Received: 1 September 2016 | Accepted: 1 September 2017

DOI: 10.1111/mafi.12173

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

ORIGINAL ARTICLE

The characteristic function of rough Heston models

Omar El Euch | Mathieu Rosenbaum

École Polytechnique, CMAP, Palaiseau, France

Correspondence
Omar El Euch, École Polytechnique, Palaiseau, 91128, France.
Email: omar.el-euch@polytechnique.edu

Abstract

It has been recently shown that rough volatility models, where the volatility is driven by a fractional Brownian motion with small Hurst parameter, provide very relevant dynamics in order to reproduce the behavior of both historical and implied volatilities. However, due to the non-Markovian nature of the fractional Brownian motion, they raise new issues when it comes to derivatives pricing. Using an original link between nearly unstable Hawkes processes and fractional volatility models, we compute the characteristic function of the log-price in rough Heston models. In the classical Heston model, the characteristic function is expressed in terms of the solution of a Riccati equation. Here, we show that rough Heston models exhibit quite a similar structure, the Riccati equation being replaced by a fractional Riccati equation.

KEYWORDS

fractional Brownian motion, fractional Riccati equation, Hawkes processes, limit theorems, rough Heston models, rough volatility models

1 | INTRODUCTION

The celebrated Heston model is a one-dimensional stochastic volatility model where the asset price S follows the following dynamic:

$$dS_t = S_t\sqrt{V_t}dW_t$$

$$dV_t = \gamma(\theta - V_t)dt + \gamma\nu\sqrt{V_t}dB_t. \quad (1.1)$$

Here the parameters γ, θ, V_0, and ν are positive, and W and B are two Brownian motions with correlation coefficient ρ, that is, $\langle dW_t, dB_t\rangle = \rho dt$.

The popularity of this model is probably due to three main reasons:

- It reproduces well several important stylized facts of low-frequency price data, namely, leverage effect, time-varying volatility, and fat tails; see Bouchaud and Potters (2003), Christie (1982), Dragulescu and Yakovenko (2002), and Mandelbrot (1997).
- It generates very reasonable shapes and dynamics for the implied volatility surface. Indeed, the "volatility of volatility" parameter ν enables us to control the smile, the correlation parameter ρ to deal with the skew, and the initial volatility V_0 to fix the at-the-money volatility level; see Forde, Jacquier, and Lee (2012), Gatheral (2011), Janek, Kluge, Weron, and Wystup (2011), and Poon (2009). Furthermore, as observed in markets and in contrast to local volatility models, in the Heston model the volatility smile moves in the same direction as the underlying and the forward smile does not flatten with time; see Gatheral (2011), Jacquier and Roome (2013, 2016), and Pascucci and Mazzon (2017).
- There is an explicit formula for the characteristic function of the asset log-price; see Heston (1993). From this formula, efficient numerical methods have been developed, allowing for instantaneous model calibration and pricing of derivatives; see Albrecher, Mayer, Schoutens, and Tistaert (2007), Carr and Madan (1999), Kahl and Jäckel (2005), and Lewis (2001).

In the classical Heston model, the volatility follows a Brownian semimartingale. However, it is demonstrated in Gatheral, Jaisson, and Rosenbaum (2018) that for a very wide range of assets, historical volatility time series exhibit a behavior that is much rougher than that of a Brownian motion. More precisely, the dynamic of log-volatility is typically very well modeled by a fractional Brownian motion with Hurst parameter of order 0.1. For example, it is shown in Gatheral et al. (2018) that in practice, the empirical moment of order $q > 0$ of log-volatility increments

$$\log(V_{t+\Delta}) - \log(V_t)$$

is proportional to Δ^{qH} with H of order 0.1, and this for any reasonable scale of interest Δ (from Δ equal to 1 day to hundreds of days). This corresponds to a rough fractional dynamic with Hurst parameter $H = 0.1$. Beyond moments, it is also established in Gatheral et al. (2018) that the empirical correlation structure of volatility is very well reproduced when using rough fractional volatility models. These findings have been confirmed by further studies; see Bennedsen, Lunde, and Pakkanen (2017a) and Livieri, Mouti, Pallavicini, and Rosenbaum (2018). Moreover, considering a fractional Brownian motion with small Hurst parameter also enables us to obtain remarkable fits for the whole volatility surface. In particular, contrary to most stochastic volatility models, rough volatility models generate an exploding term structure for the at-the-money skew when maturity goes to zero, which is very commonly observed in practice; see Bayer, Friz, and Gatheral (2016), Gatheral et al. (2018), and Section 5. Finally, convincing microstructural foundations for rough volatility models are provided in El Euch, Fukasawa, and Rosenbaum (2016) and Jaisson and Rosenbaum (2016); see also Section 2.

Hence, in this paper, we are interested in the fractional versions of the Heston model. Our main goal is to design an efficient pricing methodology for such models, in the spirit of the one introduced by Heston in the classical case. This is particularly important in fractional volatility models where the use of Monte Carlo methods can be quite intricate due to the non-Markovian nature of the fractional Brownian motion; see Bennedsen, Lunde, and Pakkanen (2017b).

We now define our so-called rough Heston model. Let us recall that a fractional Brownian motion W^H with Hurst parameter $H \in (0, 1)$ can be built through the Mandelbrot–van Ness representation

$$W_t^H = \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^{0} \left((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right) dW_s + \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-\frac{1}{2}} dW_s. \quad (1.2)$$

The kernel $(t-s)^{H-\frac{1}{2}}$ in (1.2) plays a central role in the rough dynamic of the fractional Brownian motion for $H < 1/2$. In particular, one can show that the process

$$\int_0^t (t-s)^{H-\frac{1}{2}} dW_s$$

has Hölder regularity $H - \varepsilon$ for any $\varepsilon > 0$. In order to allow for a rough behavior of the volatility in a Heston-type model, we naturally introduce the kernel $(t-s)^{\alpha-1}$ in a Heston-like stochastic volatility process as follows:

$$ds_t = S_t \sqrt{V_t} dW_t$$

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \gamma(\theta - V_s) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \gamma \nu \sqrt{V_s} dB_s. \tag{1.3}$$

The parameters γ, θ, V_0, and ν in (1.3) are positive and play the same role as in (1.1), and here also W and B are two Brownian motions with correlation ρ. The additional parameter α belongs to $(1/2, 1)$ and governs the smoothness of the volatility sample paths. More precisely, we show in this paper that the model is well defined and that the volatility trajectories have almost surely Hölder regularity $\alpha - 1/2 - \varepsilon$, for any $\varepsilon > 0$. When $\alpha = 1$, models (1.3) and (1.1) coincide, and we retrieve the classical Heston model. Therefore, it is natural to view (1.3) as a rough version of the Heston model and to call it rough Heston model. In term of Hurst parameter H, $\alpha = H + 1/2$. Nevertheless, note that other definitions of rough Heston models can make sense; see Guennoun, Jacquier, Roome, and Shi (2018) for an alternative definition and some asymptotic results.

Our aim in this work is to derive a Heston-type formula for the characteristic function of the log-price in model (1.3). In the classical case ($\alpha = 1$, model (1.1)), this formula is proved in Heston (1993). It is obtained using the fact that model (1.1) is Markovian and time-homogeneous, and applying Itô's formula to the function

$$L(t, a, V_t, S_t) = \mathbb{E}\left[e^{ia\log(S_T)} \mid \mathcal{F}_t\right], \quad \mathcal{F}_t = \sigma(W_s, B_s; s \le t), \quad a \in \mathbb{R}.$$

The process L being a martingale, the following Feynman–Kac partial differential equation for L is easily obtained

$$-\partial_t L(t, a, S, V) = \left(\gamma(\theta - V)\partial_v + \frac{1}{2}(\gamma\nu)^2 V \partial^2_{vv} + \frac{1}{2} S^2 V \partial^2_{ss} + \rho\nu\gamma S V \partial^2_{sv}\right) L(t, a, S, V),$$

with boundary condition $L(T, a, S, V) = e^{ia\log(S)}$. From this PDE, it can be checked that the characteristic function of the log-price $X_t = \log(S_t/S_0)$ satisfies

$$\mathbb{E}\left[e^{iaX_t}\right] = \exp(g(a,t) + V_0 h(a,t)),$$

where h is solution of the following Riccati equation:

$$\partial_t h(a,t) = \frac{1}{2}(-a^2 - ia) + \gamma(ia\rho\nu - 1)h(a,t) + \frac{(\gamma\nu)^2}{2} h^2(a,t), \quad h(a,0) = 0, \tag{1.4}$$

and

$$g(a,t) = \theta\gamma \int_0^t h(a,s) ds.$$

Solving this Riccati equation leads to the closed-form formula for the characteristic function of the log-price given in Heston (1993).

In the case $\alpha < 1$, the rough Heston model (1.3) is not Markovian and the variance process is no longer a semimartingale. Hence, the strategy initially used by Heston and presented above seems very hard to adapt to our setting. Here, we resort to a completely different and original approach based on point processes. Indeed, our methodology finds its root in El Euch et al. (2016) and Jaisson and Rosenbaum (2016), who provide microstructural foundations for rough volatility models. In these papers, it is shown that some well-designed microstructure models, reproducing the stylized facts of modern financial markets at high frequency, give rise in the long run to rough volatility models. These microstructure models, which we describe in more detail in Section 2, are based on so-called nearly unstable Hawkes processes. In this paper, inspired by these results and using again Hawkes processes, we design a suitable sequence of point processes, which converges to model (1.3). Exploiting the specific structure of our point processes, we derive their characteristic function, which leads us in the limit to that of the log-price in the rough Heston model (1.3).

Our main result is that, quite surprisingly, the characteristic function of the log-price in rough Heston models exhibits the same structure as the one obtained in the classical Heston model. The difference is that the Riccati equation (1.4) is replaced by a fractional Riccati equation, where a fractional derivative appears instead of a classical derivative. More precisely, we obtain

$$\mathbb{E}[e^{iaX_t}] = \exp(g_1(a,t) + V_0 g_2(a,t)),$$

where

$$g_1(a,t) = \theta\gamma \int_0^t h(a,s)ds, \quad g_2(a,t) = I^{1-\alpha} h(a,t),$$

and $h(a,.)$ is a solution of the following fractional Riccati equation:

$$D^\alpha h(a,t) = \frac{1}{2}(-a^2 - ia) + \gamma(ia\rho\nu - 1)h(a,t) + \frac{(\gamma\nu)^2}{2}h^2(a,t), \quad I^{1-\alpha}h(a,0) = 0,$$

with D^α and $I^{1-\alpha}$ the fractional derivative and integral operators defined in (4.3) and (4.4). Remark that when $\alpha = 1$, this result indeed coincides with the classical Heston's result. However, note that for $\alpha < 1$, the solutions of such Riccati equations are no longer explicit. Nevertheless, they are easily solved numerically; see Section 5.

