RESEARCH ARTICLE

Affine Heston model style with self-exciting jumps and long memory

Charles Guy Njike Leunga[1] · **Donatien Hainaut[1]**

Abstract

Classic diffusion processes fail to explain asset return volatility. Many empirical findings on asset return time series, such as heavy tails, skewness and volatility clustering, suggest decomposing the volatility of an asset's return into two components, one caused by a Brownian motion and another by a jump process. We analyze the sensitivity of European call options to memory and self-excitation parameters, underlying price, volatility and jump risks. We expand Heston's stochastic volatility model by adding to the instantaneous asset prices, a jump component driven by a Hawkes process with a kernel function or memory kernel that is a Fourier transform of a probability measure. This kernel function defines the memory of the asset price process. For instance, if it is fast decreasing, the contagion effect between asset price jumps is limited in time. Otherwise, the processes remember the history of asset price jumps for a long period. To investigate the impact of different rates of decay or types of memory, we consider four probability measures: Laplace, Gaussian, Logistic and Cauchy. Unlike Hawkes processes with exponential kernels, the Markov property is lost but stationarity is preserved; this ensures that the unconditional expected arrival rate of the jump does not explode. In the absence of the Markov property, we use the Fourier transform representation to derive a closed form expression of a European call option price based on characteristic functions. A numerical illustration shows that our extension of the Heston model achieves a better fit of the Euro Stoxx 50 option data than the standard version.

Keywords Hawkes process · Memory kernel · Stochastic volatility · Option pricing

JEL Classification C02 · C32 · C52 · C63 · G12

✉ Charles Guy Njike Leunga
   charles.njikeleunga@uclouvain.be

   Donatien Hainaut
   donatien.hainaut@uclouvain.be

1   Louvain Institute of Data Analysis and Modeling in Economics and Statistics, UCLouvain, Voie du
    Roman Pays 20, Louvain-la-Neuve 1348, Belgium

✑ Springer

1 Introduction

Over the past decade, a number of stochastic processes designed for option pricing have been introduced in the financial literature since the seminal work of Black and Scholes (1973). A common way to communicate option prices consists in computing the implied volatility (IV) resulting from the inversion of the Black and Scholes option price formula. IVs measure the uncertainty about how low or high an underlying asset might fall or rise over a given period. Unlike the historical asset's volatility, IVs are somehow expectations for future volatility. As a result, IVs play an important role in financial markets and option pricing models aim to replicate at least the main empirical features of asset time series. For a wide set of financial assets, Cont (2001) stated five stylized statistical facts about the asset's volatility. First, high volatility events tend to cluster in time. Second, there is a long memory effect in volatility which occurs when the effects of volatility shocks decay slowly. Third, the asset's volatility is negatively correlated with the asset's return. Fourth, trading volume is correlated with the asset's volatility. Fifth, there is an asymmetry in time scales of the volatility which is also known as the time reversal asymmetry of financial time series data. This has been emphasized by Zumbach and Lynch (2001) and called the "Zumbach effect" by Blanc et al. (2016). Beside these empirical findings in asset's volatility, common features of the IV surfaces are the volatility term structure effects, the at-the-money skew and the smile (see e.g. Aït-Sahalia and Lo 1998; Aït-Sahalia et al. 2021). Designing an option pricing model that includes most of these characteristics is a challenging task.

In quest of the suitable option pricing model, this article extends previous works on affine stochastic volatility (SV) models stemming from the Black and Scholes framework. While simple and easy to use, the Black and Scholes model relies on strong assumptions regarding the financial behavior observed throughout market data. Merton (1976) and Bates (1991) relax the assumption of the Black and Scholes (1973) that the asset's return is normally distributed by inserting a jump component driven by a Poisson process. Later, Heston (1993) proposes a closed form solution for an option on an asset with a SV model, which allows relaxing the assumption of constant variance of the Black and Scholes model. He also introduces a correlation between the variance and the asset price in order to replicate the volatility smile and to explain the skewness of the asset's return. The Heston's framework is extended in several directions. Among them, we can cite stochastic volatility jump (SVJ) models, introduced by Bakshi et al. (1997), Duffie et al. (2000), Bates (2000) and Eraker (2004) to mention a few pioneering works. The common point of these models, known as affine stochastic volatility models is the heteroskedastic property of the instantaneous change in variance through the square root term of the variance in the diffusion volatility. However, in the literature, other authors like Christoffersen et al. (2006) and Ignatieva et al. (2009) have studied non affine stochastic volatility models in which the instantaneous change in variance is proportional to any power of the variance in the diffusion volatility. They conclude that the choice between affine and non affine stochastic volatility models corresponds to find the right balance between fit flexibility and estimation properties. This motivate us to employ an affine SV model style for its tractability and economic interpretation documented in the literature (see e.g. Heston 1993; Bakshi et al. 1997; Duffie et al. 2000).

Notice that in view of expanding affine SVJ models, some authors like Gatheral (2012) investigated stochastic volatility models with jumps in asset prices and in the variance, denoted by SVJJ model. Although, this setting enables to maintain the clustering property of SV models, SVJJ models are less parsimonious than SVJ models, and are harder to fit to observed option prices. Another popular family of SV models considers that the log volatility behaves as a fractional Brownian motion with a small Hurst exponent. Among seminal contributions about this family of SV models we list the fractional SV model of Comte et al. (2010), and the rough volatility model of Gatheral et al. (2014) that in contrast to fractional SV, has a Hurst exponent of order less than 0.5. Even if these models are remarkably consistent with major features of financial time series, as stated by Blanc et al. (2016), they fail to induce a tail of distribution in the return fat enough even with fluctuating volatility since the asset's return in these models is conditionally Gaussian distributed. Moreover, these models are unable to explain in a realistic way how they generate fat tails, long memory in volatility and volatility clustering.

Nevertheless, most of affine SV models do not allow for the Zumbach effect. This motivates, authors like Zumbach et al. (2013), Jaisson and Rosenbaum (2015) and Blanc et al. (2016) to modify SV models in order to integrate this Zumbach effect. The two last works are based on quadratic Hawkes processes (HPs) which extend the self-exciting mechanism of HPs introduced by Hawkes (1971a). In particular, Blanc et al. (2016) reformulate the general HP in a way that intra-day price changes, have an intensity proportional to powers of past price variations. Motivated by the empirical results, authors truncate this intensity to the second exponent and ignore the first exponent term. However, recent work of Hainaut and Moraux (2018a), Hainaut and Moraux (2018b) and Njike and Hainaut (2022) have shown the role of linear HP when it comes to replicate features of asset's return time series. In combination with the self-excitation property of the linear HP, Hainaut (2021) has designed a non-markov linear HP for claims process that has memory. The author has generated this memory effect by considering a linear HP with a kernel function that admits a Fourier's transform representation. For this work, it was of interest to explore the benefit of such a setup in an option pricing model. In doing so, the linear HP will increase the probability to observe a jump of the asset price and therefore the future volatility after a jump occurrence, whilst the memory kernel of the linear HP will enable us to obtain a longer effect of past prices into future prices.

This works contributes to the financial literature in three ways. First, we introduce an option pricing model where the underlying asset price dynamics is self-excited and allows for a longer effect of past prices through a linear HP. Second, although our model is non Markov, we derive a closed-form expression for European call option prices. For hedging applications, Greeks of the three sources of risk: price risk, volatility risk and jump risk are obtained straightforwardly by taking derivatives of our formula with respect to the asset price, the variance or the arrival rate of jump. Third, we provide a calibration procedure of our model to real market option data.

The rest of the article proceeds as follows. Section 2 presents the affine Heston model style with self-exciting jumps and long memory. This section focuses on distributional properties of the arrival rate of jumps and the log-return through an infinite dimensional Markov representation of our asset price model and a discretisation scheme. Section 3

develops the option pricing model under our setting. Section 4 provides a descriptions of the Euro Stoxx 50 option data. Therein, we present our estimation procedure, discuss the estimated parameter and assess the sensitivity of option prices to memory and self-excitation parameters, underlying price, volatility and jump risks. Concluding comments are developed in Sect. 5.

2 Model specification

We define all the processes on a probability space Ω endowed with a risk neutral measure $\mathbb{Q}$ and a filtration $(\mathcal{F}_t)_{t \geq 0}$. We exclusively focus in the option pricing, and assumes that the pricing measure can be retrieve from European option prices. For sake of simplicity we assume a constant interest rate r.

Hereafter we consider a SVJ model with self-excited jump in return driven by a Hawkes process with memory kernel:

$$\frac{dS_t}{S_t} = r\,dt + \sqrt{V_t}\left(\rho dW_t^1 + \sqrt{1 - \rho^2}dW_t^2\right) \tag{1}$$

$$+d\left(\sum_{l=1}^{N_t}\left(e^{J_l} - 1\right)\right) - \lambda_t \mathbb{E}\left(e^J - 1\right)dt$$

$$dV_t = (\theta_V - k_V V_t)\,dt + \sigma_V\sqrt{V_t}dW_t^1$$

where k_V, $\frac{\theta_V}{k_V}$ and σ_V are respectively the speed of adjustment, long-run and variation coefficients of the diffusion volatility of V_t. W_t^1 and W_t^2 are two independent standard Brownian motions. ρ is the correlation between diffusion volatility of stock prices and variances. The variance process is the square root process introduced by Cox et al. (1985) and referred to as CIR. To ensure that zero is an unattainable for the diffusion volatility V_t, we assume the Feller condition $2\theta_V > \sigma_V^2$ holds. Let recall that affine stochastic model implies that the instantaneous change in variance is heteroskedastic via the $\sqrt{V_t}$ term in the diffusion volatility. We introduce memory and contagion in the dynamic of stock prices by considering an arrival intensity λ_t driven by the following equation:

$$\lambda_t|\mathcal{F}_t = \lim_{h \to 0}\frac{\mathbb{E}(N_{t+h}|\mathcal{F}_t) - N_t}{h} = \alpha + (\lambda_0 - \alpha)g(t) + \eta\int_0^t g(t - u)\,dL_u, \tag{2}$$

with α the constant reversion level, λ_0 the initial intensity at time $t = 0$, $L_t = \sum_{l=1}^{N_t}|J_l|$ the sum of the absolute values of the jump sizes up to time t. The intensity jumps by $\eta|J_l|$ when the lth shock in the asset's returns dynamics occurs. Therefore, the occurrence of an event increases the probability of a further shock. The jump sizes J are double exponentially identically and independently distributed with density $v(z)$ on $\mathbb{R}$:

$$v(z) = p\rho^+ \exp\left(-\rho^+ z\right) 1_{\{z \geq 0\}} - (1-p)\rho^- \exp\left(-\rho^- z\right) 1_{\{z < 0\}},$$

with $\rho^+ > 1$, $\rho^- < 0$ and $p \in]0, 1[$. The condition $\rho^+ > 1$ ensures that the jump size of the asset price has a finite expectation. Using the previous equation, it follows that the moment generating function (mgf) $\psi(., .)$ of the pair $(J, |J|)$ is given by:

$$\psi(z_1, z_2) := \mathbb{E}\left(e^{z_1 J + z_2 |J|}\right) = \frac{p\rho^+}{\rho^+ - (z_1 + z_2)} + \frac{(1-p)\rho^-}{\rho^- - (z_1 - z_2)}.$$

In particular, the expectation of J and $|J|$ are weighted sum of the average sizes of positive and negative shocks:

$$\mathbb{E}(J) = p\frac{1}{\rho^+} + (1-p)\frac{1}{\rho^-}, \tag{3}$$

$$\mathbb{E}(|J|) = p\frac{1}{\rho^+} - (1-p)\frac{1}{\rho^-}. \tag{4}$$

We consider double exponential jumps as in Kou (2002) in order to replicate the asymmetry of the log-return distribution. Note that our SVJ model turns back to the Heston model if the jump size is null.

In Eq. (2), the memory kernel function g, is a positive definite function such that $g(0) = 1$ and $\lim_{t \to \infty} g(t) = 0$. In this setting, the speed of reversion depends upon the decay rate of $g(.)$. The function $g(.)$ is the kernel function defining the memory of the stock prices process. If it is fast decreasing, the contagion effect between stock price jumps is limited in time. Otherwise, the processes remember the history of stock price jumps for a long period. In the present work, we consider four functions $g(.)$ that are Fourier transform of positives measures on $\mathbb{R}$:

- The first function $g_L(.)$ is the Fourier transform of the Laplace or double exponential measure denoted by $v_L(.)$. Let $\beta \in \mathbb{R}^+$ be constant. The symmetric exponential measure $v_L(.)$ defined by the following relation:

$$v_L(du) = \frac{1}{2}\left(\beta \exp(-\beta u) 1_{\{u \geq 0\}} + \beta \exp(\beta u) 1_{\{u < 0\}}\right) du, \tag{5}$$

generates the following memory kernel:

$$g_L(h) = \int_{\mathbb{R}} v_L(u) e^{ihu} du = \frac{\beta^2}{\beta^2 + h^2}. \tag{6}$$

The mean of $v_L(.)$ is $\mu_L = 0$ and its variance is $\sigma_L^2 = \frac{2}{\beta^2}$. $g_L(.)$ is a power decreasing function. Since $\int_0^\infty g_L(u) du = \frac{\pi\beta}{2}$, in this case the pair (N, λ) is a stationary HP if $\beta < \frac{2}{\pi\eta\mathbb{E}(|J|)}$.

- The second function $g_G(.)$ is the Fourier transform of a Gaussian spectral measure $v_G(.)$

$$v_G(\mathrm{d}u) = \sqrt{\frac{\beta}{\pi}} e^{-\beta u^2} \mathrm{d}u. \tag{7}$$

The mean of $v_G(.)$ is $\mu_G = 0$ and its variance is $\sigma_G^2 = \frac{1}{2\beta}$. Given that $v_G(.)$ is symmetric and $\int_{\mathbb{R}^+} e^{-\beta u^2} \cos(hu)\, \mathrm{d}u = \frac{1}{2}\sqrt{\frac{\pi}{\beta}} e^{-\frac{h^2}{4\beta}}$, the memory kernel is equal to

$$g_G(h) = e^{-\frac{h^2}{4\beta}}, \tag{8}$$

where $\beta \in \mathbb{R}^+$. Since $\int_0^\infty g_G(u)\, \mathrm{d}u = \sqrt{\pi\beta}$, in this case the pair (N, λ) is a stationary HP if $\sqrt{\beta} < \frac{1}{\sqrt{\pi}\eta\mathbb{E}(|J|)}$.

- The third function $g_{Log}(.)$ is the Fourier transform of the logistic distribution defined as follows:

$$v_{Log}(\mathrm{d}u) = \frac{e^{-\frac{u}{\beta}}}{\beta\left(1 + e^{-\frac{u}{\beta}}\right)^2} \mathrm{d}u, \tag{9}$$

where $\beta \in \mathbb{R}^+$. The mean of $v_{Log}(.)$ is $\mu_{Log} = 0$ and its variance is $\sigma_{Log}^2 = \frac{\beta^2\pi^2}{3}$. The memory kernel is equal to

$$g_G(h) = \begin{cases} \frac{\pi\beta h}{\sinh(\pi\beta h)} & \text{if } h > 0 \\ 1 & \text{if } h = 0 \end{cases}. \tag{10}$$

Since $\int_0^\infty g_{Log}(u)\, \mathrm{d}u = \frac{\pi}{4\beta}$, in this case the pair (N, λ) is a stationary HP if $\beta > \frac{\pi\eta\mathbb{E}(|J|)}{4}$.

- The last memory kernel $g_C(.)$ is the Fourier transform of a Cauchy spectral measure $v_C(.)$

$$v_C(\mathrm{d}u) = \frac{1}{\pi}\frac{\beta}{\beta^2 + u^2} \mathrm{d}u, \tag{11}$$

where $\beta \in \mathbb{R}^+$ and the moments of $v_C(.)$ are undefined. Given that this measure is symmetric, the memory kernel $g_C(.)$ is an exponentially decreasing function

$$g_C(h) = e^{-\beta|h|}, \tag{12}$$

considered in the following as the exponential kernel because the delays after the arrival of jumps are non-negative. Since $\int_0^\infty g_C(u)\, \mathrm{d}u = \frac{1}{\beta}$, in this case the pair (N, λ) is a stationary HP if $\beta > \eta\mathbb{E}(|J|)$.

Within the literature, the Hawkes process with exponential kernel has an intensity λ_t usually defined as:

$$\lambda_t|\mathcal{F}_t = \lim_{h\to 0}\frac{\mathbb{E}(N_{t+h}|\mathcal{F}_t) - N_t}{h} = \alpha + (\lambda_0 - \alpha)e^{-\beta t} + \eta\int_0^t e^{-\beta(t-u)}\mathrm{d}L_u,$$
(13)

where β is the constant rate of mean reversion. From Hawkes (1971b), the pair (N,λ) is a stationary Hawkes process if the following inequality holds $\eta\mathbb{E}(|J|)\int_0^\infty e^{-\beta u}\mathrm{d}u < 1$. This latter condition is then met by choosing the intensity parameters η and β such as $\eta\mathbb{E}(|J|) < \beta$. This stationarity property of the Hawkes process is required to ensure that the long-term expectation, noted $\lambda_\infty = \lim_{t\to\infty}\mathbb{E}(\lambda_t)$ is finite and equal to $\lambda_\infty = \frac{\beta\alpha}{\beta-\eta\mathbb{E}(|J|)}$.

