Finance and Stochastics (2022) 26:733–769
https://doi.org/10.1007/s00780-022-00489-4

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

Eduardo Abi Jaber[1]

Received: 22 September 2020 / Accepted: 17 June 2022 / Published online: 16 September 2022

Abstract
Stochastic volatility models based on Gaussian processes, like fractional Brownian motion, are able to reproduce important stylised 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. That 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 realised variance by using Fourier inversion techniques. In the context of rough volatility modelling, our results apply to the (rough) fractional Stein–Stein model and provide the first analytic formulas for option pricing known to date, generalising that of Stein–Stein, Schöbel–Zhu and a special case of Heston.

Keywords Stochastic volatility models · Non-Markovian models · Fast pricing

Mathematics Subject Classification (2020) 91G30 · 91G60

JEL Classification G13

1 Introduction

In the realm of risk management in mathematical finance, academics and practitioners have always been striving for explicit solutions to option prices and hedging strategies in their models. Undoubtedly, finding explicit expressions to a theoretical problem

✉ E. Abi Jaber
Eduardo.Abi-Jaber@univ-paris1.fr

[1] Maison des Sciences Économiques, Université Paris 1 Panthéon-Sorbonne, 112 Bd de l'Hôpital, 75013 Paris, France

can be highly satisfying in itself; but 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, and providing a better understanding of the 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.

1.1 From Black–Scholes to rough volatility

In their seminal paper, Black and Scholes [12] 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

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

with B a standard Brownian motion and σ the constant instantaneous volatility parameter. Although revolutionary, the model remains very simple; it does not provide an accurate representation of the reality of financial markets which is characterised 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 [21]. Since then, a large and growing literature has been developed to refine the Black and Scholes [12] model. One notable direction is stochastic volatility modelling where the constant volatility σ in (1.1) is replaced by a Markovian stochastic process $(\sigma_t)_{t\geq 0}$. In their celebrated paper, Stein and Stein [59] modelled $(\sigma_t)_{t\geq 0}$ by an Ornstein–Uhlenbeck process of the form

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

where W is a standard Brownian motion independent of B. Remarkably, they obtained closed-form expressions for the characteristic function of the log-price, which allowed them to recover the density as well as option prices by Fourier inversion of the characteristic function. Later on, the model has been extended by Schöbel and Zhu [54] to account for the leverage effect, i.e., an arbitrary correlation between W and B. Similar formulas to those of Stein–Stein for the characteristic function of the log-price are derived for the non-zero correlation case.

Prior to the extension by Schöbel and Zhu [54], Heston [43] 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 (squares of Brownian motion constitute the building blocks of squared Bessel processes; see Revuz and Yor [53, Chap. XI]) so that the Stein–Stein model can be recast in the form

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

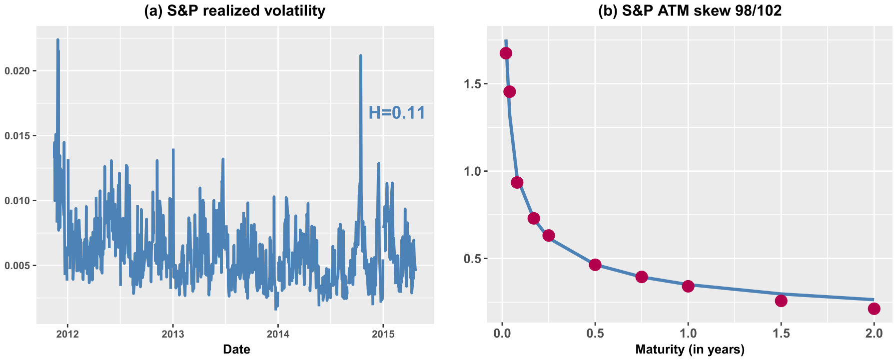

Fig. 1 (**a**) Realised volatility of the S&P downloaded from https://realized.oxford-man.ox.ac.uk/ with an estimated Hurst index of $\hat{H} = 0.11$. (**b**) Term structure of the at-the-money skew of the implied volatility $\frac{\partial \sigma_{\text{implicit}}(k,T)}{\partial k}|_{k=0}$ for the S&P index on June 20, 2018 (red dots) and a power-law fit $t \mapsto 0.35t^{-0.41}$ (blue line). Here $k := \ln(K/S_0)$ stands for the log-moneyness and T for the time to maturity

where $B = \rho W + \sqrt{1-\rho^2}W^\perp$ with $\rho \in [-1, 1]$ and $W^\perp$ a Brownian motion independent of W. That model remains tractable as shown earlier in the context of bond pricing with uncertain inflation by Cox et al. [23, Eqs. (51) and (52)]. Heston [43] 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 [43] and Schöbel and Zhu [54] share a lot of similarities and they perfectly agree when $\theta = 0$ in (1.2); see Lord and Kahl [49, Eq. (44)]. That analytical tractability motivated the development of the theory of finite-dimensional Markovian affine processes; see Duffie et al. [28].

Unfortunately, Markovian stochastic volatility models, such as the Heston and 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 [9], Ding et al. [27]. More recently, Gatheral et al. [35] and Bennedsen et al. [11] show that the sample paths of the realised volatility are rougher than standard Brownian motion at any realistic time scale as illustrated in Fig. 1 (a). From a pricing perspective, continuous semimartingale models driven by a standard Brownian motion fail to reproduce the power-law decay of the at-the-money skew of option prices shown in Fig. 1 (b); see Carr and Wu [19], Fouque et al. [31], Lee [47], Alòs et al. [8], Bayer et al. [10], Fukasawa [33, 34].

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 [20] by replacing the driving Brownian motion of the volatility process by a fractional Brownian motion W^H, given by

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

where $H \in (0,1)$ is the Hurst exponent. Here $H > 1/2$ corresponds to positively correlated returns, $H < 1/2$ to negatively correlated increments, and $H = 1/2$ reduces to the case of standard Brownian motion. The sample paths of W^H are locally Hölder-continuous of any order strictly less than H and hence less (resp. more) regular than standard Brownian motion when $H < 0.5$ (resp. $H > 0.5$). Initially, Comte and Renault [20] considered the case $H > 1/2$. However, a smaller Hurst index $H \approx 0.1$ allows one 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 (see Fig. 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. [35].

The price to pay is that, in general, these models are no longer Markovian nor semimartingales, which limits their practical use and makes their mathematical analysis quite challenging. This has initiated a thriving branch of research; for references, see https://sites.google.com/site/roughvol/home. 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 [29] who came up with a rough version of the Heston [43] model after convoluting the dynamics (1.3) with a fractional kernel to get

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

for $H \in (0, 1/2)$. Remarkably, they show that an analogous formula to that of Heston [43] for the characteristic function of the log-price continues 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. [6], Abi Jaber [2], and it can be recovered as a projection of infinite-dimensional Markovian affine processes as illustrated in Abi Jaber and El Euch [4], Cuchiero and Teichmann [25], Gatheral and Keller-Ressel [36].

Although the rough Heston model can be efficiently implemented, see for instance Abi Jaber [1], Abi Jaber and El Euch [5], Callegaro et al. [17], Gatheral and Radoičić [37], no closed-form solution for the fractional deterministic Riccati equation or for the characteristic function is known to date, which has to be contrasted with the conventional Heston [43] model. One possible explanation could be that unlike the Markovian case, squares of fractional Brownian motion have different dynamics than (1.4), so that the marginals of the process (1.4) are not chi-square-distributed except for the case $H = 1/2$.

The main objective of the present paper is to rely on squares of general Gaussian processes with arbitrary covariance structures by considering the non-Markovian extensions of the Stein and Stein [59] and Schöbel and Zhu [54] models. We show that the underlying Gaussianity makes the problem highly tractable and allows recovering analytic expressions for the joint Fourier–Laplace transform of the-log price and the integrated variance in general, which agree with those of Stein–Stein, Schöbel–Zhu and Heston under the Markovian setting. Such models have already been considered several times in the context of the non-Markovian and rough volatility literature, see Cuchiero and Teichmann [24], Gulisashvili et al. [41], Harms and Stefanovits [42], Horvath et al. [44], 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 [3] for the Laplace transform of the integrated variance and that of Abi Jaber et al. [7] where the Laplace transform of the forward covariance curve enters in the context of portfolio optimisation.