The paper is organized as follows. In Section 2, we build a sequence of Hawkes-type processes, which converges to the rough Heston model (1.3). Then, we study in Section 3 the characteristic function of these processes and show in Section 4 that it enables us to derive the characteristic function of the log-price in model (1.3). Numerical illustrations are given in Section 5 and some proofs are relegated to Section 6. Finally, some useful technical results are given in the Appendix.

2 | FROM HAWKES PROCESSES TO ROUGH HESTON MODELS

We build in this section a sequence of Hawkes-type processes, which converges to the rough Heston model (1.3). This construction is inspired by El Euch et al. (2016). In this work, microstructural foundations for rough Heston models are provided. This is done designing suitable sequences of ultra high-frequency price models, which reproduce the stylized facts of modern markets microstructure and converge in the long run to rough Heston models. These microscopic price models are based on Hawkes processes. So that the reader can understand the genesis of our original methodology to compute the characteristic function in rough Heston models, we recall here the main ideas and results in El Euch et al. (2016).

2.1 | Microstructural foundations for rough Heston models

In El Euch et al. (2016), we consider a sequence of bidimensional Hawkes processes $(N^{T,+}, N^{T,-})$ indexed by $T > 0$ going to infinity[1] and with intensity[2]

$$\lambda_t^T = \begin{pmatrix} \lambda_t^{T,+} \\ \lambda_t^{T,-} \end{pmatrix} = \mu_T \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \int_0^t a_T \phi(t-s). \begin{pmatrix} dN_s^{T,+} \\ dN_s^{T,-} \end{pmatrix}, \tag{2.1}$$

with

$$\phi = \begin{pmatrix} \varphi_1 & \varphi_3 \\ \varphi_2 & \varphi_4 \end{pmatrix}.$$

Here, the φ_i are measurable nonnegative deterministic functions and μ_T and $0 < a_T < 1$ are some deterministic sequences of positive real numbers; see Bacry, Delattre, Hoffmann, and Muzy (2013) and the references therein for more details about the definition of Hawkes processes. Then, in El Euch et al. (2016), inspired by Bacry, Delattre, Hoffmann, and Muzy (2013), Bacry et al. (2013), and Jaisson and Rosenbaum, (2015), we consider the following ultra high-frequency tick-by-tick model for the transaction price P_t^T,

$$P_t^T = N_t^{T,+} - N_t^{T,-}. \tag{2.2}$$

Hence $N_t^{T,+}$ represents the number of upward jumps of one tick of the transaction price over the period $[0, t]$ and $N_t^{T,-}$ the number of downward jumps. The relevance of this Hawkes-based modeling is that it enables us to encode very easily the most important stylized facts of high-frequency markets in term of the parameters of the Hawkes process. We now give these stylized facts and their translation in terms of the model parameters, referring to El Euch et al. (2016) for more details.

- Markets are highly endogenous: In the high-frequency trading context, most orders have no real economic motivation. They are rather sent by algorithms as reactions to other orders. In the Hawkes framework, this amounts to work with so-called *nearly unstable Hawkes processes*. This means that the stability condition

  $$\mathcal{S}\left(\int_0^\infty a_T \phi(s) ds\right) < 1,$$

  where $\mathcal{S}$ denotes the spectral radius operator, should almost be saturated and that the intensity of exogenous orders, namely, μ_T, should be small; see El Euch et al. (2016), Hardiman, Bercot, and Bouchaud (2013), and Jaisson and Rosenbaum (2015, 2016). In terms of model parameters, suitable constraints are, therefore,

  $$a_T \to 1, \quad \mathcal{S}\left(\int_0^\infty \phi(s) ds\right) = 1, \quad \mu_T \to 0.$$

- It is not an easy task to make money with high-frequency strategies on highly liquid electronic markets. Hence, some "no statistical arbitrage" mechanisms should be in force. We translate this assuming that in the long run, there are on average as many upward as downward jumps. This corresponds to the assumption

$$\varphi_1 + \varphi_3 = \varphi_2 + \varphi_4.$$

- Buying is not the same action as selling. This means that buy market orders and sell market orders are not symmetric orders. To see this, consider, for example, a market maker, with an inventory that is typically positive. After each order he receives, he modifies his bid and ask quotes, reflecting the market impact of the received order. This means that after a buy order, he will increase his ask quote and he will decrease his bid quote after a sell order. However, he typically raises the price by less following a buy order than he lowers the price following the same size sell order. Indeed, inventory becomes smaller after a buy order, which is a good thing for him, whereas it increases after a sell order. This creates a liquidity asymmetry on the bid and ask sides of the order book. This can be modeled in the Hawkes framework assuming that

$$\varphi_3 = \beta \varphi_2,$$

for some $\beta > 1$. Hence, the matrix ϕ finally takes the form

$$\phi = \begin{pmatrix} \varphi_1 & \beta\varphi_2 \\ \varphi_2 & \varphi_1 + (\beta - 1)\varphi_2 \end{pmatrix}.$$

- A significant amount of transactions is part of metaorders, which are large orders whose execution is split in time by trading algorithms. This is translated into a heavy tail assumption on the functions φ_1 and φ_2, namely, that there exists $1/2 < \alpha < 1$ (typically around 0.6 in practice; see Bacry, Jaisson, & Muzy, 2014; Hardiman et al., 2013) and $C > 0$ such that

$$\alpha x^\alpha \int_x^\infty \varphi_1(s) + \beta\varphi_2(s) ds \underset{x\to\infty}{\to} C.$$

Furthermore, it is shown in Jaisson and Rosenbaum (2016) that for a given α, there is only one way to make μ_T tends to zero and a_T tends to one so that the limit of the price is not degenerate. More precisely,

$$(1 - a_T)T^\alpha \underset{T\to\infty}{\to} \lambda^*, \qquad \mu_T T^{1-\alpha} \underset{T\to\infty}{\to} \mu,$$

for some positive λ^* and μ.

Under the above assumptions, it is proved in El Euch et al. (2016) that the properly rescaled microscopic price process

$$\sqrt{\frac{1 - a_T}{\mu T^\alpha}} P_{tT}^T, \quad t \in [0, 1],$$

where P^T is defined in (2.7), converges in law as T tends to infinity to the following macroscopic price dynamic P,

$$P_t = \frac{\sqrt{2}}{1 - \int_0^\infty (\varphi_1 - \varphi_2)} \int_0^t \sigma_s dW_s,$$

$$\sigma_t^2 = \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\gamma(1-\sigma_s^2)ds + \frac{1}{\Gamma(\alpha)}\gamma\nu\int_0^t (t-s)^{\alpha-1}\sigma_s dB_s, \tag{2.3}$$

where (W, B) is a bidimensional correlated Brownian motion with correlation

$$\rho = \frac{1-\beta}{\sqrt{2(1+\beta^2)}}$$

and

$$\nu = \sqrt{\frac{2(1+\beta^2)}{\lambda^*\mu(1+\beta)^2}}, \qquad \gamma = \lambda^*\frac{\alpha}{C\Gamma(1-\alpha)}.$$

Hence, this result shows that the main stylized facts of modern electronic markets naturally give rise to a very rough behavior of the volatility. Indeed, recall that the Hurst parameter corresponds to $\alpha - 1/2$.

Inspired by this result, our idea is to study the characteristic function of some kind of microscopic price processes in order to deduce that of our rough Heston macroscopic price of interest (1.3). However, the developments presented above cannot be directly applied and need to be adapted. Indeed, remark that in (2.3), $\sigma_0 = 0$. This does not correspond to the case of (1.3), where having a nonzero initial volatility is of course crucial for the model to be relevant in practice. Thus, we need to modify the sequence of Hawkes-type processes to obtain a nondegenerate initial volatility in the limit. This is actually a nontrivial issue. However, this can be achieved replacing μ_T in (2.1) by an inhomogeneous Poisson intensity $\hat{\mu}_T(t)$. We explain how such $\hat{\mu}_T(t)$ can be found in the next section.

2.2 | The role of the Poisson rate

We work on a sequence of probability spaces $(\Omega^T, \mathcal{F}^T, \mathbb{P}^T)$, indexed by $T > 0$ (going to infinity), on which $N^T = (N^{T,+}, N^{T,-})$ is a bidimensional Hawkes process with intensity

$$\lambda_t^T = \begin{pmatrix} \lambda_t^{T,+} \\ \lambda_t^{T,-} \end{pmatrix} = \hat{\mu}_T(t)\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \int_0^t \phi^T(t-s).dN_s^T. \tag{2.4}$$

For a given T, the probability space is equipped with the filtration $(\mathcal{F}_t^T)_{t\geq 0}$, where $\mathcal{F}_t^T$ is the σ-algebra generated by $(N_s^T)_{s\leq t}$. Because our goal is to design a sequence of processes leading in the limit to a rough Heston dynamic, we consider the same kind of assumptions on the matrix ϕ^T as those described in the previous section. However, here we can be very specific because we just need to find one convenient sequence of processes. That is why we make a particular choice for the heavy-tailed functions defining ϕ^T, using Mittag–Leffler functions; see Section A.1 in the Appendix for definition and some properties. Indeed, these functions are very convenient in order to carry out computations. More precisely, our assumptions on ϕ^T are as follows.

Definition 2.1. There exist $\beta \geq 0$, $1/2 < \alpha < 1$ and $\gamma > 0$ such that

$$a_T = 1 - \gamma T^{-\alpha}, \qquad \phi^T = \varphi^T\chi,$$

where

$$\chi = \frac{1}{\beta+1}\begin{pmatrix} 1 & \beta \\ 1 & \beta \end{pmatrix}, \qquad \varphi^T = a_T\varphi, \qquad \varphi = f^{\alpha,1},$$

with $f^{\alpha,1}$ the Mittag–Leffler density function defined in the Appendix.

