ORIGINAL ARTICLE

MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics WILEY

The rough Hawkes Heston stochastic volatility model

Alessandro Bondi[1] | **Sergio Pulido**[2] | **Simone Scotti**[3]

[1]Centre de Mathématiques Appliquées (CMAP), CNRS, École polytechnique, Institut Polytechnique de Paris, Palaiseau, France

[2]Université Paris-Saclay, CNRS, ENSIIE, Univ Evry, Laboratoire de Mathématiques et Modélisation d'Evry (LaMME), Gif-sur-Yvette, France

[3]Università di Pisa, Pisa, Italy

Correspondence
Simone Scotti, Università di Pisa, Pisa, Italy.
Email: simone.scotti@unipi.it

Funding information
Institut Louis Bachelier; Università di Pisa, Grant/Award Number: PRA_2022_21

Abstract

We study an extension of the Heston stochastic volatility model that incorporates rough volatility and jump clustering phenomena. In our model, named the rough Hawkes Heston stochastic volatility model, the spot variance is a rough Hawkes-type process proportional to the intensity process of the jump component appearing in the dynamics of the spot variance itself and the log returns. The model belongs to the class of affine Volterra models. In particular, the Fourier-Laplace transform of the log returns and the square of the volatility index can be computed explicitly in terms of solutions of deterministic Riccati-Volterra equations, which can be efficiently approximated using a multi-factor approximation technique. We calibrate a parsimonious specification of our model characterized by a power kernel and an exponential law for the jumps. We show that our parsimonious setup is able to simultaneously capture, with a high precision, the behavior of the implied volatility smile for both S&P 500 and VIX options. In particular, we observe that in our setting the usual shift in the implied volatility of VIX options is explained by a very low value of the power in the kernel. Our findings demonstrate the relevance, under an affine framework, of rough volatility and self-exciting jumps in order to capture the joint evolution of the S&P 500 and VIX.

1 | INTRODUCTION

The Black-Scholes model, where volatility is constant, and more generally classical local volatility models, where volatility is a function of time and spot asset prices, fail to reproduce the dynamics of implied volatility smiles of options written on the underlying asset. To overcome this limitation, multiple stochastic, stochastic-local, and path-dependent volatility models have been developed and studied in recent years. The complexity of volatility modeling, however, has increased with the significant growth over time of markets on volatility indices, such as the VIX. The rise in popularity of these markets is explained in part by their relevance to protect portfolios Rhoads (2011). It has therefore become fundamental to develop stochastic models able to capture the joint dynamics of the underlying prices and their volatility index. This modeling challenge, known as the joint S&P 500/VIX calibration puzzle (Guyon, 2020a; Guyon, 2020b), has inspired the introduction of more sophisticated models, for example, Abi Jaber et al. (2022a), Abi Jaber et al. (2022b), Cuchiero, Gazzani et al. (2023), Gatheral et al. (2020), Guyon (2020b), Guyon and Lekeufack (2023), Rømer (2022a), that incorporate new features to the joint dynamics of the underlying and the volatility in order to solve the problem. In this paper we tackle the challenge by proposing a tractable affine model with rough volatility and volatility jumps that cluster and that have the opposite direction but occur at the same time as the jumps of the underlying prices. In this introduction we give a brief literature review to explain the choice of our framework.

The dynamics of the VIX volatility index are highly complex. In particular, they exhibit large and systematically positive variations over very short periods, with a tendency to form clusters of spikes during difficult periods like the 2008 financial crisis and the beginning of the COVID-19 pandemic in 2020. This is accompanied by very long periods without any large fluctuation and a less important mean reversion speed. These observations are in line with an increasing number of studies that indicate the presence of jumps in the volatility (Dotsis et al., 2007; Todorov & Tauchen, 2011), on the underlying (Bates, 1996), and the fact these jumps are common to the volatility and underlying (Cont & Kokholm, 2013; Sepp, 2008).

The growing interest in volatility indices has driven the standardization of contingent claims written on the volatility indices themselves. These volatility index markets have very unique features. For VIX futures and exchange-traded products these features are studied in Avellaneda and Papanicolaou (2018). The complexity of volatility markets is also exemplified by the difficulty to jointly model the behavior of the volatility smiles of vanilla options written on the underlying and its volatility index, see for instance Alos et al. (2022), Pacati et al. (2018), Papanicolaou (2022), and Recchioni et al. (2021). This longstanding puzzle is known as the S&P 500 (SPX)/VIX calibration puzzle. A growing body of literature explains the difficulty arguing that "the state-of-the-art stochastic volatility models in the literature cannot capture the S&P 500 and VIX option prices simultaneously," see Song and Xiu (2012). As pointed out in Guyon (2020b), "all the attempts at solving the joint SPX/VIX smile calibration problem only produced imperfect, approximate fits."

The problem is that usual stochastic models either fail to reproduce one or both shapes of the implied volatility for S&P 500 and VIX options or, when both the shapes are coherent, the implied volatility levels are incorrect.

Access to high frequency data has improved our understanding of the microstructure of financial markets and the effects on volatility. In particular, recent studies indicate that non-Markovian models with rough volatility trajectories might be appropriate to better capture long time dependencies due to meta orders and the large contribution of automatic orders. This is examined in Cont (2011) which provides a general analysis of order-driven markets, the work in Comte and Renault (1998) which elucidates the memory-features of volatility, and the studies in El Euch et al. (2018), and Gatheral et al. (2018) which give a micro-structural justification to the newly developed rough volatility models.

From a modeling point of view, affine models provide a convenient framework because they are flexible and, thanks to semi-explicit formulas for the Fourier-Laplace transform, fast computations can be performed using Fourier-based techniques (Duffie et al., 2000; Duffie et al., 2003; Filipović, 2001). The most popular affine stochastic volatility model is the Heston model Heston (1993), where the spot variance is a square-root mean-reverting CIR (Cox-Ingersoll-Ross-Cox et al., 1985) process. This model is able to reproduce some stylized features like the mean-reverting property of the volatility and the leverage effect. It is, however, unable to reproduce other phenomena such as extreme paths of volatility during crisis periods (even for large values of the volatility of volatility parameter) and the at-the-money (ATM) skews of underlying options' implied volatility simultaneously for short and long maturities. These limitations, and the micro-structural behavior of markets described in the previous paragraph, motivated the introduction of the rough Heston model (El Euch & Rosenbaum, 2018; El Euch & Rosenbaum, 2019). The rough Heston model is tractable as it belongs to the class of affine Volterra models (Abi Jaber et al., 2019), and semi-explicit formulas for the Fourier-Laplace transform are still available. Unfortunately, this model cannot reproduce the features of options written on the volatility index and the underlying simultaneously (see Subsection 6.1).

In order to model the joint behavior of S&P 500 and VIX markets, consistent with empirical evidence, we add two specific features to the usual Heston model. First, we incorporate rough volatility by adding a power kernel proportional to $t^{\alpha-1}$, with $\alpha \in (1/2, 1]$, to the dynamics of the spot variance. Second, we postulate common jumps for the volatility and the underlying with a negative leverage. The presence of jumps in both underlying and variance helps to reproduce a skewed implied volatility for vanilla options as in the Barndorff-Nielsen and Shephard model (Barndorff-Nielsen & Shephard, 2001a; Barndorff-Nielsen & Shephard, 2001b). Inspired by the Hawkes framework, taking into account jump-clustering and endogeneity of financial markets, we model the spot variance to be proportional to the intensity process of the jump component appearing in the dynamics of the spot variance itself and the log returns. For these reasons, we name our model the *rough Hawkes Heston model*. Our approach shares similarities with the study in Bernis et al. (2021), which demonstrates that an exponential law for the jump size can capture upward VIX implied volatility within a Hawkes framework with an exponential kernel. Alternatively, other studies such as Jiao et al. (2021) and Nicolato et al. (2017) have explored the inclusion of more general jump measures. The advantage of the rough Hawkes Heston model over the aforementioned works is the possibility to achieve a joint calibration of the SPX/VIX smiles, while utilizing a simple and parsimonious specification for the jump distribution, namely an exponential law (see Section 6).

To keep mathematical and numerical tractability, we choose an affine specification of the model. As such, our model belongs to the class of affine Volterra processes (Abi Jaber et al.,

2019), which has been recently extended to jump processes in Bondi et al. (2024), Cuchiero and Teichmann (2019), Cuchiero and Teichmann (2020). In particular, the Fourier-Laplace transform of the log returns and the square of the volatility index can be computed explicitly in terms of solutions of deterministic Riccati-Volterra equations, see Theorems 3.3 and 4.6. We approximate the solutions of the Riccati-Volterra equations via a multi-factor scheme as in Abi Jaber and El Euch (2019b) applying a recent Gaussian quadrature method in Bayer and Breneis (2023), see Theorem 5.1 and Remark 5.2. We leave for future study the implementation and analysis in our framework of other methods such as the Adams method (Diethelm et al., 2002; Diethelm et al., 2004), asymptotic formulas inspired by forest expansions in the spirit of Alos et al. (2020), and hybrid approximation techniques for Volterra equations similar to those in Callegaro et al. (2021). It might be also interesting to investigate an adaptation to Riccati-Volterra equations of the hybrid multi-factor approach proposed in Rømer (2022b) for the discretization of stochastic Volterra equations.

The affine property is an advantage of our modeling approach compared to other recent models proposed to solve the SPX/VIX calibration problem, such as Abi Jaber et al. (2022a), Cuchiero et al. (2023), Gatheral et al. (2020), Guyon and Mustapha (2023), Rosenbaum and Zhang (2022), where pricing is done via Monte Carlo or machine learning techniques. In the case of signature-based models (Cuchiero, Gazzani et al., 2023; Cuchiero et al., 2023), their affine nature is explored in (Cuchiero et al., 2023, Section 6). However, no numerical experiments involving Fourier pricing techniques are presented, partly due to the challenging structure of the resulting Riccati equations. Our affine framework is convenient also because variance swap rates and the square VIX index have explicit affine relations to the forward curve, see Corollary 4.2 and Remark 4.3. This is a generalization, to the affine Volterra setting, of the affine relation already pointed out in Kallsen et al. (2011) within the classical affine exponential framework and empirically confirmed in Mancino et al. (2020).

Previous literature on jump-diffusion models focusing on the evolution of S&P 500 and the VIX proposes either high-dimensional models (Pacati et al., 2018; Sepp, 2008), or models based on hidden Markov chains (Goutte et al., 2017; Papanicolaou & Sircar, 2014). These models require a large number of parameters or suffer from the lack of interpretability of the random factors. It is also important to mention (Cont & Kokholm, 2013), where the authors investigate a forward variance jump model which allows them to price options on the volatility index by Fourier inversion techniques. To price options on the underlying asset, however, this approach requires a Monte Carlo simulation of the variance swap rates, and it might introduce additional parameters for every considered maturity. Our approach to model the joint SPX/VIX dynamics is different. As in Bernis et al. (2021), we keep the number of parameters low by assuming that the jump intensity is proportional to the variance process itself, and jumps are common to the volatility and underlying with opposite signs. The main new ingredient of our model, compared to Bernis et al. (2021), is the addition of a Brownian component and a power kernel to the variance process. This generates by construction a jump clustering effect and takes into account related findings in the rough volatility literature (Alos et al., 2007; Alos & Shiraya, 2019; Bayer et al., 2016; Bennedsen et al., 2021; El Euch et al., 2018; El Euch & Rosenbaum, 2019; Fukasawa, 2017; Gatheral et al., 2018; Gatheral et al., 2020; Livieri et al., 2018). In particular, our model is consistent with the so-called *Zumbach effect*. Indeed, in this paper we suggest an extension of the rough Heston model, which reproduces the Zumbach effect according to El Euch et al. (2020).

The rough Hawkes Heston model is able to reconcile the shapes and level of the S&P 500 and VIX volatility smiles. An important role is played by the parameter α characterizing the kernel. As is the case for other rough volatility models, this parameter controls the term structure of ATM skews for SPX option smiles. We show that when α is near to $1/2$, the power law is in the range $[0.5, 0.6]$. This is consistent with similar findings in the rough volatility literature (Alos & Shiraya,

2019; Bayer et al., 2016; Bennedsen et al., 2021; El Euch et al., 2018; Fukasawa, 2017; Gatheral et al., 2018; Gatheral et al., 2020). In addition, in our framework, the parameter α plays a crucial role because it controls the level of the implied volatility of VIX options for short maturities. We observe, that as α approaches $1/2$ the levels of S&P 500 and VIX smiles are coherent.

To summarize, the model that we propose in this paper shares many features with other existing models. These features are mainly: rough volatility (Bayer et al., 2016; El Euch et al., 2018; El Euch & Rosenbaum, 2019; Fukasawa, 2017; Gatheral et al., 2018; Gatheral et al., 2020), jumps (Barndorff-Nielsen & Shephard, 2001a, Barndorff-Nielsen & Shephard, 2001b; Bates, 1996; Cont & Kokholm, 2013; Grzelak, 2022; Pacati et al., 2018; Sepp, 2008), the Hawkes/branching character of volatility (Bernis et al., 2021; Brignone & Sgarra, 2019; Horst & Xu, 2022), and the affine structure (Abi Jaber et al., 2019; Bondi et al., 2024; Duffie et al., 2003; Duffie et al., 2000; Filipović, 2001; Kallsen et al., 2011; Jiao et al., 2021).

Consequently we take advantage of the low regularity and memory features of rough volatility models, the large fluctuation of jumps, the clusters of Hawkes processes and the explicit Fourier-Laplace transform of the affine setup. The specification that we adopt for the joint SPX/VIX calibration is parsimonious with only five evolution-related parameters. Moreover, all the parameters have a financial interpretation. The parameter α in the kernel controls the decay of the volatility memory, SPX ATM skews and the level of VIX smiles. We have in addition the classical parameters controlling the volatility mean reversion speed and the volatility of volatility, and two parameters related to the leverage effect that specify the correlation between Brownian motions and between the jumps in the asset and its volatility. Despite its robustness, the rough Hawkes Heston stochastic volatility model captures remarkably well the implied volatility surfaces of S&P 500 and VIX at the same time.

The paper is organized as follows. Section 2 lays out the essential hypotheses of our study and introduces the stochastic model in a general setup, that is, with a general kernel and law for the jumps. Section 3 explains the derivation of the Fourier-Laplace transform of the log returns and the application to option pricing on the underlying. Section 4 focuses on the VIX index characterizing the Fourier-Laplace transform of the VIX^2, and describes the Fourier-based formulas to price options on the VIX. Section 5 studies the multi-factor numerical scheme used in order to approximate the solutions to the Riccati-Volterra equations arising in Sections 3 and 4. Section 6 details the calibration of our model to S&P 500 and VIX options data, showing also a comparison with the (continuous) rough Heston model, see Subsection 6.1. Section 7 presents a complete and detailed sensitivity analysis of implied volatility curves with respect to the model parameters. In particular, Subsection 7.1 focuses on the effect of the jump component. Section 8 summarizes the conclusions of our study. Appendix A contains the proof of the necessary existence, uniqueness and comparison results for the Riccati-Volterra equations appearing in Section 3. Appendix B presents the proof of the Fourier-inversion formula used to price options on the underlying. To finish, in Appendix C we prove the main result related to the convergence of the multi-factor approximation scheme for the Riccati-Volterra equations.

2 | THE MODEL

We study a stochastic volatility model where the spot variance $\sigma^2 = (\sigma_t^2)_{t\geq 0}$ is a predictable process, with trajectories in $L^2_{loc}(\mathbb{R}_+)$, defined on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{Q}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0})$. We assume that the filtration $\mathbb{F}$ satisfies the usual conditions and that $\mathcal{F}_0$ is the trivial $\sigma-$ algebra.

We consider, throughout our study, a kernel K that satisfies the next requirement, see Abi Jaber (2021), Abi Jaber et al. (2021), Abi Jaber et al. (2019), Bondi et al. (2024).

Hypothesis 2.1. The kernel $K \in L^2_{loc}(\mathbb{R}_+)$ is nonnegative, nonincreasing, not identically zero and continuously differentiable on $(0, \infty)$. Furthermore, its resolvent of the first kind L is nonnegative and nonincreasing, that is, $s \mapsto L([s, s+t])$ is nonincreasing for every $t \geq 0$.

We recall that, given a kernel $K \in L^1_{loc}(\mathbb{R}_+; \mathbb{R}^{d \times d})$, an $\mathbb{R}^{d \times d}-$ valued measure L is called its (measure) resolvent of the first kind if $L * K = K * L = I$, where $I \in \mathbb{R}^{d \times d}$ is the identity matrix. The resolvent of the first kind does not always exist, but if it does then it is unique, see (Gripenberg et al., 1990, Theorem 5.2, Chapter 5). Under Hypothesis 2.1, the existence of the resolvent of the first kind is ensured by (Gripenberg et al., 1990, Theorem 5.5, Chapter 5).

Let g_0 be a function representing the initial spot variance curve. A parametric form of g_0 will be specified for the application of the model (see Section 6). At this point, however, we only make the following assumption.

Hypothesis 2.2. g_0 is continuous and nondecreasing, with $g_0(0) \geq 0$.

Fix $b \in \mathbb{R}$, $c > 0$, an $\mathbb{F}-$ Brownian motion $W_2 = (W_{2,t})_{t \geq 0}$ and a nonnegative measure ν on $\mathbb{R}_+$ such that $\nu(\{0\}) = 0$ and that $\int_{\mathbb{R}_+} |z|^2 \nu(dz) < \infty$. We assume that the spot variance σ^2 is a $\mathbb{Q} \otimes dt-$ a.e. nonnegative predictable process satisfying

$$\sigma^2 = g_0 + K * dZ, \quad \mathbb{Q} \otimes dt - \text{a.e.}, \tag{1}$$

where Z is the following semimartingale having jump measure $\mu(dt, dz)$ and compensator $\sigma_t^2 dt \otimes \nu(dz)$:

$$dZ_t = b\,\sigma_t^2 dt + \sqrt{c}\,\sigma_t\,dW_{2,t} + \int_{\mathbb{R}_+} z\,\big(\mu(dt, dz) - \sigma_t^2 dt \otimes \nu(dz)\big), \quad Z_0 = 0.$$

Therefore, the instantaneous variance σ^2 satisfies a stochastic affine Volterra equation of convolution type with jumps. From now on, we denote by $\tilde{\mu}(dt, dz) = \mu(dt, dz) - \sigma_t^2 dt \otimes \nu(dz)$ the compensated jump measure of Z. Since the intensity of the jumps of σ^2 is proportional to σ^2 itself, the spot variance is a Hawkes-type process, which is coherent with other models that incorporate endogeneity of financial markets such as Bernis et al. (2021), Callegaro et al. (2022), El Euch and Rosenbaum (2019), Gonzato and Sgarra (2021), Jiao et al. (2017). In the sequel, we denote by $\tilde{Z} = (\tilde{Z}_t)_{t \geq 0}$ the process $d\tilde{Z}_t = \sqrt{c}\,\sigma_t\,dW_{2,t} + \int_{\mathbb{R}_+} z\,\tilde{\mu}(dt, dz)$, $t \geq 0$, with starting condition $\tilde{Z}_0 = 0$. Notice that $\tilde{Z}$ is a square-integrable martingale by (Bondi et al., 2024, Lemma 1).

According to (Bondi et al., 2024, Appendix A)

$$\sigma^2 = g_0 - R_{-bK} * g_0 + E_{b,K} * d\tilde{Z}, \quad \mathbb{Q} \otimes dt - \text{a.e.}, \tag{2}$$

where R_{-bK} is the resolvent of the second kind of $-bK$ and $E_{b,K}$ is the canonical resolvent of K with parameter b. We recall that the resolvent of the second kind R_K for a kernel $K \in L^1_{loc}(\mathbb{R}_+)$ is the unique solution $R_K \in L^1_{loc}(\mathbb{R}_+)$ of the two equations $K * R_K = R_K * K = K - R_K$, see (Gripenberg et al., 1990, Theorem 3.1, Chapter 2) and the subsequent definition. The canonical resolvent $E_{\lambda,K}$ of K with parameter λ is defined by $E_{\lambda,K} = -\lambda^{-1} R_{-\lambda K}$ for $\lambda \neq 0$, whereas $E_{0,K} = K$.

Remark 2.3. If we assume that K and the shifted kernels $K(\cdot + 1/n)$, $n \in \mathbb{N}$, satisfy Hypothesis 2.1, then under Hypothesis 2.2 the (weak) existence of the spot variance process σ^2, satisfying (1), is ensured by (Abi Jaber, 2021, Theorem 2.13) and (Bondi et al., 2024, Lemma 9). Assuming weak existence, weak uniqueness is established in (Bondi et al., 2024, Corollary 12) under Hypothesis 2.1. We refer to Abi Jaber et al. (2021) and Bondi et al. (2024) for more information about stochastic Volterra equations and stochastic convolution for processes with jumps.

A useful tool for the development of the theory is the adjusted forward process, which we now define. For every $t \geq 0$, it is denoted by $(g_t(s))_{s>t}$ and it is a jointly measurable process on $\Omega \times (t, \infty)$ such that

$$g_t(s) = g_0(s) + \int_0^t K(s - r)\mathrm{d}Z_r, \quad \mathbb{Q} - \text{a.s.}, s > t. \tag{3}$$

Thanks to (Meyer, 1966, Theorem 46) and the fact that $\mathbb{F}$ satisfies the usual conditions, we can consider $g_t(\cdot)$ to be $\mathcal{F}_t \otimes \mathcal{B}(t, \infty)-$ measurable.