1.2 The Gaussian Stein–Stein model and main results

For $T > 0$, we consider the generalised version of the Stein–Stein model given by

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

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

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

Under mild assumptions on its covariance function, every Gaussian process can be written in the form (1.6) with $\kappa = 0$; see Sottinen and Viitasaari [58]. That representation is known as the *Fredholm representation*. We are chiefly interested in two classes of kernels K:

– *Symmetric kernels*, i.e., $K(t,s) = K(s,t)$ for all $s, t \le T$, 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 make sense of (1.5), (1.6) in a static sense in Sect. 2.

– *Volterra kernels*, i.e., $K(t,s) = 0$ whenever $s \ge t$, for which integration in (1.6) goes up to time t, which is more in line with standard stochastic volatility modelling and for which the stochastic integral $\int_0^\cdot X dB$ can be defined in the usual Itô sense; see Sect. 3. For instance, the conventional mean-reverting Stein–Stein model (1.2) can be recovered by setting $g_0(t) = X_0 - \kappa\theta t$, $\kappa \le 0$, 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 Riemann–Liouville fractional Brownian motion corresponds to the case

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

For suitable $u, w \in \mathbb{C}$, we provide for the conditional joint Fourier–Laplace transform of the log-price and the integrated variance the analytical expression

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

where $\langle f, h\rangle_{L^2} = \int_0^T f(s)h(s)ds$, det denotes the Fredholm [32] determinant (see Appendix A.1), g_t is the adjusted conditional mean given by

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

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

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

where $\boldsymbol{K}$ denotes the integral operator induced by K, $\boldsymbol{K}^*$ its adjoint operator (see Sect. 1.3 for detailed notations), id denotes the identity operator,

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

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

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

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

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

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

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

At first glance, the expressions for $\mathbf{\Phi}$ seem to depend on the class of the kernel, but they actually agree. Indeed, for Volterra kernels, where $K(t,s) = 0$ for $s \ge t$, we have $\det(\mathrm{id} - b\boldsymbol{K}) = \det(\mathrm{id} - b\boldsymbol{K}^*) = 1$ so that using the relation (see Simon [56, Theorem 3.8]) $\det((\mathrm{id} + \boldsymbol{F})(\mathrm{id} + \boldsymbol{G})) = \det(\mathrm{id} + \boldsymbol{F})\det(\mathrm{id} + \boldsymbol{G})$ gives

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

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

– Symmetric nonnegative kernels. We provide an elementary static derivation of (1.7) for $t = 0$ and $\kappa = 0$, based on the spectral decomposition of K which leads to the decomposition of the characteristic function as an infinite product of independent Wishart distributions. The operator $\mathbf{\Psi}_0$ appears naturally after a rearrangement of the terms. The main result is collected in Theorem 2.3.

– 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 used on the adjusted conditional mean process $(g_t(s))_{\{t\le s\}}$. The main result is stated in Theorem 3.3. This is the class of kernels which is more suited for financial applications.

From a numerical perspective, we show in Sect. 4.1 that the expression (1.7) lends itself to approximation by closed-form solutions using finite-dimensional matrices after a straightforward discretisation of the operators in the form

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

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,ν,κ,u,w) and Det is the standard determinant of a matrix; we refer to Sect. 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 Sect. 4.2.

1.3 Notations

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

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

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 $K:[0,T]^2\to\mathbb{K}$ such that

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

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

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

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 $\boldsymbol{K}$ the integral operator induced by the kernel K, that is,

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

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

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

and by $\boldsymbol{K}^*$ 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 $\boldsymbol{\Psi}$ given in (1.9). Later on, in Sect. 3, such expressions are shown to hold in the more practical case of Volterra kernels by using a dynamical approach.

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

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

Operator $\boldsymbol{K}$ is said to be *symmetric nonpositive* if $-\boldsymbol{K}$ is symmetric nonnegative.

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

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

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

$$\sum_{n\geq 1} \lambda_n < \infty. \tag{2.2}$$

Such a decomposition is possible whenever the operator $\boldsymbol{K}$ is the (nonnegative symmetric) square root of a nonnegative symmetric operator $\boldsymbol{C}$ which is generated by a continuous kernel. This is known as Mercer's theorem, see Shorack and Wellner [55, Theorem V.3.1], and leads to the so-called Kac–Siegert/Karhunen–Loève representation of the process X; see Kac and Siegert [45]. In this case, one can show that any square-integrable Gaussian process X with mean g_0 and covariance $\boldsymbol{C}$ admits the representation (1.6) with $\kappa = 0$ on some filtered probability space supporting a Brownian motion W; see Sottinen and Viitasaari [58, 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

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

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 $t \le T$ 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 $(\mathcal{F}_t^Y)_{t\ge 0}$ stands for the filtration (satisfying the usual conditions) generated by the process Y, we have that $W^\perp$ is still a Brownian motion with respect to $(\mathcal{F}_t)_{t\ge 0}$, and up to a modification, X is progressively measurable (every jointly measurable and adapted process admits a progressively measurable modification; see Ondreját and Seidler [52]) with respect to $(\mathcal{F}_t^X)_{t\ge 0}$ (and hence with respect to the filtration $(\mathcal{F}_t)_{t\ge 0}$) so that

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

is well defined as an Itô integral with respect to $(\mathcal{F}_t)_{t\ge 0}$. If $\rho = 0$ so that $\int_0^\cdot X dW$ does not appear, (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 filtration $\mathbb{F}^X \vee \mathbb{F}^W$, 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 $g_0 = \sum_{n\ge 1}\langle g_0, e_n\rangle e_n$. Making use of (2.1), we first observe that an application of Fubini's theorem (see Veraar [60, Theorem 2.2]), justified by the fact that

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

yields that $dt \otimes \mathbb{Q}$-a.e.,

$$X_t = g_0(t) + \int_0^T K(t,s)\nu dW_s = \sum_{n\ge 1}(\langle g_0, e_n\rangle + \sqrt{\lambda_n}\,\nu\xi_n)e_n(t), \qquad (2.4)$$

where $\xi_n = \int_0^T e_n(s)dW_s$ for each $n \ge 1$. Since $(e_n)_{n\ge 1}$ is an orthonormal family in L^2, the sequence $(\xi_n)_{n\ge 1}$ is a sequence of independent standard Gaussian random variables that are $\mathcal{F}_T^W$-measurable. We set

$$N_T = \sum_{n\ge 1}(\langle g_0, e_n\rangle + \sqrt{\lambda_n}\,\nu\xi_n)\xi_n. \qquad (2.5)$$

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

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

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

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

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

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

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 we have*

$$
\mathbb{E}\left[\exp\left(u \log \frac{S_T}{S_0} + w \int_0^T X_s^2 ds\right)\right] = \frac{\exp(\langle g_0, \boldsymbol{\Psi}_0 g_0\rangle_{L^2})}{\det(\boldsymbol{\Phi}_0)^{1/2}}, \tag{2.7}
$$

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

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

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

The rest of this 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 $\tilde{\boldsymbol{\Sigma}}_0$ and $\boldsymbol{\Psi}_0$ then appear naturally after a rearrangement of the terms.