Remark 2.2. From Section A.1 in the Appendix, we get that we are working in the nearly unstable heavy tail case because

$$\int_0^\infty \varphi(s)ds = 1$$

and

$$\alpha x^\alpha \int_x^\infty \varphi(t)dt \underset{x\to\infty}{\longrightarrow} \frac{\alpha}{\Gamma(1-\alpha)}.$$

We now give intuitions about the need to use a nonconstant Poisson intensity $\hat{\mu}_T(t)$. First, note that under Definition 2.1,

$$\lambda_t^{T,+} = \lambda_t^{T,-}.$$

The asymptotic behavior of the renormalized intensity processes $\lambda_t^{T,+}$ and $\lambda_t^{T,-}$ will give us that of the volatility in our limiting macroscopic price model. Thus, we need to understand the long-term limit of $\lambda_t^{T,+}$. Let us write

$$M_t^T = \left(M_t^{T,+}, M_t^{T,-}\right) = N_t^T - \int_0^t \lambda_s^T ds$$

for the martingale associated to the point process N_t^T. We easily obtain

$$\lambda_t^{T,+} = \hat{\mu}_T(t) + \int_0^t \varphi^T(t-s)\lambda_s^{T,+}ds + \frac{1}{1+\beta}\int_0^t \varphi^T(t-s)\left(dM_s^{T,+} + \beta dM_s^{T,-}\right).$$

Now let

$$\psi^T = \sum_{k\geq 1} (\varphi^T)^{*k},$$

where $(\varphi^T)^{*1} = \varphi^T$ and for $k > 1$, $(\varphi^T)^{*k}(t) = \int_0^t \varphi^T(s)(\varphi^T)^{*(k-1)}(t-s)ds$. Using Lemma A.1 in the Appendix together with Fubini theorem and the fact that $\psi^T * \varphi^T = \psi^T - \varphi^T$, we get

$$\lambda_t^{T,+} = \hat{\mu}_T(t) + \int_0^t \psi^T(t-s)\hat{\mu}_T(s)ds + \frac{1}{1+\beta}\int_0^t \psi^T(t-s)\left(dM_s^{T,+} + \beta dM_s^{T,-}\right). \tag{2.5}$$

Following El Euch et al. (2016), the inhomogeneous intensity $\hat{\mu}_T(t)$ should be of order μ_T with

$$\mu_T = \mu T^{\alpha-1},$$

where μ is some positive constant. In El Euch et al. (2016), it is shown that the right normalization for the intensity in order to get a nondegenerate limit, is to consider $(1-a_T)\lambda_{tT}^{T,+}/\mu_T$. The same applies here and thus we define the renormalized intensity

$$C_t^T = \frac{1-a_T}{\mu_T}\lambda_{tT}^{T,+}.$$

After obvious computations, this can be written as

$$C_t^T = \frac{1-a_T}{\mu_T}\hat{\mu}_T(tT) + \int_0^t T(1-a_T)\psi^T(T(t-s))\frac{\hat{\mu}_T(Ts)}{\mu_T}ds + \nu\int_0^t T(1-a_T)\psi^T(T(t-s))\sqrt{C_s^T}dB_s^T,$$

where

$$B_t^T = \int_0^{tT}\frac{dM_s^{T,+} + \beta dM_s^{T,-}}{\sqrt{T\left(\lambda_s^{T,+} + \beta^2\lambda_s^{T,-}\right)}}, \qquad \nu = \sqrt{\frac{1+\beta^2}{\gamma\mu(1+\beta)^2}}.$$

Using the fact that the Laplace transform $\hat{f}^{\alpha,\gamma}$ of the Mittag–Leffler density function $f^{\alpha,\gamma}$ is given by

$$\hat{f}^{\alpha,\gamma}(z) = \frac{\gamma}{\gamma + z^\alpha},$$

we easily obtain that

$$(1-a_T)T\psi^T(T.) = a_T f^{\alpha,\gamma}; \tag{2.6}$$

see Section A.1 in the Appendix. This leads to the following expression for C^T:

$$C_t^T = \frac{1-a_T}{\mu_T}\hat{\mu}_T(tT) + \int_0^t a_T f^{\alpha,\gamma}(t-s)\frac{\hat{\mu}_T(Ts)}{\mu_T}ds + \nu\int_0^t a_T f^{\alpha,\gamma}(t-s)\sqrt{C_s^T}dB_s^T.$$

Computing the quadratic variation of B^T, it is easy to see that it converges to a Brownian motion B. Now, if as in El Euch et al. (2016) we take $\hat{\mu}_T(t) = \mu_T$, we obtain that C^T should then give in the limit a process with starting value equal to zero. Nevertheless, we also get the intuition that a nonconstant $\hat{\mu}_T$ can lead to a nontrivial initial value. From the computations in the proof of Theorem 2.5, it will become clear that the right choice of $\hat{\mu}_T$ is as follows

Definition 2.3. The baseline intensity $\hat{\mu}_T$ is given by

$$\hat{\mu}_T(t) = \mu_T + \xi\mu_T\left(\frac{1}{1-a_T}\left(1 - \int_0^t \varphi^T(s)ds\right) - \int_0^t \varphi^T(s)ds\right),$$

with $\xi > 0$ and $\mu_T = \mu T^{\alpha-1}$ for some $\mu > 0$.

Remark 2.4. Note that $\hat{\mu}_T$ can also be written as follows:

$$\hat{\mu}_T(t) = \mu_T + \xi\mu_T\left(\frac{T^\alpha}{\gamma}\int_t^\infty \varphi(s)ds + \gamma T^{-\alpha}\int_0^t \varphi(s)ds\right).$$

This shows that $\hat{\mu}_T$ is a positive function and thus that the intensity process λ_t^T in (2.4) is well defined.

2.3 | The rough limits of Hawkes processes

We now give a rigorous statement about the limiting behavior of our specific sequence of bidimensional nearly unstable Hawkes processes with heavy tails. For $t \in [0,1]$, we define

$$X_t^T = \frac{1-a_T}{T^\alpha\mu}N_{tT}^T, \qquad \Lambda_t^T = \frac{1-a_T}{T^\alpha\mu}\int_0^{tT}\lambda_s^T ds, \qquad Z_t^T = \sqrt{\frac{T^\alpha\mu}{1-a_T}}\left(X_t^T - \Lambda_t^T\right).$$

Using a similar approach as that in El Euch et al. (2016), we obtain the following result whose proof is given in Section 6.

Theorem 2.5. *As $T \to \infty$, under Definitions 2.1 and 2.3, the process $(\Lambda_t^T, X_t^T, Z_t^T)_{t\in[0,1]}$ converges in law for the Skorokhod topology to (Λ, X, Z), where*

$$\Lambda_t = X_t = \int_0^t Y_s ds \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad Z_t = \int_0^t \sqrt{Y_s} \begin{pmatrix} dB_s^1 \\ dB_s^2 \end{pmatrix},$$

and Y is the unique solution of the rough stochastic differential equation[3]

$$Y_t = \xi + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \gamma (1 - Y_s) ds + \gamma \sqrt{\frac{1+\beta^2}{\gamma\mu(1+\beta^2)}} \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sqrt{Y_s} dB_s,$$

where

$$B = \frac{B^1 + \beta B^2}{\sqrt{1+\beta^2}}$$

and (B^1, B^2) is a bidimensional Brownian motion. Furthermore, for any $\varepsilon > 0$, Y has Hölder regularity $\alpha - 1/2 - \varepsilon$.

Hence, Theorem 2.5 shows that designing our sequence of bidimensional Hawkes processes in a suitable way, its limit is differentiable and its derivative exhibits a rough Cox–Ingersoll–Ross–like behavior, with nonzero initial value. This is exactly what we need for the limiting volatility of our microscopic price processes. Indeed, thanks to Theorem 2.5, we are now able to build such microscopic processes converging to the log-price in (1.3). More precisely, for $\theta > 0$, let us define

$$P^T = \sqrt{\frac{\theta}{2}} \sqrt{\frac{1-a_T}{T^\alpha \mu}} \left(N_{\cdot T}^{T,+} - N_{\cdot T}^{T,-} \right) - \frac{\theta}{2} \frac{1-a_T}{T^\alpha \mu} N_{\cdot T}^{T,+} = \sqrt{\frac{\theta}{2}} \left(Z^{T,+} - Z^{T,-} \right) - \frac{\theta}{2} X^{T,+}. \quad (2.7)$$

We have the following corollary of Theorem 2.5.

Corollary 2.6. *As $T \to \infty$, under Definitions 2.1 and 2.3, the sequence of processes $(P_t^T)_{t\in[0,1]}$ converges in law for the Skorokhod topology to*

$$P_t = \int_0^t \sqrt{V_s} dW_s - \frac{1}{2} \int_0^t V_s ds,$$

where V is the unique solution of the rough stochastic differential equation

$$V_t = \theta\xi + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \gamma (\theta - V_s) ds + \gamma \sqrt{\frac{\theta(1+\beta^2)}{\gamma\mu(1+\beta)^2}} \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sqrt{V_s} dB_s,$$

with (W, B) a correlated bidimensional Brownian motion whose bracket satisfies

$$d\langle W, B \rangle_t = \frac{1-\beta}{\sqrt{2(1+\beta^2)}} dt.$$

Thus, we have succeeded in building a sequence of microscopic processes P^T, defined by (2.7), which converges to (the logarithm of) our rough Heston process of interest (1.3). Now our goal is to use the result of Corollary 2.6 to compute the characteristic function of the log-price in the rough Heston model (1.3). This is done in the next two sections.

3 | THE CHARACTERISTIC FUNCTION OF MULTIVARIATE HAWKES PROCESSES

We have seen in the previous section that our sequence of Hawkes-based microscopic price processes converges to the log-price in the rough Heston model (1.3). Therefore, if we are able to compute the characteristic function for the microscopic price, its limit will give us that of the log-price in a rough Heston model. We actually provide a more general result here, deriving the characteristic function of a multivariate Hawkes process (recall that a bidimensional Hawkes process is the building block for our microscopic price process (2.7)). Hence, we extend here some results already proved in Hawkes and Oakes (1974) in the one-dimensional case.

3.1 | Cluster-based representation

To derive our characteristic function, the representation of Hawkes processes in term of clusters (see Hawkes & Oakes, 1974), is very useful. We recall it now. Let us consider a d-dimensional Hawkes process $N = (N^1, ..., N^d)$ with intensity

$$\lambda_t = \begin{pmatrix} \lambda_t^1 \\ \vdots \\ \lambda_t^d \end{pmatrix} = \mu(t) + \int_0^t \phi(t-s).dN_s, \tag{3.1}$$

where $\mu : \mathbb{R}_+ \to \mathbb{R}_+^d$ is locally integrable and $\phi : \mathbb{R}_+ \to \mathcal{M}^{\mathbf{d}}(\mathbb{R}_+)$ has integrable components such that

$$\mathcal{S}\left(\int_0^\infty \phi(s)ds\right) < 1.$$

The law of such process can be described through a population approach. Consider that there are d types of individuals and for a given type, an individual can be either a migrant or the descendant of a migrant. Then, the dynamic goes as follows from time $t = 0$:

- Migrants of type $k \in \{1, .., d\}$ arrive as a nonhomogenous Poisson process with rate $\mu_k(t)$.
- Each migrant of type $k \in \{1, .., d\}$ gives birth to children of type $j \in \{1, .., d\}$ following a nonhomogenous Poisson process with rate $\phi_{j,k}(t)$.
- Each child of type $k \in \{1, .., d\}$ also gives birth to other children of type $j \in \{1, .., d\}$ following a nonhomogenous Poisson process with rate $\phi_{j,k}(t)$.