Analogous arguments provide a version of the conditional expectation process $\mathbb{E}[\sigma^2|\mathcal{F}_t] = (\mathbb{E}[\sigma_s^2|\mathcal{F}_t])_{s>t}$ which is $\mathcal{F}_t \otimes \mathcal{B}(t, \infty)-$ measurable. In particular, from (2),

$$\mathbb{E}\left[\sigma_s^2\middle|\mathcal{F}_t\right] = g_0(s) - (R_{-bK} * g_0)(s) + \int_0^t E_{b,K}(s - r)\mathrm{d}\widetilde{Z}_r, \quad \mathbb{Q} - \text{a.s.}, s > t. \tag{4}$$

We now prescribe the dynamics of the log returns process $X = (X_t)_{t \geq 0}$ as follows:

$$\mathrm{d}X_t = -\left(\frac{1}{2} + \int_{\mathbb{R}_+} (e^{-\Lambda z} - 1 + \Lambda z)\nu(\mathrm{d}z)\right)\sigma_t^2 \, \mathrm{d}t + \sigma_t\left(\sqrt{1 - \rho^2}\, \mathrm{d}W_{1,t} + \rho \, \mathrm{d}W_{2,t}\right)$$

$$- \Lambda \int_{\mathbb{R}_+} z \, \tilde{\mu}(\mathrm{d}t, \mathrm{d}z), \quad X_0 = 0, \tag{5}$$

where $\rho \in [-1, 1]$ is a correlation parameter, $W_1 = (W_{1,t})_{t\geq 0}$ is an $\mathbb{F}-$ Brownian motion independent from W_2 and $\Lambda \geq 0$ is a leverage parameter forcing common jumps for volatility and underlying with opposite signs. This is coherent with empirical findings in Todorov and Tauchen (2011), stylized features studied in Cont (2001), and the financial/econometric literature with jumps, for example, Barndorff-Nielsen and Shephard (2001a), Bates (1996), Barndorff-Nielsen and Shephard (2001b), Bernis et al. (2021), Cont and Kokholm (2013), Curato and Sanfelici (2015), Raffaelli et al. (2022), Sepp (2008). We have assumed, for the sake of readability and without loss of generality, that interest rates and dividends are zero. The price process of the underlying asset will be $S = (S_t)_{t\geq 0} = (S_0 e^{X_t})_{t\geq 0}$, where $S_0 > 0$ represents the initial price. An application of Itô's formula shows that S is a local martingale. Indeed,

$$\frac{\mathrm{d}S_t}{S_{t-}} = -\left(\frac{1}{2} + \int_{\mathbb{R}_+} (e^{-\Lambda z} - 1 + \Lambda z)\nu(\mathrm{d}z)\right)\sigma_t^2\mathrm{d}t + \sigma_t\left(\sqrt{1 - \rho^2}\, \mathrm{d}W_{1,t} + \rho \, \mathrm{d}W_{2,t}\right)$$

$$- \Lambda \int_{\mathbb{R}_+} z \, \tilde{\mu}(\mathrm{d}t, \mathrm{d}z) + \frac{1}{2}\sigma_t^2 \, \mathrm{d}t + \int_{\mathbb{R}_+} (e^{-\Lambda z} - 1 + \Lambda z)\, \mu(\mathrm{d}t, \mathrm{d}z)$$

$$= \sigma_t \left(\sqrt{1 - \rho^2} \, dW_{1,t} + \rho \, dW_{2,t} \right) + \int_{\mathbb{R}_+} \left(e^{-\Lambda z} - 1 \right) \tilde{\mu}(dt, dz) =: dN_t,$$

where $N = (N_t)_{t \geq 0}$ is a local martingale with $N_0 = 0$. In particular, since S starts at S_0, it follows that $S = S_0 \mathcal{E}(N)$, where $\mathcal{E}$ denotes the Doléans-Dade exponential. In the next section, see Corollary 3.4, we will improve on this result by showing that, for every $T > 0$, the restriction of S to $[0, T]$ is a true martingale.

3 | THE FOURIER-LAPLACE TRANSFORM OF THE LOG RETURNS

In this section we study, for a fixed $T \geq 0$, the conditional Fourier-Laplace transform of X_T, $\mathbb{E}[e^{wX_T} \mid \mathcal{F}_t]$, $t \in [0, T]$. Here $w \in \mathbb{C}$ is subject to suitable conditions that will be specified in the sequel. In particular, we want to find a formula that allow us to compute the prices of options written on the underlying asset using Fourier-inversion techniques (Duffie et al., 2003; Duffie et al., 2000; Filipović, 2001; Grasselli, 2016). We will adopt the following notation: for $z \in \mathbb{C}$ we denote by $\operatorname{Re} z$ and $\operatorname{Im} z$ the real and imaginary parts of z, respectively. We let $\mathbb{C}_+$ [resp., $\mathbb{C}_-$] be the set of complex numbers with nonnegative [resp., nonpositive] real part.

Let us define the mapping $\mathcal{R} : \mathbb{C}_+ \times \mathbb{C}_- \to \mathbb{C}$ by

$$\mathcal{R}(u, v) = \frac{1}{2}(u^2 - u) + \left(b + \rho \sqrt{c} \, u \right) v + \frac{c}{2} v^2 + \int_{\mathbb{R}_+} \left[e^{(v - \Lambda u)z} - u \left(e^{-\Lambda z} - 1 \right) - 1 - vz \right] \nu(dz), \tag{6}$$

for every $(u, v) \in \mathbb{C}_+ \times \mathbb{C}_-$. For the development of the theory we need the following result about deterministic Riccati-Volterra equations, whose proof is postponed to Appendix A.

Theorem 3.1. *Suppose that K satisfies Hypothesis 2.1 and $w \in \mathbb{C}$ is such that $\operatorname{Re} w \in [0, 1]$.*

(i) There exists a unique continuous solution $\psi_w : \mathbb{R}_+ \to \mathbb{C}_-$ of the Riccati-Volterra equation

$$\psi_w(t) = \int_0^t K(t - s) \mathcal{R}(w, \psi_w(s)) ds = (K * \mathcal{R}(w, \psi_w(\cdot)))(t), \quad t \geq 0. \tag{7}$$

In particular, $\psi_{\operatorname{Re} w}$ is $\mathbb{R}_- -$ valued.
(ii) The following inequalities hold:

$$\operatorname{Re} \psi_w(t) \leq \psi_{\operatorname{Re} w}(t) \leq 0, \quad t \geq 0. \tag{8}$$

We also need the next preparatory lemma, which can be proven similarly to (Abi Jaber, 2021, Lemma 6.1).

Lemma 3.2. *Let f_1, f_2, f_3 : $[0, T] \to \mathbb{R}$ be bounded measurable functions such that $f_3 \leq 0$ in $[0, T]$. Then, the Doléans-Dade exponential*

$$\mathcal{E} \left(\int_0^t f_1(s) \sigma_s \, dW_{1,s} + \int_0^t f_2(s) \sigma_s \, dW_{2,s} + \int_0^t \int_{\mathbb{R}_+} \left(e^{f_3(s)z} - 1 \right) \tilde{\mu}(ds, dz) \right), \quad t \in [0, T]$$

is a martingale.

We are now ready to state the main result of this section. We introduce for every $\varepsilon \in \mathbb{R}$ the shift operator Δ_ε, which, given $I \subset \mathbb{R}$ and a function $f : I \to \mathbb{C}$, assigns the function $\Delta_\varepsilon f : I - \varepsilon \to \mathbb{C}$ defined by $\Delta_\varepsilon f(t) = f(t + \varepsilon)$, $t \in I - \varepsilon$.

Theorem 3.3. *Suppose that K satisfies Hypothesis 2.1 and that the resolvent of the first kind L is the sum of a locally integrable function and a point mass at 0. Moreover, suppose that the total variation bound*

$$\sup_{\varepsilon \in (0,\overline{T}]} \|\Delta_\varepsilon K * L\|_{TV([0,\overline{T}])} < \infty$$

holds for all $\overline{T} > 0$. Then, under Hypothesis 2.2, for every $w \in \mathbb{C}$ such that $\mathrm{Re}\, w \in [0, 1]$,

$$\mathbb{E}\left[\exp\{wX_T\}\big|\mathcal{F}_t\right] = \exp\left\{\widetilde{V}_t(w, T)\right\}, \quad \mathbb{Q} - a.s., \ t \in [0, T], \tag{9}$$

where $\widetilde{V}_t(w, T) = wX_t + \int_t^T \mathcal{R}(w, \psi_w(T - s))g_t(s)\mathrm{d}s, \ t \in [0, T]$.

Proof. Let $w \in \mathbb{C}$ be such that $\mathrm{Re}\, w \in [0, 1]$. Define the càdlàg, adapted, $\mathbb{C}-$ valued semimartingale $(V_t(w, T))_{t \in [0,T]}$ by

$$V_t(w, T) = V_0(w, T) + wX_t + \int_0^t \psi_w(T - s)\mathrm{d}\widetilde{Z}_s$$

$$- \int_0^t \left(\frac{1}{2}(w^2 - w) + \rho\sqrt{c}\, w\, \psi_w(T - s) + \frac{c}{2}\psi_w(T - s)^2\right. \tag{10}$$

$$\left. + \int_{\mathbb{R}_+} \left(e^{(-\Lambda w + \psi_w(T-s))z} - w\left(e^{-\Lambda z} - 1\right) - 1 - \psi_w(T - s)z\right)\nu(\mathrm{d}z)\right)\sigma_s^2\,\mathrm{d}s,$$

$$V_0(w, T) = \int_0^T \mathcal{R}(w, \psi_w(T - s))g_0(s)\mathrm{d}s. \tag{11}$$

The same arguments as in the proof of (Bondi et al., 2024, Theorem 5), which essentially rely on the stochastic Fubini's theorem (see, e.g., (Protter, 2005, Theorem 65, Chapter IV)), allow us to prove that

$$V_t(w, T) = \widetilde{V}_t(w, T), \quad \mathbb{Q} - a.s., \ t \in [0, T]. \tag{12}$$

We now define $H(w, T) = (H_t(w, T))_{t \in [0,T]} = (\exp\{V_t(w, T)\})_{t \in [0,T]}$. By Itô's formula and the dynamics in (5) and (10) we have, omitting (w, T) for sake of readability,

$$\frac{\mathrm{d}H_t}{H_{t-}} = \left[w\,\mathrm{d}X_t - \left(\frac{c}{2}\psi_w(T - t)^2 + \int_{\mathbb{R}_+} \left(e^{(-\Lambda w + \psi_w(T-t))z} - 1 - w\left(e^{-\Lambda z} - 1\right) - \psi_w(T - t)z\right)\nu(\mathrm{d}z)\right.\right.$$

$$\left.\left. + \frac{1}{2}(w^2 - w) + \rho\sqrt{c}\, w\, \psi_w(T - t)\right)\sigma_t^2\,\mathrm{d}t + \psi_w(T - t)\mathrm{d}\widetilde{Z}_t\right] + \frac{1}{2}\left(c\,\psi_w(T - t)^2 + w^2\right)\sigma_t^2\,\mathrm{d}t$$

$$+ \rho \sqrt{c}\, w\, \psi_w(T - t)\sigma_t^2\, \mathrm{d}t + \int_{\mathbb{R}_+} \big(e^{(-\Lambda w + \psi_w(T-t))z} - 1 - (-\Lambda w + \psi_w(T - t))z\big)\mu(\mathrm{d}t, \mathrm{d}z)$$

$$= \left[\sigma_t\left(w\sqrt{1 - \rho^2}\, \mathrm{d}W_{1,t} + \left(w\rho + \sqrt{c}\, \psi_w(T - t)\right)\mathrm{d}W_{2,t}\right) + \int_{\mathbb{R}_+} \big(e^{(-\Lambda w + \psi_w(T-t))z} - 1\big)\tilde{\mu}(\mathrm{d}t, \mathrm{d}z)\right],$$

with $H_0 = \exp(V_0)$. We define $N(w, T) = (N_t(w, T))_{t \in [0,T]}$ by $N_0(w, T) = 0$ and

$$\begin{aligned}
\mathrm{d}N_t(w, T) &= \sigma_t\left(w\sqrt{1 - \rho^2}\, \mathrm{d}W_{1,t} + \left(w\rho + \sqrt{c}\, \psi_w(T - t)\right)\mathrm{d}W_{2,t}\right) \\
&\quad + \int_{\mathbb{R}_+} \big(e^{(-\Lambda w + \psi_w(T-t))z} - 1\big)\tilde{\mu}(\mathrm{d}t, \mathrm{d}z).
\end{aligned}$$

Then $N(w, T)$ is a local martingale and the previous computations show that, omitting again (w, T), $H = \exp\{V_0\}\mathcal{E}(N)$ up to evanescence, where $\mathcal{E}$ denotes the Doléans-Dade exponential. Therefore $H(w, T)$ is a local martingale. If it is indeed a true martingale, then (9) directly follows from (12) noting also that $\widetilde{V}_T(w, T) = wX_T$.

In order to argue the martingale property of $H(w, T)$, first we observe that by Lemma 3.2 the real-valued process $H(\mathrm{Re}\, w, T) = (H_t(\mathrm{Re}\, w, T))_{t \in [0,T]} = (\exp\{V_t(\mathrm{Re}\, w, T)\})_{t \in [0,T]}$ is a true martingale. Secondly, we invoke (Bondi et al., 2024, Corollary 8) to obtain the following alternative expression for $V(w, T)$ (an analogous one holds for $V(\mathrm{Re}\, w, T)$)

$$\begin{aligned}
V_t(w, T) = wX_t &+ \int_0^{T-t} \mathcal{R}(w, \psi_w(s))g_0(T - s)\mathrm{d}s + \psi_w(T - t)L(\{0\})\big(\sigma^2 - g_0\big)(t) \\
&+ \big(\mathrm{d}\Pi_{T-t} * \big(\sigma^2 - g_0\big)\big)(t), \quad \text{for a.e. } t \in (0, T),\ \mathbb{Q} - \text{a.s.,} \quad (13)
\end{aligned}$$

where for every $\varepsilon > 0$, $\Pi_\varepsilon(t) = \int_0^\varepsilon \mathcal{R}(w, \psi_w(s))(\Delta_{\varepsilon-s}K * L)(t)\mathrm{d}s$, $t \geq 0$, is a locally absolutely continuous function. The application of this result is legitimate because the procedure carried out in Bondi et al. (2024) to infer (13) only depends on (1), (7) and the boundedness on compact intervals of $\mathbb{R}_+$ of $\mathcal{R}(w, \psi_w(\cdot))$, and does not rely on the expression of $\mathcal{R}$. A similar argument together with (8) and Hypothesis 2.2 allows us to parallel the comparison method in the proof of (Bondi et al., 2024, Theorem 11) to conclude that there is a constant $C > 0$ such that

$$|H_t(w, T)| = |\exp\{V_t(w, T)\}| = \exp\{\mathrm{Re}\, V_t(w, T)\} \leq C \exp\{V_t(\mathrm{Re}\, w, T)\} = CH_t(\mathrm{Re}\, w, T),$$

for $t \in [0, T]$, $\mathbb{Q} -$ a.s. At this point it is sufficient to invoke (Jarrow, 2018, Lemma 1.4) to claim that $H(w, T)$ is a true martingale, hence the proof is complete. $\qquad \square$

From the previous theorem we deduce the martingale property of our price process S with a direct approach (it can also be obtained by Lemma 3.2).

Corollary 3.4. *Under the hypotheses of Theorem 3.3, the price process $S = (S_t)_{t \in [0,T]}$ is a martingale.*

Proof. The computations at the end of Section 2 show that the stock price S is a nonnegative local martingale, hence it is a supermartingale. In order for it to be a martingale, it is sufficient to show

that $\mathbb{E}[S_T] = S_0$. By (9) in Theorem 3.3 with $w = 1$ we have

$$\mathbb{E}[S_T] = S_0 \exp\left\{\int_0^T \mathcal{R}(1, \psi_1(T-s))g_0(s)ds\right\}.$$

From (6)–(7), we observe that $\psi_1 \equiv 0$ in $\mathbb{R}_+$. This implies that $\mathcal{R}(1, \psi_1(\cdot)) = 0$ in $\mathbb{R}_+$, which concludes the proof. $\square$

Equation (9) in Theorem 3.3 gives a semi-explicit expression to compute the Fourier-Laplace transform Ψ^{X_T} of X_T in a suitable region of $\mathbb{C}$, namely

$$\Psi^{X_T}(w) = \exp\left\{\int_0^T \mathcal{R}(w, \psi_w(T-s))g_0(s)\,ds\right\}, \quad w \in \mathbb{C} \text{ such that } \operatorname{Re} w \in [0,1]. \tag{14}$$

As shown in the following proposition, whose proof is in Appendix B, we can use Ψ^{X_T} to price options with maturity T on the underlying asset S via Fourier-inversion techniques.

Proposition 3.5. *Fix a log strike $k > 0$. Then, under the hypotheses of Theorem 3.3, the price $C_S(k,T)$ of a call option on the underlying asset S with log strike k and maturity T is*

$$C_S(k,T) = S_0 - \frac{1}{\pi}\sqrt{S_0 e^k} \int_{\mathbb{R}_+} \operatorname{Re}\left[e^{i\lambda(\log(S_0)-k)}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)\right]\frac{1}{\frac{1}{4}+\lambda^2}\,d\lambda, \tag{15}$$

and the price $P_S(k,T)$ of a put option with the same log strike, maturity and underlying is

$$P_S(k,T) = e^k - \frac{1}{\pi}\sqrt{S_0 e^k} \int_{\mathbb{R}_+} \operatorname{Re}\left[e^{i\lambda(\log(S_0)-k)}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)\right]\frac{1}{\frac{1}{4}+\lambda^2}\,d\lambda. \tag{16}$$

Remark 3.6. The expression in (15) coincides with (Lewis, 2001, Formula (3.11)), but we have to independently prove it (see Appendix B). Indeed, in Lewis (2001) the author obtains (15) starting from the inversion of the generalized Fourier transform of the payoff function $w(x) = (e^x - e^k)^+$, $x \in \mathbb{R}$, of a call option with log strike k (here x represents the log price). Namely, for $x \in \mathbb{R}$,

$$w(x) = -\frac{1}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty} \frac{e^{k(iz+1)}}{z^2-iz}e^{-izx}dz, \quad z_i > 1.$$

If we were to follow the same approach here, then we would find a problem: we only have proved that Ψ^{X_T} is defined for complex numbers with real part in [0,1]. Therefore, in the previous expression, we would need $z_i \in [0,1]$, which is a contradiction. This setback cannot be immediately fixed by considering put options and then applying the put-call parity formula, because again the intersection between the complex strip ($z_i < 0$), where the Fourier transform for the payoff function is defined, and the strip where $\Psi^{X_T}(-i\cdot)$ is available is empty. We refer to (Schmelzle, 2010, Section 4) for a survey of pricing based on Fourier-inversion techniques.

Choosing a purely imaginary number w in Equation (14), that is, $w = it$, $t \in \mathbb{R}$, we obtain a semi-explicit formula for the characteristic function of X_T. As a result, we can approximate the price of an option on the underlying asset S with maturity T by the COS method in Fang and

Oosterlee (2009). This approach is more efficient than the exact formulas (15)–(16). In fact, in Section 6 we use the COS method to calibrate our model to market data. Its advantage relies on the reduced number of evaluations required for Ψ^{X_T}, leading to fewer numerical approximations of the Riccati-Volterra equation (7). However, given the sensitivity of the COS method with respect to the truncation bounds commonly denoted by a and b, we still need (15)–(16) as a benchmark. We refer to Section 6 for a more precise discussion on the topic.

4 | THE FOURIER-LAPLACE TRANSFORM OF VIX²

In this section the underlying asset S represents the SPX index. Then, according to the CBOE VIX white paper and Demeterfi et al. (1999), the theoretical value of $\text{VIX} = (\text{VIX}_T)_{T \geq 0}$ is

$$\text{VIX}_T = \sqrt{\left(-\frac{2}{\delta}\mathbb{E}\left[X_{T+\delta} - X_T | \mathcal{F}_T\right]\right)^+} \times 100, \quad T \geq 0. \tag{17}$$

Here $\delta = \frac{1}{12}$ and represents 30 days, the time to expiration of the log contracts involved in the computation of the index. Note that in (17), the positive part has been inserted to guarantee the good definition of the random variable VIX_T in the whole space Ω, however the radicand is nonnegative $\mathbb{Q}-$ a.s., as we are about to show.

We first derive, in the following theorem, an expression for $\mathbb{E}[X_{T+\delta} - X_T | \mathcal{F}_T]$, $T \geq 0$, in terms of the adjusted forward process at time T, $g_T(\cdot)$.

Theorem 4.1. *The log contract satisfies an infinite dimensional affine relation with respect to the adjusted forward process. More specifically,*

$$\mathbb{E}\left[X_{T+\delta} - X_T | \mathcal{F}_T\right] = c_1 \int_T^{T+\delta} \left(1 + b(E_{b,K} * 1)(T + \delta - s)\right) g_T(s) ds \leq 0, \quad \mathbb{Q} - a.s., \tag{18}$$

where $c_1 = -(\frac{1}{2} + \int_{\mathbb{R}_+} (e^{-\Lambda z} - 1 + \Lambda z)\nu(dz))$.

Proof. By (5) and the martingale property of the local martingale part of the expression (see (Bondi et al., 2024, Lemma 1)), we have

$$\mathbb{E}\left[X_{T+\delta} - X_T | \mathcal{F}_T\right] = c_1 \int_T^{T+\delta} \mathbb{E}[\sigma_s^2 | \mathcal{F}_T] ds, \quad \mathbb{Q} - a.s.$$

Recalling that $\sigma^2 \geq 0$, $\mathbb{Q} \otimes dt-$ a.e., we infer that $\mathbb{E}[\sigma_s^2 | \mathcal{F}_T] \geq 0$ for a.e. $s > T$, $\mathbb{Q}-$ a.s., hence the value of a log contract at time T is nonpositive $\mathbb{Q}-$ a.s.

By (2), (4), the stochastic Fubini's theorem – whose application is guaranteed by (Bondi et al., 2024, Lemma 1) – and a suitable change of variables, we infer that, $\mathbb{Q}-$ a.s.

$$c_1^{-1}\mathbb{E}\left[X_{T+\delta} - X_T | \mathcal{F}_T\right] = \int_0^{T+\delta} f_0(s) ds - \int_0^T \sigma_s^2 ds + \int_0^{T+\delta}\left(\int_0^T 1_{\{r \leq s\}} E_{b,K}(s - r) d\widetilde{Z}_r\right) ds$$

$$= \int_0^{T+\delta} f_0(s) ds - \int_0^T \sigma_s^2 ds + \int_0^T (E_{b,K} * 1)(T + \delta - r) d\widetilde{Z}_r$$

$$= \int_0^{T+\delta} f_0(s)\mathrm{d}s - \int_0^T \left(1 + b(E_{b,K} * 1)\right)(T+\delta-s)\sigma_s^2\,\mathrm{d}s + \int_0^T (E_{b,K} * 1)(T+\delta-r)\mathrm{d}Z_r,$$

(19)

where $f_0 = g_0 - R_{-bK} * g_0$. Notice that $E_{b,K} * 1$ is the unique, continuous (nonnegative) solution of the linear Volterra equation $\chi = K * (1 + b\chi)$. Then, another application of stochastic Fubini's theorem yields, $\mathbb{Q}-$ a.s.,

$$\int_0^T (E_{b,K} * 1)(T+\delta-r)\mathrm{d}Z_r = \int_0^T \left(\int_r^{T+\delta} K(s-r)\left(1 + b(E_{b,K} * 1)(T+\delta-s)\right)\mathrm{d}s \right)\mathrm{d}Z_r$$

$$= \int_0^{T+\delta} \left(1 + b(E_{b,K} * 1)(T+\delta-s)\right)\left(\int_0^T 1_{\{r \leq s\}}K(s-r)\mathrm{d}Z_r \right)\mathrm{d}s.$$

To conclude, we observe that by (Gripenberg et al., 1990, Theorem 2.2 (viii), Chapter 2)

$$-((R_{-bK} * g_0) * 1)(T+\delta) = b\left(\left(E_{b,K} * 1\right) * g_0\right)(T+\delta),$$

and plugging the previous two equalities in (19), together with (1), (3), we obtain the relation in (18). □

We deduce the following corollary showing an affine relation between the square of the VIX index and the adjusted forward process.

