, (3.2) holds.

*Proof.* If  $\kappa = 0$ , the existence is trivial. Fix  $\kappa \neq 0$ . An application of Lemma A.2 on the kernel  $\kappa K$  yields the existence of a resolvent  $R_T^\kappa$  satisfying (3.1). We define  $X$  as in (A.8) and we write it in compact form:

$$X = (\text{id} + \mathbf{R}_T^\kappa)(g_0) + \frac{1}{\kappa} \mathbf{R}_T^\kappa(\nu dW),$$

where we used the notation  $\mathbf{R}_T^\kappa(\nu dW)(t) = \int_0^t R_T^\kappa(t, s) \nu dW_s := N_t$ . We first observe that  $X$  admits a progressively measurable modification. Indeed, the stochastic integral  $N$  is adapted as an Itô integral and it is mean-square continuous, i.e.  $\mathbb{E}[|N_t - N_s|^2] \rightarrow 0$  as  $s \rightarrow t$  by virtue of Itô's isometry and the fact that  $R_T^\kappa$  satisfies (3.1) (see Lemma A.2). Therefore,  $N$  admits a progressively measurable modification. We now show that  $X$  solves (1.6). Using (A.6), composing both sides by  $(\text{id} + \mathbf{R}_T^\kappa)^{-1} = (\text{id} - \kappa \mathbf{K})$  and invoking stochastic Fubini's theorem yield

$$\begin{aligned} (\text{id} - \kappa \mathbf{K})(X) &= g_0 + (\text{id} - \kappa \mathbf{K}) \frac{1}{\kappa} \mathbf{R}_T^\kappa(\nu dW) \\ &= g_0 + \mathbf{K}(\nu dW), \end{aligned}$$

where we used the resolvent equation (A.5) for the last equality. This shows that

$$X_t = g_0(t) + \kappa(\mathbf{K})(X)(t) + \mathbf{K}(\nu dW)(t) = g_0(t) + \kappa \int_0^t K(t, s) X_s ds + \int_0^t K(t, s) \nu dW_s,$$

yielding that  $X$  is a strong solution of (1.6). Furthermore, (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. One can therefore define  $S$  as in (A.9) and it is immediate that  $S$  solves (1.5) by an application of Itô's formula. The uniqueness statement follows by reiterating the same argument backwards: by showing that any solution  $X$  to (1.6) is of the form (A.8) using the resolvent equation.  $\square$

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

**Lemma A.4.** *Let  $K$  satisfy (3.1) and  $L \in L^2([0, T]^2, \mathbb{R})$ . Then,  $K \star L$  satisfies (3.1). Furthermore, 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.5.** Fix  $b \in \mathbb{C}$  and a kernel  $K$  as in Definition 3.1. Then,  $(\text{id} - b\mathbf{K})$  is invertible. Furthermore, for all  $t \leq T$ ,  $\tilde{\Sigma}_t$  given by (1.11) is an integral operator of trace class with continuous kernel and can be re-written in the form

$$\tilde{\Sigma}_t = (\text{id} - b\mathbf{K}_t)^{-1} \Sigma_t (\text{id} - b\mathbf{K}_t^*)^{-1} \quad (\text{A.10})$$

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

*Proof.* Lemma A.2 yields the existence of the resolvent  $R_T^b$  of  $bK$  which is again a Volterra kernel of continuous and bounded type. Whence, (A.6) yields that  $(\text{id} - b\mathbf{K})$  is invertible with an inverse given by  $(\text{id} + \mathbf{R}_T^b)$ . To prove (A.10), we fix  $t \leq T$  and we observe that since  $\Sigma_t(s, u) = 0$  whenever  $s \wedge u \leq t$ , we have