The first order moment of the HP intensity for the kernel functions in Eqs. (6), (8) and (10) can be obtained by inverse Fourier transform. In fact, from Eq. (2), we use the properties of the expectation and apply the Fourier transform to obtain:

$$\mathcal{F}(\mathbb{E}(\lambda_t \mid \mathcal{F}_0))(\omega) = \alpha\mathcal{F}(1)(\omega) + (\lambda_0 - \alpha)\mathcal{F}(g(t))(\omega)$$
(14)
$$+\eta\mathbb{E}(|J|)\mathcal{F}\left(\int_0^t g(t-u)\mathbb{E}(\lambda_u \mid \mathcal{F}_0)\mathrm{d}u\right)(\omega)$$
$$= 2\alpha\pi\delta(\omega) + (\lambda_0 - \alpha)v_k(\omega) + \eta\mathbb{E}(|J|)v_k(\omega)\mathcal{F}(\mathbb{E}(\lambda_t \mid \mathcal{F}_0)\mathrm{d}u)(\omega)$$

where $\delta(.)$ is the Dirac Delta function. Then, given that

$$\mathcal{F}(\mathbb{E}(\lambda_t \mid \mathcal{F}_0))(\omega) = \frac{2\alpha\pi\delta(\omega) + (\lambda_0 - \alpha)v_k(\omega)}{1 - \eta\mathbb{E}(|J|)v_k(\omega)}$$
(15)
$$= \frac{2\alpha\pi\delta(\omega)}{1 - \eta\mathbb{E}(|J|)v_k(\omega)} + \frac{(\lambda_0 - \alpha)v_k(\omega)}{1 - \eta\mathbb{E}(|J|)v_k(\omega)}$$

it follows that

$$\mathbb{E}(\lambda_t \mid \mathcal{F}_0) = \alpha\mathcal{F}^{-1}\left(\frac{2\pi\delta(\omega)}{1 - \eta\mathbb{E}(|J|)v_k(\omega)}\right)(t)$$
(16)
$$+ (\lambda_0 - \alpha)\mathcal{F}^{-1}\left(\frac{v_k(\omega)}{1 - \eta\mathbb{E}(|J|)v_k(\omega)}\right)(t)$$
$$= \alpha\frac{1}{2\pi}\int_\mathbb{R} e^{-i\omega t}\frac{2\pi\delta(\omega)}{1 - \eta\mathbb{E}(|J|)v_k(\omega)}\mathrm{d}\omega$$
$$+ (\lambda_0 - \alpha)\frac{1}{2\pi}\int_\mathbb{R} e^{-i\omega t}\frac{v_k(\omega)}{1 - \eta\mathbb{E}(|J|)v_k(\omega)}\mathrm{d}\omega$$
$$= \frac{\alpha}{1 - \eta\mathbb{E}(|J|)v_k(0)} + (\lambda_0 - \alpha)\frac{1}{2\pi}\int_\mathbb{R} e^{-i\omega t}\frac{v_k(\omega)}{1 - \eta\mathbb{E}(|J|)v_k(\omega)}\mathrm{d}\omega.$$

The previous stability conditions ensure that this latter integral exists. Notice that there is no closed form expression of this expected value of the HP intensity, except for the HP intensity with exponential kernel. Through either the Laplace transform technique
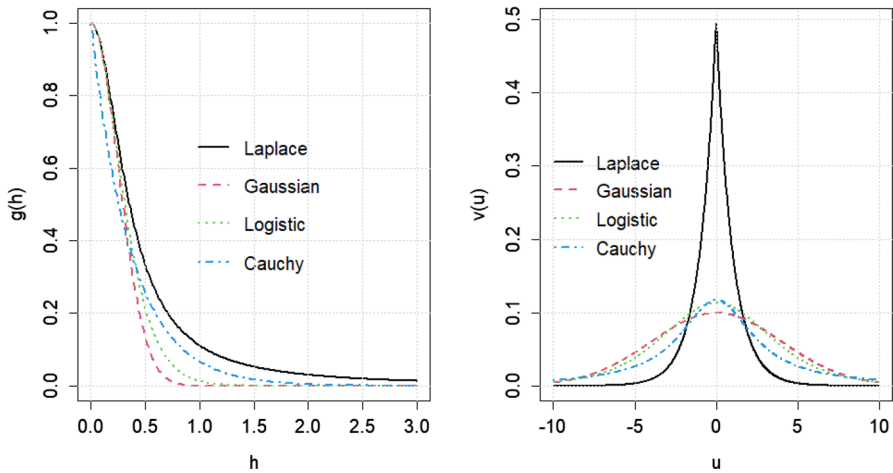
Fig. 1 Comparison of Laplace, Gaussian and Logistic kernels for $\sigma_L^2 = \sigma_G^2 = \sigma_{Log}^2 = 16$. The Cauchy kernel is computed with $\beta_C = 2.71$

or ODE technique of Fonseca and Zaatour (2013), the first order moment of the HP intensity with exponential kernel is given by,

$$\mathbb{E}(\lambda_t \mid \mathcal{F}_0) = \lambda_0 e^{(\eta \mathbb{E}(|J|) - \beta)t} - \frac{\beta\alpha}{\eta\mathbb{E}(|J|) - \beta}\left(1 - e^{(\eta\mathbb{E}(|J|) - \beta)t}\right). \qquad (17)$$

To the best of our knowledge, it is not possible to derive analytically moments of HP with non-exponential kernel listed above using the ordinary differential equation (ODE) technique of Fonseca and Zaatour (2013). As an alternative, we will derive a quasi-analytic formula of the conditional moment generating function of these HP intensities with non-exponential kernel.

To compare shapes of memory kernel we range above memory kernels by degree of decay and consider that their spectral density have the same variance σ^2. It follows that the parameter β of memory kernels $(g_k)_{k \in \{L, G, Log\}}$ is a function of σ^2. As the variance of the Cauchy spectral measure is not defined, we find an β by minimizing the sum of least square between the Log and Cauchy memory kernels. Figure 1 compares these four memory kernels when the variance $\sigma^2 = 16$. The left plot reveals that the Laplace kernel has a slower decay than Gaussian and logistic kernels. Further, in long run the exponential kernel seems to decay more slowly than the Gaussian and logistic kernels, while the opposite situation occurs in the short run. The right graph shows that the decay rate of the kernel increases with the thickness of the distribution tails.

The common point of these four memory kernels is the stability of the Hawkes intensity λ_t under a certain condition. Unlike the exponential kernel, the other kernels can enable the Hawkes intensity λ_t to remember past events for a long time according to their speed of decay. Considering the exponential kernel (12), the pair $(N_t, \lambda_t)_{t \geq 0}$ is a Markov process and the intensity λ_t is solution of the following stochastic differential equation (SDE):

$$d\lambda_t = \beta (\alpha - \lambda_t) \, dt + \eta dL_t. \tag{18}$$

Taking into account kernel functions in Eqs. (6), (8) and (10), the pair $(N_s, \lambda_s)_{s \geq 0}$ is not a Markov process and we have very little information on its behavior conditionally to the filtration $\mathcal{F}_t$ for $t \leq s$. For non-Markov processes we cannot rely on Itô calculus to determine quantity like the mgf of the arrival rate of shocks in the asset's returns dynamics. However, by using the spectral representation of kernel $g_k (.)$ for $k \in \{L, G, Log, C\}$, we reformulate $(N_s, \lambda_s)_{s \geq 0}$ as an infinite dimensional Markov process in the complex plane. To do so, we rewrite $g_k (.)$ in Eq. (2) as a Fourier transform of the measure $v_k (.)$ and change the order of integration:

$$\lambda_t = \alpha + (\lambda_0 - \alpha) \, g_k (t) + \eta \int_0^t \int_{\mathbb{R}} e^{i(t-u)\xi} v_k (d\xi) \, dL_u \tag{19}$$

$$= \alpha + (\lambda_0 - \alpha) \, g_k (t) + \eta \int_{\mathbb{R}} \left(\int_0^t e^{i(t-u)\xi} dL_u \right) v_k (d\xi) .$$

It follows that the process $Y_t^\xi = \int_0^t e^{i(t-u)\xi} dL_u$, defined on the complex plane $\mathbb{C}$ is solution of the following SDE:

$$dY_t^\xi = i\xi Y_t^\xi dt + dL_t, \tag{20}$$

and the SDE of the Hawkes intensity is given by

$$d\lambda_t = (\lambda_0 - \alpha) \frac{dg_k (t)}{dt} dt + \eta \int_{\mathbb{R}} dY_t^\xi v_k (d\xi) . \tag{21}$$

Remark that $\int_{\mathbb{R}} dY_t^\xi v_k (d\xi)$ is real since the measure $v_k (.)$ is symmetric and by rewriting Y_t^ξ as $\sum_{l=1}^{N_t^l} e^{i(t-u_l)} |J_l|$ with u_l the time of the lth shock in the asset's returns dynamics.

To benefit from the reformulation of the process $(N_t, \lambda_t)_{t \geq 0}$ like an infinite dimensional Markov process $\left(N_t, \lambda_t, \left(Y_t^\xi \right)_{\xi \in \mathbb{R}} \right)_{t \geq 0}$, we use the discretisation scheme described below, and passing to the limit we infer the mgf of the arrival rate of shocks in the asset's returns dynamics. The key point of this discretisation scheme consists of an approximation of the measure $v_k (.)$ as a discrete measures with a finite numbers of atoms. For this purpose, we consider a symmetric partition $\mathcal{E}^{(n)} := \{ -\infty < \xi_{-n}^{(n)} < \ldots < \xi_0^{(n)} = 0 < \xi_1^{(n)} < \ldots < \xi_n^{(n)} < \infty \}$, with $\xi_{-n}^{(n)} = -\xi_n^{(n)}$. For each interval $(\xi_l^{(n)}, \xi_{l+1}^{(n)})$, the barycenter is defined as

$$b_{l+1}^{(n)} = \frac{\xi_l^{(n)} + \xi_{l+1}^{(n)}}{2}, \quad \text{for } l \in \{0, \ldots, n-1\}, \tag{22}$$

$$b_l^{(n)} = \frac{\xi_l^{(n)} + \xi_{l+1}^{(n)}}{2}, \quad \text{for } l \in \{-n, \ldots, -1\},$$

while the mass of any atom with respect of the measure $v_k (.)$ is

$$m_{l+1}^{(k,n)} = \int_{\xi_l^{(n)}}^{\xi_{l+1}^{(n)}} v_k (\mathrm{d}u) , \quad \text{for } l \in \{0, \dots, n-1\} , \tag{23}$$

$$m_l^{(k,n)} = \int_{\xi_l^{(n)}}^{\xi_{l+1}^{(n)}} v_k (\mathrm{d}u) , \quad \text{for } l \in \{-n, \dots, -1\} .$$

By construction $\lim_{n \to \infty} \sum_{l=-n}^{n} m_l^{(k,n)} = 1$ and the discrete approximation of the measure $v_k (.)$ given a partition of size n, denoted by $\tilde{v}_k^{(n)} (.)$ is defined as follows

$$\tilde{v}_k^{(n)} (u) = \sum_{l=-n}^{n} m_l^{(k,n)} \delta_{b_l^{(n)}} (u) , \tag{24}$$

where $\delta_{b_l^{(n)}} (u)$ is the Dirac measure located at the barycenter $b_l^{(n)}$. According to Eqs. (22) and (23), and given that the measure $v_k (.)$ is symmetric, for $l \in \{-n, \dots, -n\} \setminus \{0\}$, $m_{-l}^{(k,n)} = m_l^{(k,n)}$ and $b_{-l}^{(n)} = -b_l^{(n)}$. It follows that the discrete measure $\tilde{v}_k^{(n)} (.)$ is also symmetric. We consider that the following assumptions holds for the partition $\mathcal{E}^{(n)}$:

(i) $\xi_{-n}^{(n)} \to -\infty$ and $\xi_n^{(n)} \to \infty$ when $n \to \infty$.
(ii) $\max |\xi_{l+1}^{(n)} - \xi_l^{(n)}| \to 0$ when $n \to \infty$.
(iii) $\mathcal{E}^{(n)} \subset \mathcal{E}^{(n+1)}$.

In this setting, for any integrable function, $f (.)$ with respect to the measure $v_k (.)$, we have that

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} f (u) \tilde{v}_k^{(n)} (\mathrm{d}u) = \int_{-\infty}^{\infty} f (u) v_k (\mathrm{d}u) . \tag{25}$$

To lighten further developments, we choose n and adopt the following notations:

- $Y_t^{(l)} := Y_t^{(b_l^{(n)})}$ for $l \in \{-n, \dots, n\}$ and $Y_t^{(l)}$ is solution of the following SDE

$$\mathrm{d}Y_t^{(l)} = i b_l^{(n)} Y_t^{(l)} \mathrm{d}t + \mathrm{d}L_t^{(n)}, \tag{26}$$

  with $L_t^{(n)} = \sum_{l=1}^{N_t^{(n)}} |J_l|$;
- the approximated intensity in the discrete framework is

$$\lambda_t^{(n)} = \alpha + (\lambda_0 - \alpha) g_k (t) + \eta \sum_{l=-n}^{n} m_l^{(k,n)} Y_t^{(l)} \tag{27}$$

and has the following SDE

$$d\lambda_t^{(n)} = (\lambda_0 - \alpha) \frac{dg_k(t)}{dt} dt + \eta \sum_{l=-n}^{n} m_l^{(k,n)} i b_l^{(n)} Y_t^{(l)} dt + \eta \sum_{l=-n}^{n} m_l^{(k,n)} dL_t^{(n)}.$$

(28)

Proposition 1 *The moment generating function of $\lambda_s^{(n)}$ conditionally to the information at time $t \leq s$ is given by*

$$\mathbb{E}\left(e^{\omega \lambda_s^{(n)}} | \mathcal{F}_t\right) = \exp\left(F(t, s, \omega) + G(t, s, \omega)\lambda_t^{(n)} + \sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} Y_t^{(l)}\right)$$

for all $\omega \in \mathbb{C}_-$ that satisfies the regularity condition $\psi\left(0, \eta\omega \sum_{l=-n}^{n} m_l^{(k,n)}\right) < \infty$. The functions H_l are defined as follows

$$H_l\left(t, s, b_l^{(n)}, \omega\right) = b_l^{(n)} \eta i \int_t^s e^{i b_l^{(n)}(u-t)} G(u, s, \omega) \, du, \quad \forall l \in \{-n, \ldots, n\},$$

F and G are time dependent functions that solve the following ODEs system

$$\frac{\partial F}{\partial t}(t, s, \omega) = -(\lambda_0 - \alpha)\frac{dg_k(t)}{dt}G(t, s, \omega)$$

(29)

$$\frac{\partial G}{\partial t}(t, s, \omega) = 1 - \psi\left(0, \eta G(t, s, \omega)\sum_{l=-n}^{n} m_l^{(k,n)} + \sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)}\right)$$

(30)

with final conditions $F(s, s, \omega) = 0$, $G(s, s, \omega) = \omega$. ψ is the moment generating function of the pair of the random jump sizes J and its absolute value: $\psi(z_1, z_2) := \mathbb{E}\left(e^{z_1 J + z_2 |J|}\right)$. $k \in \{L, G, Log, C\}$ indicates the type of HP memory kernel

Proof In Appendix A. □

Passing to the limit with respect of n and using the frequency derivative property of Fourier transform, we have the following proposition on the conditional mgf of the HP intensity.

Proposition 2 *The moment generating function of λ_s conditionally to the information at time $t \leq s$ is given by*

$$\mathbb{E}\left(e^{\omega \lambda_s} | \mathcal{F}_t\right) = \exp\left(F(t, s, \omega) + G(t, s, \omega)\lambda_t\right)$$
$$\times \exp\left(i\eta \int_t^s G(u, s, \omega) \int_{\mathbb{R}} l e^{il(u-t)} Y_t^l v_k(l) \, dl \, du\right)$$

(31)

for all $\omega \in \mathbb{C}_-$ that satisfies the regularity condition $\psi(0, \eta\omega) < \infty$. The functions F and G solve the following ODE and partial integro-differential equation (PIDE):

$$\frac{\partial F}{\partial t}(t, s, \omega) = -(\lambda_0 - \alpha) \frac{dg_k(t)}{dt} G(t, s, \omega) \tag{32}$$

$$\frac{\partial G}{\partial t}(t, s, \omega) = 1 - \psi\left(0, \eta G(t, s, \omega) + \eta \int_t^s G(u, s, \omega) \frac{dg_k(u-t)}{dt} du\right) \tag{33}$$

with final conditions $F(s, s, \omega) = 0$, $G(s, s, \omega) = \omega$. $k \in \{L, G, Log, C\}$ indicates the type of HP memory kernel.