In the sequel, to ease notation, we drop the subscript L^2 in the product $\langle \cdot, \cdot\rangle_{L^2}$. We start by computing the joint Fourier–Laplace transform of $(\int_0^T X_s^2 ds, N_T)$. Furthermore, the representation (2.4) readily leads to

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

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*

$$
\Re(\alpha) \leq 0, \qquad \Re(\beta) = 0. \tag{2.9}
$$

Then we have

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

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 μ_n and covariance matrix Σ_n given by

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

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 the Y_n and the well-known expression for the characteristic function of the Wishart distribution, see for instance Abi Jaber [3, Proposition A.1], we get

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

where I_2 stands for the 2×2 identity matrix. 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

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

and

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

Straightforward computations lead to the claimed expression (2.10). □

Relying on the spectral decomposition (2.1), we re-express the quantities entering (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 operator defined as in (1.9) with $a = \alpha + \frac{\beta^2}{2}$ and $b = \nu\beta$, i.e.,*

$$\boldsymbol{\Psi}_0^{\alpha,\beta} = (\mathrm{id} - b\boldsymbol{K}^*)^{-1} a (\mathrm{id} - 2\tilde{\boldsymbol{\Sigma}}_0 a)^{-1} (\mathrm{id} - b\boldsymbol{K})^{-1},$$

admits the decomposition

$$\boldsymbol{\Psi}_0^{\alpha,\beta} = \sum_{n\geq 1} \frac{\alpha + \frac{\beta^2}{2}}{1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n} \langle e_n, \cdot \rangle e_n \tag{2.11}$$

and

$$\det\left(\frac{1}{\alpha + \frac{\beta^2}{2}} \boldsymbol{\Psi}_0^{\alpha,\beta}\right) = \prod_{n\geq 1} \frac{1}{1 - 2\beta\nu\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n}, \tag{2.12}$$

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

$$\mathbb{E}\left[\exp\left(\alpha \int_0^T X_s^2 ds + \beta N_T\right)\right] = \det\left(\frac{1}{\alpha + \frac{\beta^2}{2}} \boldsymbol{\Psi}_0^{\alpha,\beta}\right)^{1/2} \exp(\langle g_0, \boldsymbol{\Psi}_0^{\alpha,\beta} g_0 \rangle). \tag{2.13}$$

Proof Throughout the proof, we 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 $\boldsymbol{K}$ and $\boldsymbol{L}$ in the form

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

the composition is given by

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

It follows from (2.1) that

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

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

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

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

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

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

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

This yields

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

Recalling that $a = \alpha + \frac{\beta^2}{2}$ and $b = \nu\beta$, we get

$$\left((1 - b\sqrt{\lambda_n})^2 - 2a\nu^2\lambda_n\right) = 1 - 2\nu\beta\sqrt{\lambda_n} - 2\alpha\nu^2\lambda_n.$$

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

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

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

We can now complete the proof of Theorem 2.3.

Proof of Theorem 2.3 It suffices to prove that

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

where

$$\alpha = w + \frac{1}{2}(u^2 - u) - \frac{\rho^2 u^2}{2}, \qquad \beta = \rho u.$$

Indeed, if this is the case, then

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

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. Conditionally on $\mathcal{F}_T^X \vee \mathcal{F}_T^W$, by the 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

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

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

$$L_T := \mathbb{E}\Big[\exp\Big(u\log\frac{S_T}{S_0} + w\int_0^T X_s^2 ds\Big)\Big]$$

is given by

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

leading to (2.15) due to (2.16). This ends the proof. □

3 Volterra kernels: a dynamical approach

In this section, we treat a class of Volterra kernels which are practically relevant in mathematical finance. We consider Volterra kernels of continuous and bounded type in L^2 in the terminology of Gripenberg et al. [40, Definitions 9.2.1, 9.5.1 and 9.5.2].

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

$$\sup_{t\in[0,T]}\int_0^T |K(t,s)|^2 ds < \infty,$$

$$\lim_{h\to 0}\int_0^T |K(u+h,s) - K(u,s)|^2 ds = 0, \quad u \le T. \tag{3.1}$$

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,

$$\sup_{t\in[0,T]} \int_0^T |K(t,s)|^2 ds = \sup_{t\in[0,T]} \int_0^t |k(t-s)|^2 ds \le \int_0^T |k(s)|^2 ds < \infty,$$

yielding the first part of (3.1). The second part follows from the L^2-continuity of k; see Brezis [15, Lemma 4.3].

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

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

where ${}_2F_1$ is the Gauss hypergeometric function. That 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 [26]. In this case,

$$\sup_{t\in[0,T]} \int_0^T |K(t,s)|^2 ds = \sup_{t\in[0,T]} \Sigma_0(t,t) \le T^{2H},$$

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 $h \to 0$.

(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; indeed, 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 $s,t \le T$.

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

Throughout this section, we fix a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\le T}, \mathbb{Q})$ supporting a two-dimensional Brownian motion $(W, W^\perp)$ and 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 function $g_0 \in L^2([0,T];\mathbb{R})$, there exists a progressively measurable $(\mathbb{R}\times\mathbb{R}_+)$-valued strong solution (X,S) to (1.5), (1.6) such that

$$\sup_{t\in[0,T]} \mathbb{E}[|X_t|^p] < \infty, \qquad p \ge 1; \tag{3.2}$$

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 additional assumption on the kernel that

$$\sup_{s\in[0,T]} \int_0^T |K(t,s)|^2 dt < \infty. \tag{3.3}$$

Note that in contrast to (3.1), we are integrating here over the first variable.

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 $0 \le \Re(u) \le 1$ and $\Re(w) \le 0$. Then we have*

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

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

$$\phi_t = -\int_t^T \mathrm{Tr}(\mathbf{\Psi}_t \dot{\mathbf{\Sigma}}_t) dt, \tag{3.5}$$

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

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

see Lemma B.1 below, and Tr *is the trace operator; see Appendix A.1.*

Proof We refer to Appendix B. □

The following remark establishes the link between ϕ and the Fredholm determinant.

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

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

and that $\tilde{\mathbf{\Sigma}}_t$ 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 $\mathrm{id} - 2\tilde{\mathbf{\Sigma}}_t a$; see Lemma A.6 and Simon [56, Theorem 3.9]. We set

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

Differentiation using the logarithmic derivative of the Fredholm determinant (see Gohberg and Krein [39, Chap IV, Sect. 1, Eq. (1.3)]) and (1.11) yields

$$\begin{aligned}\dot{\phi}_t &= \mathrm{Tr}\left(a(\mathrm{id} - 2\tilde{\mathbf{\Sigma}}_t a)^{-1}\dot{\tilde{\mathbf{\Sigma}}}_t\right)\\ &= \mathrm{Tr}\left(a(\mathrm{id} - 2\tilde{\mathbf{\Sigma}}_t a)^{-1}(\mathrm{id} - b\boldsymbol{K})^{-1}\dot{\mathbf{\Sigma}}_t(\mathrm{id} - b\boldsymbol{K}^*)^{-1}\right).\end{aligned}$$

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

$$\dot{\phi}_t = \mathrm{Tr}(\boldsymbol{\Psi}_t \dot{\boldsymbol{\Sigma}}_t). \tag{3.7}$$

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 discretisation of the expression (3.5). We refer to Sect. 4.1 for the numerical implementation.

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

$$g_0(t) = X_0 + \theta t, \qquad t \geq 0, \tag{3.8}$$

for some $X_0, \theta \in \mathbb{R}$, one recovers from Theorem 3.3 the well-known closed-form expressions of Stein and Stein [59] and Schöbel and Zhu [54], and that of Heston [43] 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*

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

where A, B, C solve the 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 [49, Eqs. (43)–(45)].

Proof (sketch) The characteristic function is given by (3.4). If $K(t,s) = \mathbf{1}_{\{s<t\}}$ and g_0 is as in (3.8), we obtain

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

so that taking conditional expectation yields

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

It follows that

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

with

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

Combined with (3.6) and (3.7), we obtain (3.9) with A such that $A_T = 0$ and

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

with

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

Using the operator Riccati equation satisfied by $t \mapsto \mathbf{\Psi}_t$, see Lemma B.1 below, and straightforward computations as in Abi Jaber et al. [7, Corollary 5.14] leads to the claimed system of Riccati equations for (A, B, C). □

4 Numerical illustration

In this section, we make use of the analytic expression (1.7) for the characteristic function to price options. We first present an approximation of the formula (1.7) using closed-form expressions obtained from a natural discretisation of the operators. **Throughout this section**, we consider the case $t = 0$, and we fix a Volterra kernel K as in Theorem 3.3.