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

where we defined the kernel  $R_{t,T}^b(s, u) = R_T^b(s, u)\mathbf{1}_{u \geq t}$ . Similarly,  $\Sigma_t \star (R_T^b)^* = \Sigma_t \star (R_{t,T}^b)^*$ . Using the resolvent equation (A.4) of  $R_T^b$ , it readily follows that  $R_{t,T}^b$  is the resolvent of  $bK_t$  so that  $(\text{id} - b\mathbf{K}_t)^{-1} = (\text{id} + \mathbf{R}_{t,T}^b)$ . Combining all of the above leads to

$$\begin{aligned} \tilde{\Sigma}_t &= (\text{id} - b\mathbf{K})^{-1} \Sigma_t (\text{id} - b\mathbf{K}^*)^{-1} \\ &= (\text{id} + \mathbf{R}_T^b) \Sigma_t (\text{id} + \mathbf{R}_T^b)^* \\ &= \Sigma_t + \mathbf{R}_T^b \Sigma_t + \Sigma_t (\mathbf{R}_T^b)^* + \mathbf{R}_T^b \Sigma_t (\mathbf{R}_T^b)^* \\ &= \Sigma_t + \mathbf{R}_{t,T}^b \Sigma_t + \Sigma_t (\mathbf{R}_{t,T}^b)^* + \mathbf{R}_{t,T}^b \Sigma_t (\mathbf{R}_{t,T}^b)^* \\ &= (\text{id} + \mathbf{R}_{t,T}^b) \Sigma_t (\text{id} + \mathbf{R}_{t,T}^b)^* \\ &= (\text{id} - b\mathbf{K}_t)^{-1} \Sigma_t (\text{id} + b\mathbf{K}_t^*)^{-1}, \end{aligned} \quad (\text{A.11})$$

which proves (A.10). Furthermore, it can be readily deduced from (A.11) that  $\tilde{\Sigma}_t$  is an integral operator of trace class with continuous kernel: the trace class property follows from the fact that the product of two Hilbert-Schmidt operators is of trace class; the continuity of the kernel follows from the fact that both  $K$  and  $R_T^b$  satisfy (3.1), recall Lemma A.2.  $\square$

**Lemma A.6.** Fix  $a, b \in \mathbb{C}$  such that  $\Re(a) \leq -\frac{\Im(b)^2}{2\nu^2}$ . Let  $t \leq T$  and  $K$  be a kernel as in Definition 3.1. Then,  $(\text{id} - 2\tilde{\Sigma}_t a)$  is invertible and  $\Psi_t$  given by (1.9) is well-defined. Furthermore, if  $\Im(a) = \Im(b) = 0$  then,  $\Psi_t$  is a symmetric nonpositive operator in the sense of Definition 2.1.

*Proof.* • Using Lemma A.5, we write

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

with

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

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

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

The operator  $\mathbf{I}$  is symmetric nonnegative and invertible so that  $\text{sp}(\mathbf{I}) \subset (0, \infty)$ . 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  $\text{sp}(\mathbf{II}) \in [0, \infty)$ . It follows that  $\text{sp}(\Re(\mathbf{A}_t)) \in (0, \infty)$ , showing that  $0 \notin \text{sp}(\mathbf{A}_t)$  and that  $\mathbf{A}_t$  is invertible. Whence,  $(\text{id} - 2a\tilde{\boldsymbol{\Sigma}}_t)$  is invertible. Combined with Lemma A.5, 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.4, 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,

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

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

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

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

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

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

## 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

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

so that an application of Lemma A.6 yields that  $\Psi_t$  is well-defined.

We now collect from [Abi Jaber et al. \(2021, Lemma 5.8\)](#) further properties of  $t \mapsto \Psi_t$ . In particular, its link with an operator Riccati equation. We recall that  $t \mapsto \Psi_t$  is said to be strongly differentiable at time  $t \geq 0$ , if there exists a bounded linear operator  $\dot{\Psi}_t$  from  $L^2([0, T], \mathbb{C})$  into itself such that

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

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

- (i)  $\bar{\Psi}_t := (-\text{id} + \Psi_t)$  is an integral operator induced by a symmetric kernel  $\bar{\psi}_t(s, u)$  such that