Corollary 4.2. *The square of* VIX *satisfies an infinite dimensional affine relation with respect to the adjusted forward process. More specifically*

$$\begin{aligned} \mathrm{VIX}_T^2 &= -10^4 \frac{2}{\delta}\mathbb{E}\left[X_{T+\delta} - X_T | \mathcal{F}_T\right], \quad \mathbb{Q} - a.s. \\ &= -10^4 \frac{2}{\delta}c_1 \int_T^{T+\delta} \left(1 + b(E_{b,K} * 1)(T+\delta-s)\right)g_T(s)\mathrm{d}s, \quad \mathbb{Q} - a.s., \end{aligned}$$

(20)

where $c_1 = -(\frac{1}{2} + \int_{\mathbb{R}_+} (e^{-\Lambda z} - 1 + \Lambda z)\nu(\mathrm{d}z)).$

Remark 4.3. Our framework also allows us to obtain an explicit infinite dimensional affine relation between the variance swaps and the adjusted forward process. Specifically, the variance swap rate is

$$\frac{1}{\delta}\mathbb{E}\left[[X,X]_{T+\delta} - [X,X]_T | \mathcal{F}_T\right] = \frac{c_2}{\delta} \int_T^{T+\delta} \left(1 + b(E_{b,K} * 1)(T+\delta-s)\right)g_T(s)\mathrm{d}s, \quad \mathbb{Q} - a.s., \quad (21)$$

where $c_2 = 1 + \Lambda^2 \int_{\mathbb{R}_+} |z|^2 \nu(\mathrm{d}z)$. Note that for $\Lambda = 0$ we have $c_2 = -2c_1$, hence in this case log contracts and variance swaps coincide up to the factor $-2/\delta$ (see (18)–(21)). Therefore, when there are no jumps in the dynamics of the underlying, by (20) we recover the fact that VIX^2 is a variance swap. Moreover, observe that the relation in (21) is an extension of (Kallsen et al., 2011, Lemma 4.4) in the classical affine setting. We refer to Cont and Kokholm (2013), Demeterfi et al. (1999), Mancino et al. (2020) for more details regarding the distinction between variance swaps and VIX^2.

We are now interested in finding the conditional Fourier-Laplace transform of VIX_T^2. Before addressing this question, we need some technical intermediate steps. We first recall the following functional space as defined in Abi Jaber and El Euch (2019a).

$$\mathcal{G}_K = \{g : \mathbb{R}_+ \to \mathbb{R} \text{ continuous: } g(0) \geq 0 \text{ and } \Delta_\varepsilon g - (\Delta_\varepsilon K * L)(0)g - d(\Delta_\varepsilon K * L) * g \geq 0, \varepsilon \geq 0\}. \tag{22}$$

Lemma 4.4. *Suppose that K satisfies Hypothesis 2.1. Define the function $h : \mathbb{R}_+ \to \mathbb{R}$ by*

$$h(t) = -10^4 \frac{2}{\delta} c_1 \left[1 + b(E_{b,K} * 1)(\delta - t)\right] 1_{\{t \leq \delta\}}, \quad t \geq 0.$$

Then h is a continuous nonnegative function on $[0, \delta)$ and $t \mapsto \int_{\mathbb{R}_+} h(s)K(s + t)ds$ belongs to $\mathcal{G}_K$.

Proof. The first step is to show that $1 + b(E_{b,K} * 1) \geq 0$ in $\mathbb{R}_+$, which implies that h is also nonnegative. This can be deduced from the fact that this function is the unique, continuous solution in $\mathbb{R}_+$ of the Volterra equation $\chi = 1 + bK * \chi$, which is nonnegative by (Abi Jaber and El Euch, 2019b, Theorem C.1). Secondly, h has compact support, and under Hypothesis 2.1 for every $\varepsilon \geq 0$ the function $\Delta_\varepsilon K * L$ is right-continuous nondecreasing in $\mathbb{R}_+$ and (see the proof of (Abi Jaber et al., 2019, Lemma 2.6))

$$\Delta_\varepsilon K = (\Delta_\varepsilon K * L)(0)K + d(\Delta_\varepsilon K * L) * K, \quad dt - \text{a.e. in } \mathbb{R}_+.$$

As a consequence, for every $t \geq 0$

$$\Delta_\varepsilon K(s + t) = (\Delta_\varepsilon K * L)(0)K(s + t) + (d(\Delta_\varepsilon K * L) * K)(s + t)$$

$$\geq (\Delta_\varepsilon K * L)(0)K(s + t) + \int_0^t K(s + t - u)d(\Delta_\varepsilon K * L)(u), \quad \text{for a.e. } s \in [0, \delta].$$

This implies, by Tonelli's theorem, that $t \mapsto \int_{\mathbb{R}_+} h(s)K(s + t)ds$ belongs to $\mathcal{G}_K$. □

We now define, for every $w \in \mathbb{C}_-$, the function $h_w(t) = w \cdot h(t)$, $t \geq 0$, and consider the Riccati-Volterra equation

$$\phi_w = \int_0^\infty h_w(s)K(s + \cdot)ds + K * (G(\phi_w(\cdot))), \tag{23}$$

where

$$G(u) = bu + \frac{c}{2}u^2 + \int_{\mathbb{R}_+} (e^{uz} - 1 - uz)\nu(dz), \quad u \in \mathbb{C}_-. \tag{24}$$

Lemma 4.5. *Suppose that K satisfies Hypothesis 2.1. For every $w \in \mathbb{C}_-$, there exists a unique continuous solution $\phi_w : \mathbb{R}_+ \to \mathbb{C}_-$ to (23). Moreover,*

$$Re \phi_w(t) \leq \phi_{Re\,w}(t), \quad t \geq 0. \tag{25}$$

Proof. Having in mind (Abi Jaber & El Euch, 2019b, Theorem C.1), the existence of a global solution of (23) can be deduced as in (Bondi et al., 2024, Theorem 10), whereas the uniqueness of such

ϕ_w is obtained with a procedure analogous to the proof of Theorem 3.1, see *Step III* with $\Lambda = 0$ in Appendix A. Moreover, again by analogy with (Bondi et al., 2024, Theorem 10 (ii)), the comparison result (25) holds. □

Before stating the theorem that provides the conditional Fourier-Laplace transform of VIX_T^2, we define

$$\Phi_w(t,s) = h_w(s-t)1_{\{s \geq t\}} + G(\phi_w(t-s))1_{\{s < t\}}, \quad t, s \geq 0. \tag{26}$$

Theorem 4.6. *Assume the same hypotheses as in Theorem 3.3. Then, for every* $w \in \mathbb{C}_-$,

$$\mathbb{E}\left[\exp\left\{w \cdot \mathrm{VIX}_T^2\right\} \middle| \mathcal{F}_t\right] = \exp\left\{\widetilde{U}_t(w,T)\right\}, \quad \mathbb{Q} - a.s., \, t \in [0,T], \tag{27}$$

where $\widetilde{U}_t(w,T) = \int_t^\infty \Phi_w(T,s)g_t(s)\mathrm{d}s, \, t \in [0,T]$.

Proof. Fix $w \in \mathbb{C}_-$. First of all, notice that by the definition of h_w and (20)

$$\widetilde{U}_T(w,T) = \int_T^\infty h_w(s-T)g_T(s)\mathrm{d}s = -10^4\frac{2}{\delta}c_1 w \int_T^{T+\delta} \left(1 + b\left(E_{b,K} * 1\right)(T+\delta-s)\right)g_T(s)\mathrm{d}s$$

$$= w \cdot \mathrm{VIX}_T^2, \quad \mathbb{Q} - a.s. \tag{28}$$

We introduce the process

$$\bar{g}_T(s) = \begin{cases} \sigma_s^2, & s \in [0,T], \\ g_T(s), & s > T. \end{cases}$$

Note that by (1) and (3), $\bar{g}_T(\cdot)$ is a joint measurable modification of $g_0 + \int_0^T 1_{\{r \leq \cdot\}}K(\cdot - r)\mathrm{d}Z_r$. For every $t \in [0,T]$, the stochastic Fubini's theorem, (23), (26), and suitable changes of variables, yield

$$\int_0^\infty \Phi_w(T,s)(\bar{g}_t(s) - g_0(s))\mathrm{d}s = \int_0^\infty \Phi_w(T,s)\left(\int_0^t 1_{\{u \leq s\}}K(s-u)\mathrm{d}Z_u\right)\mathrm{d}s$$

$$= \int_0^t \left(\int_0^\infty h_w(s)K(s+T-u)\mathrm{d}s + \int_0^{T-u} K(s)G(\phi_w(T-u-s))\mathrm{d}s\right)\mathrm{d}Z_u$$

$$= \int_0^t \phi_w(T-u)\mathrm{d}Z_u, \quad \mathbb{Q} - a.s. \tag{29}$$

Moreover, by (26), the following equality holds:

$$\int_0^t \Phi_w(T,s)\sigma_s^2 \, \mathrm{d}s = \int_0^t h_w(s-T)1_{\{s \geq T\}}\sigma_s^2 \, \mathrm{d}s + \int_0^t G(\phi_w(T-s))\sigma_s^2 \, \mathrm{d}s$$

$$= \int_0^t G(\phi_w(T-s))\sigma_s^2 \, \mathrm{d}s. \tag{30}$$

Recalling the definition of $\widetilde{U}_t(w,T)$, we combine (29) and (30) to write

$$\widetilde{U}_t(w,T) = \int_t^\infty \Phi_w(T,s)g_0(s)\mathrm{d}s + \int_0^\infty \Phi_w(T,s)(\bar{g}_t(s) - g_0(s))\mathrm{d}s - \int_0^t \Phi_w(T,s)\big(\sigma_s^2 - g_0(s)\big)\mathrm{d}s$$

$$= \int_0^\infty \Phi_w(T,s)g_0(s)\mathrm{d}s + \int_0^t \phi_w(T-u)\mathrm{d}Z_u - \int_0^t G(\phi_w(T-s))\sigma_s^2 \, \mathrm{d}s, \quad \mathbb{Q}-\text{a.s.} \quad (31)$$

In the sequel we denote by $U(w,T) = (U_t(w,T))_{t\in[0,T]}$ the càdlàg process defined by the rightmost side of (31). An application of Itô's formula together with (24) shows that $E(w,T) = (\exp\{U_t(w,T)\})_{t\in[0,T]}$ is a local martingale, namely $E(w,T) = \exp\{\int_0^\infty \Phi_w(T-s) g_0(s)\mathrm{d}s\}\mathcal{E}(\widetilde{N}(w,T))$, where $\mathcal{E}$ denotes the Doléans-Dade exponential and $\widetilde{N}(w,T) = (\widetilde{N}_t(w,T))_{t\in[0,T]}$ is defined by

$$\mathrm{d}\widetilde{N}_t(w,T) = \sqrt{c}\,\phi_w(T-t)\sigma_t\,\mathrm{d}W_{2,t} + \int_{\mathbb{R}_+} \big(e^{\phi_w(T-t)z} - 1\big)\tilde{\mu}(\mathrm{d}t,\mathrm{d}z), \quad \widetilde{N}_0(w,T) = 0.$$

If $E(w,T)$ is a true martingale, then (27) follows from (28) and (31). As in the proof of Theorem 3.3, we search for an expression of $U(w,T)$ which is affine on the past trajectory of σ^2. However, we cannot directly invoke (Bondi et al., 2024, Theorem 7) due to the different structure of the Riccati-Volterra equation in (23) and of the process $U(w,T)$ itself. Fortunately, we can adapt the procedure in the proof of (Bondi et al., 2024, Theorem 7). Specifically, thanks to the local boundedness of $\Phi_w(T,\cdot)$ (see (26)), $\mathbb{Q}-$ a.s.,

$$U_t(w,T) = \int_t^{T+\delta} \Phi_w(T,s)g_0(s)\mathrm{d}s + \phi_w(T-t)Z_t + \big(\pi_{T+\delta-t} * \big(\sigma^2 - g_0\big)\big)(t), \quad \text{for a.e. } t \in (0,T).$$

Here the functions

$$\pi_{T+\delta-t}(u) = \int_0^{T+\delta-t} \Phi_w(T,T+\delta-s)((\Delta_{T+\delta-t-s}K)' * L)(u)\mathrm{d}s, \quad t \in (0,T),$$

are well defined for almost every $u \in \mathbb{R}_+$ and belong to $L^1_{\mathrm{loc}}(\mathbb{R}_+)$. At this point, for every $t \in (0,T)$ we introduce the locally absolutely continuous function

$$\widetilde{\Pi}_{T+\delta-t}(u) = \int_0^u \pi_{T+\delta-t}(s)\mathrm{d}s + \phi_w(T-t)L([0,u])$$

$$= \int_0^{T+\delta-t} \Phi_w(T,T+\delta-s)((\Delta_{T+\delta-t-s}K) * L)(u)\mathrm{d}s, \quad u \geq 0,$$

where the second equality is due to (23) and a suitable change of variables. Therefore, also recalling (1), the previous formula for $U(w,T)$ can be rewritten as, $\mathbb{Q}-$ a.s., for a.e. $t \in (0,T)$,

$$U_t(w,T) = \int_t^{T+\delta} \Phi_w(T,s)g_0(s)\mathrm{d}s + \big(\mathrm{d}\widetilde{\Pi}_{T+\delta-t} * \big(\sigma^2 - g_0\big)\big)(t) + \phi_w(T-t)L(\{0\})\big(\sigma^2 - g_0\big)(t),$$

which is an affine expression in terms of the past trajectories of σ^2. Now by Lemma 3.2 the real-valued process $E(\mathrm{Re}\,w,T) = (\exp\{U_t(\mathrm{Re}\,w,T)\})_{t\in[0,T]}$ is a true martingale. Thus, thanks to (25), we

can parallel the comparison argument in the proof of (Bondi et al., 2024, Theorem 11) to deduce that

$$|\exp\{U_t(w,T)\}| = \exp\{\operatorname{Re} U_t(w,T)\} \leq C \exp\{U_t(\operatorname{Re} w,T)\}, \quad t \in [0,T], \ \mathbb{Q}-\text{a.s.},$$

for some constant $C > 0$. An application of (Jarrow, 2018, Lemma 1.4) completes the proof. $\square$

4.1 | VIX put options and futures prices

Theorem 4.6 provides a semi-explicit formula for the Fourier-Laplace transform λ_T of VIX_T^2 in $\mathbb{C}_-$, namely

$$\lambda_T(w) = \mathbb{E}\big[\exp\{w \cdot \mathrm{VIX}_T^2\}\big] = \exp\left\{\int_0^\infty \Phi_w(T,s)g_0(s)\mathrm{d}s\right\}$$

$$= \exp\left\{\int_0^\delta h_w(s)g_0(s+T)\mathrm{d}s + (g_0 * G(\phi_w(\cdot)))(T)\right\}, \quad w \in \mathbb{C}_-. \quad (32)$$

This allows us to price put options written on VIX with the Fourier-inversion technique for the bilateral Laplace transform shown in Carr and Lee (2009). More specifically, for a log strike $k \in \mathbb{R}$, the payoff function of such options defined on the whole real line is $w(x) = (e^k - \sqrt{x^+})^+$, $x \in \mathbb{R}$, where x^+ represents VIX^2. Then, denoting by $P(k,T)$ the price of a put option with maturity T (and log strike k) we have (cf. (Carr and Lee, 2009, Equations (7.6)-(7.8)))

$$P(k,T) = \mathbb{E}\left[\left(e^k - \mathrm{VIX}_T\right)^+\right] = -\frac{1}{4\sqrt{\pi}i}\int_{z_r-i\infty}^{z_r+i\infty} \frac{\operatorname{erf}\left(e^k\sqrt{z}\right)}{z^{3/2}}\lambda_T(z)\mathrm{d}z$$

$$= -\frac{1}{4\sqrt{\pi}}\int_{\mathbb{R}} \operatorname{Re}\left[\frac{\operatorname{erf}\left(e^k\sqrt{z_r+iu}\right)}{(z_r+iu)^{3/2}}\lambda_T(z_r+iu)\right]\mathrm{d}u$$

$$= -\frac{1}{2\sqrt{\pi}}\int_{\mathbb{R}_+} \operatorname{Re}\left[\frac{\operatorname{erf}\left(e^k\sqrt{z_r+iu}\right)}{(z_r+iu)^{3/2}}\lambda_T(z_r+iu)\right]\mathrm{d}u, \quad z_r < 0. \quad (33)$$

Here erf represents the error function $\operatorname{erf} z = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}\mathrm{d}t$, $z \in \mathbb{C}$, and for $z \in \mathbb{C}$ and $a \geq 0$, we consider the power $z^a = \Lambda^a e^{ia\theta}$, where $z = \Lambda e^{i\theta}$ with $\Lambda \geq 0$, $\theta \in (-\pi,\pi]$. In particular, we write $\sqrt{z} = z^{1/2}$. The last equality in (33) is due to the fact that the integrand is even. Indeed, this follows from the well-known symmetry relation $\operatorname{erf}\bar{z} = \overline{\operatorname{erf} z}$, $z \in \mathbb{C}$, as well as the identities (for $u \neq 0$)

$$\operatorname{Re}\left(\sqrt{z_r+iu}\right) = \sqrt{\frac{z_r+\sqrt{z_r^2+u^2}}{2}}, \qquad \operatorname{Im}\left(\sqrt{z_r+iu}\right) = \operatorname{sgn}(u)\sqrt{\frac{-z_r+\sqrt{z_r^2+u^2}}{2}}.$$

Moreover, we can use λ_T to determine $\mathbb{E}[\text{VIX}_T]$, that is, the futures price of VIX at time T. In order to do this, notice that for every $x \geq 0$ the function $(\sqrt{\pi s})^{-1}(e^{-xs} - 1) + \sqrt{x}\,\text{erf}(\sqrt{sx})$, $s > 0$, is an antiderivative of $(2\sqrt{\pi})^{-1}(1 - e^{-xs})s^{-3/2}$, $s > 0$. From this relation we deduce the following integral representation for the square-root function

$$\sqrt{x^+} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-sx^+}}{s^{\frac{3}{2}}}\, ds, \quad x \in \mathbb{R}.$$

An application of Tonelli's theorem yields

$$\mathbb{E}[\text{VIX}_T] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - \lambda_T(-s)}{s^{\frac{3}{2}}}\, ds. \tag{34}$$

As in the case of the log returns process X discussed in the final paragraph of Section 3, Equation (32) with $w = it$, $t \in \mathbb{R}$, provides a semi-explicit expression for the characteristic function of VIX_T^2. This enables us to approximate both the futures prices $\mathbb{E}[\text{VIX}_T]$ and VIX put option prices with maturity T via the COS method, see Fang and Oosterlee (2009). This scheme entails the computation of the integral

$$\int_0^d \sqrt{x} \cos x\, dx = \sqrt{d} \sin d - \sqrt{\frac{\pi}{2}} S\left(\sqrt{\frac{2}{\pi}d}\right), \quad d > 0,$$

where $S(x) = \int_0^x \sin(\frac{\pi}{2}u^2)\, du$, $x \geq 0$, is the Fresnel sine integral. In Section 6 we apply the COS method to calibrate our model to real market data because of its computational efficiency, which relies on a reduced number of evaluations of λ_T compared to (33)–(34). Nonetheless, we need the exact formulas (33)–(34) as a benchmark to determine the truncation bounds a and b of the COS method. Since $\text{VIX}_T^2 \geq 0$, we set $a = 0$, thereby reducing the problem to the upper bound $b > 0$. We refer to Section 6 for more details.