4.1 A straightforward approximation by closed-form expressions

The expression (1.7) lends itself to approximation by closed-form solutions by a simple discretisation as in Fredholm [32] of the operator $\mathbf{\Psi}_0$ given by (1.9). Fix $n \in \mathbb{N}$ and let $t_i = iT/n$, $i = 0, 1, \dots, n$, be a partition of $[0, T]$. Discretising the product $\star$ given in (1.13) yields an approximation for $\mathbf{\Psi}_0$ by the $n \times n$ matrix

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

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

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

and

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

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

$$\Sigma^n_{ij} = \nu^2 \int_0^T K(t_{i-1}, s) K(t_{j-1}, s) ds, \qquad 1 \le i, j \le n. \tag{4.2}$$

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

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

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

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 (4.3) to avoid complex discontinuities, either by switching the sign of the determinant each time it crosses the axis of the negative real numbers or by discretising (3.5), which would require the computation of $\boldsymbol{\Psi}_t$ for several values of t but has the advantage of being analytic on the whole domain. We refer to Mayerhofer [51] for more details on 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 discretisation of the operator and the approximation of the Fredholm determinant based on the so-called Nyström method; see for instance Bornemann [13, 14], Corlay [22], Kang et al. [46].

Remark 4.3 For the case $u = 0$ and $\kappa = 0$, the previous approximation formulas agree with the ones derived in Abi Jaber [3, Sect. 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 fractional Stein–Stein model based on the Riemann–Liouville fractional Brownian motion with the Volterra convolution kernel $K(t, s) = \mathbf{1}_{\{s<t\}} (t - s)^{H-1/2} / \Gamma(H + 1/2)$. This model is given by

$$\begin{aligned} dS_t &= S_t X_t dB_t, \qquad S_0 > 0, \\ X_t &= g_0(t) + \frac{\kappa}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} X_s ds \\ &\quad + \frac{\nu}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} dW_s, \end{aligned}$$

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 assume that the input curve g_0, which can be used in general to fit at-the-money curves observed in the market, has the parametric form

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

Remark 4.4 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 θ time-dependent for instance). The advantage is that g_0 can then be estimated from the market to match certain term structures today (e.g. the term structure of forward variances, etc.). For illustration purposes, and since a comparison with the standard Stein–Stein model is given, we restrict here to the above parametric form of g_0.

Remark 4.5 It would also have been possible to take instead of the fractional Riemann–Liouville Brownian motion the true fractional Brownian motion by considering

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

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

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

We now move on 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 as

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

but also to price derivatives on the stock price and the integrated variance by Fourier inversion techniques; see Carr and Madan [18], Fang and Oosterlee [30], Lewis [48] among many others. In the sequel, we make use of the cosine method of Fang and Oosterlee [30] to price European call options on the stock S combined with our approximation formulas of Sect. 4.1. We start by observing that the kernel Σ_0 is given

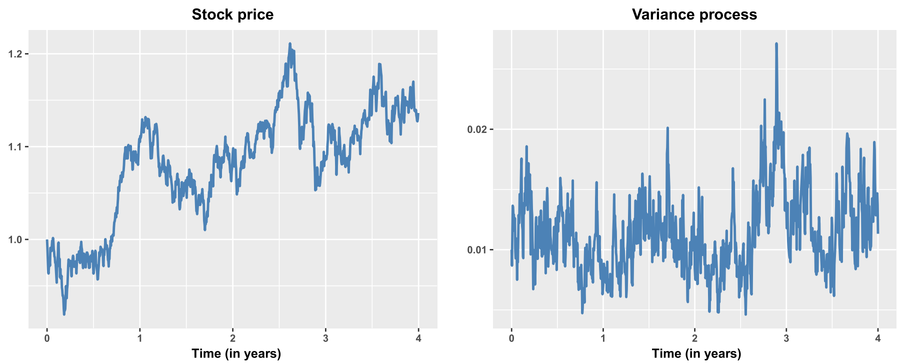

Fig. 2 One simulated sample path of the stock price S and the instantaneous variance process X^2 in the rough Stein–Stein model with parameters $X_0 = 0.1$, $\kappa = 0$, $\theta = 0.01$, $\nu = 0.02$, $\rho = -0.7$ and $H = 0.1$

in the closed form

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

where $\alpha = H + 1/2$ and ${}_2F_1$ is the Gaussian hypergeometric function; see for instance Malyarenko [50, second example below Definition 2.36]. Fix $n \in \mathbb{N}$ and a partition $0 = t_0 < t_1 < \cdots < t_n = T$. It follows that the $n\times n$ matrices (4.1), (4.2) can be computed in closed form as

$$\begin{aligned}K^n_{ij} &= \mathbf{1}_{\{j\le i-1\}}\frac{1}{\Gamma(1+\alpha)}\big((t_{i-1}-t_{j-1})^\alpha-(t_{i-1}-t_j)^\alpha\big), \qquad i,j\le n,\\ \Sigma^n_{ij} &= \frac{\nu^2}{\Gamma(\alpha)\Gamma(1+\alpha)}\frac{{t_{i-1}}^\alpha}{t_{j-1}^{1-\alpha}}\,{}_2F_1\bigg(1,1-\alpha;1+\alpha;\frac{t_{i-1}}{t_{j-1}}\bigg), \qquad i\le j\le n,\\ \Sigma^n_{ji} &= \Sigma^n_{ij}, \qquad i\le j\le n,\end{aligned}$$

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 Σ^n has zeros in its first row and first column. Note that in the case of Remark 4.5, the expression for the covariance function simplifies to $\Sigma_0(s,u) = \frac{\nu^2}{2}(s^{2H} + u^{2H} - |s-u|^{2H})$. The final ingredient to compute (4.3) is the vector g_n whose elements are given by

$$g^i_n = g_0(t_{i-1}) = X_0 + \theta\frac{{t_{i-1}}^\alpha}{\Gamma(1+\alpha)}, \qquad 1\le i\le n.$$

As a sanity check, we visualise in Fig. 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

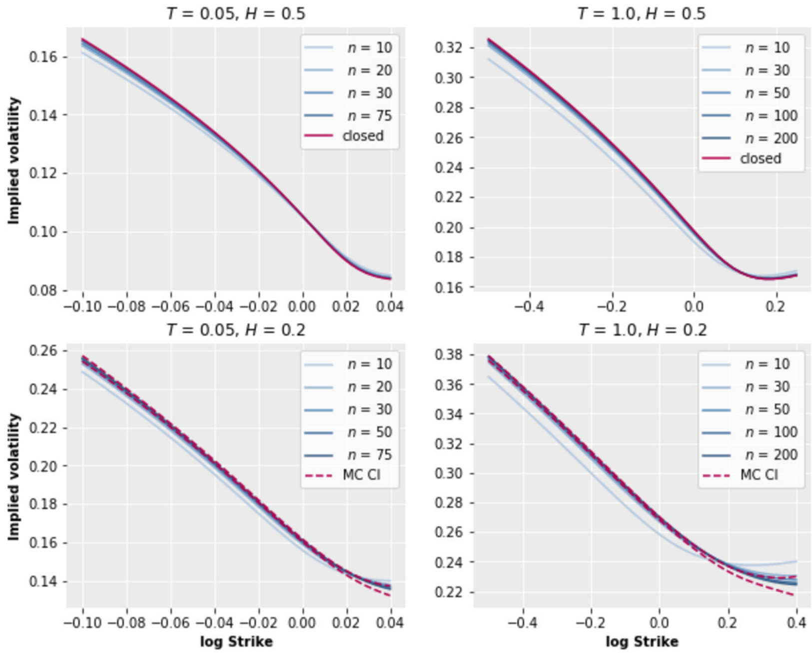

Fig. 3 Convergence of the implied volatility slices for short ($T = 0.05$ years) and long maturities ($T = 1$ year) of the operator discretisation of Sect. 4.1 towards: (i) the explicit solution of the conventional Stein–Stein model ($H = 0.5$, upper graphs); (ii) the 95% Monte Carlo confidence intervals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

closed-form expressions for the characteristic function of the conventional Stein–Stein model, see Lord and Kahl [49], and for $H = 0.2$ using Monte Carlo simulation. We see that the smaller the maturity, the faster the convergence. Other discretisation 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 this is left for future research.