Proof The approximated intensity $\lambda_t^{(n)}$ defined in Eq. (27), converges almost surely to λ_t since for fixed $w \in \Omega$, the function $z \mapsto Y_t^{(z)}(w)$ is integrable with respect to the measure v_k and for any integrable function, $f(.)$ with respect to the measure $v_k(.)$, we have that

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} f(u) \tilde{v}_k^{(n)}(du) = \int_{-\infty}^{\infty} f(u) v_k(du). \tag{34}$$

It follows that $N_t^{(n)} = \int_0^t \lambda_u^{(n)} du$ converges almost surely to N_t. For $l \in \{-n, \ldots, n\}$, $Y_t^{(l)} = \int_0^t e^{-b_l^{(n)}(t-u)} dL_u^{(n)}$. If the partition $\mathcal{E}^{(n)}$ satisfies the assumptions (i)-(iii), when n tends to ∞, a barycenter $b_l^{(n)}$ tends to be an element of $\mathbb{R}_+$. Hence the sequence $\left(Y_t^{\left(b_l^{(n)}\right)}\right)_{l=-n,\ldots,n}$ tends to $\left(Y_t^{(l)}\right)_{l\in\mathbb{R}}$ almost surely. Next, the dominated convergence provides that

$$\mathbb{E}\left(e^{\omega\lambda_s u}|\mathcal{F}_t\right) = \lim_{n\to\infty} \mathbb{E}\left(e^{\omega\lambda_s^{(n)}}|\mathcal{F}_t\right),$$

where from Proposition 1

$$\mathbb{E}\left(e^{\omega\lambda_s^{(n)}}|\mathcal{F}_t\right) = \exp\left(F(t, s, \omega) + G(t, s, \omega)\lambda_t^{(n)} + \sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} Y_t^{(l)}\right)$$

and

$$\sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} = \sum_{l=-n}^{n} b_l^{(n)} \eta i \int_t^s e^{ib_l^{(n)}(u-t)} G(u, s, \omega) du\, m_l^{(k,n)}.$$

Taking into account the definition of the discretised measure $\tilde{v}_k^{(n)}$ (24) we obtain

$$\int_{-\infty}^{\infty} H_l(t, s, l, \omega) \tilde{v}_k^{(n)}(l)\, dl = \eta i \int_{-\infty}^{\infty} \left(\int_t^s l e^{-il(u-t)} G(u, s, \omega) du\right) \tilde{v}_k^{(n)}(l)\, dl.$$

Passing to the limit we have

$$\int_{-\infty}^{\infty} H_l(t, s, l, \omega) v_k(l) \, \mathrm{d}l = \eta i \int_{-\infty}^{\infty} \left(\int_t^s l e^{-il(u-t)} G(u, s, \omega) \, \mathrm{d}u \right) v_k(l) \, \mathrm{d}l$$

$$= \eta i \int_t^s G(u, s, \omega) \int_{-\infty}^{\infty} l e^{-il(u-t)} v_k(l) \, \mathrm{d}l \mathrm{d}u$$

$$= \eta \int_t^s G(u, s, \omega) \frac{\mathrm{d}g_k(u-t)}{\mathrm{d}t} \mathrm{d}u$$

The second and last lines of the latter equation stem from the Fubini's theorem and the frequency differentiation property of the Fourier transform. Likewise, we show that the limit of $\sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}, \omega\right) Y_t^{(l)} m_l^{(k,n)}$ is

$$\int_{-\infty}^{\infty} H_l(t, s, l, \omega) Y_t^l v_k(l) \, \mathrm{d}l = i\eta \int_t^s G(u, s, \omega) \int_{\mathbb{R}} l e^{il(u-t)} Y_t^l v_k(l) \, \mathrm{d}l \, \mathrm{d}u,$$

and concludes the proof. $\qquad \square$

Note that for $t = 0$, $Y_t^l = 0$ and we evaluate the moment generating function of $\lambda_s \mid \mathcal{F}_t$ in Eq. (31) solely from λ_0. But for $0 < t < s$, we need to infer $\left(\lambda_t^{(n)}, \left(Y_t^{(n)}\right)_{l=-n,\cdots,n}\right)$ from option prices observed over the interval $[0, t]$. This can be done using the filtering technique as in Hainaut and Moraux (2018b), and Njike and Hainaut (2022). Applying a finite difference approach to the mgf of the HP intensity in Proposition 2, we draw the expected value of the HP intensity given the time in Fig. 2. From this figure, we conclude that the expected HP intensity decreases as time increases at the same rate of decay as the corresponding kernel function.

Let $X_t := \log S_t$ the log of the asset price at time t. Considering the discretisation scheme on the partition $\mathcal{E}^{(n)}$, the asset price noted $S_t^{(n)}$ is solution of the following PDE:

$$\frac{\mathrm{d}S_t^{(n)}}{S_t^{(n)}} = r\mathrm{d}t + \sqrt{V_t}\left(\rho \mathrm{d}W_t^1 + \sqrt{1-\rho^2} \mathrm{d}W_t^2\right)$$

$$+ \mathrm{d}\left(\sum_{l=1}^{N_t^{(n)}} \left(e^{J_l} - 1\right)\right) - \lambda_t^{(n)} \mathbb{E}\left(e^J - 1\right) \mathrm{d}t.$$

$$(35)$$

By Itô formula the discretised version of the process X_t, noted $X_t^{(n)} := \log S_t^{(n)}$ admits the next SDE
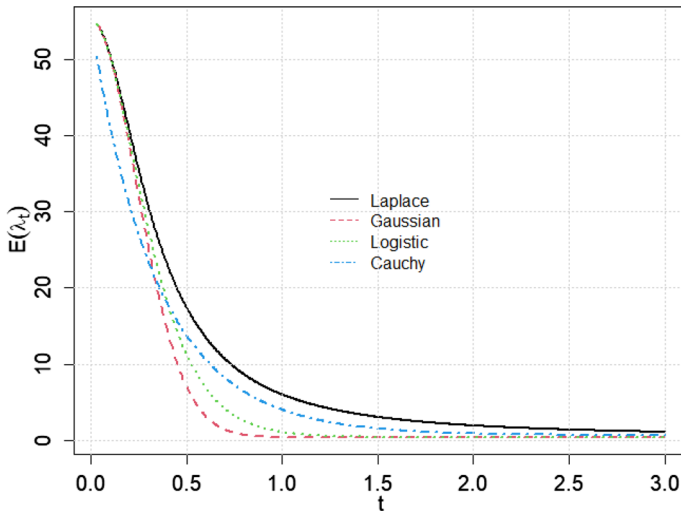
Fig. 2 Expected value of the HP intensity $\lambda_t \mid \mathcal{F}_0$. For comparison purpose, the parameter β of the Laplace, the Gaussian and the Logistic kernels are such that $\sigma_L^2 = \sigma_G^2 = \sigma_{Log}^2 = 16$. While the Cauchy kernel is computed with $\beta_C = 2.71$. The other parameters are: $\alpha = 0.4$, $\eta = 0.2$, $\lambda_0 = 50$, $p = 0.4$, $\rho^+ = 30$ and $\rho^- = -37$

$$\mathrm{d}X_t^{(n)} = \left(r - \frac{1}{2}V_t - \lambda_t^{(n)}\mathbb{E}\left(e^J - 1\right) \right)\mathrm{d}t + \sqrt{V_t}\left(\rho\mathrm{d}W_t^1 + \sqrt{1-\rho^2}\mathrm{d}W_t^2 \right)$$

$$+ \mathrm{d}\left(\sum_{l=1}^{N_t^{(n)}} J_l \right). \tag{36}$$

The process $X_t^{(n)}$ is then a Markov process with respect to the intensity $\lambda_t^{(n)}$, the $2n+1$ factors $\left(Y_t^{(l)}\right)_{l=-n,\ldots,n}$ and the counter $N_t^{(n)}$. We can therefore deal with pricing and hedging problems using standard methods developed for stochastic models. It is shown in the proof of Proposition 2 that $\left(\lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\ldots,n}, N_t^{(n)}\right)$ converges almost surely to $\left(\lambda_t, \left(Y_t^l\right)_{l=-n,\ldots,n}, N_t\right)$. Consequently, the process $X_t^{(n)}$ converges almost surely to X_t. The next proposition establishes the characteristic function of $X_s^{(n)}$ conditionally to the information at time $t \leq s$.

Proposition 3 *For all $\omega \in \mathbb{R}$ that satisfies the regularity condition $\psi\left(i\omega, \eta \sum_{l=-n}^n m_l^{(k,n)}\right) < \infty$, the characteristic function of $X_s^{(n)}$ conditionally to the information at time $t \leq s$ is given by*

$$\Upsilon_{t,s}^{(n)}(\omega) := \mathbb{E}^Q\left(e^{i\omega X_s^{(n)}} \mid \mathcal{F}_t\right)$$

$$= \exp\left(i\omega\left(X_t^{(n)} + r\,(s-t)\right)\right)$$

$$\times \exp\left(\theta_V \sigma_V^{-2}\left(k_V - i\omega\rho\sigma_V - \bar{d}\right)(s-t) - 2\log\left(\frac{1-\bar{g}e^{-\bar{d}(s-t)}}{1-\bar{g}}\right)\right)$$

$$\times \exp\left(V_t \sigma_V^{-2}\left(k_V - i\omega\rho\sigma_V - \bar{d}\right)\frac{1-e^{-\bar{d}(s-t)}}{1-\bar{g}e^{-\bar{d}(s-t)}}\right)$$

$$\times \exp\left(A_J(t,s,\omega) + \sum_{l=-n}^{n} C_l\left(t,s,b_l^{(n)},\omega\right) m_l^{(k,n)} Y_t^{(l)} + D(t,s,\omega)\lambda_t^{(n)}\right)$$

where

$$\bar{d} = \sqrt{(i\omega\rho\sigma_V - k_V)^2 + \sigma_V^2\left(i\omega + \omega^2\right)},$$

$$\bar{g} = \frac{k_V - i\omega\rho\sigma_V - \bar{d}}{\left(k_V - i\omega\rho\sigma_V + \bar{d}\right)},$$

$k \in \{L, G, Log, C\}$ indicates the type of HP memory kernel. For all $l \in \{-n, \ldots, n\}$,

$$C_l\left(t,s,b_l^{(l)},\omega\right) = ib_l^{(n)}\eta \int_t^s e^{ib_l^{(n)}(u-t)} D(u,s,\omega)\, du. \tag{37}$$

A_J and D are functions that solves the following ODEs system:

$$\begin{cases} \dfrac{\partial A_J}{\partial t}(t,s,\omega) = -D(t,s,\omega)(\lambda_0 - \alpha)\dfrac{dg_k(t)}{dt} \\[2mm] \dfrac{\partial D}{\partial t}(t,s,\omega) = i\omega(\psi(1,0)-1) - \left(\psi\left(i\omega, \sum_{l=-n}^{n} C_l\left(t,s,b_l^{(l)},\omega\right) m_l^{(k,n)}\right.\right. \\[2mm] \qquad\qquad\qquad \left.\left. + \eta D(t,s,\omega)\right) - 1\right) \end{cases} \tag{38}$$

with the terminal condition $A_J(s,s,\omega) = 0$ and $D(s,s,\omega) = 0$.

Proof In Appendix B. □

The previous proposition presents the characteristic function as the product of the exponent of an affine function related to the jump component and the characteristic function of the standard Heston model (see Schoutens et al. 2004 and Albrecher et al. 2006 for more details). Passing to the limit with respect of n, we have the following proposition on the characteristic function of the process X_s under the risk neutral measure $\mathbb{Q}$.

Proposition 4 *For all $\omega \in \mathbb{R}$ that satisfies the regularity condition $\psi(i\omega, \eta) < \infty$, the characteristic function of X_s conditionally to the information at time $t \leq s$ is given by*

$$\Upsilon_{t,s}(\omega) := \mathbb{E}^{\mathbb{Q}}\left(e^{i\omega X_s} \mid \mathcal{F}_t\right) \tag{39}$$

$$= \exp\left(i\omega\left(X_t^{(n)} + r(s-t)\right)\right)$$

$$\times \exp\left(\theta_V \sigma_V^{-2}\left(k_V - i\omega\rho\sigma_V - \bar{d}\right)(s-t) - 2\log\left(\frac{1 - \bar{g}e^{-\bar{d}(s-t)}}{1 - \bar{g}}\right)\right)$$

$$\times \exp\left(V_t \sigma_V^{-2}\left(k_V - i\omega\rho\sigma_V - \bar{d}\right)\frac{1 - e^{-\bar{d}(s-t)}}{1 - \bar{g}e^{-\bar{d}(s-t)}}\right)$$

$$\times \exp\left(A_J(t, s, \omega) + i\eta \int_t^s D(u, s, \omega) \int_{\mathbb{R}} l e^{il(u-t)} Y_t^l v_k(l)\, dl\, du + D(t, s, \omega)\lambda_t\right)$$

where

$$\bar{d} = \sqrt{(i\omega\rho\sigma_V - k_V)^2 + \sigma_V^2(i\omega + \omega^2)},$$

$$\bar{g} = \frac{k_V - i\omega\rho\sigma_V - \bar{d}}{\left(k_V - i\omega\rho\sigma_V + \bar{d}\right)}.$$

$k \in \{L, G, Log, C\}$ *indicates the type of HP memory kernel.* A_J *and* D *satisfy the next ODE and PIDE*

$$\begin{cases} \dfrac{\partial A_J}{\partial t}(t, s, \omega) = -D(t, s, \omega)(\lambda_0 - \alpha)\dfrac{dg_k(t)}{dt} \\ \dfrac{\partial D}{\partial t}(t, s, \omega) = i\omega(\psi(1, 0) - 1) \\ \qquad\qquad - \left(\psi\left(i\omega, \eta \int_t^s fracdg_k dh(u-t) D(u, s, \omega)\, du + \eta D(t, s, \omega)\right) - 1\right) \end{cases}$$

(40)

with the terminal condition $A_J(s, s, \omega) = 0$ *and* $D(s, s, \omega) = 0$

Proof This proof is similar to the one of Proposition 2, as $X_s^{(n)}$ converges to X_s almost surely and by the dominated convergence the conditional characteristic function of the log-return process is equal to the limit with respect of n of the conditional characteristic function of the approximated log-return process. By some grouping and using the frequency derivative property of Fourier transform we retrieve the result of Proposition 4. □

Our setting takes into account several features. It allows for stochastic evolution of the asset price's variance, includes a jump risk component into the asset price dynamics and replicate the jumps clustering. We also generalize some existing valuation models. For instance, we retrieve the Black and Scholes option price formula by setting $J_l = 0$ and $\theta_V = k_V = \sigma_V = 0$, the Heston option price formula by setting $J_l = 0$ and the Bates and Merton option price formula by setting $\alpha = \lambda_0$, $\eta = 0$ and the jump size J_l log-normally distributed.

For comparison purpose, Fig. 3 draws the density of the log-return at 50 days X_{50d} under the risk neutral measure for our four HP kernels. From these graphs, we conclude that the log-return density at 50 days with the HP Cauchy kernel has the thinnest tail. The log-return density at 50 days for ours extension of the Heston model
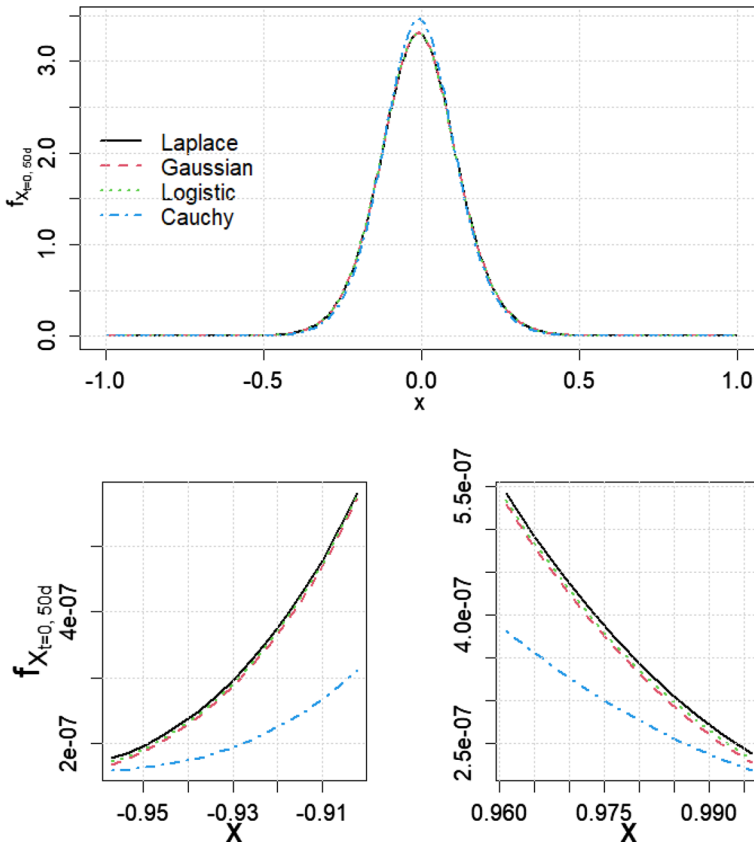
Fig. 3 Distribution of $X_{50d} \mid \mathcal{F}_0$ under the risk neutral measure. For comparison purpose, the parameter β of the Laplace, the Gaussian and the Logistic kernels are such that $\sigma_L^2 = \sigma_G^2 = \sigma_{Log}^2 = 16$. While the Cauchy kernel is computed with $\alpha_C = 2.71$. The other parameters are: $S_0 = 1$, $r = 0.01$, $\theta_V = 0.4$, $k_V = 1$, $\rho = -0.7$, $\sigma_V = 0.01$, $v_0 = 0.01$, $\alpha = 0.4$, $\eta = 0.2$, $\lambda_0 = 50$, $p = 0.4$, $\rho^+ = 30$ and $\rho^- = -37$

with the Laplace, Gaussian and Logistic kernels look identical but the zoom in reveals the differences due to the decay rate of these kernel functions. Hence, the log-return density with the Laplace kernel which has a slower decay than the Gaussian and logistic ones, has the fattest tail.