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

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

$$(\Psi_t f \mathbf{1}_t)(t) = (\text{id} + b\mathbf{K}^* \Psi_t)(f \mathbf{1}_t)(t),$$

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

- (iii)  $t \mapsto \Psi_t$  is strongly differentiable and satisfies the operator Riccati equation

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

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

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

*Proof.* The proof follows from a straightforward adaptation of the proof of [Abi Jaber et al. \(2021, Lemma 5.6\)](#).  $\square$



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

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

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

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

$$\begin{aligned} d\langle g_t, \Psi_t g_t \rangle_{L^2} &= \left( \langle g_t, \dot{\Psi}_t g_t \rangle_{L^2} - aX_t^2 - 2u\rho\nu X_t (\mathbf{K}^* \Psi_t) (g_t)(t) - \text{Tr}(\Psi_t \dot{\Sigma}_t) \right) dt \\ &\quad + 2\nu ((\mathbf{K}^* \Psi_t) g_t)(t) dW_t, \quad dt \times \mathbb{Q} - a.e \end{aligned} \quad (\text{B.2})$$

*Proof.* We first set

$$\bar{g}_t(s) = g_0(s) + \int_0^t K(s, u) \kappa X_u du + \int_0^t K(s, u) \nu dW_u, \quad (\text{B.3})$$

so that using Lemma B.1-(i), we can write

$$\langle g_t, \Psi_t g_t \rangle_{L^2} = \int_t^T (a\bar{g}_t(s)^2 ds + \bar{g}_t(s)(\bar{\Psi}_t g_t)(s)) ds.$$

The Leibniz rule yields

$$\begin{aligned} d\langle g_t, \Psi_t g_t \rangle_{L^2} &= (-a\bar{g}_t(t)^2 - \bar{g}_t(t)(\bar{\Psi}_t g_t)(t)) dt \\ &\quad + \int_t^T d(a\bar{g}_t(s)^2 + \bar{g}_t(s)(\bar{\Psi}_t g_t)(s)) ds, \quad dt \times \mathbb{Q} \text{ a.e.} \end{aligned} \quad (\text{B.4})$$

• We first compute the dynamics of  $t \mapsto a\bar{g}_t(s)^2 ds + \bar{g}_t(s)(\bar{\Psi}_t g_t)(s)$ . We fix  $s \in [0, T]$ . It follows from (B.3), that

$$d\bar{g}_t(s) = K(s, t) \kappa X_t dt + K(s, t) \nu dW_t, \quad dt \times \mathbb{Q} - a.e.$$

An application of Itô's lemma on the square yields

$$d\bar{g}_t(s)^2 = (\nu^2 K(s, t)^2 + 2\bar{g}_t(s)K(s, t)\kappa X_t) dt + 2\bar{g}_t(s)K(s, t)\nu dW_t, \quad dt \times \mathbb{Q} - a.e.$$

Furthermore, we write

$$(\bar{\Psi}_t g_t)(s) = \int_t^T \bar{\psi}_t(s, u) \bar{g}_t(u) du$$

so that an application of the Leibniz rule combined with the fact that  $\bar{g}_t(t) = X_t$  for almost every  $(t, \omega)$  and Lemma B.1-(iii) yields that  $t \mapsto (\bar{\Psi}_t g_t)(s)$  is a semimartingale on  $[0, s]$  with the following dynamics

$$\begin{aligned} d(\bar{\Psi}_t g_t)(s) &= \left( -\bar{\psi}_t(s, t)X_t + \int_t^T \dot{\bar{\psi}}_t(s, u)\bar{g}_t(u)du + \int_t^T \bar{\psi}_t(s, u)K(u, t)\kappa X_u du \right) dt \\ &\quad + \int_t^T \bar{\psi}_t(s, u)K(u, t)\nu dudW_t \\ &= \left( -X_t\bar{\psi}_t(s, t) + (\dot{\bar{\Psi}}_t g_t)(s) + X_t(\bar{\Psi}_t K(\cdot, t)\kappa)(s) \right) dt \\ &\quad + (\bar{\Psi}_t K(\cdot, t)\nu)(s)dW_t, \quad dt \times \mathbb{Q} - a.e. \end{aligned}$$