5 | NUMERICAL APPROXIMATION OF THE MODEL

According to the formulae in (15)–(16) and (33), in order to price options on S and VIX with maturity T, one needs to compute $\Psi^{X_T}(w_1)$ and $\lambda_T(w_2)$, where w_1, w_2 belong to appropriate regions of $\mathbb{C}$. In addition, the values $\lambda_T(-s)$, $s \geq 0$, are also necessary to determine the futures price of VIX at time T. Consequently, looking at the expressions of these Fourier-Laplace transforms in (14) and (32), the solutions of the Riccati-Volterra equations (7) and (23), that is, ψ_{w_1}, ϕ_{w_2} and ϕ_{-s}, have to be approximated on the interval $[0, T]$. The previous remarks also apply to the COS method, which requires the computation of the characteristic functions of X_T and VIX_T^2. Among the available numerical methods to solve (7)–(23), we choose the multi-factor scheme suggested in Abi Jaber and El Euch (2019b). Another possibility would be to use the Adams scheme Diethelm et al. (2002), Diethelm et al. (2004), hybrid schemes as in Callegaro et al. (2021), or an adaptation of the multi-factor hybrid approach in Rømer (2022b) (Figure 1).
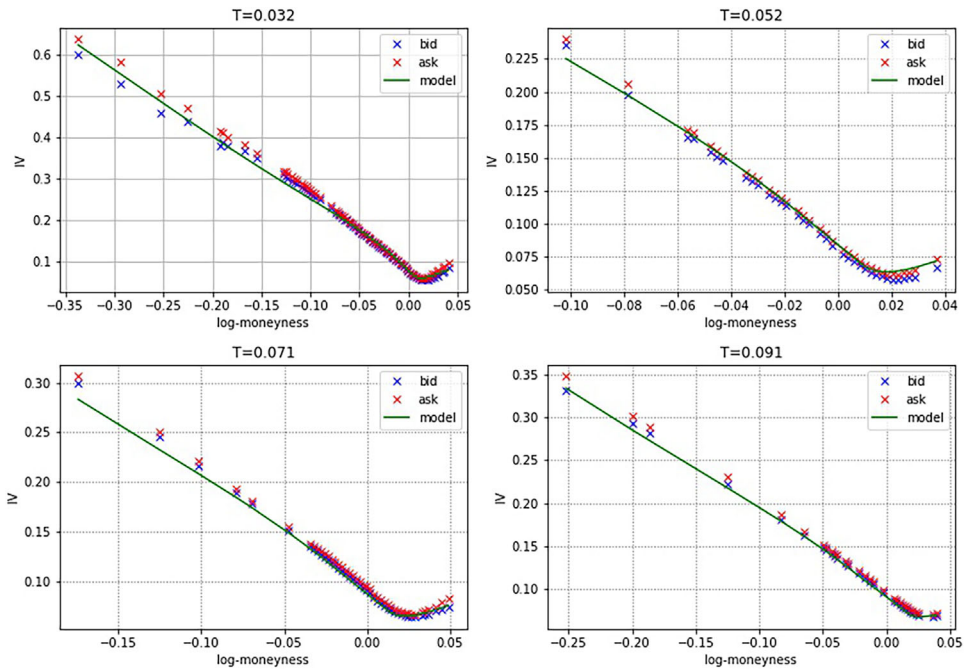
FIGURE 1 Calibrated implied volatility of SPX options on May 19, 2017, using the parameters in Table 1. The blue and red crosses are, respectively, the bid and ask of market implied volatilities. The implied volatility smiles from the model are in green. The abscissa is in log-moneyness and T is time to expiry in years. [Color figure can be viewed at wileyonlinelibrary.com]

The multi-factor scheme consists in approximating the kernel K with a weighted sum of exponentials, namely with functions K_n, $n \in \mathbb{N}$, of the form

$$K_n(t) = \sum_{j=1}^{n} m_{j,n} e^{-x_{j,n}t}, \quad t \geq 0, \tag{35}$$

where $m_{j,n}, x_{j,n} > 0$, $j = 1, \dots, n$. In what follows, we write $\mathbf{m} = \{m_{j,n} | j = 1, \dots, n, n \in \mathbb{N}\}$ and $\mathbf{x} = \{x_{j,n} | j = 1, \dots, n, n \in \mathbb{N}\}$. Notice that K_n, $n \in \mathbb{N}$, is completely monotone on $(0, \infty)$, meaning that it is nonnegative and infinitely differentiable on this interval, with nonpositive [resp., nonnegative] odd [resp., even] $k-$ derivative, $k \in \mathbb{N}$. More details about this approximation and the idea behind it can be found in Remark 5.2 below and in the references therein.

Given $n \in \mathbb{N}$ and $w \in \mathbb{C}$ such that Re $w \in [0, 1]$, we now introduce the Riccati-Volterra equation

$$\psi_{w,n}(t) = \int_0^t K_n(t-s)\mathcal{R}\big(w, \psi_{w,n}(s)\big)\mathrm{d}s = \big(K_n * \mathcal{R}\big(w, \psi_{w,n}(\cdot)\big)\big)(t), \quad t \geq 0. \tag{36}$$

Note that the existence and uniqueness of $\psi_{w,n}$ is guaranteed by Theorem 3.1 (i), because K_n satisfies Hypothesis 2.1. The advantage in considering (36) instead of (7) is that its solution $\psi_{w,n}$ can be obtained by numerically solving a system of integral equations with standard methods. More

precisely, $\psi_{w,n}(t) = \sum_{j=1}^{n} m_{j,n} \psi_{w,n}^{(j)}(t)$ for every $t \geq 0$, where

$$\psi_{w,n}^{(j)}(t) = e^{-x_{j,n}t} \int_0^t e^{x_{j,n}s} \mathcal{R}\left(w, \sum_{k=1}^{n} m_{k,n} \psi_{w,n}^{(k)}(s) \right) ds, \quad j = 1, \dots, n.$$

Analogously, for every $n \in \mathbb{N}$ and $w \in \mathbb{C}_-$, we consider the Riccati-Volterra equation

$$\phi_{w,n}(t) = \int_0^\infty h_w(s) K_n(s+t) ds + \left(K_n * \left(G(\phi_{w,n}(\cdot)) \right) \right)(t), \quad t \geq 0. \tag{37}$$

We have that $\phi_{w,n}(t) = \sum_{j=1}^{n} m_{j,n} \phi_{w,n}^{(j)}(t)$, $t \geq 0$, with

$$\phi_{w,n}^{(j)}(t) = e^{-x_{j,n}t} \left(\int_0^\infty h_w(s) e^{-x_{j,n}s} ds + \int_0^t e^{x_{j,n}s} G\left(\sum_{k=1}^{n} m_{k,n} \phi_{w,n}^{(k)}(s) \right) ds \right), \quad j = 1, \dots, n.$$

The following theorem offers an estimate on the uniform distance on $[0, T]$ between ψ_w and $\psi_{w,n}$, as well as between ϕ_w and $\phi_{w,n}$. In the former case, it generalizes (Abi Jaber & El Euch, 2019b, Theorem 4.1) to our framework with jumps. Its proof, which we postpone to Appendix C, relies on results related to Riccati-Volterra equations which are proved in Appendix A.

Theorem 5.1. *Assume that K satisfies Hypothesis 2.1. Let $T > 0$ and denote by $E_{\lambda,n}$ the canonical resolvent of K_n with parameter $\lambda \in \mathbb{R}$, $n \in \mathbb{N}$.*

(i) *Suppose that $\int_0^T |E_{b+\rho+\sqrt{c},n}(s)| ds \leq \widetilde{C}$ for every $n \in \mathbb{N}$, where $\widetilde{C} = \widetilde{C}(\rho, b, \mathbf{m}, \mathbf{x}, T) > 0$. Then there exists a constant $C = C(\rho, b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T) > 0$ such that, for every $w \in \mathbb{C}$ with $Re\,w \in [0, 1]$ and $n \in \mathbb{N}$,*

$$\sup_{t \in [0,T]} \left| \psi_w(t) - \psi_{w,n}(t) \right| \leq C \left(1 + |Im\,w|^6 \right) \int_0^T E_{C\left(1 + |Im\,w|^2\right), K}(s) ds$$

$$\times \int_0^T |K_n(s) - K(s)| ds. \tag{38}$$

In addition, if $b < 0$ and $\rho < 0$ then the constant C does not depend on $\mathbf{m}$ or $\mathbf{x}$, and the dependence on T is via $\|K\|_{L^1([0,T])}$.

(ii) *Suppose that $\int_0^{T \vee \delta} |E_{b+,n}(s)| ds \leq \widetilde{C}$ for every $n \in \mathbb{N}$, where $\widetilde{C} = \widetilde{C}(b, \mathbf{m}, \mathbf{x}, T, \delta) > 0$. Then there exists a constant $C = C(b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T, \delta) > 0$ such that, for every $w \in \mathbb{C}_-$ and $n \in \mathbb{N}$,*

$$\sup_{t \in [0,T]} \left| \phi_w(t) - \phi_{w,n}(t) \right| \leq C \left(1 + |w|^6 \right) \int_0^T E_{C\left(1 + |w|^2\right), K}(s) ds \int_0^{T \vee \delta} |K_n(s) - K(s)| ds. \tag{39}$$

Remark 5.2. Recent literature has focused on efficient multi-factor approximations of the kernel K with K_n, $n \in \mathbb{N}$, as in (35), see Bayer and Breneis (2023), Bayer and Breneis (2023), Rømer (2022b). Since in Section 6 we use the fractional kernel $K(t) = t^{\alpha-1}/\Gamma(\alpha)$ with $\alpha \in (1/2, 1]$ to calibrate our model (1)–(5) to market data, we choose the geometric Gaussian scheme in (Bayer and Breneis, 2023, Theorem 3.9). This scheme guarantees, given $T > 0$, a fast convergence $K_n \to K$ in $L^1([0, T])$

for α close to the lower bound $1/2$. A small parameter α is crucial for the joint calibration of *S&P 500* and *VIX* smiles, as shown by the numerical experiments in Sections 6–7. The fast convergence of the method enables us to consider few factors, that is, n small, for the approximation K_n of K, which improves the efficiency of our implementation. The geometric Gaussian method in Bayer and Breneis (2023) relies on a representation of the kernel as $K(t) = \int_{\mathbb{R}_+} e^{-xt} \mu(\mathrm{d}x)$, $t > 0$, where μ is a nonnegative measure on $\mathbb{R}_+$. This representation holds for general completely monotone kernels thanks to the Bernstein-Widder theorem, see, for example, (Gripenberg et al., 1990, Theorem 2.5, Chapter 5). Approximating μ with a weighted sum of Dirac measures gives K_n in (35). In order to determine **m** and **x** in (35) (Bayer and Breneis, 2023, Theorem 3.9) considers a geometric partition of $[0, \xi]$, for some $\xi > 0$, and applies, on the resulting intervals, quadrature rules of order $m \geq 1$ with respect to suitable weight functions. Previous works have proved the convergence $K_n \to K$ in $L^2_{loc}(\mathbb{R}_+)$ via quadrature rules of order 1, see (Abi Jaber, 2019, Lemma A.3) and (Abi Jaber and El Euch, 2019b, Proposition 3.3). For this type of approximated kernels K_n, $n-$ uniform $L^2_{loc}-$ estimates are also available, see (Chevalier et al., 2022, Theorem B.1). Note that, choosing $T > \delta$, the $L^1([0, T])-$ convergence of K_n to K ensures the validity of the hypotheses required in both points of Theorem 5.1 (see also (Gripenberg et al., 1990, Theorem 3.1, Chapter 2)), and therefore the convergence of the multi-factor scheme.

6 | CALIBRATION

We have shown that Fourier-based methods can be applied to the rough Hawkes Heston model in order to price options on the underlying and the corresponding volatility index. Based on these techniques, in this section we calibrate a parsimonious specification of the rough Hawkes Heston model to S&P 500 and VIX options data on May 19, 2017. This is the same data set as in Gatheral et al. (2020). In Table 3 [resp., Table 4] we report the strikes and maturities of the SPX options [resp., VIX options] considered for the calibration. Our objective is to minimize the relative RMSE (root-mean-square error) between market and theoretical implied volatilities of both SPX and VIX options. More precisely, denoting by Θ the parameters of the model, the goal of the calibration procedure is to determine

$$\arg \min_{\Theta} \sqrt{\sum_{i,j} \left(\frac{\sigma^{\mathrm{mkt}}_{\mathrm{SPX}}(T_i, K_j) - \sigma^{\Theta}_{\mathrm{SPX}}(T_i, K_j)}{\sigma^{\mathrm{mkt}}_{\mathrm{SPX}}(T_i, K_j)} \right)^2 + \sum_{i,j} \left(\frac{\sigma^{\mathrm{mkt}}_{\mathrm{VIX}}(T_i, K_j) - \sigma^{\Theta}_{\mathrm{VIX}}(T_i, K_j)}{\sigma^{\mathrm{mkt}}_{\mathrm{VIX}}(T_i, K_j)} \right)^2}. \quad (40)$$

Here $\sigma^{\mathrm{mkt}}_{\mathrm{SPX}}(T_i, K_j)$ [resp., $\sigma^{\mathrm{mkt}}_{\mathrm{VIX}}(T_i, K_j)$] is the implied volatility of an option on SPX [resp., VIX] with maturity T_i and strike K_j, obtained from market data. The symbols $\sigma^{\Theta}_{\mathrm{SPX}}(T_i, K_j)$ and $\sigma^{\Theta}_{\mathrm{VIX}}(T_i, K_j)$ denote the same quantities computed in the rough Hawkes Heston model with parameters Θ. The indexes i, j are an enumeration of the data set in Tables 3 and 4. The implied volatilities of VIX options are computed with respect to market futures prices. As we will explain below, after calibration, model futures approximate well market futures (see Figure 4).

As it is customary in rough volatility models, for our parametrization we choose a power kernel of the form $K(t) = t^{\alpha-1}/\Gamma(\alpha)$, $\alpha \in (1/2, 1]$. As explained in Remark 5.2, we approximate it with the sum of exponentials in (35) using the geometric Gaussian scheme in (Bayer and Breneis, 2023, Theorem 3.9) with $T = 0.091$: the longest maturity that we consider for the calibration. In fact, we scale the resulting approximated kernel multiplying it by a positive constant. We then determine such a constant by a fast optimization procedure in order to minimize the $L^1([0, T])-$ distance

TABLE 1 Calibrated parameters for the rough Hawkes Heston model.

α	ρ	b	c	Λ	β	σ_0^2
0.527	−0.731	−1.812	0.115	0.276	0.049	0.0079

between the multi-factor kernel and K. We use $n = 20$ factors in (35) to obtain an empirical convergence of the approximating method, noticing that fewer (~ 10) are sufficient when it comes to SPX options only. In the sequel, we denote by $\bar{K}_{20}$ the corresponding approximating kernel.

Regarding the choice of the fractional kernel K, we remark that the recent study Abi Jaber et al. (2022a), comparing different types of kernels (fractional, log-modulated, shifted and exponential) in a Gaussian Volterra setup, concludes that a conventional one factor, Markovian exponential kernel is able to jointly fit the SPX/VIX smiles outperforming its rough and non-rough path-dependent counterparts. Since the theoretical study of the previous sections covers a wide range of kernels (e.g., completely monotone), an interesting question is whether a similar behavior is exhibited in our framework. This important question could be addressed in futures studies of variants of the rough Hawkes Heston model.

By analogy with the rough Heston model introduced and studied in El Euch and Rosenbaum (2018), El Euch and Rosenbaum (2019), we consider an initial input curve g_0 of the form

$$g_0(t) = \sigma_0^2 + \beta \int_0^t K(s)\mathrm{d}s = \sigma_0^2 + \frac{\beta}{\Gamma(\alpha+1)}t^\alpha, \quad t \ge 0, \tag{41}$$

where σ_0^2, $\beta \ge 0$. Note that the structure of g_0 in (41) is quite restrictive, but it allows to keep the model parsimonious with a small number of parameters. Indeed, the initial variance curve in (41) is specified only by σ_0^2 and β. Our choice is also justified by the fact that we focus on short time-to-maturity options, so the lack of flexibility for g_0 is not a drawback in our application. More general forms of g_0 or expressions extracted from the replication formula for the log-contract as in (Abi Jaber et al., 2022b, Equation (5.1)) can be used to calibrate VIX smiles for longer times-to-maturity.

In our numerical illustration, we consider the kernel $\bar{K}_{20}$ and g_0 as in (41). These choices guarantee the well-posedness, in the weak sense, of (1), because $\bar{K}_{20}$ is completely monotone and g_0 satisfies Hypothesis 2.2 (see Remark 2.3).

For the law of the jumps, to keep the number of parameters low we choose an exponential distribution with rate 1, $\nu(\mathrm{d}z) = \exp(-z)\,\mathrm{d}z$. Our parsimonious specification of the model has therefore – other than the two parameters (β, σ_0^2) related to g_0 – five evolution-related parameters $(\alpha, \rho, b, c, \Lambda)$. Like in Gatheral et al. (2020), we concentrate on short maturities for which, as pointed out in Guyon (2020b), "VIX derivatives are most liquid and the joint calibration is most difficult."

As already mentioned in the final paragraphs of Section 3 and Subsection 4.1, we employ the COS method to calibrate our model to market data due to its efficiency. In order to determine the truncation bounds of the scheme (a_1 and b_1 for SPX options, and only the upper bound b_2 for VIX options), we rely on the exact formulas (15)–(16) and (33)–(34), which we consider as an accurate benchmark. At each maturity of the data set, these bounds possibly depend on all the parameters of the model, but we empirically observe that the resulting implied volatilities are mostly affected by Λ for SPX options. Consequently, we fix offline a grid of bounds depending on the maturities T_i and Λ for SPX options, only on T_i for VIX options, and we use these bounds in the calibration. The resulting calibrated parameters are reported in Table 1.
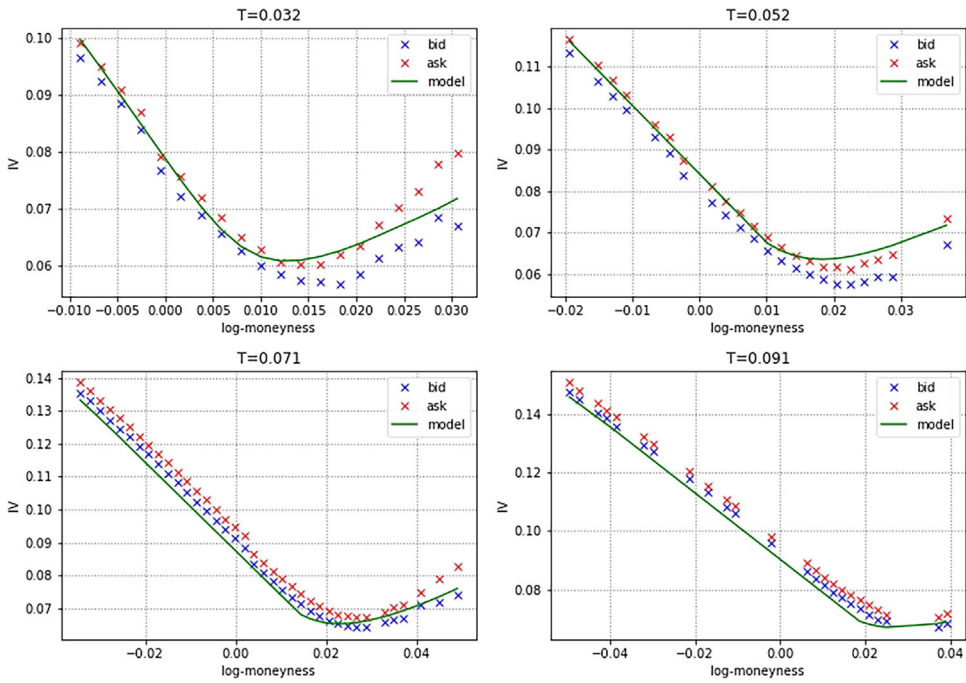
FIGURE 2 Zoom at-the-money of the calibrated implied volatility of SPX options on May 19, 2017, using the parameters in Table 1. [Color figure can be viewed at wileyonlinelibrary.com]

Starting from the values in Table 1, we then minimize the following functional of $\boldsymbol{\Theta}$, which takes into account the relative number of SPX/VIX options in the sample considered for the calibration:

$$c_1 \sqrt{\sum_{i,j} \left(\frac{\sigma_{\text{SPX}}^{\text{mkt}}(T_i, K_j) - \sigma_{\text{SPX}}^{\Theta}(T_i, K_j)}{\sigma_{\text{SPX}}^{\text{mkt}}(T_i, K_j)} \right)^2} + c_2 \sqrt{\sum_{i,j} \left(\frac{\sigma_{\text{VIX}}^{\text{mkt}}(T_i, K_j) - \sigma_{\text{VIX}}^{\Theta}(T_i, K_j)}{\sigma_{\text{VIX}}^{\text{mkt}}(T_i, K_j)} \right)^2},$$

where $c_1 = 71.4\%, \quad c_2 = 28.6\%$.

However, no significant changes in the parameters have to be reported.

We observe that the value of α in Table 1 is very close to its lower bound limit 0.5. This is coherent with previous estimates in the rough volatility literature, see for instance Alos and Shiraya (2019), Bayer et al. (2016), Bennedsen et al. (2021), El Euch et al. (2018), Fukasawa (2017), Gatheral et al. (2018), Gatheral et al. (2020). The estimation of the correlation parameter ρ is also in line with empirical estimates, for example, Cont (2001), and what is commonly known as the *leverage effect* (Curato & Sanfelici, 2015; El Euch et al., 2018; Mancino & Toscano, 2022). We notice that for the joint calibration we can keep the vol-of-vol parameter c small because an important part of the volatility fluctuation is captured by the *self-exciting* jumps controlled by the parameters α and Λ. This responds to the issue, raised in Guyon (2020b), that the "very large negative skew of short-term SPX options, which in continuous models implies a very large volatility of volatility, seems inconsistent with the comparatively low levels of VIX implied volatilities."