3 Option pricing models

In this section, we develop the option pricing model under our modeling framework outlined in the previous section.

As many authors (see e.g. Heston 1993; Bakshi et al. 1997; Gatheral 2012) we will express the call option price using the following delta probability decomposition of Black-Sholes formula (1973). For this aim, in the next proposition we introduce a convenient measure $\mathbb{Q}^\star$.

Proposition 5 *Let $\mathbb{Q}^\star$ an equivalent measures to $\mathbb{Q}$ on the measure space $(\Omega, \mathcal{F})$, and define by*

$$\frac{d\mathbb{Q}^\star}{d\mathbb{Q}}\bigg|_{\mathcal{F}_0} = \exp\left(-\frac{1}{2}\int_0^t V_u du + \int_0^t \sqrt{V_u}\left(\rho dW_u^1 + \sqrt{1-\rho^2}dW_u^2\right)\right)$$

$$\times \exp\left(-(\psi(1,0)-1)\int_0^t \lambda_u du + \sum_{l=1}^{N_u} J_l\right).$$

Then, under the measure $\mathbb{Q}^\star$,

(i) $W_t^{\star 1}$ and $W_t^{\star 2}$ are Brownian motions defined as follows:

$$dW_t^{\star 1} = dW_t^1 - \rho\sqrt{V_t}dt,$$
$$dW_t^{\star 2} = dW_t^2 - \sqrt{1-\rho^2}\sqrt{V_t}dt.$$

(ii) $N_t^\star$ is a counting process with the following intensity

$$\lambda_t^\star = \psi(1,0)\lambda_t. \tag{41}$$

Considering the discretisation scheme on the partition $\mathcal{E}^{(n)}$, this intensity noted $\lambda_t^{(n),\star}$ is solution of next SDE

$$d\lambda_t^{(n),\star} = \left(\lambda_0^\star - \alpha^\star\right)\frac{dg_k(t)}{dt}dt + \eta^\star \sum_{l=-n}^{n} m_l^{(k,n)}ib_l^{(n)}Y_t^{(l),\star}dt + \eta^\star dL_t^{(n),\star}, \tag{42}$$

where $\lambda_0^\star = \psi(1,0)\lambda_0$, $\eta^\star = \psi(1,0)\eta$, $L_t^{(n),\star} = \sum_{l=1}^{N_t^{(n),\star}}\left|J_l^\star\right|$ and for all $l \in \{-n,\ldots,n\}$

$$dY_t^{(l),\star} = ib_l^{(n)}Y_t^{(l),\star}dt + dL_t^{(n),\star}. \tag{43}$$

$J^\star$ is a double exponential distribution random variables of parameters:

$$\rho^{+,\star} = \rho^+ - 1, \tag{44}$$
$$\rho^{-,\star} = \rho^- + 1, \tag{45}$$
$$p^\star = \frac{p\rho^+\rho^{-,\star}}{p\rho^+\rho^{-,\star} + (1-p)\rho^-\rho^{+,\star}}, \tag{46}$$

with the following joint characteristic function:

$$\psi^\star(z_1, z_2) := \mathbb{E}\left(e^{z_1 J^\star + z_2|J^\star|}\right) = \frac{\psi(z_1+1, z_2)}{\psi(1,0)}. \tag{47}$$

(iii) the processes X_t and V_t satisfy the following SDEs:

$$dX_t = \left(r + \frac{1}{2}V_t - \lambda_t^\star \left(\psi^\star (1, 0) - 1 \right) \right) dt \tag{48}$$

$$+ \sqrt{V_t} \left(\rho dW_t^{\star 1} + \sqrt{1 - \rho^2} dW_t^{\star 2} \right) + d \left(\sum_{l=1}^{N_t^\star} J_l^\star \right),$$

$$dV_t = (\theta_V - (k_V - \rho \sigma_V) V_t) dt + \sigma_V \sqrt{V_t} dW_t^{\star 1}. \tag{49}$$

Proof In Appendix C. □

From to this proposition and using the characteristic function of the log-return in Proposition 3 we value an European call option of maturity T and strike K, written on the terminal spot price S_T of some underlying asset in the next proposition.

Proposition 6 *The price $C(t, T, S, V, \lambda, N, Y, K)$ of an European call option of maturity T, written on S_t is given by*

$$C(t, T, S, V, \lambda, N, Y, K) = S_t P_0(t, T, S, V, \lambda, N, Y, K) \tag{50}$$
$$- K e^{-r(T-t)} P_1(t, T, S, V, \lambda, N, Y, K),$$

where K is the strike price, $P_j \equiv P_j(t, T, S, V, \lambda, N, Y, K)$ are the probability of the call expiring in-the-money under the forward measure $\mathbb{Q}^\star$ and the risk neutral measure $\mathbb{Q}$ respectively. For $j \in \{0, 1\}$, these probabilities are of the form

$$P_0 = \frac{1}{2} + \frac{1}{\pi} \int_{\mathbb{R}_+} Re \left(\exp(-i\omega \log K) \frac{\Upsilon_{t,s}(\omega - i)}{i\omega e^{r(T-t)} S_t} \right) d\omega, \tag{51}$$

$$P_1 = \frac{1}{2} + \frac{1}{\pi} \int_{\mathbb{R}_+} Re \left(\exp(-i\omega \log K) \frac{\Upsilon_{t,s}(\omega)}{i\omega} \right) d\omega \tag{52}$$

where $\Upsilon_{t,s}$ is the characteristic function of the process X_t under the measure $\mathbb{Q}$ defined in the Proposition 4.

Proof Considering the discretisation scheme on the partition $\mathcal{E}^{(n)}$, the European call option price is obtained as

$$C\left(t, T, S^{(n)}, V, \lambda^{(n)}, N^{(n)}, Y, K \right) = C\left(t, T, X^{(n)}, V, \lambda^{(n)}, N^{(n)}, Y, K \right)$$
$$= e^{-r(T-t)} \mathbb{E} \left(\max \left(e^{X_T} - K, 0 \right) | \mathcal{F}_t \right)$$
$$= e^{X_t} \mathbb{E} \left(\frac{e^{-rT} e^{X_T}}{e^{-rt} e^{X_t}} \mathbb{I}_{X_T^{(n)} > \log K} | \mathcal{F}_t \right)$$
$$- K e^{-r(T-t)} \mathbb{E} \left(\mathbb{I}_{X_T^{(n)} > \log K} | \mathcal{F}_t \right)$$
$$= e^{X_t} \mathbb{E}^\star \left(\mathbb{I}_{X_T^{(n)} > \log K} | \mathcal{F}_t \right)$$

$$-Ke^{-r(T-t)}\mathbb{E}\left(\mathbb{I}_{X_T^{(n)}>\log K}|\mathcal{F}_t\right)$$

$$= S_t\mathbb{Q}^{\star}\left(X_T^{(n)}>\log K\right)$$

$$-Ke^{-r(T-t)}\mathbb{Q}\left(X_T^{(n)}>\log K\right)$$

$$= S_tP_0 - Ke^{-r(T-t)}P_1 \tag{53}$$

where as in the Black-Sholes formula (1973) $P_0^{(n)} \equiv P_0\left(t, T, X^{(n)}, V, \lambda^{(n)}, N^{(n)}, Y, K\right)$ and $P_1^{(n)} \equiv P_1\left(t, T, X^{(n)}, V, \lambda^{(n)}, N^{(n)}, Y, K\right)$ are the probability of the call expiring in-the-money under the measure $\mathbb{Q}^{\star}$ and the risk neutral measure $\mathbb{Q}$ respectively. Since $C\left(T, T, X^{(n)}, V, \lambda^{(n)}, N^{(n)}, Y, K\right) = \max\left(e^{X_T} - K, 0\right)$, with the terminal condition

$$P_j\left(T, T, X^{(n)}, V, \lambda^{(n)}, N^{(n)}, Y, K\right) = \begin{cases} 1 & \text{if } X_T > 0 \\ 0 & \text{otherwise} \end{cases}, \quad \forall j \in \{0, 1\}. \tag{54}$$

Given that the P_j are survival probabilities with respect of $X_T \mid \mathcal{F}_t$ under the measure $\mathbb{Q}^{\star}$ and the risk neutral measure $\mathbb{Q}$ respectively. We apply the inversion theorem of Gil-Pelaez (1951) to get:

$$P_0^{(n)} = \frac{1}{2} + \frac{1}{\pi}\int_{\mathbb{R}_+}\text{Re}\left(\frac{e^{-i\omega\log K}\Upsilon_{t,T}^{(n),\star}(\omega)}{i\omega}\right)d\omega \tag{55}$$

$$P_1^{(n)} = \frac{1}{2} + \frac{1}{\pi}\int_{\mathbb{R}_+}\text{Re}\left(\frac{e^{-i\omega\log K}\Upsilon_{t,T}^{(n)}(\omega)}{i\omega}\right)d\omega, \tag{56}$$

where $\Upsilon_{t,T}^{(n),\star}(.)$ and $\Upsilon_{t,T}^{(n)}(.)$ are the characteristic functions of $X_T \mid \mathcal{F}_t$ under the measure $\mathbb{Q}^{\star}$ and the risk neutral measure $\mathbb{Q}$ respectively. From the definition of the measure $\mathbb{Q}^{\star}$ we have

$$\Upsilon_{t,T}^{(n),\star}(\omega) = \mathbb{E}^{\star}\left(e^{i\omega X_T^{(n)}} \mid \mathcal{F}_t\right)$$

$$= \mathbb{E}\left(\frac{e^{-rT}e^{X_T^{(n)}}}{e^{-rt}e^{X_t^{(n)}}}e^{i\omega X_T^{(n)}} \mid \mathcal{F}_t\right)$$

$$= \frac{1}{e^{r(T-t)}S_t}\mathbb{E}\left(e^{i(\omega-i)X_T^{(n)}} \mid \mathcal{F}_t\right)$$

$$= \frac{\Upsilon_{t,T}^{(n)}(\omega - i)}{e^{r(T-t)}S_t}.$$

Replacing this into Eq. (55), passing to the limit with respect of n and using the frequency derivative property of Fourier transform, allows us to complete this proof.

$\square$

The main problem for implementing this latter call option formula is the numerical approximation of integrals in Eqs. (51) and (52). It worth noting that a wrong implementation involves a weak quality of fitness on real market option data. One may think about Fourier transform, but the characteristic function $\Upsilon_{t,s}$ is only evaluated numerically with possible discontinuities and the integrands have a point of singularity at $\omega = 0$. However, these integrals can be evaluate efficiently via a trapezoidal quadrature approximation. Note that the Lemma 1 in Witkovský (2001) enables one to compute the limits of integrands in Eqs. (51) and (52) when ω vanish:

$$\text{Re}\left(\frac{e^{-i\omega \log K}}{i\omega} \frac{\Upsilon_{t,T}(\omega - i)}{e^{r(T-t)}S_t}\right) \to \mathbb{E}^{\star}(X_T \mid \mathcal{F}_t) - \log K \tag{57}$$

$$\text{Re}\left(\frac{e^{-i\omega \log K}}{i\omega} \Upsilon_{t,T}(\omega)\right) \to \mathbb{E}(X_T \mid \mathcal{F}_t) - \log K \tag{58}$$

where

$$\mathbb{E}^{\star}(X_T \mid \mathcal{F}_t) = r(T-t) + \frac{1}{2}\frac{\theta_V}{k_V - \rho\sigma_V}(T-t) \tag{59}$$
$$+ \left(V_t - \frac{\theta_V}{k_V - \rho\sigma_V}\right)\frac{1}{2(k_V - \rho\sigma_V)}\left(1 - e^{-(k_V - \rho\sigma_V)(T-t)}\right),$$

$$\mathbb{E}(X_T \mid \mathcal{F}_t) = r(T-t) - \frac{1}{2}\frac{\theta_V}{k_V}(T-t) - \frac{1}{2}\left(V_t - \frac{\theta_V}{k_V}\right)\frac{1}{k_V}\left(1 - e^{-k_V(T-t)}\right). \tag{60}$$

are results proven in Appendix D.

The trapezoidal quadrature approximation of the probabilities of the call expiring in-the-money for M discrete log-strikes $\log K_u := k_u = -k_{max} + (u-1)\Delta_k$, $u = 1, \ldots, M$, equally spaced $\Delta_k = \frac{2k_{max}}{M}$ are given by:

$$P_0 \approx \frac{1}{2} + \frac{1}{\pi}\sum_{l=1}^{M}\gamma_l\text{Re}\left(\frac{e^{-i\omega_l k_u}}{i\omega_l}\frac{\Upsilon_{t,T}(\omega_l - i)}{e^{r(T-t)}S_t}\right)\Delta_\omega$$

$$P_1 \approx \frac{1}{2} + \frac{1}{\pi}\sum_{l=1}^{M}\gamma_l\text{Re}\left(\frac{e^{-i\omega_l k_u}}{i\omega_l}\Upsilon_{t,T}(\omega_l)\right)\Delta_\omega$$

where $\gamma_l = \frac{1}{2}1_{\{l=1\}} + 1_{\{l\neq 1\}}$ is the trapezoidal quadrature weights, $\Delta_\omega = \frac{2\pi}{M\Delta_K}$ ensures that the absolute value of the integrand function is sufficiently small for all $\omega > \omega_M$ with $\omega_l = (l-1)\Delta_\omega$ for $l = 1, \ldots, M$.

From our call option price formula (50) we can derive Greeks for hedging purpose given the three sources of risk: the price risk S_t the volatility risk V_t and the jump risk λ_t. The analytical expressions of the three deltas are given by

$$\Delta S = \frac{\partial C}{\partial S} = P_0 + S_t\frac{\partial P_0}{\partial S} - Ke^{-r(T-t)}\frac{\partial P_1}{\partial S}, \tag{61}$$

$$\Delta V = \frac{\partial C}{\partial V} = S_t \frac{\partial P_0}{\partial V} - Ke^{-r(T-t)} \frac{\partial P_1}{\partial V},$$

$$\Delta \lambda = \frac{\partial C}{\partial \lambda} = S_t \frac{\partial P_0}{\partial \lambda} - Ke^{-r(T-t)} \frac{\partial P_1}{\partial \lambda},$$

where for $j \in \{0, 1\}$

$$\frac{\partial P_0}{\partial S} = \frac{1}{\pi} \int_{\mathbb{R}_+} \text{Re}\left(\frac{e^{-i\omega \log K} \Upsilon_{t,T}(\omega - i)}{e^{r(T-t)} S_t^2} \right) d\omega,$$

$$\frac{\partial P_1}{\partial S} = \frac{1}{\pi} \int_{\mathbb{R}_+} \text{Re}\left(\frac{e^{-i\omega \log K} \Upsilon_{t,T}(\omega)}{S_t} \right) d\omega,$$

$$\frac{\partial P_0}{\partial V} = \frac{1}{\pi} \int_{\mathbb{R}_+} \text{Re}\left(B(t, T, \omega - i) \frac{e^{-i\omega \log K} \Upsilon_{t,T}(\omega - i)}{i\omega e^{r(T-t)} S_t} \right) d\omega,$$

$$\frac{\partial P_1}{\partial V} = \frac{1}{\pi} \int_{\mathbb{R}_+} \text{Re}\left(B(t, T, \omega) \frac{e^{-i\omega \log K} \Upsilon_{t,T}(\omega)}{i\omega} \right) d\omega,$$

$$\frac{\partial P_0}{\partial \lambda} = \frac{1}{\pi} \int_{\mathbb{R}_+} \text{Re}\left(D(t, T, \omega - i) \frac{e^{-i\omega \log K} \Upsilon_{t,T}(\omega - i)}{i\omega e^{r(T-t)} S_t} \right) d\omega,$$

$$\frac{\partial P_1}{\partial \lambda} = \frac{1}{\pi} \int_{\mathbb{R}_+} \text{Re}\left(D(t, T, \omega) \frac{e^{-i\omega \log K} \Upsilon_{t,T}(\omega)}{i\omega} \right) d\omega,$$

with

$$B(t, T, \omega) = \sigma_V^{-2} \left(k_V - i\omega \rho \sigma_V - \bar{d} \right) \frac{1 - e^{-\bar{d}(s-t)}}{1 - \bar{g} e^{-\bar{d}(s-t)}},$$

and D which satisfies the ODE in Eq. (40). Similarly we can obtain the second and third order Greeks letter, gamma and vega which are respectively the second and third order partial derivatives of the call option price with respect of the asset price, the variance and the arrival rate of jumps. Thus, the sensitivity of our call option price with respect of these three state variables has an analytic form that relies on the computation of integrals.