where we used that  $\dot{\bar{\Psi}}_t = \dot{\bar{\Psi}}_t$  and that  $K(u, t) = 0$  for all  $u \leq t$ . Moreover, the quadratic covariation between  $t \mapsto \bar{g}_t(s)$  and  $t \mapsto (\bar{\Psi}_t g_t)(s)$  is given by

$$\begin{aligned} d[\bar{g}_t(s), (\bar{\Psi}_t g_t)(s)]_t &= \nu^2 \int_0^T \bar{\psi}_t(s, u)K(u, t)K(s, t)dudt \\ &= - \int_0^T \bar{\psi}_t(s, u)\dot{\Sigma}_t(u, s)dudt \\ &= -(\bar{\Psi}_t \dot{\Sigma}_t(\cdot, s))(s)dt. \end{aligned}$$

Whence, combining the previous three identities, we get the dynamics of  $U_t(s) := a\bar{g}_t(s)^2 + (\bar{g}_t(s)(\bar{\Psi}_t g_t)(s))$ :

$$\begin{aligned} dU_t(s) &= ad\bar{g}_t(s)^2 + d\bar{g}_t(s)(\bar{\Psi}_t g_t)(s) + \bar{g}_t(s)d(\bar{\Psi}_t g_t)(s) + d[\bar{g}_t(s), (\bar{\Psi}_t g_t)(s)]_t \\ &= a\nu^2 K(s, t)^2 dt + 2a\bar{g}_t(s)K(s, t)\kappa X_t dt \\ &\quad + X_t \kappa K(s, t)(\bar{\Psi}_t g_t)(s)dt + \bar{g}_t(s)(\dot{\bar{\Psi}}_t g_t)(s)dt \\ &\quad - \bar{g}_t(s)\bar{\psi}_t(s, t)X_t dt + \bar{g}_t(s)X_t(\bar{\Psi}_t K(\cdot, t)\kappa)(s)dt \\ &\quad - (\bar{\Psi}_t \dot{\Sigma}_t(\cdot, s))(s)dt \\ &\quad + (2a\bar{g}_t(s)K(s, t)\nu + \nu K(s, t)(\bar{\Psi}_t g_t)(s) + \bar{g}_t(s)(\bar{\Psi}_t K(\cdot, t)\nu)(s))dW_t \\ &= [\mathbf{I}(s) + \mathbf{II}(s) + \mathbf{III}(s) + \mathbf{IV}(s) + \mathbf{V}(s) + \mathbf{VI}(s) + \mathbf{VII}(s)]dt \\ &\quad + (\mathbf{VIII}(s) + \mathbf{IX}(s) + \mathbf{X}(s))dW_t, \quad dt \times \mathbb{Q} - a.e. \end{aligned} \tag{B.5}$$

• We now integrate in  $s$  to obtain the right hand side in (B.4). We let  $\mathcal{N} = \{(t, \omega) : \exists s \in [0, T] \text{ such that (B.5) does not hold}\}$ . Then,  $\mathcal{N}$  is a null set and we fix  $(t, \omega) \in [0, T] \times \Omega \setminus \mathcal{N}$ . In the sequel, all the equalities are written for this particular  $\omega$ . First, using that  $\dot{\Sigma}_t(s, s) = -\nu^2 K(s, t)^2$  and recalling that

$$\Psi = aid + \bar{\Psi} \tag{B.6}$$

we obtain that<sup>10</sup>

$$\int_t^T (\mathbf{I}(s) + \mathbf{VII}(s)) ds = -\text{Tr}(\Psi_t \dot{\Sigma}_t).$$

Combining (B.6) with Lemma B.1-(ii) and the fact that  $\bar{\Psi}^* = \bar{\Psi}$  we obtain that

$$\int_t^T [\mathbf{II}(s) + \mathbf{III}(s) + \mathbf{VI}(s)] ds = 2\kappa X_t \int_0^T K(s, t) (\Psi_t g_t)(s) ds = 2\kappa X_t (\mathbf{K}^* \Psi_t g_t)(t).$$