The calibrated implied volatility smiles for S&P 500 and VIX options are plotted in Figures 1 and 3, respectively. We zoom the calibration of the S&P 500 options at-the-money in Figure 2. Figure 4 focuses on the VIX term structure, which we do not include in the functional (40) used for the
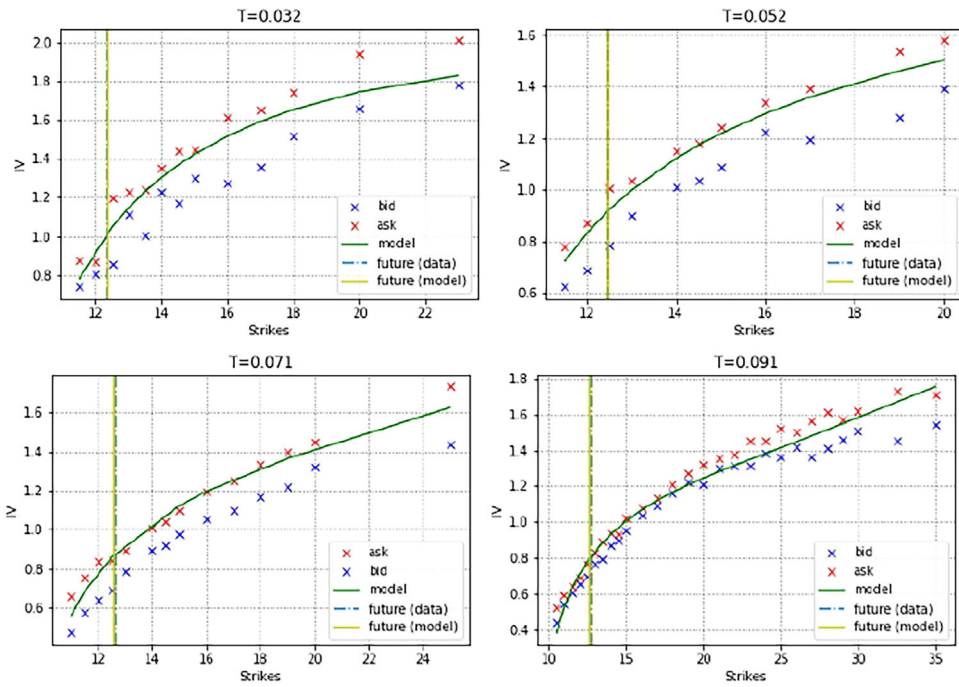
FIGURE 3 Calibrated implied volatility of VIX options on May 19, 2017, using the parameters in Table 1. The blue and red crosses are respectively the bid and ask of market implied volatilities. The implied volatility smiles from the model are in green. The abscissa is in strikes and T is time to expiry in years. [Color figure can be viewed at wileyonlinelibrary.com]
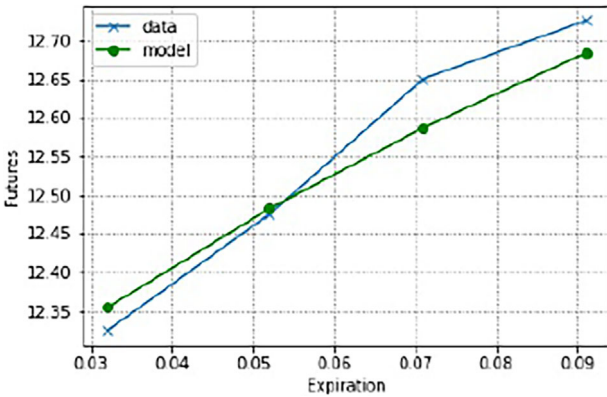

FIGURE 4 VIX term structure. [Color figure can be viewed at wileyonlinelibrary.com]

calibration. We remark that the term structure of the model is not flexible enough to perfectly reproduce the shape of the market VIX futures, mainly due to a change of convexity. However, the maximal relative distance between market and model data is $\sim 0.5\%$, which is reasonable considering that we have implicitly assumed a "flat" initial volatility curve, see (41). Furthermore, Figure 5 shows the implied volatility smile for S&P 500 options with expiration $T = 0.179$ (July 21, 2017): 1 month after the last, calibrated maturity $T = 0.091$. Despite the fact that $T = 0.179$ is not included in (40), the model is able to replicate market implied volatility smiles with an accuracy

FIGURE 5 Implied volatility of SPX options on May 19, 2017, with time to expiry $T = 0.179$ years. [Color figure can be viewed at wileyonlinelibrary.com]
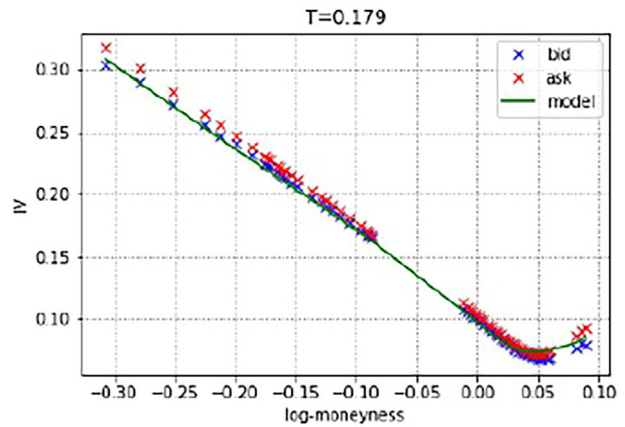
TABLE 2 Calibrated parameters for the rough Heston model.

| α	ρ	b	c	β	σ_0^2
0.501	−0.774	−2.185	0.153	0.055	0.0092

comparable to the other maturities considered for S&P 500 options (cf. Figure 1). Overall, these graphs show that the model fits remarkably well both S&P 500 and VIX implied volatilities.

The shape of the smile at-the-money for S&P 500 options is well-captured and the distance to the bid-ask corridor – across the maturities – is at most of approximately one bid-ask spread. For the two shortest maturities, most of the model implied volatilities at-the-money are actually inside the bid-ask corridor. The fit is not perfect for very negative log-moneyness. This is also seen – possibly to a less extent – in the quadratic rough Heston model Gatheral et al. (2020). We conjecture that, at the cost of increasing the complexity of the model, even better results could be obtained if we replace the exponential law for the jumps by a law with Pareto tails as suggested in Cont (2001), Jiao et al. (2021) and the references therein. Regarding the VIX implied volatilities, we observe that – even for options deep out-of-the-money – the model implied volatilities stay almost systematically within the bid-ask corridor, whether it is calculated using call or for put options.

6.1 | Calibration with the rough Heston model

The model (1)–(5) proposed in this paper is an extension of the rough Heston model (see El Euch & Rosenbaum, 2018, 2019), obtained by adding a jump component with intensity proportional to the spot variance σ^2. In particular, the rough Heston model can be recovered from (1) to (5) by setting $\nu(dz) = 0$. In Introduction 1 and Section 2, we justify the presence of self-exciting jumps common to the underlying and the volatility based on empirical evidence of jump-clustering phenomena and endogeneity of financial markets. The purpose of this subsection (see also Subsection 7.1) is to investigate the impact of jumps with numerical experiments. To do this, we perform a calibration exercise with the rough Heston model using the same dataset as before (cf. Section 6) and considering the same functional (40) to minimize. The calibrated parameters are reported in Table 2.

The resulting implied volatility smiles of SPX/VIX options for two selected maturities ($T = 0.032$, $T = 0.091$) are shown in Figure 6. Here we see that the rough Heston model is able to
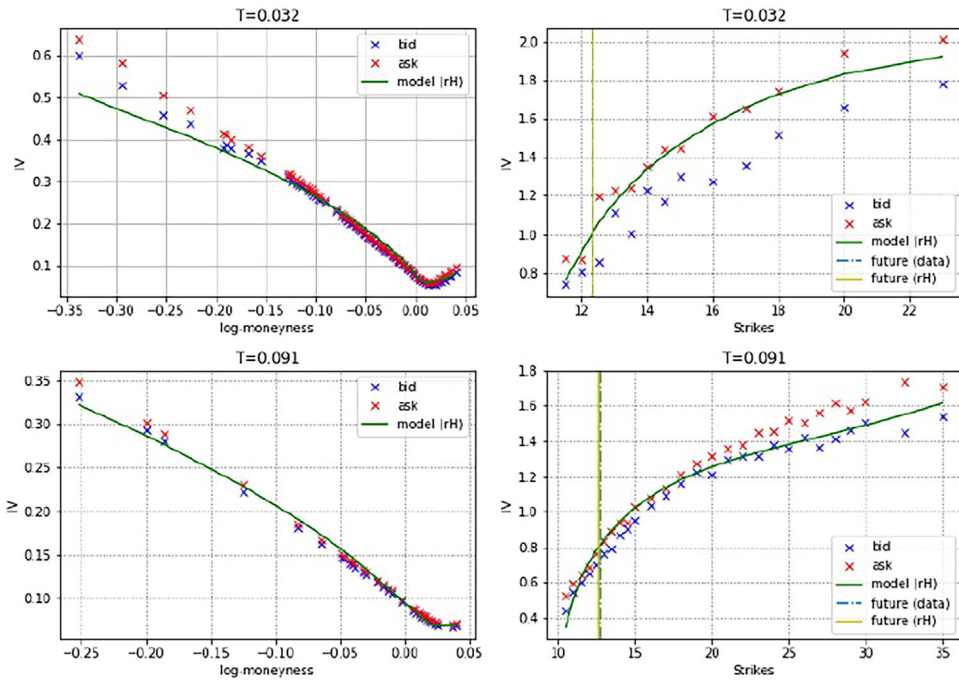
FIGURE 6 Calibrated implied volatility of SPX (left) and VIX (right) options on May 19, 2017, using the rough Heston model. The parameters are reported in Table 2. The abscissa is in log-moneyness for SPX options and in strikes for VIX options. The time to expiry in years is $T = 0.032$ in the first line and $T = 0.091$ in the second line. [Color figure can be viewed at wileyonlinelibrary.com]

reproduce quite well the level and shape of VIX options smiles. However, the rough Heston model struggles to capture the out-of-the-money skew of S&P 500 options, especially for $T = 0.032$. In an attempt to understand whether this effect is due to the VIX options in (40), we run another calibration exercise with the aim of minimizing the following functional in Θ:

$$\sqrt{\sum_{i,j} \left(\frac{\sigma_{\text{SPX}}^{\text{mkt}}(T_i, K_j) - \sigma_{\text{SPX}}^{\Theta}(T_i, K_j)}{\sigma_{\text{SPX}}^{\text{mkt}}(T_i, K_j)} \right)^2}. \tag{42}$$

The strikes and maturities used in this example are listed in Tables 3–5. We perform the same optimization for the rough Hawkes Heston model and display the outcomes in Figure 7. These results confirm that the rough Heston model does not produce a correct shape for the left-tails of SPX implied volatility smiles, and that in this aspect it is consistently outperformed by the extension (1)–(5) proposed in this paper.

On the basis of these experiments, we conclude that the introduction of a jump component in the rough Heston model is significant and does not cause redundancy. In particular, the jumps allow to better reproduce the skew of out-of-the-money S&P 500 options, especially for short maturities.

TABLE 3   SPX options data on May 19, 2017, considered for the calibration. The listed maturities correspond to $T = 0.032,\ 0.052,\ 0.071,\ 0.091$, respectively.

Maturity	log-moneyness
May 31, 2017	−0.3369, −0.2938, −0.2524, −0.2257, −0.1921, −0.1895, −0.1845, −0.1669, −0.1546, −0.128, −0.1256, −0.1232, −0.1185, −0.1161, −0.1138, −0.1091, −0.1067, −0.1044, −0.1021, −0.0998, −0.0974, −0.0951, −0.0905, −0.0791, −0.0746, −0.0723, −0.07, −0.0678, −0.0655, −0.0611, −0.0588, −0.0544, −0.0522, −0.05, −0.0478, −0.0456, −0.0434, −0.0412, −0.039, −0.0368, −0.0346, −0.0325, −0.0303, −0.0281, −0.026, −0.0238, −0.0217, −0.0195, −0.0174, −0.0153, −0.0131, −0.011, −0.0089, −0.0068, −0.0047, −0.0026, −0.0005, 0.0016, 0.0037, 0.0058, 0.0079, 0.01, 0.0121, 0.0141, 0.0162, 0.0183, 0.0203, 0.0224, 0.0244, 0.0265, 0.0285, 0.0306, 0.0346, 0.0367, 0.0407.
June 7, 2017	−0.1019, −0.0789, −0.0565, −0.0543, −0.0476, −0.0454, −0.0432, −0.0345, −0.0323, −0.0302, −0.0258, −0.0237, −0.0215, −0.0194, 0.0101, 0.0122, 0.0143, 0.0164, 0.0184, 0.0205, 0.0225, 0.0246, 0.0266, 0.0287, 0.0368.
June 14, 2017	−0.1741, −0.1253, −0.1018, −0.0788, −0.0698, −0.0475, −0.0343, −0.0322, −0.03, −0.0278, −0.0257, 0.0144, 0.0165, 0.0186, 0.0206, 0.0227, 0.0247, 0.0268, 0.0288, 0.0329, 0.0349, 0.037, 0.041, 0.045, 0.049.
June 21, 2017	−0.2519, −0.1993, −0.1866, −0.1252, −0.0832, −0.0651, −0.0495, −0.0473, −0.0429, −0.0408, −0.0386, 0.0187, 0.0208, 0.0228, 0.0249, 0.0371, 0.0391.

TABLE 4   VIX options data on May 19, 2017, considered for the calibration. The listed maturities correspond to $T = 0.032,\ 0.052,\ 0.071,\ 0.091$, respectively.

Maturity	Strikes
May 31, 2017	11.5, 12.5, 13, 13.5, 14, 14.5, 15, 16, 17, 18, 20, 23.
June 7, 2017	11.5, 12, 12.5, 13, 14, 14.5, 15, 16, 17, 19, 20.
June 14, 2017	11, 11.5, 12.5, 15, 16, 19, 25.
June 21, 2017	10.5, 11, 11.5, 12, 12.5, 13, 13.5, 14, 14.5, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32.5, 35.

TABLE 5   SPX options data on May 19, 2017, considered for the calibration of (42) in addition to Table 3. The maturity corresponds to $T = 0.179$.

Maturity	log-moneyness
July 21, 2017	−0.3071, −0.2789, −0.2515, −0.2248, −0.2118, −0.1989, −0.1861, −0.176, −0.1735, −0.171, −0.1661, −0.1636, −0.1611, −0.1587, −0.1537, −0.1488, −0.1367, −0.1295, −0.1247, −0.12, −0.1129, −0.1059, −0.0966, −0.092, −0.0897, −0.0874, −0.0851, −0.0828, −0.0123, −0.008, −0.0059, −0.0038, −0.0017, 0.0004, 0.0025, 0.0046, 0.0088, 0.0109, 0.015, 0.0171, 0.0191, 0.0212, 0.0233, 0.0253, 0.0274, 0.0294, 0.0314, 0.0335, 0.0355, 0.0375, 0.0396, 0.0416, 0.0436, 0.0456, 0.0476, 0.0496, 0.0516, 0.0536, 0.0556, 0.0576, 0.0596, 0.0811, 0.085, 0.0888.

7 | SENSITIVITIES OF THE IMPLIED VOLATILITIES

In this section we study the sensitivity of the implied volatilities of S&P 500 and VIX options to the parameters of the rough Hawkes Heston model. Starting from the calibrated parameters presented in Table 1, we analyze the impact of a change in the evolution-related parameters (α, ρ, b, c, Λ) and
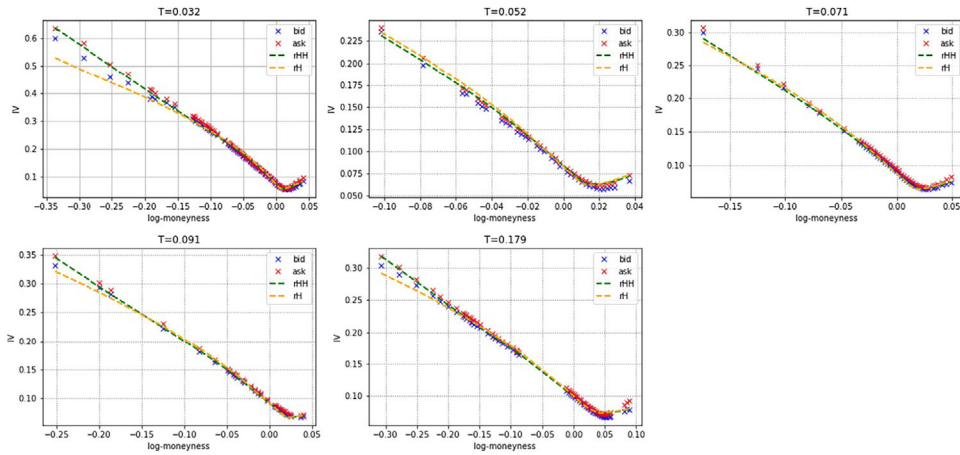
FIGURE 7 SPX implied volatility smiles obtained by minimizing the functional in (42) for the rough Heston model (orange) and the rough Hawkes Heston model (green). The blue and red crosses are respectively the bid and ask of market implied volatilities on May 19, 2017. The abscissa is in log-moneyness and T is the time to expiry in years. [Color figure can be viewed at wileyonlinelibrary.com]
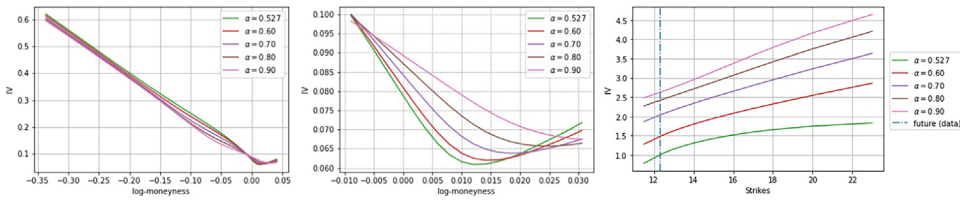
FIGURE 8 Sensitivity of implied volatility for SPX (left and center) and VIX (right) options with respect to the kernel power α for the shortest maturity. [Color figure can be viewed at wileyonlinelibrary.com]

the initial curve parameters (β, σ_0^2) on the implied volatilities for the shortest maturity, and for the shortest and longest maturities, respectively.

We begin with the sensitivity with respect to the parameter $\alpha \in (0.5, 1]$, which as we will see plays a crucial role in our model. We can observe in Figure 8 – as is the case for other rough volatility models – that modifications of the parameter α change the ATM skew of the implied volatility of S&P 500 options. A good convexity and ATM skew, for the maturities considered in the calibration, can be obtained with very low values of the parameter α, confirming the findings in the rough volatility literature. To elucidate the influence of the parameter α on the ATM skews, we plot in Figure 9 the log-log plots of ATM skews as a function of maturity, for the calibrated parameters and different values of α. We observe that a perfect power decay, for the given maturities, is captured by $\alpha = 0.527$, but not by higher values of α. For $\alpha = 0.527$, the linear fit is almost perfect with a -0.587 power decay and an unquestionable coefficient of determination $R^2 = 0.99901$. It is important to mention at this point the recent works (Delemotte et al., 2023; Guyon & El Amrani, 2023) which point out that the linear fit is no longer optimal when considering a larger range of maturities. Our findings for the maturities considered in the calibration are coherent with the results in the rough volatility literature, for example, Bayer et al. (2016), Gatheral et al. (2018), indicating a power law for the ATM skew as a function of maturity given approximately by $T^{-\frac{1}{2}}$. For other
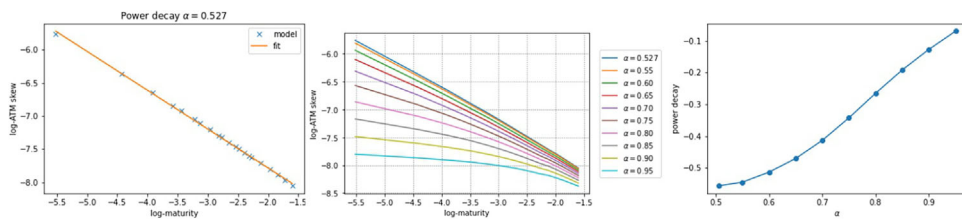
FIGURE 9 Power decay of the ATM volatility skew. On the left, the log-log plot of ATM volatility skew for the calibrated parameters of Table 1. At the center, the log-log plot of ATM volatility skew for different values of α; the other parameters are as in Table 1. On the right, the fitted power decay of the ATM volatility skew as function of α; the power decay is estimated using the five shortest maturities, that is, $\log(T) \in [-5.5, -3.5]$. [Color figure can be viewed at wileyonlinelibrary.com]

values of α, the linear fit with a different slope is also observed for the shortest maturities. In Figure 9, we plot the estimated power decay for the short maturities as a function of α. This plot shows that the relationship between the power decay and α is approximately linear for the short maturities.

More importantly, within the joint calibration framework, the parameter α has a big impact on the level and shape of implied volatilities of VIX options, see Figure 8. As α decreases the implied volatilities shift downwards. This feature is fundamental to bring down the VIX implied volatilities maintaining the correct skew for SPX implied volatilities, explaining therefore the shift mentioned in Guyon (2020a), Guyon (2020b). We ratify therefore – within the affine framework – the relevance of rough non-Markovian volatility to jointly calibrate SPX and VIX smiles.

We now analyze the dependency of the implied volatilities with respect to the other parameters. Figure 10 shows the sensitivities with respect to the evolution-related parameters (b, c, ρ, Λ). We notice that – unless we zoom at-the-money – the sensitivity of the SPX smiles with respect to (b, c) is relatively small. The main effect of an increment in the reverting speed $-b$ is a shift slightly downwards of the SPX implied volatility and a more pronounced upward shift and a reduction of the concavity on the VIX implied volatility. The impact of the volatility of volatility c is similar for SPX options, with a slight change of concavity, and a more pronounced and less symmetric effect on the level and concavity of implied volatility of VIX options. As usual, the correlation parameter ρ plays a big role by moving the minimum value to the left ($\rho > 0$) or to the right ($\rho < 0$). Obviously, the VIX smiles do not depend on the correlation ρ. The effect of the (jump) leverage Λ is relatively small on SPX implied volatilities, being primarily concentrated in the left-tails of the smile, but fundamental on the VIX implied volatilities. For SPX implied volatilities, the impact of Λ could be reduced to a rotation with the at-the-money value as pivot. The parameter Λ also controls the level of VIX implied volatility out-of-the-money. As Λ increases this level goes down, achieving the correct shift for the calibrated parameter. This effect is similar to the one observed for the vol-of-vol c, but the sensitivity is larger, and it allows us to keep a low value of c for the joint calibration. This explains the importance in our model of self-exciting jumps in opposite directions for the underlying and volatility.