4 Illustration

This section enables us to achieve three goals. Firstly, we present the method for calibrating our Heston model extension to an implied volatility surface. Secondly, we show that our model achieves a better fit of implied volatilities than the standard Heston. Finally, we study the sensitivity of European call option price to the memory parameter β (Eqs. 6, 8, 10 and 12).
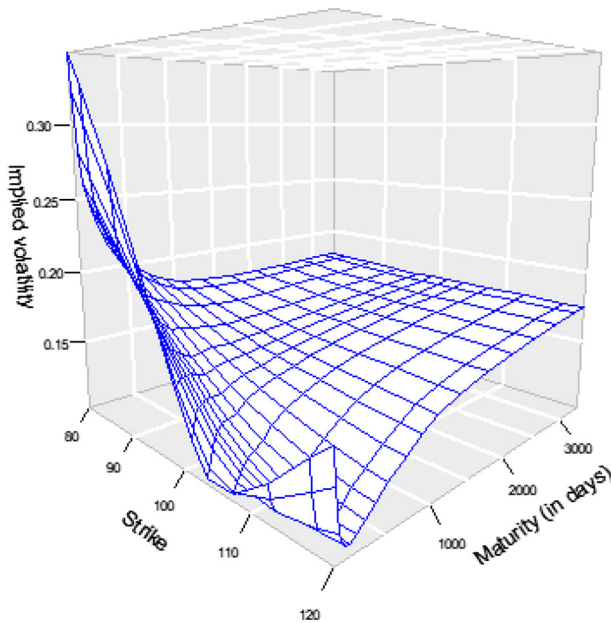
Fig. 4 Implied volatility surface of the Euro Stoxx 50 as of the 26th of september 2019, with the spot price $S_0 = 3541$ and the dividend rate $q = 0.00225$

4.1 Data description

The data used in this section is a result of a filtering process applied to the implied volatility surface of the Euro Stoxx 50 as of the 26th of September 2019 in Fig. 4, with the spot price $S_0 = 3541$ and the dividend rate $q = 0.00225$. To our raw data obtain from the Bloomberg's Option Monitor we exclude options with maturity greater than one as there are less sensitive to volatility. The constant risk free rate r is approximated by the three-month instantaneous forward rate on the 26th of the September 2019 implied in the yield curve for the euro area and taking into account issuers whose rating is triple A.

4.2 Calibration and sensitivity analysis

To calibrate our SVJ models and the Heston model under the risk neutral measure, we use the standard approach that consists in minimizing the sum of relative squares errors (SRSE) between IVs derived from our model and market IVs. More specifically, If we denote by Θ the parameters set of our SVJ model, one evaluate the vector of parameters $\hat{\Theta}$ for which our model gives the closest prices to those observed in the market as follows:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \left(\operatorname{SRSE}(\Theta) + \gamma \max \left(\sigma_V^2 - 2\theta_V, 0 \right) \right), \tag{62}$$

Table 1 Estimated parameters with $k_{max} = 1$, $M = 2^8$ and $\gamma = 10000$ used for the trapezoidal quadrature approximation of the probabilities of the call expiring in-the-money (Eqs. 51 and 52)
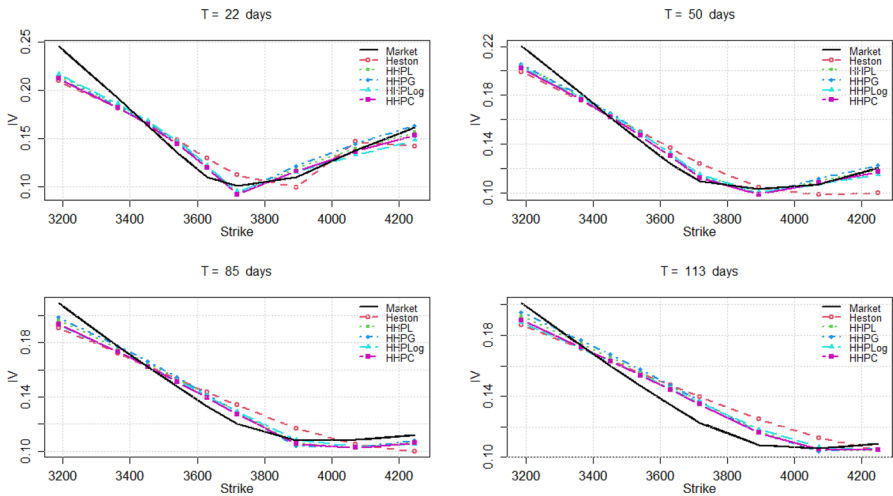
	Heston	HHPL	HHPG	HHPLog	HHPC
θ_V	0.2183	0.2475	0.2223	0.2744	0.2320
k_V	7.3697	8.0511	6.7373	9.7406	7.8421
σ_V	0.6575	0.70201	0.6632	0.7406	0.6808
ρ	-0.8136	-0.9983	-0.9969	-0.9989	-0.9913
V_0	0.0217	0.0208	0.0201	0.0218	0.0202
λ_0		1.5484	1.1158	1.3813	1.3013
α		2.48734	8.2820	2.7708	4.5728
η		3.1635	0.2860	9.4444	4.3008
β		0.1362	0.3683	7.9078	4.9346
p		0.5106	0.5452	0.7765	0.4081
ρ^+		53.5226	46.6534	64.0890	55.3728
ρ_-		-144.0291	-137.6665	-97.4525	-149.4079
SRSE	0.5256	0.2595	0.2647	0.2786	0.3005

$$
\text{SRSE}(\Theta) = \sum_s^{\#C} \left(\frac{\sigma_{IV}^{Market}(t, T_s, S, V, \lambda, N, Y, K_s) - \sigma_{IV}(t, T_s, S, V, \lambda, N, Y, K_s, \Theta)}{\sigma_{IV}^{Market}(t, T_s, S, V, \lambda, N, Y, K_s)} \right)^2,
$$

(63)

where K_s is the strike price of the s-th call option, T_s is the maturity of the call option, σ_{IV}^{Market} is the market implied volatility of the observed call option and $\#C$ is the number of call options used. Note that the objective function is not convex and is not of any particular structure. Thus, by using a gradient based optimizing procedure we are not sure to obtain a global minimum. As proposed by Hamida and Cont (2005), we add a convex penalization term $\gamma \max(\sigma_V^2 - 2\theta_V, 0)$ to the objective function (63) in order to ensure that a gradient based algorithm can be used.

Based on the filtered IV surface of the Euro Stoxx 50 as of the 26th September 2019, we find the set of model parameters minimizing the objective function (63). Table 1 presents the result of our calibration procedure applied to the Heston model, and to its extensions with a self-exciting jumps induce by a Laplace (6) (HHPL), Gaussian (8) (HHPG), Logistic (10) (HHPLog) or Cauchy (12) (HHPC) kernels. In Table 1 the SRSE measures the quality of fit of theses models. We find that our Heston models with jumps in the asset price dynamics are better, with SRSEs lower than the SRSE of the Heston model. In Fig. 5, we plot the market implied volatility and the volatility replicated through the Heston, the HHPL, HHPG, HHPLog and HHPC models at maturities T equal to 22, 50, 85 and 113 days. We see that the implied volatilities derived from the HHPL, HHPG, HHPLog and HHPC models are the closest to market one and they outperform the Heston model, particularly for small maturities. From the Heston model to its extensions we observe that the correlation ρ becomes close to -1. This results from the contribution of the jump component to the explanation the log-return volatility.

Fig. 5 Market implied Volatilities versus replicated implied volatilities from the Heston, the HHPL, HHPG, HHPLog and HHPC models at maturities T equal to 22, 50, 85 and 113 days

Another important quality check, when it comes to calibrate an implied volatility surface, is the replication of the term structure of at-the-money (ATM) volatility skews. Denoted by $\Psi(T)$, the ATM volatility skew at expiry time $T \geq t$ is the derivative of implied volatility with respect to the money strike:

$$\Psi(T) := \left| \frac{\partial \sigma_{IV}(t, T, S, V, \lambda, N, Y, K, \Theta)}{\partial K} \right|_{K=S_t}$$

For each model and for ϵ small enough, we approximate $\Psi(T)$ by

$$\hat{\Psi}(T) = \left| \frac{\sigma_{IV}\left(t, T, S, V, \lambda, N, Y, K+\epsilon, \hat{\Theta}\right) - \sigma_{IV}\left(t, T, S, V, \lambda, N, Y, K-\epsilon, \hat{\Theta}\right)}{2\epsilon} \right|_{K=S_t},$$

where the approximation holds for small enough ϵ. Empirical studies (see e.g. Gatheral et al. 2014) show that the observed term structure of ATM volatility skew is well approximated by power-law functions of the form $T^{H-\frac{1}{2}}$ with $H \in \left(0, \frac{1}{2}\right)$ the Hurst exponent. Gatheral et al. (2014) shows that a value of H close to zero generates a volatility surface with a reasonable shape. From the Table 2, we deduce that all models considered have an estimated Hurst exponent $\hat{H}$ between 0.09 and 0.14. The HHPLog model exhibits the smallest estimated Hurst exponent with an R-squared of the linear regression model that explains $\log \hat{\Psi}(T)$ given $\log T$ equal to 0.88. From Fig. 6, we conclude that compared to the Heston model, the other models better replicate the empirical short term explosion of the ATM volatility skew.

We now focus on the differences between our extensions of the Heston model. To distinguish them in term of memory length, we draw in Fig. 7 their corresponding kernel functions given the fitted parameters in Table 1. We observe that the Gaussian

Table 2 Estimated Hurst exponent $\hat{H}$ of the ATM volatility skew and the R-squared of the linear regression model that explains the log-ATM volatility skew $\log \hat{\Psi}(T)$ given $\log T$

	Heston	HHPL	HHPG	HHPLog	HHPC
$\hat{H}$	0.13	0.12	0.14	0.09	0.12
R^2	0.87	.87	0.85	0.88	0.87
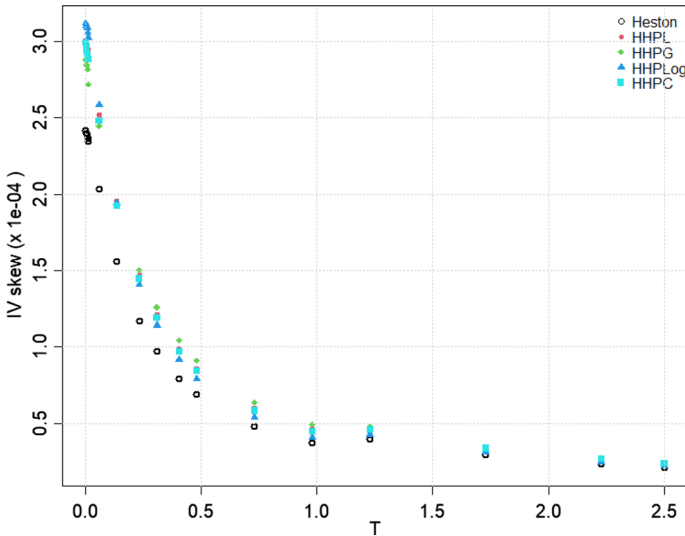
Fig. 6 The slope of implied volatility skew as function of the maturity over a range from one day to 2.5 years. Model Parameters used are from Table 1

kernel has the slowest decay, while the Logistic kernel has the fastest decay. Thus, the HHPG and HHPLog are respectively the models with the longest and shortest memory among our extensions of the Heston model. From the Fig. 7 also, we find that the memory of the HHPL and HHPG models increases with the parameter β. Whereas, the one of HHPLog and HHPC models increases when the parameter β decreases. The Fig. 8 plot the probability density function of the log-return at 113 days. This figure allows us to make link between the memory length of models and the thickness of the tail of the log-return distribution. In agreement of the result of the Sect. 2, the thickness of the tail of the log-return distribution for our extension of the Heston model are ordered by their memory lengths.

The impact of the memory parameter β of our affine Heston style models on option pricing is studied through the implied volatilities and the price risk ΔS (61). To inspect how recalling past price changes for a long period influences option prices, we vary the memory parameter β of our extensions of the Heston model. Figure 9 reveals that an increase in memory indicated by a slower decaying kernel function $g.(.)$, induces an increase in implied volatilities of out-of-the-money (OTM) call options ($S < K$). This increase in implied volatilities becomes more pronounced as the expiry rises. Figure 10 confirms this findings and brings out two points. Firstly, the impact of the
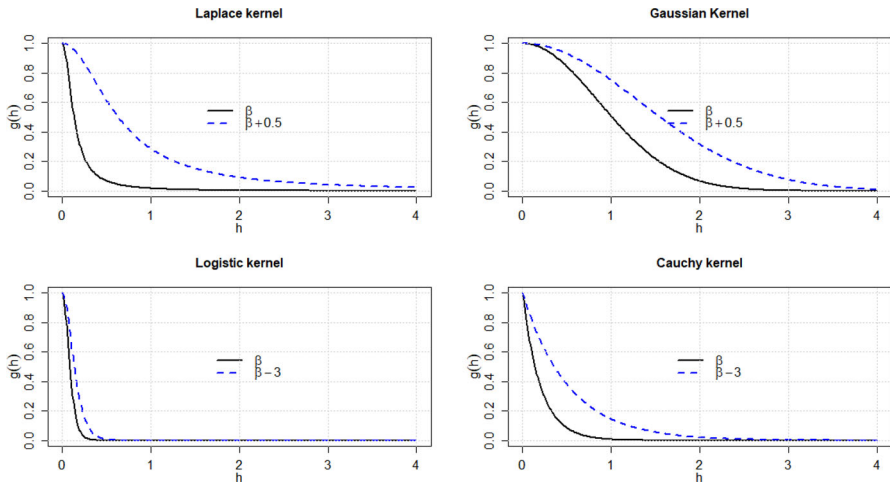
Fig. 7 Sensitivity of kernel functions to the memory parameter β. The initial value of β for each model is in Table 1

memory parameter β is negligible in the short term, but increases with the expiry time T. Secondly, for the HHPLog model with the shortest memory, the impact of a small increase or decrease in memory remains negligible in the long term.

The price risk ΔS measures the sensitivity of the option to the stock price. Let us recall that for options it should be interpreted as the amount of stocks to hold ($\Delta S > 0$) or to short sale ($\Delta S < 0$) for hedging an option. Called the "delta hedging", this is a very common strategy to do the arbitrage and minimize risk of portfolio in the option market. From the Fig. 11, the price risk ΔS calculated from the Heston model is slightly higher than the one given by the other models for OTM call options, while, for in-the-money (ITM) call options, the price risk ΔS in the Heston model is almost the same. The Fig. 12 highlights the sensitivity of the price risk ΔS to the self-excitation parameter λ_0. We find that the increase in self-excitation caused by a rising λ_0, implies a higher price risk ΔS for OTM call options and a lower one for ITM call options. Unlike the price risk ΔS, from Fig. 13, the implied volatilities derived from our extensions of the Heston model are more sensitive to the self-excitation parameter λ_0 for OTM call options.