Going back to real market data, we calibrate the fractional Stein–Stein model first to the at-the-money skew of Fig. 1 (b). Keeping the parameters $X_0 = 0.44$, $\theta = 0.3$, $\kappa = 0$ fixed, the calibrated parameters are given by

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

This power-law behaviour of the at-the-money skew observed in the market is perfectly captured by the fractional Stein–Stein model as illustrated in Fig. 4 with only three parameters. We then calibrate to the implied volatility surface of the S&P across several maturities to obtain Fig. 5.

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

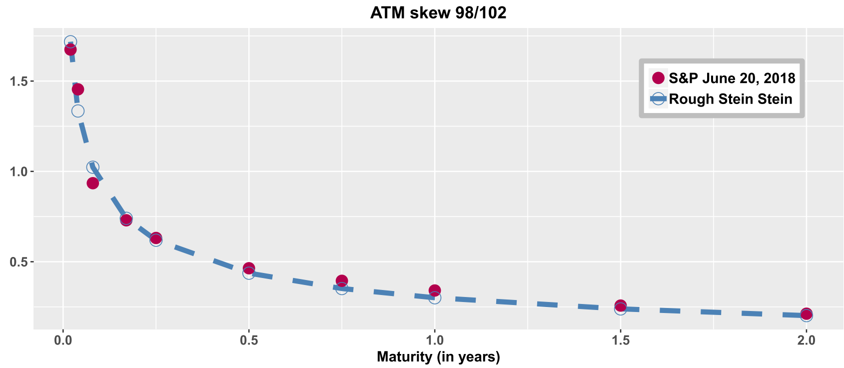

Fig. 4 Term structure of the at-the-money skew for the S&P index on June 20, 2018 (red dots) and for the rough Stein–Stein model with calibrated parameters (4.5) (blue circles with dashed line)

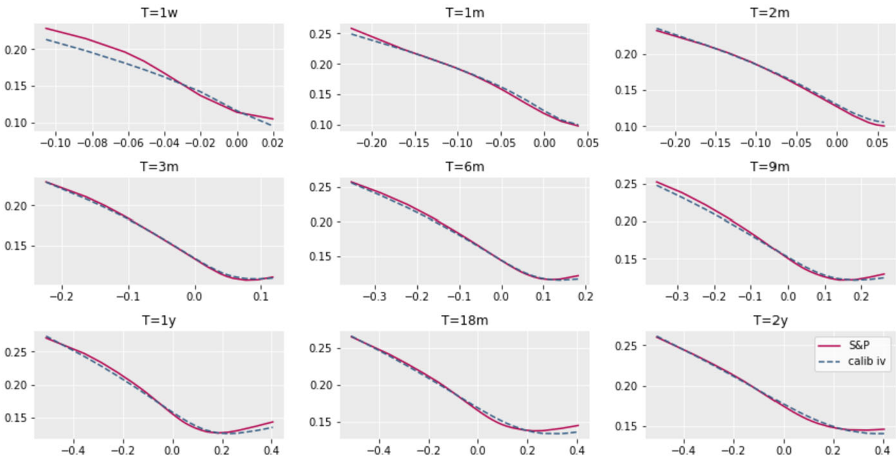

Fig. 5 The implied volatility surface of the S&P index (red) and the calibrated fractional Stein–Stein model (blue) with parameters $\hat{X}_0 = 0.113$, $\hat{\theta} = -0.044$, $\hat{\kappa} = -8.9 \times 10^{-5}$, $\hat{\nu} = 0.176$, $\hat{\rho} = -0.704$ and $\hat{H} = 0.279$

Appendix 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 [39, Chap. IV], Gohberg et al. [38, Chap. II], Simon [56] and also Bornemann [13, Sects. 2 and 3]. Let $\boldsymbol{A}$ be a linear compact operator acting on $L^2([0,T];\mathbb{C})$. Then the operator $\boldsymbol{A}$ has a countable spectrum. We recall that the spectrum of an operator, denoted by $\mathrm{sp}(\boldsymbol{A}) = (\lambda_n(\boldsymbol{A}))_{n \le N(\boldsymbol{A})}$, where $N(\boldsymbol{A})$ is either a finite integer or infinity, is the set of points $\lambda \in \mathbb{C}$ for which there does not exist a bounded inverse operator $(\lambda \mathrm{id} - \boldsymbol{A})^{-1}$. Whenever $\boldsymbol{A}$ is a linear operator induced by a kernel

$A \in L^2([0,T]^2;\mathbb{C})$, $\boldsymbol{A}$ is a Hilbert–Schmidt operator from $L^2([0,T];\mathbb{C})$ to itself and is in particular compact.

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

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

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 it is called the *trace* of the operator $\boldsymbol{A}$. Furthermore, Lidskii's theorem, see Simon [57, Theorem 3.7], ensures that

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

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

$$\operatorname{Tr}(\boldsymbol{K}\boldsymbol{L}) = \int_0^T (K\star L)(s,s)ds; \tag{A.2}$$

see Brislawn [16, Proposition 3].

Furthermore, the equivalence

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

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

$$\det(\mathrm{id}+z\boldsymbol{A}) = \prod_{n=1}^{N(\boldsymbol{A})}\big(1+z\lambda_n(\boldsymbol{A})\big)$$

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

$$\det(\mathrm{id}+z\boldsymbol{A}) = \sum_{n\geq 0}\frac{z^n}{n!}\int_0^T\cdots\int_0^T \operatorname{Det}\Big(\big(A(s_i,s_j)\big)_{1\leq i,j\leq n}\Big)ds_1\cdots ds_n. \tag{A.3}$$

The determinant (A.3) is named after Fredholm [32] 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 $R_T \in L^2([0,T]^2;\mathbb{C})$ as the unique solution to

$$R_T = K + K \star R_T, \qquad K \star R_T = R_T \star K. \tag{A.4}$$

In terms of integral operators, this translates into

$$\boldsymbol{R}_T = \boldsymbol{K} + \boldsymbol{K}\boldsymbol{R}_T, \qquad \boldsymbol{K}\boldsymbol{R}_T = \boldsymbol{R}_T\boldsymbol{K}. \tag{A.5}$$

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

$$(\mathrm{id} - \boldsymbol{K})^{-1} = \mathrm{id} + \boldsymbol{R}_T. \tag{A.6}$$

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 We first argue the existence of a resolvent R_T. It follows from (3.1) and the Cauchy–Schwarz inequality that

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

meaning that the function $\mathcal{K} : t \mapsto (s \mapsto K(t,s))$ is bounded and continuous from $[0,T]$ to $L^1([0,T],\mathbb{R})$. Since $[0,T]$ is a compact set, $\mathcal{K}$ is uniformly continuous. Therefore, the kernel K is of bounded and uniformly continuous type in the terminology of Gripenberg et al. [40, Definition 9.5.2]. An application of Gripenberg et al. [40, Theorem 9.5.5-(ii)], yields that K admits a resolvent kernel R_T which is again a Volterra kernel that satisfies

$$|R_T|_{L^1} := \sup_{t\in[0,T]} \int_0^T |R_T(t,s)|ds < \infty.$$

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

$$\int_0^T \int_0^T |R_T(t,s)|^2 dtds < \infty. \tag{A.7}$$

An application of Jensen's inequality on the normalised measure

$$|R_T(t,z)|dz \Big/ \int_0^T |R_T(t,z')|dz'$$

yields

$$\begin{aligned}
&\int_0^T \int_0^T |(R_T \star K)(t,s)|^2 dtds \\
&\le |R_T|_{L^1} \sup_{r\in[0,T]} \int_0^T |K(r,s)|^2 ds \int_0^T \int_0^T |R_T(t,z)| dtdz < \infty.
\end{aligned}$$

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

$$\int_0^T |(K \star R_T)(t,u)|^2 du \leq \sup_{t' \in [0,T]} \int_0^T |K(t',z)|^2 dz \int_0^T \int_0^T |R_T(u,z)|^2 du dz < \infty$$

for all $t \leq T$, which combined with (3.1) and (A.7) gives

$$\sup_{t \in [0,T]} \int_0^T |R_T(t,s)|^2 ds < \infty,$$

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

$$\begin{aligned}\int_0^T |(K \star R_T)(t+h,s) - (K \star R_T)(t,s)|^2 ds \leq & \int_0^T \int_0^T |R_T(u',s)|^2 du' ds \\ & \times \int_0^T |K(t+h,u) - K(t,u)|^2 du,\end{aligned}$$