Then, for $k \in \{1, .., d\}$, N_t^k can be taken as the number up to time t of migrants and children born with type k. Indeed, the population approach above and the theoretical characterization (3.1) define the same point process law.

3.2 | The result

Let $L(a,t)$ be the characteristic function of the Hawkes process N,

$$L(a,t) = \mathbb{E}[\exp(ia.N_t)], \qquad t \geq 0, \qquad a \in \mathbb{R}^d,$$

where $a.N_t$ stands for the scalar product of a and N_t. The cluster-based representation of multivariate Hawkes processes enables us to show the following result, proved in Section 3.3, for their characteristic function.

Theorem 3.1. *We have*

$$L(a,t) = \exp\left(\int_0^t (C(a,t-s) - \mathbf{1}).\mu(s)ds\right),$$

where $C : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{C}^d$ is solution of the following integral equation:

$$C(a,t) = \exp\left(ia + \int_0^t \phi^*(s).(C(a,t-s) - \mathbf{1})ds\right),$$

with $\phi^*(s)$ the transpose of $\phi(s)$.

From Theorem 3.1, we are able to derive in Section 4 the characteristic function of rough Heston models.

3.3 | Proof of Theorem 3.1

We now give the proof of Theorem 3.1, exploiting the population construction presented in Section 3.1. We start by defining d auxiliary independent d-dimensional point processes $(\tilde{N}^{k,j})_{1\leq j\leq d}$, $k \in \{1, ..., d\}$, defined as follows for each given $k \in \{1, .., d\}$:

- Migrants of type $j \in \{1, ..., d\}$ arrive as a nonhomogenous Poisson process with rate $\phi_{j,k}(t)$.
- Each migrant of type $j \in \{1, .., d\}$ gives birth to children of type $l \in \{1, .., d\}$ following a nonhomogenous Poisson process with rate $\phi_{l,j}(t)$.
- Each child of type $j \in \{1, .., d\}$ also gives birth to other children of type $l \in \{1, .., d\}$ following a nonhomogenous Poisson process with rate $\phi_{l,j}(t)$.

For a given $k \in \{1, .., d\}$, $\tilde{N}_t^{k,j}$ corresponds to the number, up to time t, of migrants and children with type j. A simple but crucial remark is that $(\tilde{N}^{k,j})_{1\leq j\leq d}$ is actually also a multivariate Hawkes process with migrant rate $(\phi_{j,k})_{1\leq j\leq d}$ and kernel matrix ϕ. We write $L_k(a,t)$ for its characteristic function

$$L_k(a,t) = \mathbb{E}\left[\exp\left(ia.\left(\tilde{N}_t^{k,j}\right)_{1\leq j\leq d}\right)\right], \quad t \geq 0, \quad a \in \mathbb{R}^d.$$

Now let us come back to the initial Hawkes process of interest N defined by (3.1). For each $k \in \{1, ..., d\}$ and $t \geq 0$, let $N_t^{0,k}$ be the number of its migrants of type k arrived up to time t. Recall that the $N^{0,k}$, $1 \leq k \leq d$, are independent Poisson processes with rates $\mu_k(t)$. We also define $T_1^k < \cdots < T_{N_t^{0,k}}^k \in [0,t]$ the arrival times of migrants of type k of the Hawkes process N, up to time t. Using the population approach presented in Section 3.1, it is clear that at time t, the number of descendants of different types of a migrant of type k arrived at time T_u^k has the same law as $(\tilde{N}^{k,j}_{t-T_u^k})_{1\le j\le d}$, where $\tilde{N}$ is taken independent from N. Consequently,

$$N_t^k \underset{law}{=} N_t^{0,k} + \sum_{1\le j\le d} \sum_{1\le l\le N_t^{0,j}} \tilde{N}^{j,k,(l)}_{t-T_l^j}, \tag{3.2}$$

where the $(\tilde{N}^{j,k,(l)})_{1\le k\le d}$, $1\le j\le d$, $l\in\mathbb{N}$ are independent copies of $(\tilde{N}^{j,k})_{1\le k\le d}$, $1\le j\le d$, also independent of $N^0=(N^{0,k})_{1\le k\le d}$.

From (3.2), we derive that conditional on N^0,

$$\begin{aligned}\mathbb{E}\left[\exp(ia.N_t)|N^0\right] &= \exp\left(ia.N_t^0\right)\prod_{1\le j\le d}\prod_{1\le l\le N_t^{0,j}}\mathbb{E}\left[\exp\left(ia.\left(\tilde{N}^{j,k,(l)}_{t-T_l^j}\right)_{1\le k\le d}|N^0\right)\right]\\ &= \exp(ia.N_t^0)\prod_{1\le j\le d}\prod_{1\le l\le N_t^{0,j}} L_j\left(a,t-T_l^j\right).\end{aligned}$$

Now, for a given $k\in\{1,...,d\}$, conditional on $N_t^{0,k}$, it is well known that $(T_1^k,...,T^k_{N_t^{0,k}})$ has the same law as $(X_{(1)},...,X_{(N_t^{0,k})})$ the order statistics built from iid variables $(X_1,..,X_{N_t^{0,k}})$ with density $\frac{\mu_k(s)1_{s\le t}}{\int_0^t\mu_k(s)ds}$. Thus, we get

$$\mathbb{E}\left[\exp(ia.N_t)|N_t^0\right] = \exp\left(ia.N_t^0\right)\prod_{1\le j\le d}\left(\int_0^t L_j(a,t-s)\frac{\mu_j(s)}{\int_0^t\mu_j(s)ds}ds\right)^{N_t^{0,j}}.$$

Therefore,

$$L(a,t)=\prod_{1\le j\le d}\exp\left(\left(\int_0^t e^{ia_j}L_j(a,t-s)\frac{\mu_j(s)}{\int_0^t\mu_j(s)ds}ds-1\right)\int_0^t\mu_j(s)ds\right).$$

Thus, we finally obtain

$$L(a,t)=\exp\left(\sum_{1\le j\le d}\int_0^t\left(e^{ia_j}L_j(a,t-s)-1\right)\mu_j(s)ds\right). \tag{3.3}$$

In the same way, because $(\tilde{N}^{k,j})_{1\le j\le d}$ is a multivariate Hawkes process with migrant rate $(\phi_{j,k})_{1\le j\le d}$ and kernel matrix ϕ, we get

$$L_k(a,t)=\exp\left(\sum_{1\le j\le d}\int_0^t\left(e^{ia_j}L_j(a,t-s)-1\right)\phi_{j,k}(s)ds\right). \tag{3.4}$$

Let us define

$$C(a,t)=\left(e^{ia_j}L_j(a,t)\right)_{1\le j\le d}.$$

From (3.3), we have that

$$L(a,t) = \exp\left(\int_0^t (C(a,t-s) - \mathbf{1}).\mu(s)ds\right)$$

and from (3.4), we deduce that C is solution of the following integral equation:

$$C(a,t) = \exp\left(ia + \int_0^t \phi^*(s).(C(a,t-s) - \mathbf{1})ds\right).$$

This ends the proof of Theorem 3.1.

4 | THE CHARACTERISTIC FUNCTION OF ROUGH HESTON MODELS

We give in this section our main theorem, that is the characteristic function for the log-price in rough Heston models (1.3). It is obtained combining the convergence result for Hawkes processes stated in Corollary 2.6 together with the characteristic function for multivariate Hawkes processes derived in Theorem 3.1. We start with some intuitions about the result.

4.1 | Intuition about the result

We consider the rough Heston model (1.3). The parameters of the dynamic in (1.3) are here given in terms of those of the sequence of processes P^T defined in (2.7). More precisely, we set

$$V_0 = \xi\theta, \qquad \rho = \frac{1-\beta}{\sqrt{2(1+\beta^2)}}, \qquad \nu = \sqrt{\frac{\theta(1+\beta^2)}{\gamma\mu(1+\beta)^2}},$$

and γ and θ are the same as those in the dynamic of P^T. Remark that the fact that $\beta \geq 0$ implies that $\rho \in (-1/\sqrt{2}, 1/\sqrt{2}]$[4]. We also write $P_t = \log(S_t/S_0)$. From Corollary 2.6, we know that

$$P^T = \sqrt{\frac{\gamma\theta}{2\mu}} T^{-\alpha}\left(N^{T,+}_{.T} - N^{T,-}_{.T}\right) - \frac{\gamma\theta}{2\mu}T^{-2\alpha}N^{T,+}_{.T}$$

converges in law to P as T tends to infinity, where $N^T = (N^{T,+}, N^{T,-})$ is a sequence of bidimensional Hawkes processes satisfying Definitions 2.1 and 2.3. Let us write $L^T((a,b),u)$ for the characteristic function of the process N^T at time u at point (a,b) and L_p for the characteristic function of P. The convergence in law implies that of $L^T((a^+_T, a^-_T), tT)$ toward $L_p(a,t)$, where

$$a^+_T = a\sqrt{\frac{\gamma\theta}{2\mu}}T^{-\alpha} - a\frac{\gamma\theta}{2\mu}T^{-2\alpha}, \qquad a^-_T = -a\sqrt{\frac{\gamma\theta}{2\mu}}T^{-\alpha}.$$

From Theorem 3.1, we know that

$$L^T\left(\left(a^+_T, a^-_T\right), tT\right) = \exp\left(\int_0^{tT} \hat{\mu}_T(s)\left((C^{T,+}\left(\left(a^+_T, a^-_T\right), tT-s\right) - 1) + (C^{T,-}\left(\left(a^+_T, a^-_T\right), tT-s\right) - 1)\right) ds\right),$$

where $C^T\left(\left(a_T^+,a_T^-\right),t\right)=(C^{T,+}((a_T^+,a_T^-),t),C^{T,-}((a_T^+,a_T^-),t))\in\mathcal{M}^{1\times 2}(\mathbb{C})$ is solution of

$$C^T\left(\left(a_T^+,a_T^-\right),t\right)=\exp\left(i\left(a_T^+,a_T^-\right)+\int_0^t\left(C^T\left(\left(a_T^+,a_T^-\right),t-s\right)-(1,1)\right).\phi^T(s)ds\right).$$

Now let

$$Y^T(a,.)=\left(Y^{T,+}(a,.),Y^{T,-}(a,.)\right)=C^T\left(\left(a_T^+,a_T^-\right),.T\right)\,:\,[0,1]\to\mathcal{M}^{1\times 2}(\mathbb{C}).$$

Using a change of variables, we easily get that $Y^T(a,.)$ is solution of the equation

$$Y^T(a,t)=\exp\left(i\left(a_T^+,a_T^-\right)+T\int_0^t\left(Y^T(a,t-s)-(1,1)\right).\phi^T(Ts)ds\right)\tag{4.1}$$

and that

$$L^T\left(a_T^+,a_T^-,tT\right)=\exp\left(\int_0^t\left(T^\alpha(Y^{T,+}(a,t-s)-1)+T^\alpha(Y^{T,-}(a,t-s)-1)\right)\left(T^{1-\alpha}\hat{\mu}(sT)\right)ds\right).\tag{4.2}$$

Thanks to Remarks 2.2 and 2.4, it is easy to see that

$$\begin{aligned}T^{1-\alpha}\hat{\mu}(sT)&=T^{1-\alpha}\mu_T+\xi T^{1-\alpha}\mu_T\left(\frac{T^\alpha}{\gamma}\int_{sT}^\infty\varphi(u)du+\gamma T^{-\alpha}\int_0^{sT}\varphi(u)du\right)\\&=\mu\left(1+\frac{\xi}{\gamma}s^{-\alpha}(sT)^\alpha\int_{sT}^\infty\varphi(u)du\right)+\mu\xi\gamma T^{-\alpha}\int_0^{sT}\varphi(u)du\\&\underset{T\to\infty}{\longrightarrow}\mu\left(1+\frac{\xi}{\gamma\Gamma(1-\alpha)}s^{-\alpha}\right).\end{aligned}$$

Then the convergence of $T^\alpha(Y^T(a,t)-(1,1))$ to some functions $(c(a,t),d(a,t))$ solutions of Volterra-type equations is proved in Section 6.2. It is based on a Taylor expansion from (4.1). This will lead us to the expression of $L_p(a,t)$.