We now turn to the parameters (β, σ_0^2) of the initial curve $g_0(t) = \sigma_0^2 + \beta \int_0^t K(s)\,ds$, $t \geq 0$ (see (41)). Figure 11 shows the SPX and VIX implied volatility sensitivities for the shortest and longest maturity. The impact of both parameters is similar for SPX and VIX options. When σ_0^2 or β increase the SPX implied volatilities move up and to the right, while the VIX implied volatilities move down and the concavity increases.
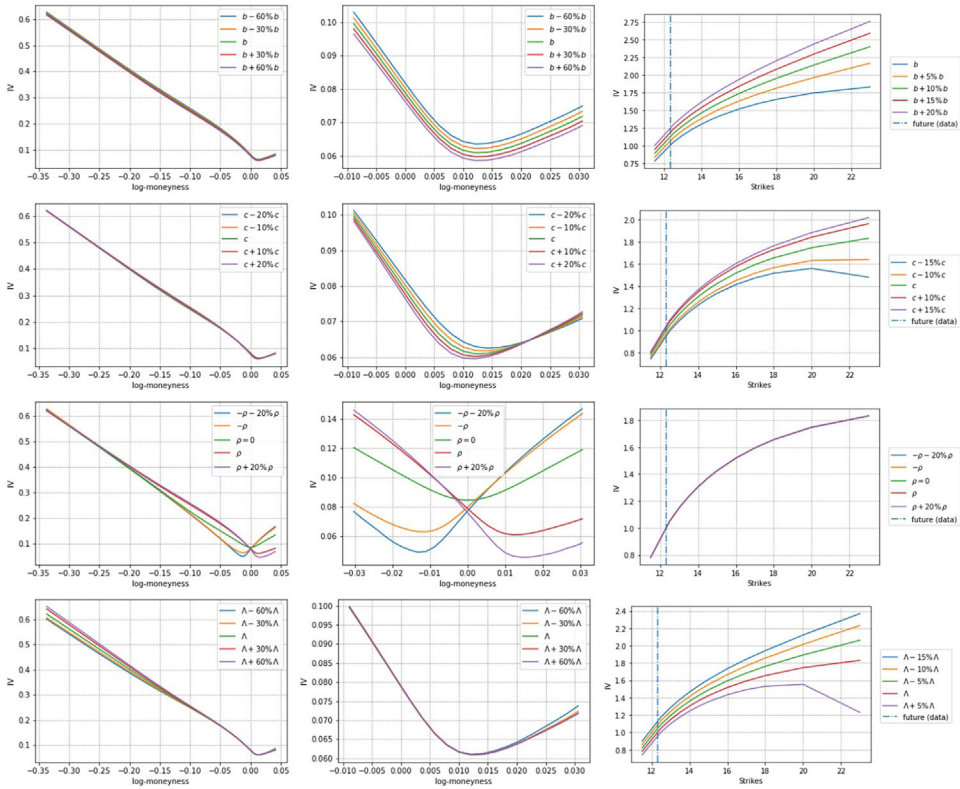
FIGURE 10 Sensitivity of implied volatility for SPX (left, center) and VIX (right) options for the shortest maturity with respect to: the mean reversion speed parameter b (first line), the volatility of volatility c (second line), the correlation ρ (third line), and the jump-leverage Λ (fourth line). [Color figure can be viewed at wileyonlinelibrary.com]
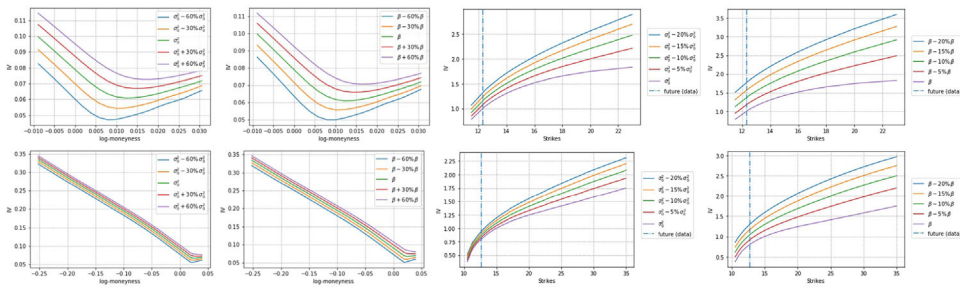
FIGURE 11 Sensitivity of implied volatility for SPX (left, left-center) and VIX (right-center, right) options for the shortest (first line) and longest maturity (second line) with respect to the initial spot variance curve, that is, intercept σ_0^2, and proportional coefficient β. [Color figure can be viewed at wileyonlinelibrary.com]

7.1 | Comparison with the rough Heston model

We now continue the discussion started in Subsection 6.1 regarding the relevance of jumps in the implementation of the rough Hawkes Heston model. Contrary to Subsection 6.1, here we do not
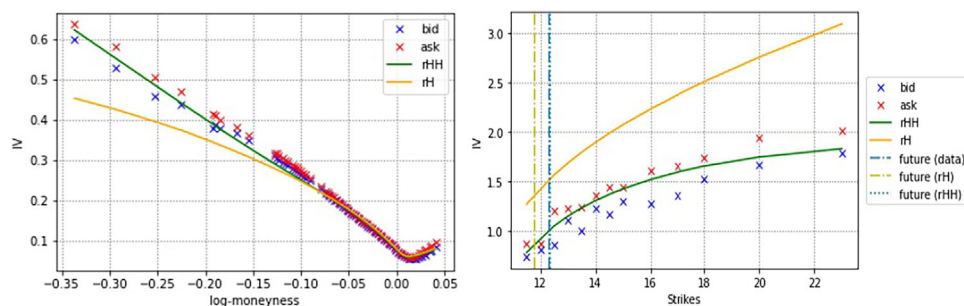
FIGURE 12 Effect of the jump component on the implied volatility of SPX (left) and VIX (right) options. The rough Hawkes Heston (rHH) is in green and the rough Heston (rH) in orange. The blue and red crosses are respectively the bid and ask of market implied volatilities for $T = 0.032$. [Color figure can be viewed at wileyonlinelibrary.com]

focus on the calibration to a particular dataset, but we take a more general point of view. More precisely, we are interested in understanding how the jumps affect the implied volatility curves of SPX/VIX options keeping all the other parameters constant. To do this, we take the values in Table 1 for the rough Hawkes Heston model and simply remove the jump component by setting $\nu(\mathrm{d}z) = 0$, recovering then the rough Heston model. Figure 12 clearly shows that the jumps have a significant impact on implied volatility smiles. In fact, they change the shape of S&P 500 curves and the level of VIX curves. Hence we conclude that the model (1)–(5) suggested in this article is a parsimonious extension of the rough Heston model which does provide a considerably richer framework. Speaking of the joint calibration, the level and shape of SPX/VIX implied volatility curves constitute two main issues to reconcile in order to successfully tackle the problem, see Introduction 1. Since the self-exciting jumps affect them both, it appears that the rough Hawkes Heston model has an important advantage over its continuous counterpart (rough Heston), which is also coherent with the experiments in Subsection 6.1.

8 | CONCLUSION

We develop and study a new stochastic volatility model named the rough Hawkes Heston model. It is a tractable affine Volterra model with rough volatility and volatility jumps that cluster and that have the opposite direction but occur at the same time as the jumps of the underlying prices. This model shares many features with other existing models, mainly the Heston Heston (1993), Barndorff-Nielsen and Shephard Barndorff-Nielsen and Shephard (2001b), and rough Heston El Euch and Rosenbaum (2019) models. It takes advantage of the low regularity and memory features of rough volatility models, the large fluctuation of jumps, the clusters of Hawkes processes and the explicit Fourier-Laplace transform of the affine setup. By combining the modeling advantages of these approaches, it is able to better capture the joint dynamics of underlying prices and their volatility index in a tractable fashion. The addition of a singular kernel in the dynamics of the volatility, together with jumps, incorporates not only the rough volatility feature but also a jump-clustering component. The presence of common jumps in the underlying and the volatility in opposite directions is coherent with previous studies such as Cont and Kokholm (2013), Todorov and Tauchen (2011). Moreover, the introduction of jumps that cluster – as in Bernis et al. (2021) – is in accordance with empirical findings, for example, Cont (2001), Cont (2011). Similar to Barndorff-Nielsen and Shephard (2001b), El Euch and Rosenbaum (2019), Heston (1993), the rough Hawkes

Heston model is parsimonious with only five evolution-related parameters, and it belongs to the class of affine Volterra models (Abi Jaber et al., 2019; Bondi et al., 2024). The affine Volterra nature of the model is appealing because it enables fast option pricing using the COS method and efficient low-factor approximations of the kernel with a sum of exponentials.

The parameter α describing the power kernel in the volatility dynamics controls – as in the rough Heston model – the underlying implied volatilities ATM skews. Our calibration example indicates that this value is close to 0.5, which agrees with previous estimates in the rough volatility literature (Bayer et al., 2016; Gatheral et al., 2018). This is not, however, the only role played by the parameter α in our setup, because the power kernel also affects the jump-clustering feature of the model. As a consequence, the parameter α plays a crucial role in controlling the level of VIX implied volatilities. Together with the jump-leverage parameter Λ, the power kernel allows us to bring down the VIX implied volatilities maintaining the correct skew for SPX implied volatilities, consequently capturing the shift mentioned in Guyon (2020a), Guyon (2020b). This confirms the relevance, in our affine framework, of rough volatility and clustering jumps to model simultaneously the S&P 500 and VIX dynamics.

The affine relation between variance swap rates and forward variance – which generalizes the affine relation between variance swap rates and spot variance in the classical framework Kallsen et al. (2011) – is a by-product of our affine Volterra framework. This affine relation has been confirmed empirically in Mancino et al. (2020).

To conclude, the rough Hawkes Heston model is able – in a tractable and parsimonious fashion – to jointly calibrate S&P 500 and VIX options. The parsimonious character of our model is an advantage compared to other models that jointly calibrate SPX/VIX options with either a large number of parameters (Guyon & Lekeufack, 2023) or based on martingale transport considerations (Guyon, 2020b; Guyon, 2023; Guyon & Bourgey, 2024). The affine character of the rough Hawkes Heston model allows fast pricing using Fourier-techniques, instead of Monte Carlo or machine learning methods as those used for instance in Gatheral et al. (2020), Rosenbaum and Zhang (2022). Moreover, all the parameters in our model have a financial interpretation, and a complete sensitivity analysis shows that they are not redundant since each of them controls a different feature of the S&P 500 and VIX volatility smiles.

DATA AVAILABILITY STATEMENT

Market data were purchased from the CBOE website https://datashop.cboe.com/.

ORCID

Alessandro Bondi https://orcid.org/0000-0002-8246-6992
Sergio Pulido https://orcid.org/0000-0001-7629-8585
Simone Scotti https://orcid.org/0000-0003-4365-6539

REFERENCES

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

Abi Jaber, E. (2021). Weak existence and uniqueness for affine stochastic Volterra equations with L^1-kernels. *Bernoulli*, *27*(3), 1583–1615.

Abi Jaber, E., Illand, C., & Li, S. (2022a). Joint SPX-VIX calibration with Gaussian polynomial volatility models: Deep pricing with quantization hints. *Preprint arXiv:2212.08297*.

Abi Jaber, E., Illand, C., & Li, S. (2022b). The quintic Ornstein-Uhlenbeck volatility model that jointly calibrates SPX & VIX smiles. *Preprint arXiv:2212.10917*.

Abi Jaber, E., Cuchiero, C., Larsson, M., & Pulido, S. (2021). A weak solution theory for stochastic Volterra equations of convolution type. *The Annals of Applied Probability*, *31*(6), 2924–2952.

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

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

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

Alos, E., Garcia-Lorite, D., & Gonzalez, A. M. (2022). On smile properties of volatility derivatives: Understanding the VIX skew. *SIAM Journal on Financial Mathematics*, *13*(1), 32–69.

Alos, E., Gatheral, J., & Radoičić, R. (2020). Exponentiation of conditional expectations under stochastic volatility. *Quantitative Finance*, *20*(1), 13–27.

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

Alos, E., & Shiraya, K. (2019). Estimating the Hurst parameter from short term volatility swaps: A Malliavin calculus approach. *Finance and Stochastics*, *23*(2), 423–447.

Avellaneda, M., & Papanicolaou, A. (2018). Statistics of VIX futures and applications to trading volatility exchange-traded products. *Journal of Investment Strategies*, *7*(2), 1–33.

Barndorff-Nielsen, O., & Shephard, N. (2001a). Modelling by Lévy processes for financial econometrics. In *Lévy processes*, Birkhauser, 283–318.

Barndorff-Nielsen, O., & Shephard, N. (2001b). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistics Society: Series B (Statistical Methodology)*, *63*(2), 167–241.

Bates, D. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. *Review of Financial Studies*, *9*, 69–107.

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

Bayer, C., & Breneis, S. (2023). Weak Markovian approximations of rough Heston. *Preprint arXiv:2309.07023*.

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

Beesack, P. R. (1969). Comparison theorems and integral inequalities for Volterra integral equations. *Proceedings of the American Mathematical Society*, *20*(1), 61–66.

Bennedsen, M., Lunde, A., & Pakkanen, M. S. (2021). Decoupling theshort-and long-term behavior of stochastic volatility. *Journal of Financial Econometrics*.

Bernis, G., Brignone, R., Scotti, S., & Sgarra, C. (2021). A Gamma Ornstein-Uhlenbeck model driven by a Hawkes process. *Mathematics and Financial Economics*, *15*(4), 747–773.

Bondi, A., Livieri, G., & Pulido, S. (2024). Affine Volterra processes with jumps. *Stochastic Processes and their Applications*, *168*, 104264.

Brignone, R., & Sgarra, C. (2019). Asian options pricing in Hawkes-type jump-diffusion models. *Annals of Finance*, *16*(1), 101–119.

Callegaro, G., Grasselli, M., & Pages, G. (2021). Fast hybrid schemes for fractional Riccati equations (rough is not so tough). *Mathematics of Operations Research*, *46*(1), 221–254.

Callegaro, G., Mazzoran, A., & Sgarra, C. (2022). A self-exciting modeling framework for forward prices in power markets. *Applied Stochastic Models in Business and Industry*, *38*(1), 27–48.

Carr, P., & Lee, R. (2009). Robust Replication of Volatility Derivatives. https://math.uchicago.edu/%7Erl/rrvd.pdf

Chevalier, E., Pulido, S., & Zúñiga, E. (2022). American options in the Volterra Heston model. *SIAM Journal on Financial Mathematics*, *13*(2), 426–458.

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

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, *1*(2), 223–236.

Cont, R. (2011). Statistical modeling of high frequency financial data: Facts, models and challenges. *IEEE Signal Processing*, *28*(5), 16–25.

Cont, R., & Kokholm, T. (2013). A consistent pricing model for index options and volatility derivatives. *Mathematical Finance*, *23*(2), 248–274.

Cox, J., Ingersoll, J., & Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica*, *53*, 385–408.

Cuchiero, C., Gazzani, G., Möller, J., & Svaluto-Ferro, S. (2023). Joint calibration to SPX and VIX options with signature-based models. *Preprint arXiv:2301.13235*.

Cuchiero, C., Gazzani, G., & Svaluto-Ferro, S. (2023). Signature-based models: Theory and calibration. *SIAM journal on financial mathematics*, *14*(3), 910–957.

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

Cuchiero, C., & Teichmann, J. (2020). Generalized Feller processes and Markovian lifts of stochastic Volterra processes: The affine case. *Journal of evolution equations*, *20*(4), 1301–1348.

Curato, I. V., & Sanfelici, S. (2015). Measuring the leverage effect in a high frequency trading framework. *Handbook of High Frequency Trading*, 425–446.

Delemotte, J., De Marco, S., & Segonne, F. (2023). Yet another analysis of the SP500 at-the-money skew: Crossover of different power-law behaviours. *SSRN 4428407*.

Demeterfi, K., Derman, E., Kamal, M., & Zou, J. (1999). More than you ever wanted to know about volatility swaps. *Goldman Sachs Quantitative Strategies Research Notes*.

Diethelm, K., Ford, N. J., & Freed, A. D. (2002). A predictor-corrector approach for the numerical solution of fractional differential equations. *Nonlinear Dynamics*, *29*(1), 3–22.

Diethelm, K., Ford, N. J., & Freed, A. D. (2004). Detailed error analysis for a fractional Adams method. *Numerical Algorithms*, *36*(1), 31–52.

Dotsis, G., Psychoyios, D., & Skiadopoulos, G. (2007). An empirical comparison of continuous-time models of implied volatility indices. *Journal of Banking & Finance*, *31*(12), 3584–3603.

Duffie, D., Filipović, D., & Schachermayer, W. (2003). Affine processes and applications in finance. *The Annals of Applied Probability*, *13*(3), 984–1053.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, *68*(6), 1343–1376.

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

El Euch, O., & Rosenbaum, M. (2018). Perfect hedging in rough Heston models. *The Annals of Applied Probability*, *28*(6), 3813–3856.

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

El Euch, O., Gatheral, J., Radoičić, R., & Rosenbaum, M. (2020). The Zumbach effect under rough Heston. *Quantitative Finance*, *20*(2), 235–241.

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

Filipović, D. (2001). A general characterization of one factor affine term structure models. *Finance and Stochastics*, *5*(3), 389–412.

Fukasawa. M,. (2017). Short-time at-the-money skew and rough fractional volatility. *Quantitative Finance*, *17*(2), 189–198.

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

Gatheral, J., Jusselin, P., & Rosenbaum, M. (2020). The quadratic rough Heston model and the joint S&P 500/VIX smile calibration problem. *Risk, cutting edge*, https://www.risk.net/cutting-edge/banking/7530461/the-quadratic-rough-heston-model-and-the-joint-sp-500vix-smile-calibration-problem

Gonzato, L., & Sgarra, C. (2021). Self-exciting jumps in the oil market: Bayesian estimation and dynamic hedging. *Energy Economics*, *99*, 105279.

Goutte, S., Ismail, A., & Pham, H. (2017). Regime-switching stochastic volatility model: Estimation and calibration to VIX options. *Applied Mathematical Finance*, *24*(1), 38–75.

Grasselli, M. (2016). The 4/2 stochastic volatility model: A unified approach for the Heston and the 3/2 model. *Mathematical Finance*, *27*(4), 1013–1034.

Gripenberg, G., Londen, S. O., & Staffans, O. (1990). *Volterra integral and functional equations* (No. 34). Cambridge University Press.

Grzelak, L.A. (2022). On randomization of affine diffusion processes with application to pricing of options on VIX and S&P 500. *Preprint arXiv:2208.12518*.

Guyon, J. (2020a). Inversion of convex ordering in the VIX market. *Quantitative Finance*, *20*(10), 1597–1623.

Guyon, J. (2020b). The joint S&P 500/VIX smile calibration puzzle solved. *Risk, April*.

Guyon, J. (2023). Dispersion-constrained martingale Schrödinger problems and the exact joint S&P 500/VIX smile calibration puzzle. *Finance and Stochastics*, 1–53.

Guyon, J., & Bourgey, F. (2024). Fast exact joint S&P 500/VIX smile calibration in discrete and continuous time. *Risk, Cutting Edge*, https://www.risk.net/cutting-edge/7958955/fast-exact-joint-sp-500vix-smile-calibration-in-discrete-and-continuous-time

Guyon, J., & El Amrani, M. (2023). Does the term-structure of equity at-the-money skew really follow a power law? *Risk, Cutting Edge*, https://www.risk.net/cutting-edge/7957304/does-the-term-structure-of-the-at-the-money-skew-really-follow-a-power-law

Guyon, J., & Lekeufack, J. (2023). Volatility is (mostly) path-dependent. *Quantitative Finance*, *23*(9), 1221–1258.

Guyon, J., & Mustapha, S. (2023). Neural joint S&P 500/VIX smile calibration. *Risk, Cutting Edge*, https://www.risk.net/cutting-edge/7958175/neural-joint-sp-500vix-smile-calibration

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

Horst, U., & Xu, W. (2022). The microstructure of stochastic volatility models with self-exciting jump dynamics. *The Annals of Applied Probability*, *32*(6), 4568–4610.

Kallsen, J., Muhle-Karbe, J., & Voss, M. (2011). Pricing options on variance in affine stochastic volatility models. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, *21*(4), 627–641.

Jarrow, R. A. (2018). *Continuous-time asset pricing theory*. Springer.

Jiao, Y., Ma, C., & Scotti, S. (2017). Alpha-CIR model with branching processes in sovereign interest rate modeling. *Finance and Stochastics*, *21*(3), 789–813.

Jiao, Y., Ma, C., Scotti, S., & Zhou, C. (2021). The Alpha-Heston stochastic volatility model. *Mathematical Finance*, *31*(3), 943–978.

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

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

Mancino, M. E., Scotti, S., & Toscano, G. (2020). Is the variance swap rate affine in the spot variance? Evidence from S&P500 Data. *Applied Mathematical Finance*, *27*(4), 288–316.

Mancino, M. E., & Toscano, G. (2022). Rate-efficient asymptotic normality for the Fourier estimator of the leverage process. *Statistics and Its Interface*, *15*(1), 73–89.

Meyer, P. A. (1966). *Probability and potentials* (Vol. 1318). Blaisdell Publishing Company.

Nicolato, E., Pisani, C., & Sloth, D. (2017). The impact of jump distributions on the implied volatility of variance. *SIAM Journal on Financial Mathematics*, *8*(1), 28–53.

Pacati, C., Pompa, G., & Renò, R. (2018). Smiling twice: the Heston++ model. *Journal of Banking & Finance*, *96*, 185–206.

Papanicolaou, A. (2022). Consistent time-homogeneous modeling of SPX and VIX derivatives. *Mathematical Finance*.

Papanicolaou, A., & Sircar, R. (2014). A regime-switching Heston model for VIX and S&P 500 implied volatilities. *Quantitative Finance*, *14*(10), 1811–1827.

Protter, P. E. (2005). *Stochastic integration and differential equations*. Springer, Berlin, Heidelberg.

Raffaelli, I., Scotti, S., & Toscano, G. (2022). Hawkes-driven stochastic volatility models: Goodness-of-fit testing of alternative intensity specifications with S&P500 data. *Annals of Operations Research*, 1–19.

Recchioni, M. C., Iori, G., Tedeschi, G., & Ouellette, M. S. (2021). The complete Gaussian kernel in the multi-factor Heston model: Option pricing and implied volatility applications. *European Journal of Operational Research*, *293*(1), 336–360.

Rhoads, R. (2011). *Trading VIX derivatives: Trading and hedging strategies using VIX futures, options, and exchange-traded notes*. Wiley.

Rømer, S. E. (2022a). Empirical analysis of rough and classical stochastic volatility models to the SPX and VIX markets. *Quantitative Finance*, *22*(10), 1805–1838.

Rømer, S. E. (2022b). Hybrid multifactor scheme for stochastic Volterra equations with completely monotone kernels. *SSRN 3706253*.

Rosenbaum, M., & Zhang, J. (2022). Deep calibration of the quadratic rough Heston model. *Risk, cutting-hedge*, https://www.risk.net/cutting-edge/7954656/deep-calibration-of-the-quadratic-rough-heston-model

Rudin, W. (1987). *Real and complex analysis* (3rd ed.). McGraw-Hill International Editions: Mathematics series.

Schmelzle, M. (2010). Option pricing formulae using Fourier transform: Theory and application. *Preprint*, https://pfadintegral.com/docs/Schmelzle2010%20Fourier%20Pricing.pdf

Sepp, A. (2008). Pricing options on realized variance in the Heston model with jumps in returns and volatility. *Journal of Computational Finance*, *11*, 33–70.

Song, Z., & Xiu, D. (2012). A tale of two option markets: State-price densities implied from S&P 500 and VIX option prices. *Unpublished working paper. Federal Reserve Board and University of Chicago*.

Todorov, V., & Tauchen, G. (2011). Volatility jumps. *Journal of Business & Economic Statistics*, *29*(3), 356–371.

APPENDIX A: PROOF OF THEOREM 3.1

In this appendix we prove Theorem 3.1 regarding the Riccati-Volterra equation (6)–(7) used to study the Fourier-Laplace transform of the log returns $(X_t)_{t \geq 0}$. We use the following notation: given $u, v \in \mathbb{C}$, let $[u, v]$ be the segment in $\mathbb{C}$ having u and v as endpoints, that is, $[u, v] = \{z \in \mathbb{C} : z = (1 - t)u + tv, t \in [0, 1]\}$, and denote by $u \vee v = \operatorname{Re} u \vee \operatorname{Re} v + i \operatorname{Im} u \vee \operatorname{Im} v$.

Proof. Fix $w \in \mathbb{C}$ with $\operatorname{Re} w \in [0, 1]$.