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

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

which yields the second condition in (3.1) for R_T. □

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*

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

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

where R_T^κ is the resolvent kernel of κK 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 κK yields the existence of a resolvent R_T^κ satisfying (3.1). We define X as in (A.8) and write it in compact form as

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

where we used the notation $\boldsymbol{R}_T^\kappa(\nu dW)(t) = \int_0^t R_T^\kappa(t,s)\nu dW_s$. We observe that X is measurable and adapted as the sum of Lebesgue and Itô integrals. It therefore admits a progressively measurable modification by virtue of Ondreját and Seidler [52]. We now show that X solves (1.6). Using (A.6), composing both sides by $(\mathrm{id} + \boldsymbol{R}_T^\kappa)^{-1} = \mathrm{id} - \kappa\boldsymbol{K}$ and invoking the stochastic Fubini theorem yields

$$(\mathrm{id} - \kappa\boldsymbol{K})(X) = g_0 + (\mathrm{id} - \kappa\boldsymbol{K})\frac{1}{\kappa}\boldsymbol{R}_T^\kappa(\nu dW) = g_0 + \boldsymbol{K}(\nu dW),$$

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

$$\begin{aligned} X_t &= g_0(t) + \kappa(\boldsymbol{K})(X)(t) + \boldsymbol{K}(\nu dW)(t) \\ &= g_0(t) + \kappa\int_0^t K(t,s)X_s ds + \int_0^t K(t,s)\nu dW_s, \end{aligned}$$

yielding that X is a strong solution of (1.6). Furthermore, (3.2) follows from the fact that $\sup_{s\in[0,T]}\int_0^T |R_T^\kappa(s,u)|^2 du < \infty$ 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 of the solution follows via similar computations, by showing that any solution X to (1.6) is of the form (A.8) using the resolvent equation. □

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

Lemma A.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 [3, Lemma 3.2]. □

Lemma A.5 *Fix $b \in \mathbb{C}$ and a kernel K as in Definition 3.1. Then $\mathrm{id} - b\boldsymbol{K}$ is invertible. Furthermore, for all $t \le T$, $\tilde{\boldsymbol{\Sigma}}_t$ given by (1.11) is an integral operator of trace class with continuous kernel and can be rewritten in the form*

$$\tilde{\boldsymbol{\Sigma}}_t = (\mathrm{id} - b\boldsymbol{K}_t)^{-1}\boldsymbol{\Sigma}_t(\mathrm{id} - b\boldsymbol{K}_t^*)^{-1}, \tag{A.10}$$

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

Proof Lemma A.2 yields the existence of the resolvent R_T^b of bK which is again a Volterra kernel that satisfies (3.1). Thus (A.6) yields that $\mathrm{id} - b\boldsymbol{K}$ is invertible with an inverse given by $\mathrm{id} + \boldsymbol{R}_T^b$. To prove (A.10), we fix $t \le T$ and observe that since $\Sigma_t(s,u) = 0$ whenever $s \wedge u \le t$, we have

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

where we defined the kernel $R^b_{t,T}(s,u) = R^b_T(s,u)\mathbf{1}_{\{u\geq t\}}$. In a similar way, we obtain $\Sigma_t \star (R^b_T)^* = \Sigma_t \star (R^b_{t,T})^*$. Using the resolvent equation (A.4) of R^b_T, it readily follows that $R^b_{t,T}$ is the resolvent of bK_t so that $(\mathrm{id} - b\boldsymbol{K_t})^{-1} = \mathrm{id} + \boldsymbol{R}^b_{t,T}$. Combining all of the above leads to

$$\begin{aligned}
\tilde{\boldsymbol{\Sigma}}_t &= (\mathrm{id} - b\boldsymbol{K})^{-1}\boldsymbol{\Sigma}_t(\mathrm{id} - b\boldsymbol{K}^*)^{-1} \\
&= (\mathrm{id} + \boldsymbol{R}^b_T)\boldsymbol{\Sigma}_t(\mathrm{id} + \boldsymbol{R}^b_T)^* \\
&= \boldsymbol{\Sigma}_t + \boldsymbol{R}^b_T\boldsymbol{\Sigma}_t + \boldsymbol{\Sigma}_t(\boldsymbol{R}^b_T)^* + \boldsymbol{R}^b_T\boldsymbol{\Sigma}_t(\boldsymbol{R}^b_T)^* \\
&= \boldsymbol{\Sigma}_t + \boldsymbol{R}^b_{t,T}\boldsymbol{\Sigma}_t + \boldsymbol{\Sigma}_t(\boldsymbol{R}^b_{t,T})^* + \boldsymbol{R}^b_{t,T}\boldsymbol{\Sigma}_t(\boldsymbol{R}^b_{t,T})^* \qquad (A.11) \\
&= (\mathrm{id} + \boldsymbol{R}^b_{t,T})\boldsymbol{\Sigma}_t(\mathrm{id} + \boldsymbol{R}^b_{t,T})^* \\
&= (\mathrm{id} - b\boldsymbol{K}_t)^{-1}\boldsymbol{\Sigma}_t(\mathrm{id} + b\boldsymbol{K}^*_t)^{-1},
\end{aligned}$$

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

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 $\mathrm{id} - 2\tilde{\boldsymbol{\Sigma}}_t a$ is invertible and $\boldsymbol{\Psi}_t$ given by (1.9) is well defined. Furthermore, if $\Im(a) = \Im(b) = 0$, then $\boldsymbol{\Psi}_t$ is a symmetric nonpositive operator in the sense of Definition 2.1.*

Proof 1) Using Lemma A.5, we write

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

with

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

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

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

The operator $\mathbf{I}$ is symmetric, nonnegative and invertible so that $\mathrm{sp}(\mathbf{I}) \subseteq (0,\infty)$. Furthermore, since $2\Re(a) + \frac{\Im(b)^2}{\nu^2} \leq 0$ by assumption and $\boldsymbol{\Sigma}_t$ is symmetric nonnegative, we have $\mathrm{sp}(\mathbf{II}) \subseteq [0,\infty)$. It follows that $\mathrm{sp}(\Re(\boldsymbol{A}_t)) \subseteq (0,\infty)$, showing that

$0 \notin \mathrm{sp}(\boldsymbol{A}_t)$ and that $\boldsymbol{A}_t$ is invertible. Thus $\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_t$ is invertible. Together with Lemma A.5, we obtain that $\boldsymbol{\Psi}_t$ is well defined.

2) Now assume that $\Im(a) = \Im(b) = 0$. Clearly, $\tilde{\boldsymbol{\Sigma}}_t$ defined as in (1.11) is a symmetric nonnegative operator with a continuous kernel on $[0, T]^2$; recall Lemma A.4. So an application of Mercer's theorem (see Shorack and Wellner [55, Chap. V, Sect. 3, Theorem 1]) 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

$$\tilde{\boldsymbol{\Sigma}}_t = \sum_{n\geq 1} \lambda_n \langle e_n, \cdot \rangle e_n.$$

This yields

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

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

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

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

$$\langle f, \boldsymbol{\Psi}_t f \rangle = a \langle \tilde{f}, (\mathrm{id} - 2a\tilde{\boldsymbol{\Sigma}}_t)^{-1} \tilde{f} \rangle \leq 0,$$

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

Appendix 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

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

so that an application of Lemma A.6 yields that $\boldsymbol{\Psi}_t$ is well defined.