4.2 | Main result

We define the fractional integral of order $r\in(0,1]$ of a function f as

$$I^rf(t)=\frac{1}{\Gamma(r)}\int_0^t(t-s)^{r-1}f(s)ds,\tag{4.3}$$

whenever the integral exists, and the fractional derivative of order $r\in[0,1)$ as

$$D^rf(t)=\frac{1}{\Gamma(1-r)}\frac{d}{dt}\int_0^t(t-s)^{-r}f(s)ds,\tag{4.4}$$

whenever it exists. The following theorem, proved in Section 6, is the main result of the paper.

Theorem 4.1. *Consider the rough Heston model (1.3) with a correlation between the two Brownian motions ρ satisfying $\rho\in(-1/\sqrt{2},1/\sqrt{2}]$. For all $t\geq 0$ and fixed $a\in\mathbb{R}$, we have*

$$L_p(a,t)=\exp\left(\theta\gamma I^1h(a,t)+V_0I^{1-\alpha}h(a,t)\right),\tag{4.5}$$

where $h(a,.)$ is solution of the fractional Riccati equation

$$D^\alpha h(a,t) = \frac{1}{2}(-a^2 - ia) + \gamma(ia\rho\nu - 1)h(a,s) + \frac{(\gamma\nu)^2}{2}h^2(a,s), \quad I^{1-\alpha}h(a,0) = 0, \tag{4.6}$$

which admits a unique continuous solution.

Thus, we have been able to obtain a semiclosed formula for the characteristic function in rough Heston models. This means that pricing of European options becomes an easy task in this model; see Section 5. For $\alpha = 1$, we retrieve the classical Heston formula. For $\alpha < 1$, the formula is almost the same. The difference is essentially only in that in the Riccati equation, the classical derivative is replaced by a fractional derivative. The drawback is that such fractional Riccati equations do not have explicit solutions. However, they can be solved numerically almost instantaneously; see Section 5. Finally, note that this strong link between Hawkes processes and (rough) Heston models is probably natural because both of them exhibit some kind of affine structure (although infinite-dimensional).

5 | NUMERICAL APPLICATION

5.1 | Numerical scheme

We explain in this section how to compute numerically the characteristic function of the log-price in a rough Heston model. By Theorem 4.1, $L_p(a,t)$ is entirely defined through the fractional Riccati equation (4.6)

$$D^\alpha h(a,t) = F(a, h(a,t)), \qquad I^{1-\alpha}h(a,0) = 0,$$

where

$$F(a,x) = \frac{1}{2}(-a^2 - ia) + \gamma(ia\rho\nu - 1)x + \frac{(\gamma\nu)^2}{2}x^2.$$

Several schemes for solving (4.6) numerically can be found in the literature. Most of them are based on the idea that (4.6) implies the following Volterra equation:

$$h(a,t) = \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}F(a, h(a,s))\,ds. \tag{5.1}$$

Then one develops numerical schemes for (5.1). Here, we choose the well-known fractional Adams method investigated in Diethelm, Ford, and Freed (2002, 2004) and Diethelm and Freed (1999). The idea goes as follows. Let us write $g(a,t) = F(a, h(a,t))$. Over a regular discrete time-grid $(t_k)_{k\in\mathbb{N}}$ with mesh Δ ($t_k = k\Delta$), we estimate

$$h(a,t_{k+1}) = \frac{1}{\Gamma(\alpha)}\int_0^{t_{k+1}} (t_{k+1}-s)^{\alpha-1}g(a,s)ds$$

by

$$\frac{1}{\Gamma(\alpha)}\int_0^{t_{k+1}} (t_{k+1}-s)^{\alpha-1}\hat{g}(a,s)ds,$$

where

$$\hat{g}(a,t) = \frac{t_{j+1}-t}{t_{j+1}-t_j}\hat{g}(a,t_j) + \frac{t-t_j}{t_{j+1}-t_j}\hat{g}(a,t_{j+1}), \quad t\in[t_j,t_{j+1}), \quad 0\le j\le k.$$

This corresponds to a trapezoidal discretization of the fractional integral and leads to the following scheme:

$$\hat{h}(a, t_{k+1}) = \sum_{0 \le j \le k} a_{j,k+1} F(a, \hat{h}(a, t_j)) + a_{k+1,k+1} F(a, \hat{h}(a, t_{k+1})), \tag{5.2}$$

with

$$a_{0,k+1} = \frac{\Delta^\alpha}{\Gamma(\alpha+2)} \left(k^{\alpha+1} - (k-\alpha)(k+1)^\alpha \right),$$

$$a_{j,k+1} = \frac{\Delta^\alpha}{\Gamma(\alpha+2)} ((k-j+2)^{\alpha+1} + (k-j)^{\alpha+1} - 2(k-j+1)^{\alpha+1}), \quad 1 \le j \le k, \tag{5.3}$$

and

$$a_{k+1,k+1} = \frac{\Delta^\alpha}{\Gamma(\alpha+2)}.$$

However, $\hat{h}(a, t_{k+1})$ being on both sides of (5.2), this scheme is implicit. Thus, in a first step, we compute a pre-estimation of $\hat{h}(a, t_{k+1})$ based on a Riemann sum that we then plug into the trapezoidal quadrature. This pre-estimation, called predictor and that we denote by $\hat{h}^P(a, t_{k+1})$, is defined by

$$\hat{h}^P(a, t_{k+1}) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \tilde{g}(a, s) ds,$$

with

$$\tilde{g}(a, t) = \hat{g}(a, t_j), \quad t \in [t_j, t_{j+1}), \quad 0 \le j \le k.$$

Therefore,

$$\hat{h}^P(a, t_{k+1}) = \sum_{0 \le j \le k} b_{j,k+1} F\left(a, \hat{h}(a, t_j)\right),$$

where

$$b_{j,k+1} = \frac{\Delta^\alpha}{\Gamma(\alpha+1)} \left((k-j+1)^\alpha - (k-j)^\alpha\right), \quad 0 \le j \le k.$$

Thus, the final explicit numerical scheme is given by

$$\hat{h}(a, t_{k+1}) = \sum_{0 \le j \le k} a_{j,k+1} F(a, \hat{h}(a, t_j)) + a_{k+1,k+1} F(a, \hat{h}^P(a, t_j)), \quad \hat{h}(a, 0) = 0,$$

where the weights $a_{j,k+1}$ are defined in (5.3). Theoretical guarantees for the convergence of this scheme are provided in Li and Tao (2009). In particular, it is shown that for given $t > 0$ and $a \in \mathbb{R}$,

$$\max_{t_j \in [0,t]} |\hat{h}(a, t_j) - h(a, t_j)| = o(\Delta)$$

and

$$\max_{t_j \in [\varepsilon,t]} |\hat{h}(a, t_j) - h(a, t_j)| = o(\Delta^{2-\alpha}),$$

for any $\varepsilon > 0$.

5.2 | Numerical illustrations

To compute $L_p(a, t)$, we use the numerical scheme presented above to solve fractional Riccati equations and then plug the numerical solutions into (4.5). Once the characteristic function is obtained, classical methods are available to obtain call prices

$$C(K, T) = \mathbb{E}[(S_T - K)_+];$$

see Carr and Madan (1999), Itkin (2005), Lewis (2001), and the survey, Schmelzle (2010). In our case, we use the Lewis method; see Lewis (2001).

The most costly operation is the computation of $(\hat{h}(a_l, t_j), 1 \leq l \leq N_a, 1 \leq j \leq n)$ from the scheme in Section 5.1, where $n = T/\Delta$ is the number of time steps and N_a is the number of space steps a_l used for the Fourier-type method finally leading to $C(T, K)$. Hence, the complexity of call price computation is $O(N_a n^2)$. Note that for the classical Heston model ($\alpha = 1$), $h(a, t)$ has an explicit form and the complexity is then reduced to $O(N_a)$.

We now give a calibration result on the S&P implied volatility surface of January 7, 2010. We obtain the following optimal parameters for the rough Heston model (1.3):

$$\alpha = 0.62, \quad \gamma = 0.1, \quad \rho = -0.681, \quad V_0 = 0.0392, \quad \nu = 0.331, \quad \theta = 0.3156.$$

We compare model- and market-implied volatility surfaces in Figure 5.1.