(i) The proof of this point is divided into three steps. In the first step, we show the existence of a noncontinuable solution ψ_w of (7). In the second step, we prove that ψ_w does not explode in finite time, that is, that it is global solution. To conclude, in the third and last step, we prove the uniqueness of ψ_w.

Step I. Let us compute from (6), for every $v \in \mathbb{C}_-$,

$$
\begin{aligned}
\operatorname{Re} \mathscr{R}(w, v) = {} & \frac{1}{2}\left(|\operatorname{Re} w|^2 - \operatorname{Re} w\right) + \left(b + \rho\sqrt{c}\,\operatorname{Re} w\right)\operatorname{Re} v + \frac{c}{2}|\operatorname{Re} v|^2 \\
& - \frac{1}{2}\left(|\operatorname{Im} w|^2 + c|\operatorname{Im} v|^2 + 2\rho\sqrt{c}\,\operatorname{Im} w \operatorname{Im} v\right) \\
& + \int_{\mathbb{R}_+}\left[e^{(\operatorname{Re} v - \Lambda\operatorname{Re} w)z}\cos((\operatorname{Im} v - \Lambda\operatorname{Im} w)z) - \operatorname{Re} w\left(e^{-\Lambda z} - 1\right) - 1 - \operatorname{Re} v z\right]\nu(\mathrm{d}z).
\end{aligned}
$$

(A.1)

Since $|\rho| \leq 1$ we have $|\rho\sqrt{c}\, \mathrm{Im}\, w \mathrm{Im}\, v| \leq \sqrt{c}|\mathrm{Im}\, w||\mathrm{Im}\, v|$, which implies

$$-\frac{1}{2}\Big(|\mathrm{Im}\, w|^2 + c|\mathrm{Im}\, v|^2 + 2\rho\sqrt{c}\, \mathrm{Im}\, w \mathrm{Im}\, v\Big) \leq -\frac{1}{2}\Big(|\mathrm{Im}\, w| - \sqrt{c}|\mathrm{Im}\, v|\Big)^2 \leq 0. \qquad (A.2)$$

Recalling that $\mathrm{Re}\, w \in [0,1]$, we then obtain

$$\mathrm{Re}\, \mathscr{R}(w,v) \leq \Big(b + \rho\sqrt{c}\, \mathrm{Re}\, w\Big)\mathrm{Re}\, v + \frac{c}{2}|\mathrm{Re}\, v|^2 + \int_{\mathbb{R}_+} \big[e^{-\Lambda \mathrm{Re}\, wz} - \mathrm{Re}\, w\big(e^{-\Lambda z} - 1\big) - 1\big]\nu(\mathrm{d}z)$$

$$+ \int_{\mathbb{R}_+} \big[e^{(\mathrm{Re}\, v - \Lambda \mathrm{Re}\, w)z} - e^{-\Lambda \mathrm{Re}\, wz} - \mathrm{Re}\, vz\big]\nu(\mathrm{d}z)$$

$$\leq \Big(b + \rho\sqrt{c}\, \mathrm{Re}\, w + \int_{\mathbb{R}_+} z(e^{-\Lambda \mathrm{Re}\, wz} - 1)\nu(\mathrm{d}z)\Big)\mathrm{Re}\, v + \frac{c}{2}|\mathrm{Re}\, v|^2 + \int_{\mathbb{R}_+} e^{-\Lambda \mathrm{Re}\, wz}(e^{\mathrm{Re}\, vz} - 1 - \mathrm{Re}\, vz)\nu(\mathrm{d}z), \qquad (A.3)$$

where for the second inequality we use

$$e^{-\Lambda \mathrm{Re}\, wz} - \mathrm{Re}\, w\big(e^{-\Lambda z} - 1\big) - 1 \leq 0, \quad z \geq 0. \qquad (A.4)$$

Let $h : \mathbb{R}_+ \times \mathbb{R}_- \to \mathbb{R}_-$ be the continuous function defined by

$$h(x,y) = \begin{cases} \frac{1}{y}\int_{\mathbb{R}_+} e^{-\Lambda xz}(e^{yz} - 1 - yz)\nu(\mathrm{d}z), & y < 0 \\ 0, & y = 0 \end{cases}, \quad x \geq 0,$$

and note that $y \cdot h(x,y) = \int_{\mathbb{R}_+} e^{-\Lambda xz}(e^{yz} - 1 - yz)\nu(\mathrm{d}z)$. At this point, we can use (A.3) to show that

$$\mathrm{Re}\, \mathscr{R}(w,v) \leq \Big(C_w + \frac{c}{2}\mathrm{Re}\, v + h(\mathrm{Re}\, w, \mathrm{Re}\, v)\Big)\mathrm{Re}\, v, \quad v \in \mathbb{C}_-, \qquad (A.5)$$

where $C_w = b + \rho\sqrt{c}\, \mathrm{Re}\, w + \int_{\mathbb{R}_+} z(e^{-\Lambda \mathrm{Re}\, wz} - 1)\nu(\mathrm{d}z)$.

We now introduce the function $\widetilde{\mathscr{R}}_w : \mathbb{C} \to \mathbb{C}$ given by

$$\widetilde{\mathscr{R}}_w(v) = \mathscr{R}(w, -\mathrm{Re}\, v^- + i\mathrm{Im}\, v) + C_w \mathrm{Re}\, v^+, \quad v \in \mathbb{C}.$$

Observe that, by construction (see also (A.5))

$$\mathrm{Re}\, \widetilde{\mathscr{R}}_w(v) \leq \Big(C_w - \frac{c}{2}\mathrm{Re}\, v^- + h(\mathrm{Re}\, w, -\mathrm{Re}\, v^-)\Big)\mathrm{Re}\, v, \quad v \in \mathbb{C}.$$

Since $\widetilde{\mathscr{R}}_w$ is continuous, we can invoke (Gripenberg et al., 1990, Therorem 1.1, Chapter 12) to assert the existence of a continuous, noncontinuable solution $\psi_w : [0, T_{\max}) \to \mathbb{C}$ of the equation

$$\chi = K * \widetilde{\mathscr{R}}_w(\chi(\cdot)), \quad t \in [0, T_{\max}), \qquad (A.6)$$

for some $T_{\max} \in (0, \infty]$. If we can show that $\mathrm{Re}\, \psi_w \leq 0$ in $[0, T_{\max})$, then we conclude that ψ_w is indeed a noncontinuable solution of (7), as well. To this end, consider the continuous function $\zeta(t) = C_w - \frac{c}{2}\mathrm{Re}\, \psi_w(t)^- + h(\mathrm{Re}\, w, -\mathrm{Re}\, \psi_w(t)^-)$ defined for $t \in [0, T_{\max})$. Taking the real part in

(A.6), for every $T \in (0, T_{\max})$, we obtain

$$\operatorname{Re} \psi_w(t) = -\gamma_T(t) + \int_0^t K(t - s) \zeta(s) \operatorname{Re} \psi_w(s) ds, \quad t \in [0, T],$$

where $\gamma_T(t) = \int_0^t K(t - s) 1_{\{s \leq T\}} (\zeta(s) \operatorname{Re} \psi_w(s) - \operatorname{Re} \widetilde{\mathcal{R}}_w(\psi_w(s))) ds$. By (Abi Jaber and El Euch, 2019b, Remark B.6) $\gamma_T \in \mathcal{G}_K$ (recall (22)), and we can invoke (Abi Jaber and El Euch, 2019b, Theorem C.1) to infer that $\operatorname{Re} \psi_w \leq 0$ in $[0, T]$. Given that T was arbitrary, such an inequality holds in the whole interval $[0, T_{\max})$, completing the first step of the proof.

Step II. Our goal here is to show that $T_{\max} = \infty$. Let us fix again a generic $T \in (0, T_{\max})$. Taking the imaginary part in (6) and (7) we have, on the interval $[0, T]$,

$$\operatorname{Im} \psi_w = K * \left[\left(\operatorname{Re} w - \frac{1}{2} \right) \operatorname{Im} w + \left(b + \rho \sqrt{c} \operatorname{Re} w \right) \operatorname{Im} \psi_w + \rho \sqrt{c} \operatorname{Im} w \operatorname{Re} \psi_w + c \operatorname{Re} \psi_w \operatorname{Im} \psi_w \right.$$
$$\left. + \int_{\mathbb{R}_+} \left(e^{\operatorname{Re}(\psi_w - \Lambda w) \cdot z} \sin \left(\operatorname{Im}(\psi_w - \Lambda w) \cdot z \right) - \operatorname{Im} w \left(e^{-\Lambda z} - 1 \right) - \operatorname{Im} \psi_w \cdot z \right) \nu(dz) \right]. \quad (A.7)$$

Consider the function $d : \mathbb{R}_- \times \mathbb{R} \to \mathbb{R}$ defined as follows

$$d(x, y) = \begin{cases} \frac{1}{y} \int_{\mathbb{R}_+} e^{xz} (\sin(y z) - y z) \nu(dz), & y \neq 0 \\ 0, & y = 0 \end{cases}, \quad x \leq 0.$$

Note that d is continuous and nonpositive in its domain. Moreover, by construction

$$y \cdot d(x, y) = \int_{\mathbb{R}_+} e^{xz} (\sin(y z) - y z) \nu(dz), \quad (x, y) \in \mathbb{R}_- \times \mathbb{R}.$$

To shorten the notation we define $\widetilde{\psi_w} = \psi_w - \Lambda w$. Using the function d we rewrite (A.7) as

$$\operatorname{Im} \psi_w + \frac{\rho^+}{\sqrt{c}} \operatorname{Im} w = \frac{\rho^+}{\sqrt{c}} \operatorname{Im} w + K * \left[\left(\operatorname{Re} w - \frac{1}{2} - \int_{\mathbb{R}_+} (e^{-\Lambda z} - 1 + \Lambda z) \nu(dz) - \frac{\rho^+}{\sqrt{c}} \left(b + \rho \sqrt{c} \operatorname{Re} w \right) \right) \operatorname{Im} w \right.$$
$$+ \left(-\rho^- \sqrt{c} \operatorname{Re} \psi_w - \left(\Lambda + \frac{\rho^+}{\sqrt{c}} \right) \int_{\mathbb{R}_+} z \left(e^{\operatorname{Re} \widetilde{\psi_w} \cdot z} - 1 \right) \nu(dz) - \left(\Lambda + \frac{\rho^+}{\sqrt{c}} \right) d \left(\operatorname{Re} \widetilde{\psi_w}, \operatorname{Im} \widetilde{\psi_w} \right) \right) \operatorname{Im} w$$
$$\left. + \left(\left(b + \rho \sqrt{c} \operatorname{Re} w \right) + c \operatorname{Re} \psi_w + \int_{\mathbb{R}_+} z \left(e^{\operatorname{Re} \widetilde{\psi_w} \cdot z} - 1 \right) \nu(dz) + d \left(\operatorname{Re} \widetilde{\psi_w}, \operatorname{Im} \widetilde{\psi_w} \right) \right) \left(\operatorname{Im} \psi_w + \frac{\rho^+}{\sqrt{c}} \operatorname{Im} w \right) \right]$$

$$=: \frac{\rho^+}{\sqrt{c}} \operatorname{Im} w$$

$$+ K * \left[\left(C_1 - \frac{\rho^+}{\sqrt{c}} \left(b + \rho \sqrt{c} \operatorname{Re} w \right) \right) \operatorname{Im} w + f_1(\cdot) \operatorname{Im} w + \left(b + \rho \sqrt{c} \operatorname{Re} w + f_2(\cdot) \right) \left(\operatorname{Im} \psi_w + \frac{\rho^+}{\sqrt{c}} \operatorname{Im} w \right) \right],$$

which holds on $[0, T]$. In particular, note that $f_1 \geq 0$ and $f_2 \leq 0$ in $[0, T]$. We want to find a continuous function $u : \mathbb{R}_+ \to \mathbb{R}_+$ such that $|\operatorname{Im} \psi_w| \leq u$ on $[0, T]$. To do this, we argue by cases on $\operatorname{Im} w$. In the following, we denote $\widetilde{\Lambda} = \max\{\rho^- c^{-1/2}, \Lambda\}$. All the claims regarding the sign of solutions to linear Volterra equations are justified by (Abi Jaber & El Euch, 2019b, Theorem C.1).

If $\operatorname{Im} w \geq 0$, then we can consider the unique, nonnegative, continuous solution $l_1 : [0, T] \to \mathbb{R}_+$ of the linear equation

$$l_1 = \frac{\rho^+}{\sqrt{c}} \operatorname{Im} w + K * \left[\left| C_1 - \frac{\rho^+}{\sqrt{c}}\left(b + \rho\sqrt{c}\operatorname{Re} w\right)\right| \operatorname{Im} w + \left(\left(b + \rho\sqrt{c}\operatorname{Re} w\right) + f_2\right) l_1\right].$$

Since the function $\operatorname{Im}\psi_w + \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w + l_1$ satisfies – in $[0, T]$ – the linear equation

$$\chi = 2\frac{\rho^+}{\sqrt{c}}\operatorname{Im} w + K * \left[2\left(C_1 - \frac{\rho^+}{\sqrt{c}}\left(b + \rho\sqrt{c}\operatorname{Re} w\right)\right)^+ \operatorname{Im} w + f_1 \operatorname{Im} w + \left(\left(b + \rho\sqrt{c}\operatorname{Re} w\right) + f_2\right)\chi\right],$$

we deduce that $\operatorname{Im}\psi_w \geq -l_1 - \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w$ on $[0, T]$. Next, we introduce the unique, nonnegative, continuous solution $\overline{l_1} : \mathbb{R}_+ \to \mathbb{R}_+$ of the linear equation

$$\overline{l_1} = \frac{\rho^+}{\sqrt{c}}|\operatorname{Im} w| + K * \left[\left| C_1 - \frac{\rho^+}{\sqrt{c}}\left(b + \rho\sqrt{c}\operatorname{Re} w\right)\right| |\operatorname{Im} w| + \left(b + \rho\sqrt{c}\operatorname{Re} w\right)\overline{l_1}\right] \tag{A.8}$$

and observe that $\overline{l_1} - l_1 \geq 0$ on $[0, T]$, because $\overline{l_1} - l_1$ solves on $[0, T]$

$$\chi = K * \left[-f_2 l_1 + \left(b + \rho\sqrt{c}\operatorname{Re} w\right)\chi\right].$$

Hence, $\operatorname{Im}\psi_w \geq -\overline{l_1} - \frac{\rho^+}{\sqrt{c}}|\operatorname{Im} w|$ on $[0, T]$. We now focus on the upper bound. Observe that

$$\operatorname{Im}\psi_w - \widetilde{\Lambda}\operatorname{Im} w = -\widetilde{\Lambda}\operatorname{Im} w + K * \left[\left(C_1 + \left(b + \rho\sqrt{c}\operatorname{Re} w\right)\widetilde{\Lambda}\right)\operatorname{Im} w + \left(b + \rho\sqrt{c}\operatorname{Re} w + f_2\right)\left(\operatorname{Im}\psi_w - \widetilde{\Lambda}\operatorname{Im} w\right)\right.$$

$$\left. + \left(\left(\widetilde{\Lambda}c + \rho\sqrt{c}\right)\operatorname{Re}\psi_w + (\widetilde{\Lambda} - \Lambda)\left(\int_{\mathbb{R}_+} z\left(e^{\operatorname{Re}\widetilde{\psi_w}\cdot z} - 1\right)\nu(dz) + d\left(\operatorname{Re}\widetilde{\psi_w}, \operatorname{Im}\widetilde{\psi_w}\right)\right)\right)\operatorname{Im} w\right].$$

We then take the unique, nonnegative, continuous solution $u_1 : [0, T] \to \mathbb{R}_+$ of the linear equation

$$u_1 = \widetilde{\Lambda}\operatorname{Im} w + K * \left[\left|C_1 + \left(b + \rho\sqrt{c}\operatorname{Re} w\right)\widetilde{\Lambda}\right|\operatorname{Im} w + \left(b + \rho\sqrt{c}\operatorname{Re} w + f_2\right)u_1\right].$$

We infer that $u_1 - (\operatorname{Im}\psi_w - \widetilde{\Lambda}\operatorname{Im} w) \geq 0$ since $\widetilde{\Lambda}c + \rho\sqrt{c}$, $\widetilde{\Lambda} - \Lambda \geq 0$, and $u_1 - (\operatorname{Im}\psi_w - \widetilde{\Lambda}\operatorname{Im} w)$ satisfies (on $[0, T]$)

$$\chi = 2\widetilde{\Lambda}\operatorname{Im} w + K * \left[2\left(C_1 + \left(b + \rho\sqrt{c}\operatorname{Re} w\right)\widetilde{\Lambda}\right)^- \operatorname{Im} w + \left(b + \rho\sqrt{c}\operatorname{Re} w + f_2\right)\chi\right.$$

$$\left. - \left(\left(\widetilde{\Lambda}c + \rho\sqrt{c}\right)\operatorname{Re}\psi_w + (\widetilde{\Lambda} - \Lambda)\left(\int_{\mathbb{R}_+} z\left(e^{\operatorname{Re}\widetilde{\psi_w}\cdot z} - 1\right)\nu(dz) + d\left(\operatorname{Re}\widetilde{\psi_w}, \operatorname{Im}\widetilde{\psi_w}\right)\right)\right)\operatorname{Im} w\right].$$

To end, we introduce the unique, nonnegative, continuous solution $\overline{u_1} : \mathbb{R}_+ \to \mathbb{R}_+$ of the linear equation

$$\overline{u_1} = \widetilde{\Lambda} |\text{Im } w| + K * \left[\left| C_1 + \left(b + \rho \sqrt{c} \, \text{Re } w \right) \widetilde{\Lambda} \right| |\text{Im } w| + \left(b + \rho \sqrt{c} \, \text{Re } w \right) \overline{u_1} \right], \qquad \text{(A.9)}$$

and since $\overline{u_1} - u_1$ satisfies the linear equation $\chi = K * [-f_2 \, u_1 + (b + \rho \sqrt{c} \, \text{Re } w)\chi]$ on $[0, T]$, we conclude that $\overline{u_1} \geq u_1$ on the same interval. Therefore, $\text{Im } \psi_w \leq \overline{u_1} + \widetilde{\Lambda} \, \text{Im } w$ on $[0, T]$.

In the case $\text{Im } w \leq 0$ the argument is analogous, but the upper and lower bounds are inverted. Specifically, with the same steps as the ones just carried out, we have $-\overline{u_1} - \widetilde{\Lambda} |\text{Im } w| \leq \text{Im } \psi_w \leq \overline{l_1} + \frac{\rho^+}{\sqrt{c}} |\text{Im } w|$ on $[0, T]$.

Therefore, defining the continuous function $u : \mathbb{R}_+ \to \mathbb{R}_+$ by $u = \overline{l_1} + \overline{u_1} + (\widetilde{\Lambda} + \frac{\rho^+}{\sqrt{c}})|\text{Im } w|$, we have

$$|\text{Im } \psi_w(t)| \leq u(t), \quad 0 \leq t \leq T. \qquad \text{(A.10)}$$

Taking the real part in (7) and using (A.1) we deduce that

$$\text{Re } \psi_w = K * \left[\frac{1}{2}\left(|\text{Re } w|^2 - \text{Re } w\right) + \left(b + \rho \sqrt{c} \, \text{Re } w \right)\text{Re } \psi_w + \frac{c}{2}|\text{Re } \psi_w|^2 \right.$$

$$- \frac{1}{2}\left(|\text{Im } w|^2 + c \, |\text{Im } \psi_w|^2 + 2\rho \sqrt{c} \, \text{Im } w \, \text{Im } \psi_w \right) - \left| \int_{\mathbb{R}_+} e^{\text{Re } \widetilde{\psi_w} \cdot z}\left(\cos\left(\text{Im } \widetilde{\psi_w} \cdot z\right) - 1\right)\nu(\mathrm{d}z) \right|$$

$$\left. + \int_{\mathbb{R}_+} \left(e^{\text{Re } \psi_w \cdot z}\left(e^{-\Lambda \text{Re } wz} - 1\right) - \text{Re } w\left(e^{-\Lambda z} - 1\right)\right)\nu(\mathrm{d}z) + \int_{\mathbb{R}_+}\left(e^{\text{Re } \psi_w \cdot z} - 1 - \text{Re } \psi_w \cdot z\right)\nu(\mathrm{d}z) \right]$$

on $[0, T]$. Since $|\cos(x) - 1| = 1 - \cos(x) \leq x^2/2$, $x \in \mathbb{R}$, by (A.10) we have

$$\left| \int_{\mathbb{R}_+} e^{\text{Re } \widetilde{\psi_w} \cdot z}\left(\cos\left(\text{Im } \widetilde{\psi_w} \cdot z\right) - 1\right)\nu(\mathrm{d}z) \right| \leq \frac{1}{2}\left(\int_{\mathbb{R}_+} |z|^2 \nu(\mathrm{d}z)\right)\left|\text{Im } \widetilde{\psi_w}\right|^2$$

$$\leq \left(\int_{\mathbb{R}_+} |z|^2 \nu(\mathrm{d}z)\right)\left(u^2 + \Lambda^2|\text{Im } w|^2\right), \quad \text{on } [0, T]. \qquad \text{(A.11)}$$

Moreover, notice that by (A.10), since $|\rho| \leq 1$

$$\frac{1}{2}\left| |\text{Im } w|^2 + c \, |\text{Im } \psi_w|^2 + 2\rho \sqrt{c} \, \text{Im } w \, \text{Im } \psi_w \right| \leq \frac{1}{2}\left(|\text{Im } w| + \sqrt{c}|\text{Im } \psi_w| \right)^2 \leq |\text{Im } w|^2 + cu^2. \qquad \text{(A.12)}$$

These facts coupled with (A.4) suggest to consider the linear equation

$$l = K * \left[\frac{1}{2}\left(|\text{Re } w|^2 - \text{Re } w - 2|\text{Im } w|^2 \right) + \int_{\mathbb{R}_+}\left(e^{-\Lambda \text{Re } wz} - 1 - \text{Re } w\left(e^{-\Lambda z} - 1\right)\right)\nu(\mathrm{d}z) - c \, u^2 \right.$$

$$\left. - \left(\int_{\mathbb{R}_+} |z|^2 \nu(\mathrm{d}z)\right)\left(u^2 + \Lambda^2|\text{Im } w|^2\right) + \left(b + \rho \sqrt{c} \, \text{Re } w \right)l \right], \qquad \text{(A.13)}$$

which has a unique, continuous, nonpositive solution l defined on the whole $\mathbb{R}_+$. At this point, observe that the difference $\operatorname{Re}\psi_w - l$ satisfies the linear equation

$$
\chi = K * \left[\left(b + \rho\sqrt{c}\operatorname{Re}w \right)\chi + \frac{c}{2}|\operatorname{Re}\psi_w|^2 + \left(|\operatorname{Im}w|^2 + cu^2 - \frac{1}{2}\left(|\operatorname{Im}w|^2 + c|\operatorname{Im}\psi_w|^2 + 2\rho\sqrt{c}\operatorname{Im}w\operatorname{Im}\psi_w \right) \right) \right.
$$

$$
+ \int_{\mathbb{R}_+} \left(e^{\operatorname{Re}\psi_w \cdot z} - 1 - \operatorname{Re}\psi_w \cdot z \right)\nu(dz) + \int_{\mathbb{R}_+} \left(e^{\operatorname{Re}\psi_w \cdot z} - 1 \right)\left(e^{-\Lambda\operatorname{Re}wz} - 1 \right)\nu(dz)
$$

$$
+ \left(\left(\int_{\mathbb{R}_+} |z|^2\nu(dz) \right)\left(u^2 + \Lambda^2|\operatorname{Im}w|^2 \right) - \left| \int_{\mathbb{R}_+} e^{\operatorname{Re}\widetilde{\psi}_w \cdot z}\left(\cos\left(\operatorname{Im}\widetilde{\psi}_w \cdot z \right) - 1 \right)\nu(dz) \right| \right) \right].
$$

It admits a unique, continuous solution on $[0, T]$ which is nonnegative by (A.11), (A.12) and the fact that $e^x - 1 - x \geq 0$, $x \in \mathbb{R}$. Since $T \in (0, T_{\max})$ was chosen arbitrarily, we infer that

$$
l(t) \leq \operatorname{Re}\psi_w(t) \leq 0 \quad \text{and} \quad |\operatorname{Im}\psi_w(t)| \leq u(t), \quad 0 \leq t < T_{\max}.
$$

Recalling that l and u are continuous on $\mathbb{R}_+$, and in particular bounded on every compact interval, we conclude that $T_{\max} = \infty$, as desired.