We now collect from Abi Jaber et al. [7, Lemma 5.8] further properties of $t \mapsto \boldsymbol{\Psi}_t$, and in particular its link with an operator Riccati equation. We recall that $t \mapsto \boldsymbol{\Psi}_t$ is said to be *strongly differentiable at time* $t \geq 0$ if there exists a bounded linear operator $\dot{\boldsymbol{\Psi}}_t$ from $L^2([0, T]; \mathbb{C})$ into itself such that

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

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

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

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

(ii) *For any $f \in L^2([0,T];\mathbb{R})$, with $1_t : s \mapsto \mathbf{1}_{\{t\le s\}}$, we have*

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

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

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

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

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

Proof The proof follows from a straightforward adaptation of the proof of Abi Jaber et al. [7, Lemma 5.6]. □

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

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

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

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

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

Proof We first set

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

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

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

The Leibniz rule yields

$$d\langle g_t, \boldsymbol{\Psi}_t g_t\rangle_{L^2} = \big(-a\bar{g}_t(t)^2 - \bar{g}_t(t)(\bar{\boldsymbol{\Psi}}_t g_t)(t)\big)dt \\ + \int_t^T d\big(a\bar{g}_t(s)^2 + \bar{g}_t(s)(\bar{\boldsymbol{\Psi}}_t g_t)(s)\big)ds \qquad dt \times \mathbb{Q}\text{-a.e.} \tag{B.4}$$

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

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

An application of Itô's lemma on the square yields

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

Furthermore, we write

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

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

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

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

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

Combining the previous three identities, we get the dynamics of

$$U_t(s) := a\bar{g}_t(s)^2 + \big(\bar{g}_t(s)(\bar{\boldsymbol{\Psi}}_t g_t)(s)\big)$$

as

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

We now integrate in s to get 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 nullset and we fix $(t,\omega) \in ([0,T] \times \Omega) \setminus \mathcal{N}$. In the sequel, all equalities are written for this particular ω. First, using $\dot{\Sigma}_t(s,s) = -\nu^2 K(s,t)^2$ and recalling that

$$\boldsymbol{\Psi} = a\mathrm{id} + \bar{\boldsymbol{\Psi}}, \tag{B.6}$$

we obtain that

$$\int_t^T \big(\mathbf{I}(s) + \mathbf{VII}(s)\big) ds = -\mathrm{Tr}(\boldsymbol{\Psi}_t \dot{\boldsymbol{\Sigma}}_t).$$

Indeed, the operator $\boldsymbol{\Psi}_t \dot{\boldsymbol{\Sigma}}_t = a\dot{\boldsymbol{\Sigma}}_t + \bar{\boldsymbol{\Psi}}_t \dot{\boldsymbol{\Sigma}}_t$ is of trace class because (i) $\dot{\boldsymbol{\Sigma}}_t$ is of trace class since it can be written as the product of two Hilbert–Schmidt integral operators $\dot{\boldsymbol{\Sigma}}_t = \tilde{\boldsymbol{K}}_t \tilde{\boldsymbol{K}}_t^*$ with $\tilde{K}_t(s,z) = K(s,t)/\sqrt{T}$, so that by (A.2), $\mathrm{Tr}(\dot{\boldsymbol{\Sigma}}_t) = \int_0^T \dot{\Sigma}_t(s,s) ds$, and (ii) $\bar{\boldsymbol{\Psi}}_t \dot{\boldsymbol{\Sigma}}_t$ is of trace class as the product of two Hilbert–Schmidt integral operators so that (A.2) yields $\mathrm{Tr}(\bar{\boldsymbol{\Psi}}_t \dot{\boldsymbol{\Sigma}}_t) = \int_0^T \int_0^T \bar{\psi}_t(s,z) \dot{\Sigma}_t(z,s) dz ds$. Combining (B.6) with Lemma B.1 (ii) and the fact that $\bar{\boldsymbol{\Psi}}^* = \bar{\boldsymbol{\Psi}}$, we obtain that

$$\begin{aligned}
\int_t^T \big(\mathbf{II}(s) + \mathbf{III}(s) + \mathbf{VI}(s)\big) ds &= 2\kappa X_t \int_0^T K(s,t)(\boldsymbol{\Psi}_t g_t)(s) ds \\
&= 2\kappa X_t (\boldsymbol{K}^* \boldsymbol{\Psi}_t g_t)(t).
\end{aligned}$$

On the other hand, we have

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

$$\int_0^T \mathbf{V}(s) ds = -X_t (\bar{\boldsymbol{\Psi}}_t g_t)(t),$$

$$\int_0^T \big(\mathbf{VIII}(s) + \mathbf{IX}(s) + \mathbf{X}(s)\big) ds = 2\nu (\boldsymbol{K}^* \boldsymbol{\Psi}_t)(g_t)(t).$$

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 $\bar{g}_t(t) = X_t$ yields

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

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

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

$$\dot{\phi}_t = \mathrm{Tr}(\boldsymbol{\Psi}_t \dot{\boldsymbol{\Sigma}}_t). \tag{B.7}$$

Proof of Theorem 3.3 It suffices to prove that (3.4) holds for all $0 \le u \le 1$ and $w \le 0$ to obtain the claimed expression by analytic continuation. 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 [61, 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 $0 \le u \le 1$ and $w \le 0$, it remains valid on $\{(u, w) \in \mathbb{C}^2 : 0 \le \Re(u) \le 1 \text{ and } \Re(v) \le 0\}$ by analytic continuation. Fix $u \in [0, 1]$, $w \in \mathbb{R}_-$. Set

$$U_t = u \log S_t + w \int_0^t X_s^2 ds + \phi_t + \langle g_t, \boldsymbol{\Psi}_t g_t\rangle_{L^2} \tag{B.8}$$

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 $t \le T$ yields (3.4).

We first prove that M is a local martingale by explicitly writing its dynamics. Observe that

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

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

$$dU_t = \Big(\langle g_t, \dot{\boldsymbol{\Psi}}_t g_t \rangle_{L^2} - \frac{u^2}{2} X_t^2 - 2u\rho\nu X_t (\boldsymbol{K}^* \boldsymbol{\Psi}_t)(g_t)(t) + \dot{\phi}_t - \mathrm{Tr}(\boldsymbol{\Psi}_t \dot{\boldsymbol{\Sigma}}_t) \Big) dt$$

$$+ \big(\rho u X_t + 2\nu (\boldsymbol{K}^* \boldsymbol{\Psi}_t)(g_t)(t) \big) dW_t + u\sqrt{1-\rho^2} X_t dW_t^{\perp},$$

so that

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

Observing that

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

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

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

by virtue of the Riccati equations (B.1) and (B.7). This shows that M is a local martingale.

It remains to argue that M is a true martingale. To this end, we fix $t \le T$. An application of the second part of Lemma A.6 yields that $\boldsymbol{\Psi}_t$ is a symmetric nonpositive operator so that, recall (B.7),

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

Therefore, since $w \le 0$ and $0 \le u \le 1$, it follows from (B.8) that

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

Therefore,

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

with $N_t = S_0^u \exp(-\frac{u^2}{2} \int_0^t X_s^2 ds + u \int_0^t X_s dB_s)$ which can be shown to be a true martingale by a similar argument to that used in Abi Jaber et al. [6, Lemma 7.3]. So the local martingale M is bounded by a martingale, which gives that M is also a true martingale. The proof is complete. □

Acknowledgement I should like to thank Shaun Li for interesting discussions.