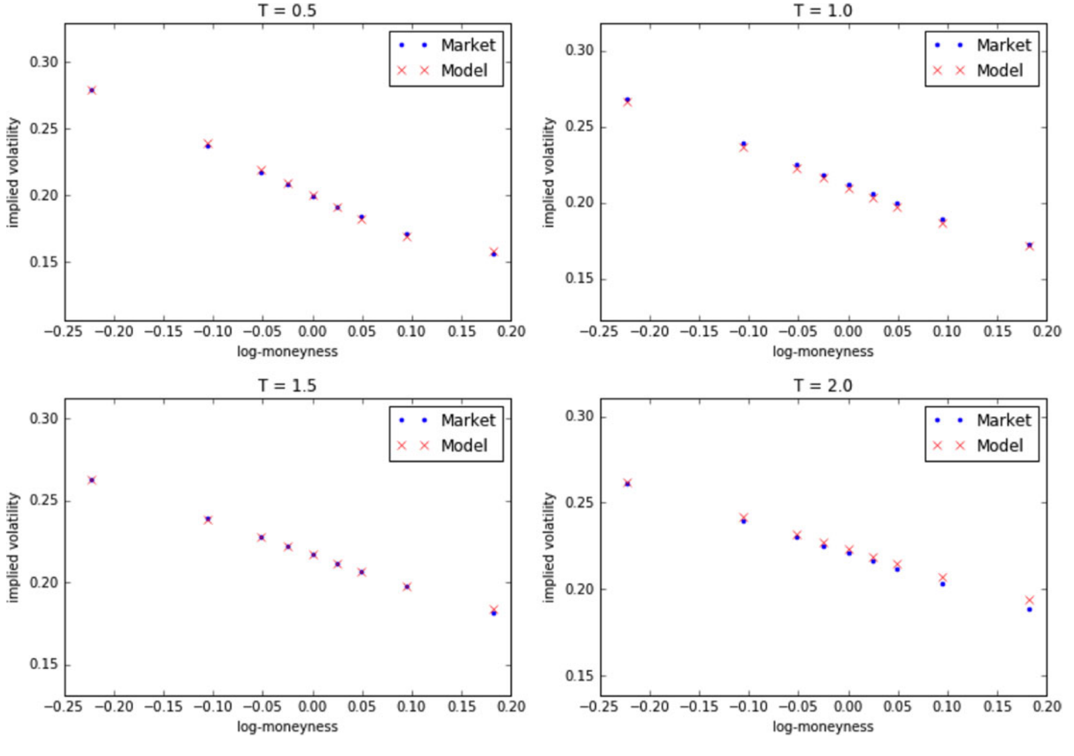

FIGURE 5.1 Implied volatility surface calibration with a rough Heston model [Color figure can be viewed at wileyonlinelibrary.com]

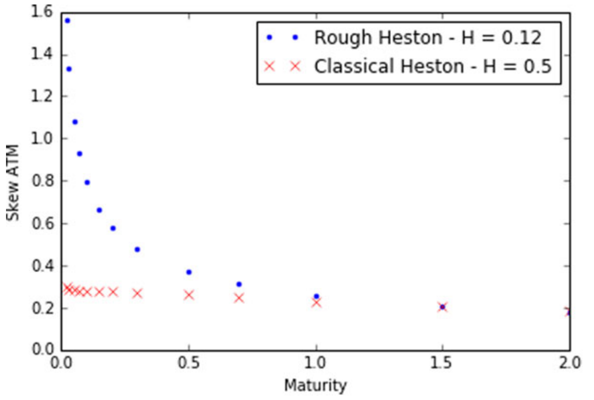

FIGURE 5.2 At-the-money skew as a function of maturity for $\alpha = 1$ and $\alpha = 0.62$ [Color figure can be viewed at wileyonlinelibrary.com]

We see that the rough Heston model provides remarkable fits for the smile, for all the considered maturities. Again, one very important point here is that the model volatility surface can be computed very efficiently, thanks to our procedure.

Finally, we display in Figure 5.2 the term structure of the at-the-money skew, that is, the derivative of the implied volatility with respect to log-strike for at-the-money calls. We compute it for $\alpha = 1$ (classical Heston) and $\alpha = 0.62$ (rough Heston with optimal Hurst parameter equal to 0.12).

In the rough case, the skew explodes when maturity goes to zero, whereas it remains flat with the classical Heston model. This is a remarkable feature of rough volatility models because this exploding behavior is commonly observed on real data and very important for practical applications; see Bayer et al. (2016), Fukasawa (2011), and Jaisson and Rosenbaum (2016).

6 | PROOFS

In the sequel, c denotes a constant that may vary from line to line.

6.1 | Proof of Theorem 2.5

The proof of Theorem 2.5 is close to the one given in El Euch et al. (2016) for the convergence of a microscopic price model to a Heston-like dynamic. The main difference is that we have to deal here with a time-varying baseline intensity $\hat{\mu}_T$, which we have introduced to get a nonzero initial volatility in the limit. As in El Euch et al. (2016), we start by showing the C-tightness of (Λ^T, X^T, Z^T).

6.1.1 | C-tightness of (Λ^T, X^T, Z^T)

We have the following proposition.

Proposition 6.1. *Under Definitions 2.1 and 2.3, the sequence (Λ^T, X^T, Z^T) is C-tight and*

$$\sup_{t\in[0,1]} \left\|\Lambda_t^T - X_t^T\right\| \underset{T\to\infty}{\longrightarrow} 0$$

in probability. Moreover, if (X, Z) is a possible limit point of (X^T, Z^T), then Z is a continuous martingale with $[Z, Z] = diag(X)$.

Proof. C-tightness of X^T and Λ^T: Recall that as in (2.5), we can write

$$\lambda_t^{T,+} = \lambda_t^{T,-} = \hat{\mu}_T(t) + \int_0^t \psi^T(t-s)\hat{\mu}_T(s)ds + \frac{1}{\beta+1}\int_0^t \psi^T(t-s)\left(dM_s^{T,+} + \beta dM_s^{T,-}\right),$$

where

$$M_t^T = \left(M_t^{T,+}, M_t^{T,-}\right) = N_t^T - \int_0^t \lambda_s^T ds$$

is a martingale. Using that $\int_0^{\cdot}(f * g) = (\int_0^{\cdot} f) * g$, we get

$$\mathbb{E}\left[N_T^{T,+}\right] = \mathbb{E}\left[N_T^{T,-}\right] = \mathbb{E}\left[\int_0^T \lambda_s^{T,+} ds\right] = \int_0^T \hat{\mu}_T(s)ds + \int_0^T \psi^T(T-s)\left(\int_0^s \hat{\mu}_T(u)du\right)ds.$$

Consequently, $\hat{\mu}$ being a positive function and using that

$$1 + \int_0^\infty \psi^T(s)ds = 1 + \sum_{k\geq 1}\int_0^\infty (\varphi^T)^{*k} = \sum_{k\geq 0}(a_T)^k = \frac{T^\alpha}{\gamma},$$

we obtain

$$\mathbb{E}\left[N_T^{T,+}\right] \leq \int_0^T \hat{\mu}_T(s)ds\left(1 + \int_0^\infty \psi^T(s)ds\right) \leq \frac{1}{\gamma}T^{\alpha+1}\int_0^1 \hat{\mu}_T(Ts)ds.$$

Moreover, from the definition of $\hat{\mu_T}$ and Remark 2.2, we have

$$\int_0^1 \hat{\mu}_T(Ts)ds = \mu T^{\alpha-1}\left(1 + \xi\int_0^1 s^{-\alpha}\frac{(sT)^\alpha}{\gamma}\int_{sT}^\infty \varphi(u)duds + \gamma T^{-\alpha}\int_0^1\int_0^{sT}\varphi(u)duds\right) \leq cT^{\alpha-1}.$$

Hence, $\mathbb{E}[N_T^{T,+}] \leq cT^{2\alpha}$ and therefore

$$\mathbb{E}\left[X_1^T\right] = \mathbb{E}\left[\Lambda_1^T\right] \leq c,$$

for each component. Each component of X^T and Λ^T being increasing, we deduce the tightness of each component of (X^T, Λ^T). Furthermore, the maximum jump size of X^T and Λ^T being $\frac{1-a_T}{T^\alpha \mu}$ which goes to zero, the C-tightness of (X^T, Λ^T) is obtained from proposition VI-3.26 in Jacod and Shiryaev (2013).

C-tightness of Z^T: It is easy to check that

$$\langle Z^T, Z^T\rangle = diag(\Lambda^T),$$

which is C-tight. From theorem VI-4.13 in Jacod and Shiryaev (2013), this gives the tightness of Z^T. The maximum jump size of Z^T vanishing as T goes to infinity, we obtain that Z^T is C-tight. Convergence of $X^T - \Lambda^T$ We have

$$X_t^T - \Lambda_t^T = \frac{1-a_T}{T^\alpha \mu}M_{tT}^T.$$

From Doob's inequality, we get that for each component

$$\mathbb{E}\left[\sup_{t\in[0,1]}\left|\Lambda_t^T - X_t^T\right|^2\right] \leq cT^{-4\alpha}\mathbb{E}\left[M_T^T\right]^2.$$

Because $[M^T, M^T] = N^T$, we deduce

$$\mathbb{E}\left[\sup_{t\in[0,1]} |\Lambda_t^T - X_t^T|^2\right] \le cT^{-4\alpha}\mathbb{E}\left[N_T^T\right] \le cT^{-2\alpha}.$$

This gives the uniform convergence to zero in probability of $X^T - \Lambda^T$.

Limit of Z^T: Let (X, Z) be a limit point of (X^T, Z^T). We know that (X, Z) is continuous and from corollary IX-1.19 in Jacod and Shiryaev (2013), Z is a local martingale. Moreover, because

$$[Z^T, Z^T] = diag(X^T),$$

using theorem VI-6.26 in Jacod and Shiryaev (2013), we get that $[Z, Z]$ is the limit of $[Z^T, Z^T]$ and $[Z, Z] = diag(X)$. By Fatou's lemma, the expectation of $[Z, Z]$ is finite and therefore Z is a martingale. □

6.1.2 | Convergence of X^T and Z^T

First remark that because

$$\sup_{t\in[0,1]} \left|\Lambda_t^T - X_t^T\right| \underset{T\to\infty}{\longrightarrow} 0$$

and

$$\Lambda_t^{T,+} = \Lambda_t^{T,-},$$

we get

$$\sup_{t\in[0,1]} \left|X_t^{T,+} - X_t^{T,-}\right| \underset{T\to\infty}{\longrightarrow} 0.$$

Therefore, if a subsequence of $X_t^{T,+}$ converges to some X, then the associated subsequence of $X_t^{T,-}$ converges to the same X. We have the following proposition for the limit points of $X_t^{T,+}$ and $X_t^{T,-}$.

Proposition 6.2. *If (X, X, Z^+, Z^-) is a possible limit point for $(X^{T,+}, X^{T,-}, Z^{T,+}, Z^{T,-})$, then (X_t, Z_t^+, Z_t^-) can be written as*

$$X_t = \int_0^t Y_s ds, \qquad Z_t^+ = \int_0^t \sqrt{Y_s} dB_s^1, \qquad Z_t^- = \int_0^t \sqrt{Y_s} dB_s^2,$$

where (B_1, B_2) is a bidimensional Brownian motion and Y is solution of

$$Y_t = \xi(1 - F^{\alpha,\gamma}(t)) + F^{\alpha,\gamma}(t) + \sqrt{\frac{1+\beta^2}{\gamma\mu(1+\beta)^2}} \int_0^t f^{\alpha,\gamma}(t-s)\sqrt{Y_s} dB_s, \tag{6.1}$$

with

$$B = \frac{B^1 + \beta B^2}{\sqrt{1+\beta^2}}.$$

Furthermore, for any $\varepsilon > 0$, Y has Hölder regularity $\alpha - 1/2 - \varepsilon$.