Although call option prices from the Heston model and its extensions have roughly the same sensitivity to changes in underlying price, we find that these models induce different shapes of sensitivity to volatility and jump risks in the option price. The upper graphs in Fig. 14 draw the sensitivity of call option prices expiring at 50 and 113 days to the volatility V. Known as the greek "vega", for all models considered, it takes its greatest value around the ATM. Except for the Heston model, the vega remains positive and decreases exponentially to zero as the strike moves away from the spot price. In the lower graphs, we observe a similar shape of the call option price change in case of a jump. Denoted by $\Delta\lambda$, it takes its greatest value at the OTM region and its curve widens when the expiry increases. It is worth noting that a higher memory length does not necessary results in a wider $\Delta\lambda$ curve.
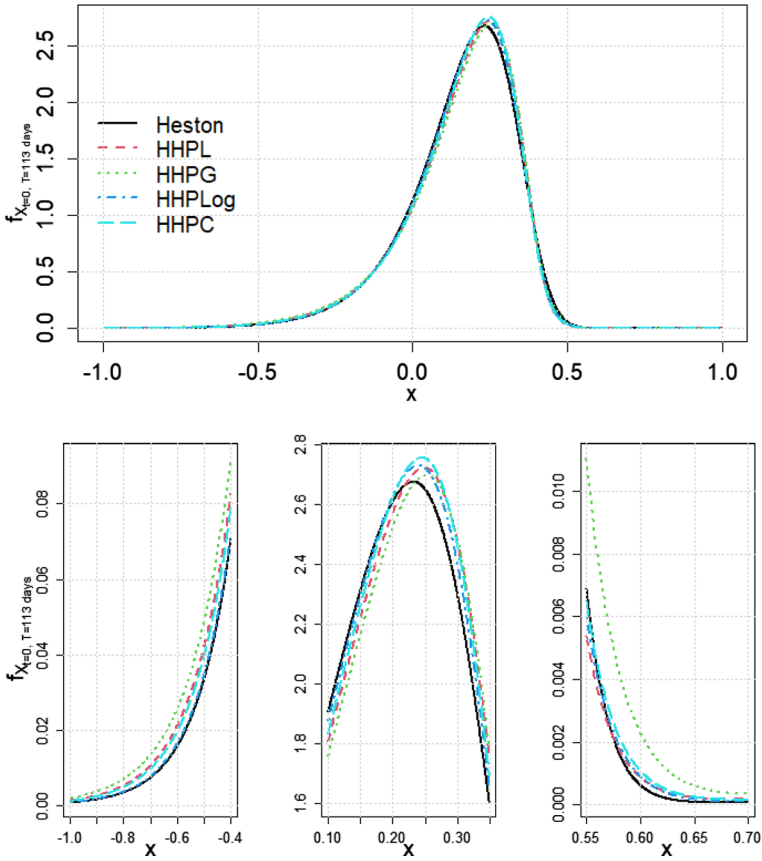
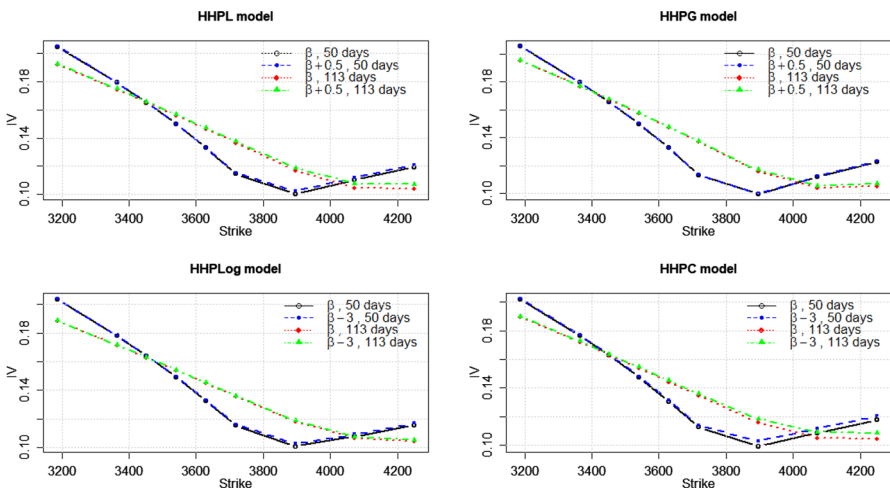**Fig. 8** Probability density function of the log-return $f_{X_0, 113\,\text{days}}$

Fig. 9 Sensitivity to the memory parameter β of the HHPL model. Model parameters used are in Table 1
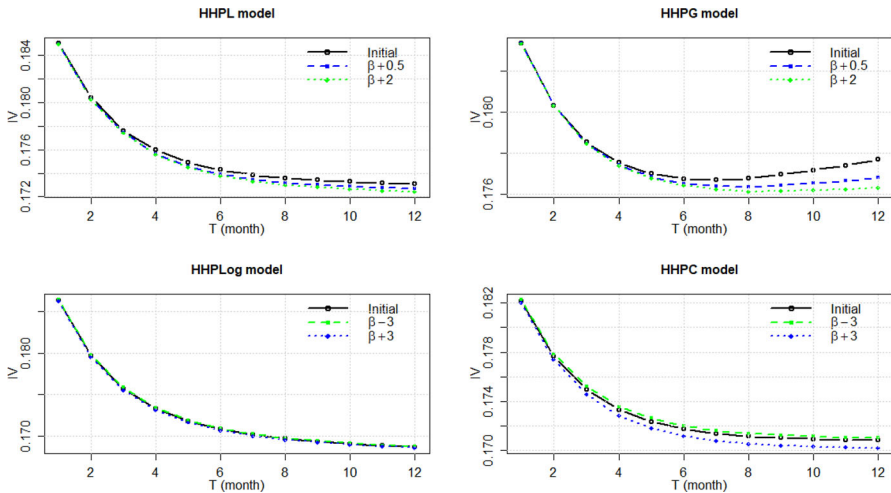
Fig. 10 Sensitivity of the implied volatility to the memory parameter β regarding the expiry time T in month. Model parameters used are in Table 1
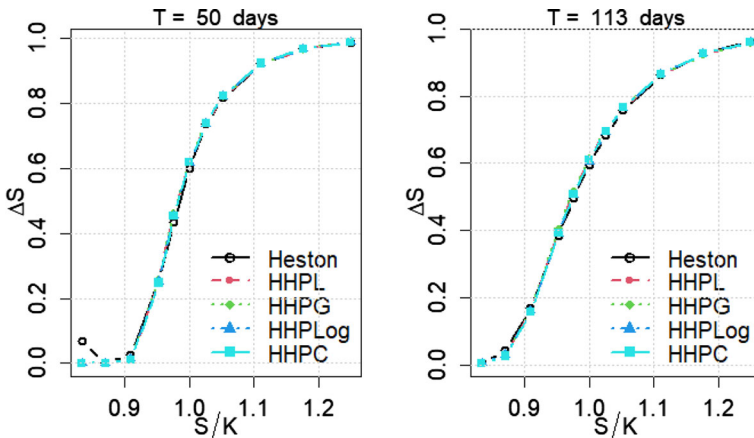

Fig. 11 The price risk ΔS for the Heston, HHPL, HHPG, HHPLog and HHPC models. Model parameters used are in Table 1

5 Conclusions

In this article, we propose an extension of the Heston model that allows for self-excited jumps in asset price and long memory of past price changes. This expansion corresponds to an asset price process with a stochastic volatility and a jump component driven by a linear Hawkes process (HP). The memory feature of our model stems from the decay rate of the HP kernel function. Considering HP kernel function that are Fourier transform of the Laplace, Gaussian, Logistic, and Cauchy measures, we generate different memory ranges, from the longest to the shortest. The self-excitation and memory properties of our framework enables us to have a realistic option pricing
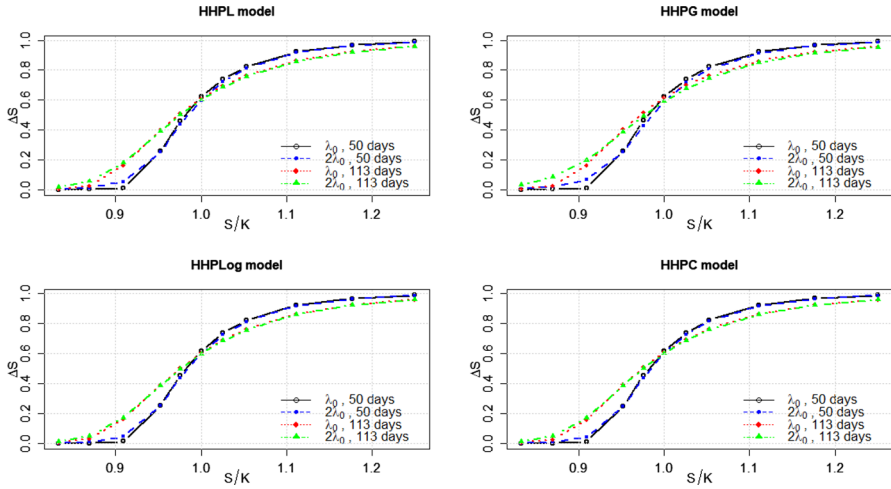
Fig. 12 Sensitivity of the price risk ΔS to the initial intensity λ_0 of Hawkes processes. Model parameters used are in Table 1
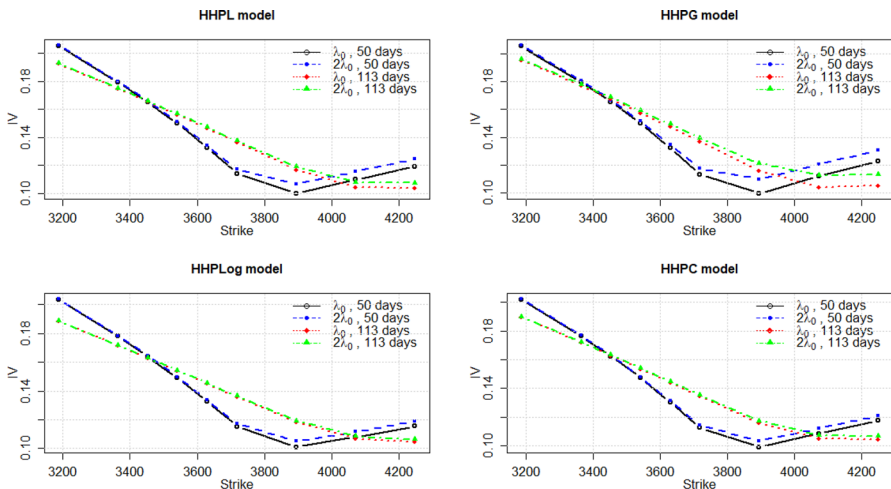

Fig. 13 Sensitivity to the memory parameter λ_0 of the HHPL model. Model parameters used are in Table 1

model that encompass major characteristic of asset's return times series such as the volatility clustering and the long memory in volatility. However, the memory effect leads to a dependence on the past that makes our asset price process non Markov.

The HP with kernel function which is a Fourier transform of the Cauchy measure corresponds to the tractable exponential kernel. Except the stochastic volatility model with jumps self-excited by this exponential kernel, our Heston extensions are non Markov processes. We use the Fourier transform representation of their kernel functions to establish a closed form expression of the conditional moment generating functions of the HP's intensity and the log-returns. Based on these results, we write the
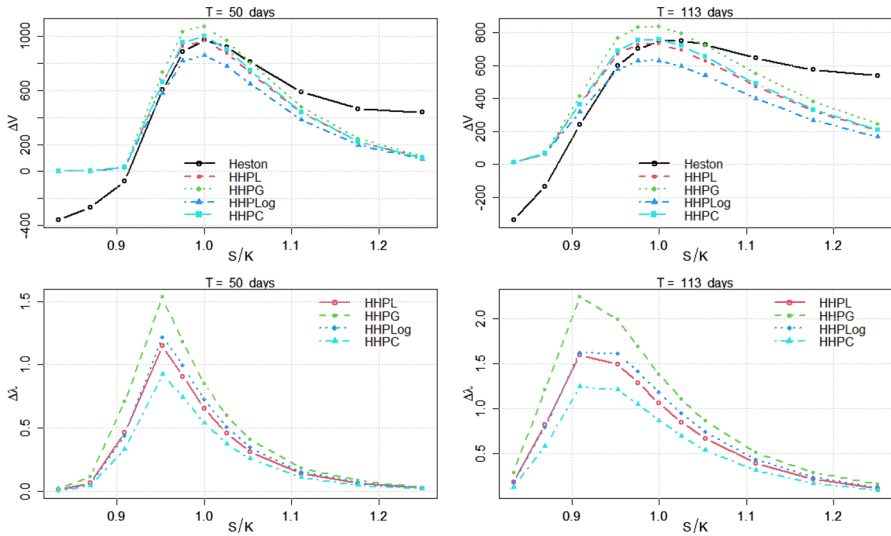
Fig. 14 Vega and jump risk sensitivity of call prices expiring at 50 and 113 days. Model parameters used are in Table 1

call option price according to the delta probability decomposition and derive Greeks for hedging purpose given the three sources of risk: the change in underlying price, volatility and the occurrence of a jump. By inverse Fourier transform of the conditional moment generating function of the log-returns, we find that with equal parameters, the thickness of the tails of its probability density function is fatter if the length of the memory of the HP kernel function is greater.

An estimation of our models to implied volatilities of the Euro Stoxx 50 as of the 26th of September 2019 achieves better calibration performance than the one with the standard Heston model, particularly in the short term. They also provide a good fit of the term structure of at-the-money volatility skew, approximted by power-law functions of the form $T^{\hat{H}-\frac{1}{2}}$ with an estimated Hurst exponent $\hat{H}$ between 0.09 and 0.14. Considering the estimated parameters found and varying the memory parameter, we see that implied volatilities derived from our extensions of the Heston model, increase as the memory length expands. This sensitivity to the memory length appears more pronounced for out-of-the money call option prices. A similar conclusion is drawn, with respect to the sensitivity of the implies volatilities to other self-excitation parameters. The analysis of the Greeks delta, vega and the jump risk reveals that our extensions of the Heston model have roughly the same delta curve than the standard Heston. But these models induce different shapes of sensitivity to volatility and jump risks in the option price, which does not seem to be related only to their memory lengths.

A next step to this work will be to explore hedging strategies in our setting, to manage options whose underlying asset may experience some clustering of jumps. From a broader perspective, it would be interesting to extend this work to the pricing of bond, currency, exotic options or exotic derivatives which are path dependent. From

an empirical point of view, it would be worthy to study the empirical relevance of such Hawkes memory kernels and develop accurated calibration procedure to financial time series.

Appendix A

Proof of proposition 1 For fixed time s and $\omega \in \mathbb{C}_-$ with $\mathbb{C}_-$ the set of complex numbers with non-positive real part, we define the conditional moment generating function by

$$
f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right) := \mathbb{E}\left(e^{\omega \lambda_s^{(n)}} | \mathcal{F}_t\right),
$$

where f is a complex-valued function on $[0, s] \times G$ with $G = \mathbb{R}_+ \times \mathbb{R}_+^n \times \mathbb{N}$ the state space of the Markov process $\left(\lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right)_{t \geq 0}$. By the tower property, for any $\tau \geq t$,

$$
f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right) = \mathbb{E}\left(f\left(\tau, s, \lambda_\tau^{(n)}, \left(Y_\tau^{(l)}\right)_{l=-n,\dots,n}, N_\tau^{(n)}\right) | \mathcal{F}_t\right),
$$

and

$$
\lim_{\tau \to t} \frac{\mathbb{E}\left(f\left(\tau, s, \lambda_\tau^{(n)}, \left(Y_\tau^{(l)}\right)_{l=-n,\dots,n}, N_\tau^{(n)}\right) | \mathcal{F}_t\right) - f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right)}{\tau - t} = 0.
$$

By definition, the infinitesimal generator $\mathcal{A}$ of the Markov process $\left(\lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right)$ evaluated on f is equal to

$$
\lim_{\tau \to t} \frac{\mathbb{E}\left(f\left(\tau, s, \lambda_\tau^{(n)}, \left(Y_\tau^{(l)}\right)_{l=-n,\dots,n}, N_\tau^{(n)}\right) | \mathcal{F}_t\right) - f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right)}{l - t}.
$$

It follows that

$$
\mathcal{A}f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right) = 0. \tag{A1}
$$

Since $\mathcal{A}f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right)$ exists the function f is in the domain of $\mathcal{A}$ which contains the set of twice differentiable functions such that all partial derivatives of up to order two are continuous function (see e.g. Duffie et al. (2003)).

On the other hand, the Itô's lemma for the semi-martingales allows us to infer that f satisfies the following differential equation:

$$
0 = \frac{\partial f}{\partial t} + i \sum_{l=-n}^{n} \frac{\partial f}{\partial Y^{(l)}} b_l^{(n)} Y_t^{(l)} + \left((\lambda_0 - \alpha) \frac{dg_k(t)}{dt} + \eta \sum_{l=-n}^{n} m_l^{(k,n)} i b_l^{(n)} Y_t^{(l)} \right) \frac{\partial f}{\partial \lambda^{(n)}}
$$
$$
+ \lambda_t^{(n)} \int_{\mathbb{R}} \left(f \left(t, s, \lambda_t^{(n)} + \eta |z| \sum_{l=-n}^{n} m_l^{(k,n)}, \left(Y_t^{(l)} + |z| \right)_{l=-n,\ldots,n}, N_t^{(n)} + 1 \right) - f(.) \right) \nu(dz).
$$
$$
\tag{A2}
$$

Given that $\left(\lambda_t^{(n)}, \left(Y_t^{(l)} \right)_{l=-n,\ldots,n}, N_t^{(n)} \right)$ is an affine jump process, from Duffie et al. (2000) and Errais et al. (2010), under technical regularity condition, we conjecture that f is an exponential affine function of $\lambda_t^{(n)}$ and $\left(Y_t^{(l)} \right)_{l=-n,\ldots,n}$:

$$
f \left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)} \right)_{l=-n,\ldots,n}, N_t^{(n)} \right) := \exp \left(F(t, s, \omega) + G(t, s, \omega) \lambda_t^{(n)} \right)
$$
$$
\times \exp \left(\sum_{l=-n}^{n} H_l \left(t, s, b_l^{(n)}, \omega \right) m_l^{(k,n)} Y_t^{(l)} \right)
$$

where F and G are time dependent functions such that $F(s, s, \omega) = 0$ and $G(s, s, \omega) = \omega$. While H_l is a function that depends on the time and the position $l \in \{-n, \ldots, n\}$ within the partition $\mathcal{E}^{(n)}$. Its final value $H_l \left(s, s, b_l^{(n)}, \omega \right)$ is equal to zero. From (Hubalek et al. 2016), the technical regularity condition needed is equivalent to the integrability of the jumps, that is

$$
\left. \frac{\partial G}{\partial t}(t, s, \omega) \right|_{t=s} = \mathbb{E} \left(e^{\eta \omega \sum_{l=-n}^{n} m_l^{(k,n)} |J|} \right) < \infty,
$$

where the random jump size J has density $\nu(z)$ on $\mathbb{R}$. The partial derivatives of f are then given by:

$$
\frac{\partial f}{\partial t} = \left(\frac{\partial F}{\partial t}(t, s, \omega) + \frac{\partial G}{\partial t}(t, s, \omega) \lambda_t^{(n)} + \sum_{l=-n}^{n} \frac{\partial H_l}{\partial t} \left(t, s, b_l^{(n)}, \omega \right) m_l^{(k,n)} Y_t^{(l)} \right) f,
$$
$$
\frac{\partial f}{\partial \lambda^{(n)}} = G(t, s, \omega) f, \quad \frac{\partial f}{\partial Y^{(l)}} = H_l \left(t, s, b_l^{(n)}, \omega \right) m_l^{(k,n)} f \quad \forall l \in \{-n, \ldots, n\},
$$

and the variation of f when a jump of size z occurs is equal to

$$
\left(\exp \left(|z| \left(\eta G(t, s, \omega) \sum_{l=-n}^{n} m_l^{(k,n)} + \sum_{l=-n}^{n} H_l \left(t, s, b_l^{(n)}, \omega \right) m_l^{(k,n)} \right) \right) - 1 \right) f.
$$