References

1. Abi Jaber, E.: Lifting the Heston model. Quant. Finance **19**, 1995–2013 (2019)
2. Abi Jaber, E.: Weak existence and uniqueness for affine stochastic Volterra equations with L^1-kernels. Bernoulli **27**, 1583–1615 (2021)
3. Abi Jaber, E.: The Laplace transform of the integrated Volterra Wishart process. Math. Finance **32**, 309–348 (2022)
4. Abi Jaber, E., El Euch, O.: Markovian structure of the Volterra Heston model. Stat. Probab. Lett. **149**, 63–72 (2019)
5. Abi Jaber, E., El Euch, O.: Multifactor approximation of rough volatility models. SIAM J. Financ. Math. **10**, 309–349 (2019)
6. Abi Jaber, E., Larsson, M., Pulido, S.: Affine Volterra processes. Ann. Appl. Probab. **29**, 3155–3200 (2019)
7. Abi Jaber, E., Miller, E., Pham, H.: Markowitz portfolio selection for multivariate affine and quadratic Volterra models. SIAM J. Financ. Math. **12**, 369–409 (2021)
8. Alòs, E., León, J.A., Vives, J.: On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. Finance Stoch. **11**, 571–589 (2007)
9. Andersen, T.G., Bollerslev, T.: Intraday periodicity and volatility persistence in financial markets. J. Empir. Finance **4**, 115–158 (1997)
10. Bayer, C., Friz, P., Gatheral, J.: Pricing under rough volatility. Quant. Finance **16**, 887–904 (2016)
11. Bennedsen, M., Lunde, A., Pakkanen, M.S.: Decoupling the short-and long-term behavior of stochastic volatility. J. Financ. Econom. (2022). https://doi.org/10.1093/jjfinec/nbaa049. Forthcoming
12. Black, F., Scholes, M.: The pricing of options and corporate liabilities. J. Polit. Econ. **81**, 637–654 (1973)
13. Bornemann, F.: On the numerical evaluation of distributions in random matrix theory: a review. Markov Process. Relat. Fields **16**, 803–866 (2010)
14. Bornemann, F.: On the numerical evaluation of Fredholm determinants. Math. Comput. **79**, 871–915 (2010)
15. Brezis, H.: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer, Berlin (2010)
16. Brislawn, C.: Kernels of trace class operators. Proc. Am. Math. Soc. **104**, 1181–1190 (1988)
17. Callegaro, G., Grasselli, M., Pages, G.: Fast hybrid schemes for fractional Riccati equations (rough is not so tough). Math. Oper. Res. **46**, 221–254 (2021)
18. Carr, P., Madan, D.: Option valuation using the fast Fourier transform. J. Comput. Finance **2**(4), 61–73 (1999)
19. Carr, P., Wu, L.: The finite moment log stable process and option pricing. J. Finance **58**, 753–777 (2003)
20. Comte, F., Renault, E.: Long memory in continuous-time stochastic volatility models. Math. Finance **8**, 291–323 (1998)
21. Cont, R.: Empirical properties of asset returns: stylized facts and statistical issues. Quant. Finance **1**, 223–236 (2001)
22. Corlay, S.: The Nyström method for functional quantization with an application to the fractional Brownian motion. Preprint (2010). Available online at https://arxiv.org/abs/1009.1241
23. Cox, J.C., Ingersoll, J.E. Jr., Ross, S.A.: An intertemporal general equilibrium model of asset prices. Econometrica **53**, 363–384 (1985)
24. Cuchiero, C., Teichmann, J.: Markovian lifts of positive semidefinite affine Volterra type processes. Decis. Econ. Finance **42**, 407–448 (2019)
25. Cuchiero, C., Teichmann, J.: Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case. J. Evol. Equ. **20**, 1301–1348 (2020)
26. Decreusefond, L., Üstünel, A.S.: Stochastic analysis of the fractional Brownian motion. Potential Anal. **10**, 177–214 (1999)
27. Ding, Z., Granger, C.W., Engle, R.F.: A long memory property of stock market returns and a new model. J. Empir. Finance **1**, 83–106 (1993)
28. Duffie, D., Filipović, D., Schachermayer, W.: Affine processes and applications in finance. Ann. Appl. Probab. **13**, 984–1053 (2003)
29. El Euch, O., Rosenbaum, M.: The characteristic function of rough Heston models. Math. Finance **29**, 3–38 (2019)

30. Fang, F., Oosterlee, C.W.: A novel pricing method for European options based on Fourier-cosine series expansions. SIAM J. Sci. Comput. **31**, 826–848 (2009)
31. Fouque, J.-P., Papanicolaou, G., Sircar, R., Solna, K.: Multiscale stochastic volatility asymptotics. Multiscale Model. Simul. **2**, 22–42 (2003)
32. Fredholm, I.: Sur une classe d'équations fonctionnelles. Acta Math. **27**, 365–390 (1903)
33. Fukasawa, M.: Asymptotic analysis for stochastic volatility: martingale expansion. Finance Stoch. **15**, 635–654 (2011)
34. Fukasawa, M.: Volatility has to be rough. Quant. Finance **21**, 1–8 (2021)
35. Gatheral, J., Jaisson, T., Rosenbaum, M.: Volatility is rough. Quant. Finance **18**, 933–949 (2018)
36. Gatheral, J., Keller-Ressel, M.: Affine forward variance models. Finance Stoch. **23**, 501–533 (2019)
37. Gatheral, J., Radoičić, R.: Rational approximation of the rough Heston solution. Int. J. Theor. Appl. Finance **22**, 1950010 (2019)
38. Gohberg, I., Goldberg, S., Krupnik, N.: Traces and Determinants of Linear Operators. Birkhäuser, Basel (2012)
39. Gohberg, I., Krein, M.G.: Introduction to the Theory of Linear Nonselfadjoint Operators. American Mathematical Soc., Providence (1978)
40. Gripenberg, G., Londen, S.-O., Staffans, O.: Volterra Integral and Functional Equations. Cambridge University Press, Cambridge (1990)
41. Gulisashvili, A., Viens, F., Zhang, X.: Extreme-strike asymptotics for general Gaussian stochastic volatility models. Ann. Finance **15**, 59–101 (2019)
42. Harms, P., Stefanovits, D.: Affine representations of fractional processes with applications in mathematical finance. Stoch. Process. Appl. **129**, 1185–1228 (2019)
43. Heston, S.L.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. **6**, 327–343 (1993)
44. Horvath, B., Jacquier, A., Lacombe, C.: Asymptotic behaviour of randomised fractional volatility models. J. Appl. Probab. **56**, 496–523 (2019)
45. Kac, M., Siegert, A.J.: On the theory of noise in radio receivers with square law detectors. J. Appl. Phys. **18**, 383–397 (1947)
46. Kang, S.-Y., Koltracht, I., Rawitscher, G.: Nyström–Clenshaw–Curtis quadrature for integral equations with discontinuous kernels. Math. Comput. **72**, 729–756 (2003)
47. Lee, R.W.: Implied volatility: statics, dynamics, and probabilistic interpretation. In: Baeza-Yates, R., et al. (eds.) Recent Advances in Applied Probability, pp. 241–268. Springer, Boston (2005)
48. Lewis, A.L.: A simple option formula for general jump-diffusion and other exponential Lévy processes. Preprint (2001). Available online at https://ssrn.com/abstract=282110
49. Lord, R., Kahl, C.: Why the rotation count algorithm works. Tinbergen Institute Discussion Paper No. 2006-065/2 (2006). Available online at https://ssrn.com/abstract=921335
50. Malyarenko, A.: Invariant Random Fields on Spaces with a Group Action. Springer, Berlin (2012)
51. Mayerhofer, E.: Reforming the Wishart characteristic function. Preprint (2019). Available online at https://arxiv.org/abs/1901.09347
52. Ondreját, M., Seidler, J.: On existence of progressively measurable modifications. Electron. Commun. Probab. **18**, 1–6 (2013)
53. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Springer, Berlin (1999)
54. Schöbel, R., Zhu, J.: Stochastic volatility with an Ornstein–Uhlenbeck process: an extension. Rev. Finance **3**, 23–46 (1999)
55. Shorack, G.R., Wellner, J.A.: Empirical Processes with Applications to Statistics. SIAM, Philadelphia (2009)
56. Simon, B.: Notes on infinite determinants of Hilbert space operators. Adv. Math. **24**, 244–273 (1977)
57. Simon, B.: Trace Ideals and Their Applications, 2nd edn. American Mathematical Soc., Providence (2005)
58. Sottinen, T., Viitasaari, L.: Stochastic analysis of Gaussian processes via Fredholm representation. Int. J. Stoch. Anal. **2016**, 8694365 (2016)
59. Stein, E.M., Stein, J.C.: Stock price distributions with stochastic volatility: an analytic approach. Rev. Financ. Stud. **4**, 727–752 (1991)
60. Veraar, M.: The stochastic Fubini theorem revisited. Stoch. Int. J. Probab. Stoch. Process. **84**, 543–551 (2012)
61. Widder, D.V.: The Laplace Transform. Princeton University Press, Princeton (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.