Proof. A convenient equality We first show the following equality

$$\hat{\mu}_T(t) + \int_0^t \psi^T(t-s)\hat{\mu}_T(s)ds = \mu_T + \xi\mu_T \frac{1}{1-a_T} + \mu_T(1-\xi)\int_0^t \psi^T(t-s)ds. \tag{6.2}$$

To obtain this result, we consider (6.2) as an equation with unknown $\hat{\mu}_T$. From Lemma A.1, it admits a solution. We now look for necessary condition for this solution, showing in the end that the specific $\hat{\mu}_T$ given in Definition 2.3 is the only possible choice. Using convolution by φ^T and the fact that $\psi^T * \varphi^T = \psi^T - \varphi^T$, we obtain from the left-hand side of (6.2):

$$\begin{aligned}
&\int_0^t \hat{\mu}_T(s)\varphi^T(t-s)ds + \int_0^t \int_0^s \psi^T(s-u)\hat{\mu}_T(u)du\varphi^T(t-s)ds \\
&= \int_0^t \hat{\mu}_T(s)\varphi^T(t-s)ds + \int_0^t \int_0^{t-u} \psi^T(s)\varphi^T(t-u-s)ds\hat{\mu}_T(u)du \\
&= \int_0^t \hat{\mu}_T(s)\varphi^T(t-s)ds + \int_0^t \left(\psi^T(t-u) - \varphi^T(t-u)\right)\hat{\mu}_T(u)du \\
&= \int_0^t \psi^T(t-s)\hat{\mu}_T(s)ds.
\end{aligned}$$

From the right-hand side of (6.2), we get

$$\begin{aligned}
&\int_0^t \varphi^T(t-s)\left(\mu_T + \xi\mu_T\frac{1}{1-a_T}\right)ds + \mu_T(1-\xi)\int_0^t \varphi^T(t-s)\int_0^s \psi^T(s-u)du\, ds \\
&= \mu_T\left(1+\xi\frac{1}{1-a_T}\right)\int_0^t \varphi^T(t-s)ds + \mu_T(1-\xi)\int_0^t\int_0^{t-u}\psi^T(s)\varphi^T(t-u-s)ds\, du \\
&= \mu_T\left(1+\xi\frac{1}{1-a_T}\right)\int_0^t \varphi^T(t-s)ds + \mu_T(1-\xi)\int_0^t \left(\psi^T(t-u) - \varphi^T(t-u)\right)du.
\end{aligned}$$

Consequently, we necessarily have

$$\int_0^t \psi^T(t-s)\hat{\mu}_T(s)ds = \mu_T\xi\left(\frac{1}{1-a_T}+1\right)\int_0^t \varphi^T(t-s)ds + \mu_T(1-\xi)\int_0^t \psi^T(t-s)ds.$$

This last equation together with (6.2) gives that the only possible choice is

$$\hat{\mu}_T(t) = \mu_T + \xi\mu_T\frac{1}{1-a_T}\left(1 - \int_0^t \varphi^T(t-s)ds\right) - \mu_T\xi\int_0^t \varphi^T(t-s)ds.$$

End of the proof of Proposition 6.2 Recall that $\lambda_t^{T,+} = \lambda_t^{T,-}$. Note that using similar computations as in Section 2.2 together with (6.2) we can write

$$\begin{aligned}
\lambda_t^{T,+} &= \mu_T + \mu_T\int_0^t \psi^T(t-s)ds + \xi\mu_T\left(\frac{1}{1-a_T} - \int_0^t \psi^T(t-s)ds\right) \\
&\quad + \frac{1}{\beta+1}\int_0^t \psi^T(t-s)\left(dM_s^{T,+} + \beta dM_s^{T,-}\right).
\end{aligned}$$

Then using Fubini theorem together with the fact that $\int_0^{\cdot}(f * g) = (\int_0^{\cdot} f) * g$, we get

$$\int_0^t \lambda_s^{T,+} ds = \mu_T t + \mu_T \int_0^t \psi^T(t-s) s ds + \xi \mu_T \left(\frac{t}{1-a_T} - \int_0^t \psi^T(t-s) s ds \right)$$
$$+ \frac{1}{\beta+1} \int_0^t \psi^T(t-s) \left(M_s^{T,+} + \beta M_s^{T,-} \right) ds.$$

Therefore, for $t \in [0,1]$, we have the decomposition

$$\Lambda_t^{T,+} = \Lambda_t^{T,-} = T_1 + T_2 + T_3, \tag{6.3}$$

with

$$T_1 = (1-a_T)t,$$
$$T_2 = T(1-a_T) \int_0^t \psi^T\left(T(t-s)\right) s ds + \xi \left(t - T(1-a_T) \int_0^t \psi^T\left(T(t-s)\right) s ds \right),$$
$$T_3 = \frac{1}{\sqrt{\gamma \mu (1+\beta)^2}} \int_0^t T(1-a_T) \psi^T\left(T(t-s)\right) \left(Z_s^{T,+} + \beta Z_s^{T,-} \right) ds.$$

Now recall that we have shown in (2.6) that

$$T(1-a_T)\psi(T.) = a_T f^{\alpha,\gamma}.$$

Thus,

$$T_2 \underset{T\to\infty}{\longrightarrow} \int_0^t f^{\alpha,\gamma}(t-s) s ds + \xi \left(t - \int_0^t f^{\alpha,\gamma}(t-s) s ds \right)$$

and

$$T_3 \underset{T\to\infty}{\longrightarrow} \frac{1}{\sqrt{\gamma \mu (1+\beta)^2}} \int_0^t f^{\alpha,\gamma}(t-s)(Z_s^+ + \beta Z_s^-) ds.$$

Therefore, letting T go to infinity in (6.3), we obtain using Proposition 6.1 that X satisfies

$$X_t = \int_0^t f^{\alpha,\gamma}(t-s) s ds + \xi \left(t - \int_0^t f^{\alpha,\gamma}(t-s) s ds \right) + \frac{1}{\sqrt{\gamma \mu (1+\beta)^2}} \int_0^t f^{\alpha,\gamma}(t-s) \left(Z_s^+ + \beta Z_s^- \right) ds.$$

In the same way as for the proof of theorem 3.2 in Jaisson and Rosenbaum (2016), we show that

$$X_t = \int_0^t Y_s ds,$$

where Y satisfies

$$Y_t = F^{\alpha,\gamma}(t) + \xi \left(1 - F^{\alpha,\gamma}(t)\right) + \frac{1}{\sqrt{\gamma \mu (1+\beta)^2}} \int_0^t f^{\alpha,\gamma}(t-s) \left(dZ_s^+ + \beta dZ_s^- \right).$$

Because, by Proposition 6.1,

$$[Z, Z] = \int_0^t Y_s ds \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

we can apply theorem V-3.9 in Revuz and Yor (1999) to show the existence of a bidimensional Brownian motion $(\boldsymbol{B}^1, \boldsymbol{B}^2)$ such that

$$Z_t^+ = \int_0^t \sqrt{Y_s} dB_s^1, \qquad Z_t^- = \int_0^t \sqrt{Y_s} dB_s^2.$$

Finally, we define the following Brownian motion:

$$\boldsymbol{B} = \frac{\boldsymbol{B}^1 + \beta \boldsymbol{B}^2}{\sqrt{1+\beta^2}}.$$

Then, in the same way as for the proof of theorem 3.2 in Jaisson and Rosenbaum (2016), we get that Y satisfies

$$Y_t = F^{\alpha,\gamma}(t) + \xi(1 - F^{\alpha,\gamma}(t)) + \sqrt{\frac{1+\beta^2}{\gamma\mu(1+\beta)^2}} \int_0^t f^{\alpha,\gamma}(t-s)\sqrt{Y_s} dB_s,$$

and has Hölder regularity $\alpha - 1/2 - \varepsilon$ for any $\varepsilon > 0$. □

6.1.3 | End of the proof of Theorem 2.5

We now recall the following proposition stating that the process Y is uniquely defined by equation (6.1) and that this equation is equivalent to that given in Theorem 2.5. The proof of this result can be found in El Euch et al. (2016). Theorem 2.5 is readily obtained from this proposition together with Propositions 6.1 and 6.2.

Proposition 6.3. *Let γ, ν, θ and V_0 be positive constants, $\alpha \in (1/2, 1)$ and $\boldsymbol{B}$ be a Brownian motion. The process V is solution of the following fractional stochastic differential equation:*

$$V_t = V_0(1 - F^{\alpha,\gamma}(t)) + \theta F^{\alpha,\gamma}(t) + \nu \int_0^t f^{\alpha,\gamma}(t-s)\sqrt{V_s} dB_s$$

if and only if it is solution of

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1}\gamma(\theta - V_s) ds + \frac{\gamma\nu}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1}\sqrt{V_s} dB_s.$$

Furthermore, both equations admit a unique strong solution.

6.1.4 | Proof of Corollary 2.6

From Theorem 2.5, we know that P^T converges in law for the Skorokhod topology to the process P given by

$$P_t = \sqrt{\frac{\theta}{2}} \int_0^t \sqrt{Y_s}\left(dB_s^1 - dB_s^2\right) - \frac{\theta}{2} \int_0^t Y_s ds.$$

Let $V_t = \theta Y_t$ and $W_t = \frac{1}{\sqrt{2}}(B_t^1 - B_t^2)$. Then,

$$P_t = \int_0^t \sqrt{V_s} dW_s - \frac{1}{2}\int_0^t V_s ds,$$

where

$$V_t = \xi\theta + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\gamma(\theta - V_s)ds + \gamma\sqrt{\frac{\theta(1+\beta^2)}{\gamma\mu(1+\beta)^2}}\frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\sqrt{V_s}dW'_s$$

and (W, B) is a correlated bidimensional Brownian motion with

$$d\langle W, B\rangle_t = \frac{1-\beta}{\sqrt{2(1+\beta^2)}}dt.$$

6.2 | Proof of Theorem 4.1

We now give the proof of Theorem 4.1. We do it for $t \in [0, 1]$ but the proof can obviously be extended for any $t \geq 0$. We start by controlling the process $Y^T(a, t) - (1, 1)$. In the sequel, $c(a)$ denotes a positive constant independent of t and T that may vary from line to line.

6.2.1 | Control of $Y^T(a, t) - (1, 1)$

We have the following proposition.