Injecting all these expressions into Eq. (A2) lead to:

$$
\begin{aligned}
0 = & \frac{\partial F}{\partial t}(t, s, \omega) + \frac{\partial G}{\partial t}(t, s, \omega)\lambda_t^{(n)} + \sum_{l=-n}^{n} \frac{\partial H_l}{\partial t}\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} Y_t^{(l)} \\
& + i \sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} b_l^{(n)} Y_t^{(l)} \\
& + \left((\lambda_0 - \alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t} + \eta \sum_{l=-n}^{n} m_l^{(k,n)} i b_l^{(n)} Y_t^{(l)}\right) G(t, s, \omega) \\
& + \lambda_t^{(n)}\left(\psi\left(0, \eta G(t, s, \omega)\sum_{l=-n}^{n} m_l^{(k,n)} + \sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)}\right) - 1\right)
\end{aligned}
$$
$$\tag{A3}$$

where $\psi(z_1, z_2) := \mathbb{E}\left(e^{z_1 J + z_2 |J|}\right)$. From which we find the following ODE system:

$$
\begin{cases}
0 = \frac{\partial F}{\partial t}(t, s, \omega) + (\lambda_0 - \alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t} G(t, s, \omega) \\
0 = \frac{\partial G}{\partial t}(t, s, \omega) + \psi\left(0, \eta G(t, s, \omega)\sum_{l=-n}^{n} m_l^{(k,n)} + \sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)}\right) - 1 \\
0 = \frac{\partial H_l}{\partial t}\left(t, s, b_l^{(n)}, \omega\right) + i b_l^{(n)} H_l\left(t, s, b_l^{(n)}, \omega\right) + \eta i b_l^{(n)} G(t, s, \omega), \ \forall l \in \{-n, \dots, n\}.
\end{cases}
$$

For all $l \in \{-n, \dots, n\}$, the last ODE admits the following solution:

$$
H_l\left(t, s, b_l^{(n)}, \omega\right) = b_l^{(n)} \eta i \int_t^s e^{i b_l^{(n)}(u-t)} G(u, s, \omega)\, \mathrm{d}u,
$$

and the proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Appendix B

Proof of Theorem 3 Let $\omega \in \mathbb{R}$. As in Proposition 1, if we note

$$
f\left(t, s, X_t^{(n)}, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right) := \mathbb{E}\left(e^{i\omega X_s^{(n)}} \mid \mathcal{F}_t\right)
$$

and under the regularity condition $\psi\left(i\omega, \eta \sum_{l=-n}^{n} m_l^{(k,n)}\right) < \infty$, we conjecture that f is an exponential affine function of $X_t^{(n)}$, $\lambda_t^{(n)}$ and $\left(Y_t^{(l)}\right)_{l=-n,\dots,n}$,

$$f\left(t, s, X_t^{(n)}, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\ldots,n}, N_t^{(n)}\right) := \exp\left(i\omega X_t^{(n)} + A(t, s, \omega) + B(t, s, \omega) V_t\right)$$

$$\times \exp\left(\sum_{l=-n}^{n} C_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} Y_t^{(l)}\right)$$

$$\times \exp\left(D(t, s, \omega) \lambda_t^{(n)}\right) \tag{B1}$$

where $A(s, s, \omega) = 0$, $B(s, s, \omega) = 0$, $D(s, s, \omega) = 0$ and $C_l\left(s, s, b_l^{(n)}, \omega\right) = 0$, for all $l \in \{-n, \cdots, n\}$. Then, by the Itô's lemma for the semi-martingales, f solves the following differential equation:

$$0 = \frac{\partial f}{\partial t} + \left(r - \frac{1}{2}V_t - \lambda_t^{(n)}(\psi(1,0) - 1)\right)\frac{\partial f}{\partial X^{(n)}} + \frac{1}{2}V_t\frac{\partial^2 f}{\partial (X^{(n)})^2} + \rho\sigma_V V_t\frac{\partial^2 f}{\partial X^{(n)}\partial V}$$

$$+ (\theta_V - k_V V_t)\frac{\partial f}{\partial V} + \frac{1}{2}\sigma_V^2 V_t\frac{\partial^2 f}{\partial V^2} + \sum_{l=-n}^{n}\frac{\partial f}{\partial Y^{(l)}}ib_l^{(n)}Y_t^{(l)} \tag{B2}$$

$$+ \frac{\partial f}{\partial \lambda^{(n)}}\left((\lambda_0 - \alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t} + \eta\sum_{l=-n}^{n}m_l^{(k,n)}ib_l^{(n)}Y_t^{(l)}\right)$$

$$+ \lambda_t^{(n)}\int_{\mathbb{R}}\Delta f(z)\nu(\mathrm{d}z),$$

where

$$\frac{\partial f}{\partial t} = \left(\frac{\partial A}{\partial t} + \frac{\partial B}{\partial t}V_t + \sum_{l=-n}^{n}\frac{\partial C_l}{\partial t}m_l^{(k,n)}Y_t^{(l)} + \frac{\partial D}{\partial t}\lambda_t^{(n)}\right)f, \quad \frac{\partial f}{\partial X^{(n)}} = i\omega f,$$

$$\frac{\partial^2 f}{\partial (X^{(n)})^2} = -\omega^2 f, \qquad\qquad\qquad \frac{\partial^2 f}{\partial X^{(n)}\partial V} = i\omega B f,$$

$$\frac{\partial f}{\partial V} = Bf, \qquad\qquad\qquad\qquad \frac{\partial^2 f}{\partial V^2} = B^2 f,$$

$$\frac{\partial f}{\partial Y^{(l)}} = C_l m_l^{(k,n)} f, \qquad\qquad\qquad \frac{\partial f}{\partial \lambda^{(n)}} = Df,$$

and $\Delta f(z)$ is the variation of the function f when a jump of size z occurs

$$\Delta f(z) = f\left(t, s, X_t^{(n)} + z, V_t, \lambda^{(n)} + \eta|z|\sum_{l=-n}^{n}m_l^{(k,n)}, \left(Y_t^{(l)}\right)_{l=-n,\ldots,n} + |z|, N_t^{(n)} + 1\right)$$

$$- f\left(t^-, s, X_{t^-}^{(n)}, V_{t^-}, \lambda_{t^-}^{(n)}, \left(Y_{t^-}^{(l)}\right)_{l=-n,\ldots,n}, N_{t^-}^{(n)}\right) \tag{B3}$$

$$= \left(\exp\left(i\omega z + \left(\sum_{l=-n}^{n}C_l\left(t, s, b_l^{(n)}, \omega\right)m_l^{(k,n)} + \eta D(t, s, \omega)\right)|z|\right) - 1\right)f. \tag{B4}$$

Injecting all these expressions into Eq. (B2) lead to:

$$
\begin{cases}
\dfrac{\partial A}{\partial t} = -i\omega r - \theta_V B - D\left(\lambda_0 - \alpha\right) \dfrac{\mathrm{d}g_k\,(t)}{\mathrm{d}t} \\[2ex]
\dfrac{\partial D}{\partial t} = i\omega\left(\psi\,(1,0) - 1\right) - \displaystyle\int_{\mathbb{R}} \left(\exp\left(i\omega z + \left(\sum_{l=-n}^{n} C_l m_l^{(k,n)} + \eta D\right)|z|\right) - 1\right)\nu\,(\mathrm{d}z) \\[2ex]
\dfrac{\partial B}{\partial t} = \bar{a} - \bar{b}B + \bar{c}B^2 \\[2ex]
\dfrac{\partial C_l}{\partial t} = -\left(C_l + \eta D\right) i b_l^{(n)}, \forall l \in \{-n, \dots n\}
\end{cases}
\tag{B5}
$$

where $\bar{a} = \frac{\omega}{2}\left(\omega + i\right), \bar{b} = -k_V + i\omega\rho\sigma_V, \bar{c} = -\frac{1}{2}\sigma_V^2$. In order to separate the part of the characteristic function linked to the jump component we assume that $A\,(t, s, \omega) = A_D\,(t, s, \omega) + A_J\,(t, s, \omega)$, with terminal condition $A_D\,(s, s, \omega) = 0 = A_J\,(s, s, \omega)$. It follows that the ODEs system (B5) can be subdivided into two ODEs systems:

$$
\begin{cases}
\dfrac{\partial A_D}{\partial t} = -i\omega r - \theta_V B \\[2ex]
\dfrac{\partial B}{\partial t} = \bar{a} - \bar{b}B + \bar{c}B^2
\end{cases}
\tag{B6}
$$

and

$$
\begin{cases}
\dfrac{\partial A_J}{\partial t} = -D\left(\lambda_0 - \alpha\right)\dfrac{\mathrm{d}g_k\,(t)}{\mathrm{d}t} \\[2ex]
\dfrac{\partial C_l}{\partial t} = -\left(C_l + \eta D\right) i b_l^{(n)}, \forall l \in \{-n, \dots n\} \\[2ex]
\dfrac{\partial D}{\partial t} = i\omega\left(\psi\,(1,0) - 1\right) - \displaystyle\int_{\mathbb{R}} \left(\exp\left(i\omega z + \left(\sum_{l=-n}^{n} C_l m_l^{(k,n)} + \eta D\right)|z|\right) - 1\right)\nu\,(\mathrm{d}z)
\end{cases}
\tag{B7}
$$

The ODEs system (B6) corresponds at the one of the standard Heston model. From Schoutens et al. (2004)), we have:

$$
\begin{aligned}
\exp\left(i\omega X_t^{(n)} + A_D\,(t, s, \omega) + B\,(t, s, \omega)\,V_t\right) &= \exp\left(i\omega X_t^{(n)} + r\,(s - t)\right) \\
&\times \exp\left(\theta_V \sigma_V^{-2}\left(k_V - i\omega\rho\sigma_V - \bar{d}\right)(s - t)\right. \\
&\left. -2\log\left(\frac{1 - \bar{g}e^{-\bar{d}(s-t)}}{1 - \bar{g}}\right)\right) \\
&\times \exp\left(V_t \sigma_V^{-2}\left(k_V - i\omega\rho\sigma_V - \bar{d}\right)\right. \\
&\left.\times \frac{1 - e^{-\bar{d}(s-t)}}{1 - \bar{g}e^{-\bar{d}(s-t)}}\right)
\end{aligned}
$$

where

$$\bar{d} = \sqrt{(i\omega\rho\sigma_V - k_V)^2 + \sigma_V^2 \left(i\omega + \omega^2\right)},$$

$$\bar{g} = \frac{k_V - i\omega\rho\sigma_V - \bar{d}}{\left(k_V - i\omega\rho\sigma_V + \bar{d}\right)}.$$

Albrecher et al. (2006) prove that this form of the characteristic function of the standard Heston model is stable for Fourier inversion or numerical integration.

Let us focus on the ODEs system (B7). Its last ODE admits the following solution:

$$C_l \left(t, s, b_l^{(n)}, \omega\right) = i\eta b_l^{(n)} \int_t^s e^{ib_l^{(n)}(u-t)} D (u, s, \omega) \, \mathrm{d}u. \tag{B8}$$

This last formula allows us to conclude the proof. $\qquad\square$

Appendix C

Proof of Theorem 5 The derivative

$$\left.\frac{\mathrm{d}\mathbb{Q}^\star}{\mathrm{d}\mathbb{Q}}\right|_{\mathcal{F}_0} = \exp\left(-\frac{1}{2} \int_0^t V_u \mathrm{d}u + \int_0^t \sqrt{V_u} \left(\rho \mathrm{d}W_u^1 + \sqrt{1-\rho^2}\mathrm{d}W_u^2\right)\right.$$

$$\left. - (\psi(1,0) - 1) \int_0^t \lambda_u \mathrm{d}u + \sum_{l=1}^{N_u} J_l\right)$$

is an exponential martingale as a solution of a driftless SDE and given that $\mathbb{E}\left(\int_0^\tau \lambda_s \mathrm{d}s\right) < \infty$ and $\mathbb{E}\left(\int_0^\tau V_s^2 \mathrm{d}s\right) < \infty$ for all $\tau \geq 0$.

i) is straightforward. Hence, we start by the proof of ii). The moment generating function of $\lambda_s^{(n),\star}$ conditionally to the information at time $t \leq s$ is given by

$$\mathbb{E}^\star \left(e^{\omega \lambda_s^{(n),\star}}|\mathcal{F}_t\right) = e^{U_t^{(n)}} \mathbb{E} \left(e^{U_s^{(n)} + \omega\psi(1,0)\lambda_s^{(n)}}|\mathcal{F}_t\right),$$

$$\exp\left(F(t,s) + G(t,s)\lambda_t^{(n)} + \sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}\right) m_l^{(k,n)} Y_t^{(l)}\right)$$

where

$$U_t^{(n)} = -\frac{1}{2}\int_0^t V_u \mathrm{d}u + \int_0^t \sqrt{V_u}\left(\rho \mathrm{d}W_u^1 + \sqrt{1-\rho^2}\mathrm{d}W_u^2\right)$$

$$-\mathbb{E}\left(e^J - 1\right)\int_0^t \lambda_u^{(n)}\mathrm{d}u + \sum_{l=1}^{N_u^{(n)}} J_l^{(n)}$$

is the approximated exponent of the exponential martingale $\frac{d\mathbb{Q}^\star}{d\mathbb{Q}}\Big|_{\mathcal{F}_0}$.

If we note

$$f\left(t,s,U_t^{(n)},\lambda_t^{(n)},\left(Y_t^{(l)}\right)_{l=-n,\dots,n},N_t^{(n)}\right) := \mathbb{E}\left(e^{U_s^{(n)}+\omega\psi(1,0)\lambda_s^{(n)}}|\mathcal{F}_t\right)$$

and we conjecture that f is an exponential affine function of $U_t^{(n)}$ $\lambda_t^{(n)}$ and $\left(Y_t^{(v)}\right)_{v=1,\dots,n}$,

$$f\left(t,s,U_t^{(n)},\lambda_t^{(n)},\left(Y_t^{(l)}\right)_{l=-n,\dots,n},N_t^{(n)}\right) := \exp\left(A(t,s,\omega)+\psi(1,0)B(t,s,\omega)\lambda_t^{(n)}\right)$$

$$\times \exp\left(\sum_{l=-n}^{n}C_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}Y_t^{(l)}\right.$$

$$+ D(t,s,\omega)U_t^{(n)}\Big)$$

where $A(t,s,\omega)=0$, $B(s,s,\omega)=\omega$, $C_l\left(s,s,b_l^{(n)},\omega\right)=0$ for all $l\in\{-n,\dots,n\}$ and $D(s,s,\omega)=1$. According to the Itô's lemma, f solves the following differential equation:

$$0 = \frac{\partial f}{\partial t}+\frac{\partial f}{\partial U^{(n)}}\left(-(\psi(1,0)-1)\lambda_t^{(n)}-\frac{1}{2}V_t\right) \tag{C1}$$

$$+\frac{1}{2}\frac{\partial^2 f}{\partial U^{(n)2}}V_t+i\sum_{l=-n}^{n}\frac{\partial f}{\partial Y^{(l)}}b_l^{(n)}Y_t^{(l)}$$

$$+\frac{\partial f}{\partial\lambda^{(n)}}\left((\lambda_0-\alpha)\frac{dg_k(t)}{dt}+\eta\sum_{l=-n}^{n}m_l^{(k,n)}ib_l^{(n)}Y_t^{(l)}\right)$$

$$+\lambda_t^{(n)}\int_{-\infty}^{\infty}\left(f\left(t,s,U_t^{(n)}+z,\lambda_t^{(n)}+\eta|z|\sum_{l=-n}^{n}m_l^{(k,n)},\left(Y_t^{(l)}+|z|\right)_{l=-n,\dots,n},N_t^{(n)}+1\right)\right.$$

$$-f(.)\Big)\nu(dz)$$

where $\frac{\partial f}{\partial U^{(n)}}=D(t,s,\omega)f$, $\frac{\partial f}{\partial\lambda^{(n)}}=\psi(1,0)B(t,s,\omega)f$, $\frac{\partial f}{\partial Y^{(l)}}=C_l\left(t,s,b_l^{(n)},\omega\right)$ $m_l^{(k,n)}f$, $\frac{\partial^2 f}{\partial U^{(n)2}}=D(t,s,\omega)^2 f$,

$$\frac{\partial f}{\partial t}=\left(\frac{\partial A}{\partial t}(t,s,\omega)+\psi(1,0)\frac{\partial B}{\partial t}(t,s,\omega)\lambda_t^{(n)}+\sum_{l=-n}^{n}\frac{\partial C_l}{\partial t}\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}Y_t^{(v)}\right.$$

$$+\frac{\partial D}{\partial t}(t,s,\omega)U_t^{(n)}\Big)f,$$

and the variation of f when a jump of size z occurs is

$$\left(\exp\left(D\left(t, s, \omega\right) z + \left(\eta \psi\left(1, 0\right) B\left(t, s, \omega\right) \sum_{l=-n}^{n} m_l^{(k,n)} \right. \right. \right.$$
$$\left. \left. \left. + \sum_{l=-n}^{n} C_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} \right) |z| \right) - 1 \right) f$$