On the other hand, we have

$$\begin{aligned} \int_0^T \mathbf{IV}(s) ds &= \langle g_t, \dot{\Psi}_t g_t \rangle_{L^2}, \quad \int_0^T \mathbf{V}(s) ds = -X_t (\bar{\Psi}_t g_t)(t), \\ \int_0^T [\mathbf{VIII}(s) + \mathbf{IX}(s) + \mathbf{X}(s)] ds &= 2\nu (\mathbf{K}^* \Psi_t)(g_t)(t) dW_t. \end{aligned}$$

Therefore, summing the above, plugging in (B.4), using Lemma B.1-(ii) and recalling (B.6) and that  $b = \kappa + u\rho\nu$  and  $\bar{g}_t(t) = X_t$  yield

$$\begin{aligned} d\langle g_t, \Psi_t g_t \rangle_{L^2} &= (-aX_t^2 + 2\kappa X_t (\mathbf{K}^* \Psi_t g_t)(t) - 2X_t (\bar{\Psi}_t g_t)(t)) dt \\ &\quad + \left( \langle g_t, \dot{\Psi}_t g_t \rangle_{L^2} - \text{Tr}(\Psi_t \dot{\Sigma}_t) \right) dt + 2\nu (\mathbf{K}^* \Psi_t)(g_t)(t) dW_t \\ &= \left( -aX_t^2 - 2u\rho\nu X_t (\mathbf{K}^* \Psi_t)(g_t)(t) - \text{Tr}(\Psi_t \dot{\Sigma}_t) + \langle g_t, \dot{\Psi}_t g_t \rangle_{L^2} \right) dt \\ &\quad + 2\nu (\mathbf{K}^* \Psi_t)(g_t)(t) dW_t \end{aligned}$$

leading to the claimed dynamics (B.2). □

We can now complete the proof of Theorem 3.3. We recall that  $\phi$  given in (3.5) solves

$$\dot{\phi}_t = \text{Tr}(\Psi_t \dot{\Sigma}_t). \tag{B.7}$$

*Proof of Theorem 3.3.* It suffices to prove that (3.4) holds for all  $0 \leq u \leq 1$  and  $w \leq 0$  to obtain the claimed expression by analytic continuation. Indeed, the left hand side in (3.4) is analytic in  $(u, w)$  in an open region  $(\Re(u), \Re(w)) \in (u_-, u_+) \times (w_-, w_+)$  by general results on the analyticity of characteristic functions, see Widder (2015, Theorem II.5a). The right hand side is also analytic in  $(u, w)$  since resolvents are analytic: they are given by power series. Therefore, if (3.4) holds for all  $0 \leq u \leq 1$  and  $w \leq 0$ , then by analytic continuation

---

<sup>10</sup>The operator  $\Psi_t \dot{\Sigma}_t = a\dot{\Sigma}_t + \bar{\Psi}_t \dot{\Sigma}_t$  is of trace class: (i)  $\dot{\Sigma}_t$  is of trace class since it can be written as product of two Hilbert-Schmidt integral operators  $\dot{\Sigma}_t = \tilde{K}_t \tilde{K}_t^*$  with  $\tilde{K}_t(s, z) = K(s, t)/\sqrt{T}$ , so that (A.2) yields  $\text{Tr}(\dot{\Sigma}_t) = \int_0^T \dot{\Sigma}_t(s, s) ds$ ; (ii)  $\bar{\Psi}_t \dot{\Sigma}_t$  is of trace class as product of two Hilbert-Schmidt integral operators so that (A.2) yields  $\text{Tr}(\bar{\Psi}_t \dot{\Sigma}_t) = \int_0^T \int_0^T \bar{\psi}_t(s, z) \dot{\Sigma}_t(z, s) dz ds$ .