Step III. Consider two global solutions ψ_w, ψ'_w of (7), and let $\delta = \psi_w - \psi'_w$ and $\tilde{\delta} = \psi'_w \vee \psi_w$. Then, for every $t \geq 0$,

$$
\delta(t) = \int_0^t K(t-s)\left[\left(b + \rho\sqrt{c}w + \frac{c}{2}\left(\psi_w + \psi'_w \right)(s) + \int_{\mathbb{R}_+} z\left(e^{(-\Lambda w + \tilde{\delta}(s))z} - 1 \right)\nu(dz) \right)\delta(s) \right.
$$

$$
\left. + \int_{\mathbb{R}_+} e^{(-\Lambda w + \tilde{\delta}(s))z}\left(e^{(\psi_w - \delta)(s)z} - e^{(\psi'_w - \delta)(s)z} - \delta(s)z \right)\nu(dz) \right]ds. \tag{A.14}
$$

We introduce the function $k_w : \mathbb{C}_- \times \mathbb{C}_- \to \mathbb{C}$ defined for $(u, v) \in \mathbb{C}_- \times \mathbb{C}_-$ by

$$
k_w(u, v) = \begin{cases} \frac{1}{v-u}\int_{\mathbb{R}_+} e^{(-\Lambda w + u\vee v)z}\left(e^{(v - u\vee v)z} - e^{(u - u\vee v)z} - (v - u)z \right)\nu(dz), & u \neq v \\ 0, & \text{otherwise} \end{cases}. \tag{A.15}
$$

We claim that k_w is continuous on its domain. This is a consequence of an application of the mean value theorem to the functions $f_z(u) = e^{uz} - uz$, $u \in \mathbb{C}_-$, with the parameter $z \in \mathbb{R}_+$. Indeed, using the inequality $|1 - \cos x| \leq x^2$, $x \in \mathbb{R}$,

$$
|f_z(v) - f_z(u)| \leq z \sup_{\xi\in[u,v]} \left| e^{\xi z} - 1 \right| |v - u|
$$

$$
\leq z \sup_{\xi\in[u,v]} \left(\left| e^{\operatorname{Re}\xi \cdot z} - 1 \right| + \sqrt{2}e^{\frac{1}{2}\operatorname{Re}\xi \cdot z}(1 - \cos(\operatorname{Im}\xi \cdot z))^{\frac{1}{2}} \right) |v - u|
$$

$$
\leq z\left(\left(1 - e^{(\operatorname{Re}u \wedge \operatorname{Re}v)z} \right) + \sqrt{2}(|\operatorname{Im}u| \vee |\operatorname{Im}v|)|z| \right)|v - u|, \quad u, v \in \mathbb{C}_-, z \in \mathbb{R}_+. \tag{A.16}
$$

Consequently, the continuity of k_w follows from

$$
|f_z(v - u \vee v) - f_z(u - u \vee v)| \leq |z|^2\left(1 + \sqrt{2} \right)|v - u|^2, \quad u, v \in \mathbb{C}_-, z \in \mathbb{R}_+. \tag{A.17}
$$

Coming back to (A.14) we have (on $\mathbb{R}_+$)

$$\delta = K * \left[\left(b + \rho\sqrt{c}\, w + \frac{c}{2}(\psi_w + \psi'_w)(\cdot) + \int_{\mathbb{R}_+} z\left(e^{(-\Lambda w + \tilde{\delta}(\cdot))z} - 1\right)\nu(\mathrm{d}z) + k_w\left(\psi'_w(\cdot), \psi_w(\cdot)\right)\right)\delta \right],$$
(A.18)

which is a linear equation admitting the zero function as its unique solution. Hence $\psi'_w = \psi_w$ on $\mathbb{R}_+$, completing the proof of this step.

The fact that $\psi_{\mathrm{Re}\,w}$ is $\mathbb{R}_-$− valued follows from (A.10), because in this case $u \equiv 0$. This concludes the proof of the statement in *3.1*.

(ii) From (A.1) and (A.2) we deduce that $\mathrm{Re}\;\mathcal{R}(w, v) \leq \mathcal{R}(\mathrm{Re}\,w,\;\mathrm{Re}\,v)$ for every $v \in \mathbb{C}_-$. Taking the real part in (7) and recalling that – under Hypothesis 2.1 – the kernel K is nonnegative on $(0, \infty)$ we obtain

$$\mathrm{Re}\,\psi_w(t) \leq \int_0^t K(t-s)\mathcal{R}(\mathrm{Re}\,w,\;\mathrm{Re}\,\psi_w(s))\mathrm{d}s, \quad t \geq 0.$$

We can then introduce a nonnegative function $\widetilde{\gamma} : \mathbb{R}_+ \to \mathbb{R}_+$ defined by the relation

$$\mathrm{Re}\,\psi_w(t) = -\widetilde{\gamma}(t) + \int_0^t K(t-s)\mathcal{R}(\mathrm{Re}\,w,\;\mathrm{Re}\,\psi_w(s))\mathrm{d}s, \quad t \geq 0.$$
(A.19)

Using (7), one can rewrite $\widetilde{\gamma}$ as

$$\widetilde{\gamma}(t) = \int_0^t K(t-s)(\mathcal{R}(\mathrm{Re}\,w,\;\mathrm{Re}\,\psi_w(s)) - \mathrm{Re}\;\mathcal{R}(w, \psi_w(s)))\mathrm{d}s, \quad t \geq 0.$$

Thus $\widetilde{\gamma} \in \mathcal{G}_K$ by (Abi Jaber and El Euch, 2019b, Remark B.6). At this point we subtract (A.19) from (7) (with $\mathrm{Re}\,w$ instead of w) to deduce that $\delta = \psi_{\mathrm{Re}\,w} - \mathrm{Re}\,\psi_w$ satisfies

$$\delta(t) = \widetilde{\gamma}(t) + \int_0^t K(t-s)(\mathcal{R}(\mathrm{Re}\,w, \psi_{\mathrm{Re}\,w}(s)) - \mathcal{R}(\mathrm{Re}\,w,\;\mathrm{Re}\,\psi_w(s)))\mathrm{d}s, \quad t \geq 0.$$
(A.20)

If we denote by $\tilde{\delta} = \mathrm{Re}\,\psi_w \vee \psi_{\mathrm{Re}\,w}$, we then need to study (on $\mathbb{R}_+$)

$$\mathcal{R}(\mathrm{Re}\,w, \psi_{\mathrm{Re}\,w}) - \mathcal{R}(\mathrm{Re}\,w,\;\mathrm{Re}\,\psi_w) = \left(b + \rho\sqrt{c}\,\mathrm{Re}\,w + \frac{c}{2}(\mathrm{Re}\,\psi_w + \psi_{\mathrm{Re}\,w}) + \int_{\mathbb{R}_+} z\left(e^{(-\Lambda\mathrm{Re}\,w+\delta)z} - 1\right)\nu(\mathrm{d}z)\right)\delta$$

$$+ \int_{\mathbb{R}_+} e^{(-\Lambda\mathrm{Re}\,w+\delta)z}\left(e^{(\psi_{\mathrm{Re}\,w}-\tilde{\delta})z} - e^{(\mathrm{Re}\,\psi_w-\tilde{\delta})z} - \delta z\right)\nu(\mathrm{d}z)$$

$$=: (w_1(\cdot) + k_{\mathrm{Re}\,w}(\mathrm{Re}\,\psi_w(\cdot), \psi_{\mathrm{Re}\,w}(\cdot)))\delta,$$

with $k_{\mathrm{Re}\,w}$ as in (A.15). Going back to (A.20),

$$\delta(t) = \widetilde{\gamma}(t) + \int_0^t K(t-s)(w_1(s) + k_{\mathrm{Re}\,w}(\mathrm{Re}\,\psi_w(s), \psi_{\mathrm{Re}\,w}(s)))\delta(s)\mathrm{d}s, \quad t \geq 0.$$

We can now apply (Abi Jaber and El Euch, 2019b, Theorem C.1) in order to conclude that $\delta \geq 0$ on $\mathbb{R}_+$. This yields (8) and concludes the proof of *3.1*. □

APPENDIX B: PROOF OF PROPOSITION 3.5

This section is devoted to the proof of Proposition 3.5, a result which allows to price options on the underlying asset S with maturity $T > 0$.

Proof. Let us define the function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(m) = \mathbb{E}\left[e^{X_T} - \left(e^{X_T} - e^m\right)^+\right]e^{-\frac{1}{2}m} = \mathbb{E}\left[e^{X_T}1_{\{X_T \le m\}} + e^m 1_{\{m < X_T\}}\right]e^{-\frac{1}{2}m}, \quad m \in \mathbb{R}. \quad (B.1)$$

Denote by μ_T the probability distribution of X_T on $\mathbb{R}$ and note that $f \in L^1(\mathbb{R})$, because, thanks to Tonelli's theorem,

$$\int_{\mathbb{R}} e^{-\frac{1}{2}m}\left[\int_{\mathbb{R}} \left(e^x 1_{\{x \le m\}} + e^m 1_{\{m < x\}}\right)\mu_T(dx)\right]dm = 4\int_{\mathbb{R}} e^{\frac{1}{2}x}\mu_T(dx) = 4\mathbb{E}\left[e^{\frac{1}{2}X_T}\right] < \infty. \quad (B.2)$$

Therefore we can compute the Fourier transform of f as follows

$$\hat{f}(\lambda) = \int_{\mathbb{R}} e^{\left(-\frac{1}{2}+i\lambda\right)m}\left[\int_{\mathbb{R}} \left(e^x 1_{\{x \le m\}} + e^m 1_{\{m < x\}}\right)\mu_T(dx)\right]dm$$

$$= \int_{\mathbb{R}} \left[e^x \int_x^\infty e^{\left(-\frac{1}{2}+i\lambda\right)m}dm + \int_{-\infty}^x e^{\left(\frac{1}{2}+i\lambda\right)m}dm\right]\mu_T(dx) = \frac{1}{\frac{1}{4}+\lambda^2}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right), \quad \lambda \in \mathbb{R},$$

where in the second equality we are allowed to use Fubini's theorem by (B.2).

Since $|\Psi^{X_T}(\frac{1}{2}+i\lambda)| \le \mathbb{E}[e^{\frac{1}{2}X_T}] < \infty$ and, by dominated convergence, f in continuous on $\mathbb{R}$, we invoke the Fourier inversion theorem, see for instance (Rudin, 1987, Theorem 9.11), to obtain

$$f(m) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-im\lambda}\frac{1}{\frac{1}{4}+\lambda^2}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)d\lambda, \quad m \in \mathbb{R}. \quad (B.3)$$

Combining (B.1) and (B.3) and recalling Corollary 3.4 we deduce that

$$\mathbb{E}\left[\left(e^{X_T} - e^m\right)^+\right] = 1 - \frac{1}{2\pi}\int_{\mathbb{R}} e^{\left(\frac{1}{2}-i\lambda\right)m}\frac{1}{\frac{1}{4}+\lambda^2}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)d\lambda, \quad m \in \mathbb{R}. \quad (B.4)$$

Now, for every $k \in \mathbb{R}$, we can determine the price $C_S(k, T)$ of a call option written on S with log strike k and maturity T. Indeed, taking $m = k - \log(S_0)$ in (B.4) we have

$$C_S(k, T) = \mathbb{E}\left[\left(S_T - e^k\right)^+\right] = S_0 - \frac{1}{2\pi}\sqrt{S_0 e^k}\int_{\mathbb{R}} e^{i\lambda(\log(S_0)-k)}\frac{1}{\frac{1}{4}+\lambda^2}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)d\lambda$$

$$= S_0 - \frac{1}{\pi}\sqrt{S_0 e^k}\int_{\mathbb{R}_+} \mathrm{Re}\left[e^{i\lambda(\log(S_0)-k)}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)\right]\frac{1}{\frac{1}{4}+\lambda^2}d\lambda,$$

which coincides with (15). The expression (16) for the price $P_S(k, T)$ of a put option with the same underlying, log strike and maturity as before, follows from (15), Corollary 3.4, and the put-call parity formula. This completes the proof. □

APPENDIX C: PROOF OF THEOREM 5.1

This section is devoted to the proof of Theorem 5.1, a result providing estimates for the multi-factor approximation of the Riccati-Volterra equations appearing in the Fourier-Laplace transform of the log returns and VIX^2.

Proof. Fix $T > 0$. We first prove Point 5.1. Take $w \in \mathbb{C}$ such that $\mathrm{Re}\, w \in [0,1]$ and $n \in \mathbb{N}$, and observe that $|\psi_{w,n}| \le \overline{l_{1,n}} + \overline{u_{1,n}} - l_n + (\widetilde{\Lambda} + \frac{\rho^+}{\sqrt{c}})|\mathrm{Im}\, w|$ on $\mathbb{R}_+$. Here $\widetilde{\Lambda} = \max\{\rho^- c^{-1/2}, \Lambda\}$ and $\overline{l_{1,n}}$ [resp., $\overline{u_{1,n}}$, l_n] is the unique, continuous solution of (A.8) [resp., (A.9), (A.13)] in Appendix A with K_n instead of K. (Abi Jaber and El Euch, 2019b, Corollary C.4) guarantees the existence of a positive constant $C_1 = C_1(\rho, b, c, \Lambda, \nu)$ such that

$$\overline{l_{1,n}}(t) + \overline{u_{1,n}}(t) + \left(\widetilde{\Lambda} + \frac{\rho^+}{\sqrt{c}}\right)|\mathrm{Im}\, w| \le C_1\left(1 + \int_0^T \left|E_{b+\rho^+\sqrt{c},n}(s)\right|ds\right)|\mathrm{Im}\, w|, \quad t \in [0,T].$$

Then, recalling the hypothesis of boundedness for $(\int_0^T |E_{b+\rho^+\sqrt{c},n}(s)|ds)_n$ and using (A.13), another application of (Abi Jaber and El Euch, 2019b, Corollary C.4) provides the existence of a constant $C_2 = C_2(\rho, b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T) > 0$ such that $|l_n(t)| \le C_2(1 + |\mathrm{Im}\, w|^2)$, $t \in [0,T]$. This implies, given that $n \in \mathbb{N}$ is arbitrary, that

$$\sup_{n\in\mathbb{N}} \sup_{t\in[0,T]} |\psi_{w,n}(t)| \le C_3\left(1 + |\mathrm{Im}\, w|^2\right), \quad \text{for some } C_3 = C_3(\rho, b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T) > 0. \quad (C.1)$$

Since the same argument works for ψ_w, without loss of generality, we assume that the upper bound in (C.1) holds also for ψ_w. Now, from (7) and (36) we have (on $\mathbb{R}_+$)

$$\psi_w - \psi_{w,n} = (K - K_n) * \mathcal{R}(w, \psi_{w,n}(\cdot)) + K * (\mathcal{R}(w, \psi_w(\cdot)) - \mathcal{R}(w, \psi_{w,n}(\cdot))), \quad n \in \mathbb{N}.$$

For every $v \in \mathbb{C}_-$, recalling the inequality $e^x - 1 - x \le x^2/2$, $x \le 0$, and thanks to the computations in Appendix A (see (A.16))

$$\left| \int_{\mathbb{R}_+} \left[e^{(v-\Lambda w)z} - w\left(e^{-\Lambda z} - 1\right) - 1 - vz\right]\nu(dz) \right|$$

$$\le 4\sqrt{2}\left[\frac{\Lambda^2}{2}(1 + |\mathrm{Im}\, w|) + |v|^2 + \Lambda^2\left(1 + |\mathrm{Im}\, w|^2\right)\right]\int_{\mathbb{R}_+} |z|^2 \nu(dz).$$

Then by (C.1) and (6) we deduce that there exists a constant $C_4 = C_4(\rho, b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T) > 0$ such that

$$\sup_{t\in[0,T]} \left|((K - K_n) * (\mathcal{R}(w, \psi_{w,n}(\cdot))))(t)\right| \le C_4\left(1 + |\mathrm{Im}\, w|^4\right)\int_0^T |K_n(s) - K(s)|\, ds, \quad n \in \mathbb{N}.$$
$$(C.2)$$

In what follows, we denote by $h_n = (K - K_n) * \mathcal{R}(w, \psi_{w,n}(\cdot))$, that is, the function that we have just bounded. Next, computations analogous to those carried out to obtain the Volterra equation (A.18) in Appendix A, allow us to write (on $\mathbb{R}_+$)

$$\mathcal{R}(w, \psi_w) - \mathcal{R}(w, \psi_{w,n}) = \left(b + \rho\sqrt{c}\, w + \frac{c}{2}(\psi_w + \psi_{w,n}) + \int_{\mathbb{R}_+} z\left(e^{(-\Lambda w + \psi_{w,n}\vee\psi_w)z} - 1\right)\nu(dz)\right.$$

$$+ k_w\big(\psi_{w,n}, \psi_w\big)\big)\big(\psi_w - \psi_{w,n}\big),$$

where k_w is the continuous function in (A.15). Therefore, since $|k_w(u,v)| \leq (1 + \sqrt{2})(\int_{\mathbb{R}_+} |z|^2 \nu(dz))|v - u|$ for every $u, v \in \mathbb{C}_-$ (see (A.17)) and recalling (C.1)-(C.2), an application of (Abi Jaber and El Euch, 2019b, Corollary C.4) yields

$$\sup_{t \in [0,T]} |\psi_w(t) - \psi_{w,n}(t) - h_n(t)| \leq C_5\Big(1 + |\mathrm{Im}\, w|^6\Big) \frac{\int_0^T E_{b^+ + \rho^+ \sqrt{c} + c_\nu C_3\big(1 + |\mathrm{Im}\, w|^2\big), K}(s)ds}{\int_0^T \big|E_{b^+ + \rho^+ \sqrt{c}, K}(s)\big|ds}$$

$$\times \int_0^T |K_n(s) - K(s)|ds, \quad n \in \mathbb{N}. \quad (C.3)$$

for some $C_5 = C_5(\rho, b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T) > 0$ and where $c_\nu = 2(1 + \sqrt{2})(\int_{\mathbb{R}_+} |z|^2 \nu(dz))$. Notice that by (Gripenberg et al., 1990, Proposition 8.1, Chapter 9) and Hypothesis 2.1, $E_{b^+ + \rho^+ \sqrt{c} + c_\nu C_3(1 + |\mathrm{Im}\, w|^2), K} \geq 0$. Consequently, thanks to (Abi Jaber and El Euch, 2019b, Theorem C.1, Remark B.6), $E_{b^+ + \rho^+ \sqrt{c}, K} \leq E_{b^+ + \rho^+ \sqrt{c} + c_\nu C_3(1 + |\mathrm{Im}\, w|^2), K}$ a.e. in $\mathbb{R}_+$. Hence the ratio in (C.3) is greater or equal to 1. Combining (C.3) with (C.2) yields (38).

In order to prove the final remark about the independence of the constant C in (38) with respect to $\mathbf{m}$ and $\mathbf{x}$, note that in the previous argument such a dependence is only due to $\widetilde{C}$, the positive constant given by the hypothesis controlling the sequence $(\int_0^T |E_{b + \rho^+ \sqrt{c}, n}(s)|ds)_n$. When $b < 0$, the kernels $-bK_n$ inherit the property of complete monotonicity from K_n. If in addition $\rho < 0$, we can use (Gripenberg et al., 1990, Theorem 3.1, Chapter 5) to infer that $\int_0^T |E_{b + \rho^+ \sqrt{c}, n}(s)|ds = \int_0^T |E_{b,n}(s)|ds \leq |b|^{-1}$ for every $n \in \mathbb{N}$, and $\int_0^T |E_{b^+ + \rho^+ \sqrt{c}, K}(s)|ds = \|K\|_{L^1([0,T])}$. In particular, in this case C depends on T only via the L^1- norm of K in $[0, T]$ (see (C.1)-(C.3)).

The proof of Point 5.1 follows by an analogous argument. In this case we use the estimates in (Bondi et al., 2024, Appendix B.1) and the fact that $\int_0^\delta K_n(s)ds \leq \int_0^{T \vee \delta} E_{b^+, n}(s)ds \leq \widetilde{C}, n \in \mathbb{N}$. We also combine (Abi Jaber and El Euch, 2019b, Corollary C.4), the comparison result for linear Volterra equations in (Beesack, 1969, Theorem 2), and the inequality

$$\int_0^\delta h(s)K_n(s + t)ds \leq \int_0^\delta h(s)K_n(s)ds, \quad t \geq 0,$$

which holds also for K by Hypothesis 2.1. $\qquad \square$