Injecting all these expressions into Eq. (C1) lead to:

$$0 = \frac{\partial A}{\partial t}\left(t, s, \omega\right) + \psi\left(1, 0\right) \frac{\partial B}{\partial t}\left(t, s, \omega\right) \lambda_t^{(n)} + \sum_{l=-n}^{n} \frac{\partial C_l}{\partial t}\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} Y_t^{(v)}$$
$$+ \frac{\partial D}{\partial t}\left(t, s, \omega\right) U_t^{(n)} + D\left(t, s, \omega\right) \left(-\left(\psi\left(1, 0\right) - 1\right) \lambda_t^{(n)} - \frac{1}{2} V_t \right) + \frac{1}{2} D\left(t, s, \omega\right)^2 V_t$$
$$+ i \sum_{l=-n}^{n} C_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} b_l^{(n)} Y_t^{(l)}$$
$$+ \psi\left(1, 0\right) B\left(t, s, \omega\right) \left(\left(\lambda_0 - \alpha\right) \frac{dg_k\left(t\right)}{dt} + \eta \sum_{l=-n}^{n} m_l^{(k,n)} i b_l^{(n)} Y_t^{(l)} \right) + \lambda_t^{(n)}$$
$$\times \left(\psi\left(D\left(t, s, \omega\right), \eta \psi\left(1, 0\right) B\left(t, s, \omega\right) \sum_{l=-n}^{n} m_l^{(k,n)} + \sum_{l=-n}^{n} C_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} \right) - 1 \right)$$

or

$$0 = \frac{\partial A}{\partial t}\left(t, s, \omega\right) - \frac{1}{2} D\left(t, s, \omega\right) V_t + \frac{1}{2} D\left(t, s, \omega\right)^2 V_t \qquad (C2)$$
$$+ \psi\left(1, 0\right) B\left(t, s, \omega\right) \left(\lambda_0 - \alpha\right) \frac{dg_k\left(t\right)}{dt}$$

$$0 = \psi\left(1, 0\right) \frac{\partial B}{\partial t}\left(t, s, \omega\right) - D\left(t, s, \omega\right) \left(\psi\left(1, 0\right) - 1\right) \qquad (C3)$$
$$+ \psi\left(D\left(t, s, \omega\right), \eta \psi\left(1, 0\right) B\left(t, s, \omega\right) \sum_{l=-n}^{n} m_l^{(k,n)} + \sum_{l=-n}^{n} C_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} \right) - 1,$$

$$0 = \frac{\partial C_l}{\partial t}\left(t, s, b_l^{(n)}, \omega\right) + i \left(C_l\left(t, s, b_l^{(n)}, \omega\right) + \eta \psi\left(1, 0\right) B\left(t, s, \omega\right) \right) b_l^{(n)}, \qquad (C4)$$
$$\forall l \in \{-n, \dots, n\}$$

$$0 = \frac{\partial D}{\partial t}\left(t, s, \omega\right). \qquad (C5)$$

$D\left(t, s, \omega\right) = 1$ solves the last PDE as $D\left(s, s, \omega\right) = 1$. Given $D\left(t, s, \omega\right) = 1$, the functions A and B satisfy the following PDE system:

$$\begin{cases} \dfrac{\partial A}{\partial t}\,(t,s,\omega) = -\psi\,(1,0)\,B\,(t,s,\omega)\,(\lambda_0 - \alpha)\,\dfrac{\mathrm{d}g_k\,(t)}{\mathrm{d}t}, \\[4mm] \dfrac{\partial B}{\partial t}\,(t,s,\omega) = 1 - \dfrac{\psi\left(1,\eta\psi\,(1,0)\,B\,(t,s,\omega)\sum_{l=-n}^{n} m_l^{(k,n)} + \sum_{l=-n}^{n} C_l\left(t,s,b_l^{(n)}\right) m_l^{(k,n)}\right)}{\psi\,(1,0)}. \end{cases}$$

$$\tag{C6}$$

Let $\lambda_0^\star = \psi\,(1,0)\,\lambda_0$, $\alpha^\star = \psi\,(1,0)\,\alpha$, $\eta^\star = \psi\,(1,0)\,\eta$ and

$$\psi^\star\,(z_1, z_2) = \frac{\psi\,(z_1 + 1, z_2)}{\psi\,(1,0)}. \tag{C7}$$

Taking into account these definitions, the latter ODE system is similar to the one obtained under risk neutral measure $\mathbb{Q}$ in Proposition 1. On the other hand, for all $l \in \{-n, \ldots, n\}$,

$$C_l\left(t,s,b_l^{(n)},\omega\right) = i b_v^{(n)} \eta^\star \int_t^s e^{i b_l^{(n)}(u-t)} B\,(u,s,\omega)\,\mathrm{d}u,$$

solves Eq. C4. By a suitable rearrangement, we can show that this definition of the joint moment generating function $\psi^\star$ corresponds at the one of the pair $J^\star$ and $|J^\star|$ where the jump size $J^\star$ follows a double exponential with parameters

$$\rho^{+,\star} = \rho^+ - 1,$$
$$\rho^{-,\star} = \rho^- - 1,$$
$$p^\star = \frac{p\rho^+ \rho^{-,\star}}{p\rho^+ \rho^{-,\star} + \rho^{+,\star}\,(1 - p)\,\rho^-}.$$

All the previous results allows us to complete the proof of ii). The SDE of the variance V_t under $\mathbb{Q}^\star$ is derived from i). While for the process X_t, similarly to the proof of ii), the dynamics of X under the measure $\mathbb{Q}^\star$ is obtained by comparison with the results of Proposition 3. $\qquad\square$

Appendix D Expected value of $X_T \mid \mathcal{F}_t$

Given the dynamics of X_T under the measure $\mathbb{Q}$, we have

$$\mathbb{E}\,(X_T \mid \mathcal{F}_t) = \mathbb{E}\left(\int_t^T \left(r - \frac{1}{2}V_u\right) \mid \mathcal{F}_t\right) \tag{D1}$$
$$= r\,(T - t) - \frac{1}{2}\mathbb{E}\left(\int_t^T V_u \mathrm{d}u \mid \mathcal{F}_t\right),$$

where the solution of the SDE of the variance process is given by

$$V_T = \frac{\theta_V}{k_V} + \left(V_t - \frac{\theta_V}{k_V}\right) e^{-k_V(T-t)} + \sigma_V \int_t^T e^{-k_V(T-s)}\sqrt{V_s}\mathrm{d}W_s^1. \tag{D2}$$

Applying the Fubini rule leads to

$$\int_t^T V_u du = \frac{\theta_V}{k_V}(T-t) + \left(V_t - \frac{\theta_V}{k_V}\right)\frac{1}{k_V}\left(1 - e^{-k_V(T-t)}\right)$$

$$+ \frac{\sigma_V}{k_V}\int_t^u \sqrt{V_s}\left(e^{-k_V(t-s)} - e^{-k_V(T-s)}\right)dW_s^1.$$

(D3)

Hence

$$\mathbb{E}(X_T \mid \mathcal{F}_t) = r(T-t) - \frac{1}{2}\frac{\theta_V}{k_V}(T-t) - \frac{1}{2}\left(V_t - \frac{\theta_V}{k_V}\right)\frac{1}{k_V}\left(1 - e^{-k_V(T-t)}\right).$$

(D4)

Under the measure $\mathbb{Q}^\star$, the SDE of X_t and V_t are given by:

$$dX_t = \left(r + \frac{1}{2}V_t - \lambda_t^\star\psi^\star(1,0)\right)dt$$

(D5)

$$+ \sqrt{V_t}\left(\rho dW_t^{\star 1} + \sqrt{1-\rho^2}dW_t^{\star 2}\right) + d\left(\sum_{l=1}^{N_t^\star} J_l^\star\right),$$

$$dV_t = (\theta_V - (k_V - \rho\sigma_V)V_t)dt + \sigma_V\sqrt{V_t}dW_t^{\star 1}.$$

(D6)

It follows that

$$\mathbb{E}^\star(X_T \mid \mathcal{F}_t) = \mathbb{E}^\star\left(\int_t^T\left(r + \frac{1}{2}V_u\right) \mid \mathcal{F}_t\right)$$

(D7)

$$= r(T-t) + \frac{1}{2}\mathbb{E}^\star\left(\int_t^T V_u du \mid \mathcal{F}_t\right)$$

$$= r(T-t) + \frac{1}{2}\frac{\theta_V}{k_V - \rho\sigma_V}(T-t)$$

$$+ \left(V_t - \frac{\theta_V}{k_V - \rho\sigma_V}\right)\frac{1}{2(k_V - \rho\sigma_V)}\left(1 - e^{-(k_V - \rho\sigma_V)(T-t)}\right)$$

where as under the measure $\mathbb{Q}$

$$\int_t^T V_u du = \frac{\theta_V}{k_V - \rho\sigma_V}(T-t)$$

$$+ \left(V_t - \frac{\theta_V}{k_V - \rho\sigma_V}\right)\frac{1}{k_V - \rho\sigma_V}\left(1 - e^{-(k_V - \rho\sigma_V)(T-t)}\right)$$

$$+ \frac{\sigma_V}{k_V - \rho\sigma_V}\int_t^u \sqrt{V_s}\left(e^{-(k_V - \rho\sigma_V)(t-s)} - e^{-(k_V - \rho\sigma_V)(T-s)}\right)dW_s^{\star 1}.$$

(D8)

References

Aït-Sahalia, Y., Lo, A.W.: Nonparametric estimation of state-price densities implicit in financial asset prices. J Financ **53**(2), 499–547 (1998). https://doi.org/10.1111/0022-1082.215228

Aït-Sahalia, Y., Li, C., Li, C.X.: Closed-form implied volatility surfaces for stochastic volatility models with jumps. J Econom **222**(1), 364–392 (2021). https://doi.org/10.1016/j.jeconom.2020.07.006

Albrecher, H., Mayer, P.A., Schoutens, W., et al.: The Little Heston Trap, pp. 83–92. Wilmott (2006)

Bakshi, G., Cao, C., Chen, Z.: Empirical performance of alternative option pricing models. J Financ **52**(5), 2003–2049 (1997). https://doi.org/10.1111/j.1540-6261.1997.tb02749.x

Bates, D.S.: The crash of '87: Was it expected? The evidence from options markets. J Financ **46**(3), 1009–1044 (1991). https://doi.org/10.1111/j.1540-6261.1991.tb03775.x

Bates, D.S.: Post-'87 crash fears in the S&P 500 futures option market. J Econom **94**(1–2), 181–238 (2000). https://doi.org/10.1016/s0304-4076(99)00021-4

Black, F., Scholes, M.: The pricing of options and corporate liabilities. J Polit Econ **81**(3), 637–654 (1973). https://doi.org/10.1086/260062

Blanc, P., Donier, J., Bouchaud, J.P.: Quadratic Hawkes processes for financial prices. Quant Financ **17**(2), 171–188 (2016). https://doi.org/10.1080/14697688.2016.1193215

Christoffersen, P., Jacobs, K., Mimouni, K.: An empirical comparison of affine and non-affine models for equity index options. SSRN Electron. J. (2006). https://doi.org/10.2139/ssrn.891127

Comte, F., Coutin, L., Renault, E.: Affine fractional stochastic volatility models. Ann Financ **8**(2–3), 337–378 (2010). https://doi.org/10.1007/s10436-010-0165-3

Cont, R.: Empirical properties of asset returns: stylized facts and statistical issues. Quant Financ **1**(2), 223 (2001). https://doi.org/10.1080/713665670

Cox, J.C., Ingersoll, J.J.E., Ross, S.A.: A theory of the term structure of interest rates. Econometrica **53**, 385–408 (1985). https://doi.org/10.2307/1911242

Duffie, D., Filipović, D., Schachermayer, W.: Affine processes and applications in finance. Ann Appl Probab **13**(3), 1 (2003). https://doi.org/10.1214/aoap/1060202833

Duffie, D., Pan, J., Singleton, K.: Transform analysis and asset pricing for affine jump diffusions. Econometrica **68**, 1343–1376 (2000). https://doi.org/10.1111/1468-0262.00164

Eraker, B.: Do stock prices and volatility jump? Reconciling evidence from spot and option prices. J Financ **59**(3), 1367–1403 (2004). https://doi.org/10.1111/j.1540-6261.2004.00666.x

Errais, E., Giesecke, K., Goldberg, L.R.: Affine point processes and portfolio credit risk. SIAM J Financ Math **1**(1), 642–665 (2010). https://doi.org/10.1137/090771272

Fonseca, J.D., Zaatour, R.: Hawkes process: fast calibration, application to trade clustering and diffusive limit. SSRN Electron J (2013). https://doi.org/10.2139/ssrn.2294112

Gatheral, J.: The Volatility Surface. Wiley (2012). https://doi.org/10.1002/9781119202073

Gatheral, J., Jaisson, T., Rosenbaum, M.: Volatility is rough. Quant Financ **18**, 933–949 (2014). https://doi.org/10.1080/14697688.2017.1393551

Gil-Pelaez, J.: Note on the inversion theorem. Biometrika **38**(3–4), 481–482 (1951). https://doi.org/10.1093/biomet/38.3-4.481

Hainaut, D.: Moment generating function of non-Markov self-excited claims processes. Insurance: Math Econ **101**, 406–424 (2021). https://doi.org/10.1016/j.insmatheco.2021.08.013

Hainaut, D., Moraux, F.: Hedging of options in the presence of jump clustering. J Comput Financ (2018). https://doi.org/10.21314/jcf.2018.354

Hainaut, D., Moraux, F.: A switching self-exciting jump diffusion process for stock prices. Ann Financ **15**(2), 267–306 (2018). https://doi.org/10.1007/s10436-018-0340-5

Hamida, S.B., Cont, R.: Recovering volatility from option prices by evolutionary optimization. J Comput Financ **8**(4), 43–76 (2005). https://doi.org/10.21314/jcf.2005.130

Hawkes, A.G.: Point spectra of some mutually exciting point processes. J R Stat Soc Ser B **33**, 438–443 (1971). https://doi.org/10.1111/j.2517-6161.1971.tb01530.x

Hawkes, A.G.: Spectra of some self-exciting and mutually exciting point processes. Biometrika **58**, 83–90 (1971). https://doi.org/10.1093/biomet/58.1.83

Heston, S.L.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev Financ Stud **6**(2), 327–343 (1993). https://doi.org/10.1093/rfs/6.2.327

Hubalek, F., Keller-Ressel, M., Sgarra, C.: Geometric Asian option pricing in general affine stochastic volatility models with jumps. Quant Financ **17**(6), 873–888 (2016). https://doi.org/10.1080/14697688.2016.1256495

Ignatieva, K., Rodrigues, P., Seeger, N.: Stochastic volatility and jumps: exponentially affine yes or no? An empirical analysis of S&P500 dynamics. SSRN Electron J (2009). https://doi.org/10.2139/ssrn.1363959

Jaisson, T., Rosenbaum, M.: Limit theorems for nearly unstable Hawkes processes. Ann Appl Probab **25**(2), 600 (2015). https://doi.org/10.1214/14-aap1005

Kou, S.G.: A jump-diffusion model for option pricing. Manag Sci **48**(8), 1086–1101 (2002). https://doi.org/10.1287/mnsc.48.8.1086.166

Merton, R.C.: Option pricing when underlying stock returns are discontinuous. J Financ Econ **3**(1–2), 125–144 (1976). https://doi.org/10.1016/0304-405x(76)90022-2

Njike, C.G.L., Hainaut, D.: Valuation of annuity guarantees under a self-exciting switching jump model. Methodol Comput Appl Probab (2022). https://doi.org/10.1007/s11009-022-09931-8

Schoutens, W., Simons, E., Tistaert, J.: A perfect calibration! Now what? Wilmott **2004**, 66–78 (2004)

Witkovský, V.: On the exact computation of the density and of the quantiles of linear combinations of t and F random variables. J Stat Plan Inference **94**(1), 1–13 (2001). https://doi.org/10.1016/s0378-3758(00)00208-1

Zumbach, G., Lynch, P.: Heterogeneous volatility cascade in financial markets. Physica A: Stat Mech Appl **298**(3–4), 521–529 (2001). https://doi.org/10.1016/s0378-4371(01)00249-7

Zumbach, G., Fernández, L., Weber, C.: Processes for stocks capturing their statistical properties from one day to one year. Quant Financ **14**(5), 849–861 (2013). https://doi.org/10.1080/14697688.2013.765956