(3.4) remains valid on  $\{(u, w) \in \mathbb{C}^2 : 0 \leq \Re(u) \leq 1 \text{ and } \Re(v) \leq 0\}$ .

Fix  $u \in [0, 1]$ ,  $w \in \mathbb{R}_-$ . Set

$$U_t = u \log S_t + w \int_0^t X_s^2 ds + \phi_t + \langle g_t, \Psi_t g_t \rangle_{L^2}, \quad (\text{B.8})$$

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  $t \leq T$ , yields (3.4).

*Step 1. We prove that  $M$  is a local martingale by expliciting its dynamics.* We first observe that

$$dM_t = M_t (dU_t + \frac{1}{2} d\langle U \rangle_t). \quad (\text{B.9})$$

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

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

so that

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

Observing that

$$4\nu^2 ((\mathbf{K}^* \Psi_t)(g_t)(t))^2 = -4 \langle g_t, \Psi_t \dot{\Sigma}_t \Psi_t g_t \rangle_{L^2},$$

we get that the drift part in (B.9) is given by

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

by virtue of the Riccati equations (B.1) and (B.7). 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.6 yields that  $\Psi_t$  is a symmetric nonpositive operator so that, recall (B.7),

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

Whence, since  $w \leq 0$  and  $0 \leq u \leq 1$ , it follows from (B.8) 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.  $\square$

## References

- Abi Jaber, E. (2019a). The Laplace transform of the integrated Volterra Wishart process. *arXiv:1911.07719, to appear in Mathematical Finance*.
- Abi Jaber, E. (2019b). Lifting the Heston model. *Quantitative Finance*, 19(12):1995–2013.
- Abi Jaber, E. (2021). Weak existence and uniqueness for affine stochastic Volterra equations with L1-kernels. *Bernoulli*, 27(3):1583–1615.
- 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. (2021). Markowitz portfolio selection for multivariate affine and quadratic Volterra models. *SIAM Journal on Financial Mathematics*, 12(1):369–409.
- 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.
- Brezis, H. (2010). *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media.
- Brislaw, C. (1988). Kernels of trace class operators. *Proceedings of the American Mathematical Society*, 104(4):1181–1190.
- 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.
- Carr, P. and Wu, L. (2003). The finite moment log stable process and option pricing. *The journal of finance*, 58(2):753–777.
- 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.
- Fouque, J.-P., Papanicolaou, G., Sircar, R., and Solna, K. (2003). Multiscale stochastic volatility asymptotics. *Multiscale Modeling & Simulation*, 2(1):22–42.
- Fredholm, I. (1903). Sur une classe d'équations fonctionnelles. *Acta mathematica*, 27(1):365–390.
- Fukasawa, M. (2011). Asymptotic analysis for stochastic volatility: martingale expansion. *Finance and Stochastics*, 15(4):635–654.
- Fukasawa, M. (2021). Volatility has to be rough. *Quantitative Finance*, 21(1):1–8.
- 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.
- Gohberg, I., Goldberg, S., and Krupnik, N. (2012). *Traces and determinants of linear operators*, volume 116. Birkhäuser.
- Gohberg, I. and Krein, M. G. (1978). *Introduction to the theory of linear nonselfadjoint operators*, volume 18. American Mathematical Soc.
- 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.
- Lee, R. W. (2005). Implied volatility: Statics, dynamics, and probabilistic interpretation. *Recent advances in applied probability*, pages 241–268.
- 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.
- Mayerhofer, E. (2019). Reforming the wishart characteristic function. *arXiv preprint arXiv:1901.09347*.
- Ondreját, M. and Seidler, J. (2013). On existence of progressively measurable modifications. *Electronic Communications in Probability*, 18:1–6.
- 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.
- Simon, B. (2005). *Trace ideals and their applications*. Number 120. American Mathematical Soc.
- 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.



Veraar, M. (2012). The stochastic Fubini theorem revisited. *Stochastics An International Journal of Probability and Stochastic Processes*, 84(4):543–551.

Widder, D. V. (2015). *Laplace transform*. Princeton university press.