Proposition 6.4. *For any $t \in [0, 1]$,*

$$T^\alpha \|Y^T(a, t) - (1, 1)\| \leq c(a).$$

Proof. Let us show that

$$T^\alpha |Y^{T,+}(a, t) - 1| \leq c(a).$$

Recall that $Y^T(a, t)$ is defined in Section 4.1 for $a \in \mathbb{R}$ by

$$Y^T(a, t) = (Y^{T,+}(a, t), Y^{T,-}(a, t)) = \left(C^{T,+}\left(\left(a_T^+, a_T^-\right), tT\right), C^{T,-}\left(\left(a_T^+, a_T^-\right), tT\right)\right),$$

with

$$a_T^+ = a\sqrt{\frac{\gamma\theta}{2\mu}}T^{-\alpha} - a\frac{\gamma\theta}{2\mu}T^{-2\alpha}, \qquad a_T^- = -a\sqrt{\frac{\gamma\theta}{2\mu}}T^{-\alpha}.$$

Using the elements in the proof of Theorem 3.1 in Section 3.3, we get that

$$C^{T,+}((a, b), t) = \mathbb{E}\left[\exp\left(ia + ia\tilde{N}_t^{T,+} + ib\tilde{N}_t^{T,-}\right)\right],$$

where $\tilde{N}^{T,\cdot} = (\tilde{N}^{T,+}, \tilde{N}^{T,-})$ is a bidimensional Hawkes process with intensity $(\tilde{\lambda}^T, \tilde{\lambda}^T)$ given by

$$\tilde{\lambda}_t^T = \frac{1}{\beta+1}\varphi^T(t) + \frac{1}{\beta+1}\int_0^t \varphi^T(t-s)\left(d\tilde{N}_s^{T,+} + \beta d\tilde{N}_s^{T,-}\right).$$

As already seen, using Lemma A.1, we can rewrite the intensity under the following form:

$$\tilde{\lambda}_t^T = \frac{1}{\beta+1}\psi^T(t) + \frac{1}{\beta+1}\int_0^t \psi^T(t-s)\left(d\tilde{M}_s^{T,+} + \beta d\tilde{M}_s^{T,-}\right),$$

where $\tilde{M}^T = (\tilde{M}^{T,+}, \tilde{M}^{T,-}) = \tilde{N}^T - \int_0^{\cdot} \tilde{\lambda}^T(s)ds(1,1)$ is a martingale. Using Fubini theorem, we get

$$\int_0^{tT} \tilde{\lambda}_s^T ds = \frac{1}{\beta+1}T\int_0^t \psi^T(Ts)ds + \frac{1}{\beta+1}\int_0^t T\psi^T\left(T(t-s)\right)\left(\tilde{M}_{sT}^{T,+} + \beta\tilde{M}_{sT}^{T,-}\right)ds.$$

Then, from (2.6), we derive

$$\int_0^{tT} \tilde{\lambda}_s^T ds = \frac{1}{\gamma(\beta+1)}a_T T^\alpha F^{\alpha,\gamma}(t) + \frac{1}{\gamma(\beta+1)}a_T T^\alpha \int_0^t f^{\alpha,\gamma}(t-s)\left(\tilde{M}_{sT}^{T,+} + \beta\tilde{M}_{sT}^{T,-}\right)ds. \tag{6.4}$$

Consequently,

$$\mathbb{E}\left[\int_0^{tT} \tilde{\lambda}_s^T ds\right] \le \frac{1}{\gamma(\beta+1)}F^{\alpha,\gamma}(1)T^\alpha.$$

Let us now set $\tilde{X}_t^T = a_T^+\tilde{N}_{tT}^{T,+} + a_T^-\tilde{N}_{tT}^{T,-}$. Using the last inequality, we deduce

$$\left|\mathbb{E}\tilde{X}_t^T\right| \le c|a|T^{-\alpha}F^{\alpha,\gamma}(1).$$

Now recall that

$$T^\alpha(Y^{T,+}(a,t)-1) = T^\alpha\left(\mathbb{E}\left[\exp\left(ia_T^+ + ia_T^+\tilde{N}_{tT}^{T,+} + ia_T^-\tilde{N}_{tT}^{T,-}\right)\right]-1\right).$$

Using the fact that there exists $c>0$ such that for any $x\in\mathbb{R}$,

$$|\exp(ix) - 1 - ix| \le c|x|^2,$$

we obtain

$$\begin{aligned}
T^\alpha|Y^{T,+}(a,t)-1| &= T^\alpha\left|\mathbb{E}\left[\exp\left(ia_T^+ + i\tilde{X}_t^T\right) - 1 - i\tilde{X}_t^T - ia_T^+ + i\tilde{X}_t^T + ia_T^+\right]\right| \\
&\le T^\alpha\left|\mathbb{E}\left[\tilde{X}_t^T\right]\right| + T^\alpha\left|a_T^+\right| + T^\alpha\mathbb{E}\left[\left|\exp\left(ia_T^+ + i\tilde{X}_t^T\right) - 1 - i\tilde{X}_t^T - ia_T^+\right|\right] \\
&\le c(a)\left(1 + T^\alpha\left(a_T^+\right)^2 + T^\alpha\mathbb{E}\left[\left(\tilde{X}_t^T\right)^2\right]\right) \\
&\le c(a)\left(1 + T^\alpha\mathbb{E}\left[\left(\tilde{X}_t^T\right)^2\right]\right).
\end{aligned}$$

Then, using that

$$\tilde{X}_t^T = a\sqrt{\frac{\gamma\theta}{2\mu}}T^{-\alpha}\left(\tilde{N}_{tT}^{T,+} - \tilde{N}_{tT}^{T,-}\right) - a\frac{\gamma\theta}{2\mu}T^{-2\alpha}\tilde{N}_{tT}^{T,+}$$

together with the fact that $\tilde{N}^{T,+} - \tilde{N}^{T,-} = \tilde{M}^{T,+} - \tilde{M}^{T,-}$, we deduce

$$T^{\alpha}\mathbb{E}\left[\left(\tilde{X}_t^T\right)^2\right] \le ca^2 T^{-\alpha}\mathbb{E}\left[\left(\tilde{M}_{tT}^{T,+} - \tilde{M}_{tT}^{T,-}\right)^2\right] + ca^2 T^{-3\alpha}\mathbb{E}\left[\left(\tilde{N}_{tT}^{T,+}\right)^2\right].$$

Because $[\tilde{M}^{T,+} - \tilde{M}^{T,-}, \tilde{M}^{T,+} - \tilde{M}^{T,-}] = \tilde{N}^{T,+} + \tilde{N}^{T,-}$, we get

$$\begin{aligned}
T^{\alpha}\mathbb{E}\left[\left(\tilde{X}_t^T\right)^2\right] &\le ca^2 T^{-\alpha}\mathbb{E}\left[\tilde{N}_{tT}^{T,+} + \tilde{N}_{tT}^{T,-}\right] + ca^2 T^{-3\alpha}\mathbb{E}\left[\left(\tilde{N}_{tT}^{T,+}\right)^2\right] \\
&\le ca^2\left(T^{-\alpha}\mathbb{E}\left[\int_0^{tT} \tilde{\lambda}_s^T ds\right] + T^{-3\alpha}\mathbb{E}\left[\left(\tilde{N}_{tT}^{T,+}\right)^2\right]\right) \\
&\le ca^2\left(1 + T^{-3\alpha}\mathbb{E}\left[\left(\tilde{N}_{tT}^{T,+}\right)^2\right]\right).
\end{aligned}$$

In order to control the term $\mathbb{E}[(\tilde{N}_{tT}^{T,+})^2]$, we now compute a bound for $\mathbb{E}[(\int_0^{tT} \tilde{\lambda}_s^T ds)^2]$. Using (6.4), this last quantity is equal to

$$\frac{1}{\gamma^2(\beta+1)^2}a_T^2 T^{2\alpha}\left(F^{\alpha,\gamma}(t)\right)^2 + \frac{1}{\gamma^2(\beta+1)^2}a_T^2 T^{2\alpha}\mathbb{E}\left[\left(\int_0^t f^{\alpha,\gamma}(t-s)\left(\tilde{M}_{sT}^{T,+} + \beta\tilde{M}_{sT}^{T,-}\right)ds\right)^2\right],$$

which is smaller than

$$c(a)T^{2\alpha}\left(1 + \mathbb{E}\left[\int_0^t \left(f^{\alpha,\gamma}(t-s)\right)^2 (\tilde{M}_{sT}^{T,+} + \beta\tilde{M}_{sT}^{T,-})^2 ds\right]\right).$$

Because $[\tilde{M}^{T,+} + \beta\tilde{M}^{T,-}, \tilde{M}^{T,+} + \beta\tilde{M}^{T,-}] = \tilde{N}^{T,+} + \beta^2\tilde{N}^{T,-}$, we obtain

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^{tT} \tilde{\lambda}_s^T ds\right)^2\right] &\le c(a)T^{2\alpha}\left(1 + \int_0^t \left(f^{\alpha,\gamma}(t-s)\right)^2 \mathbb{E}\left[\tilde{N}_{sT}^{T,+} + \beta^2\tilde{N}_{sT}^{T,-}\right] ds\right) \\
&\le c(a)T^{2\alpha}\left(1 + \int_0^t \left(f^{\alpha,\gamma}(t-s)\right)^2 \mathbb{E}\left[\int_0^{sT} \tilde{\lambda}_u^T du\right] ds\right) \\
&\le c(a)T^{2\alpha}\left(1 + T^{\alpha}\int_0^1 \left(f^{\alpha,\gamma}(s)\right)^2 ds\right) \\
&\le c(a)T^{3\alpha}.
\end{aligned}$$

Thus,

$$\mathbb{E}\left[\left(\tilde{N}_{tT}^{T,+}\right)^2\right] \le 2\mathbb{E}\left[\left(\tilde{M}_{tT}^{T,+}\right)^2\right] + 2\mathbb{E}\left[\left(\int_0^{tT} \tilde{\lambda}_s^T ds\right)^2\right] \le c(a)T^{3\alpha}.$$

Finally, $T^{\alpha}\mathbb{E}[(\tilde{X}_t^T)^2] \le c(a)$ and therefore

$$T^{\alpha}|Y^{T,+}(a,t) - 1| \le c(a).$$

The fact that

$$T^\alpha |Y^{T,-}(a,t) - 1| \leq c(a)$$

is proved similarly. □

6.2.2 | Convergence of $T^\alpha(Y^T - (1,1))$

Let $\kappa = \gamma\theta/(2\mu)$. We have the following proposition.

Proposition 6.5. *The sequence $T^\alpha(Y^T(a,t) - (1,1))$ converges uniformly in $t \in [0,1]$ to $(c(a,t), d(a,t))$, where (c,d) are solutions of*

$$c(a,t) = ia\sqrt{\kappa} - ia\frac{\kappa}{\gamma(\beta+1)}F^{\alpha,\gamma}(t) + \frac{1}{2\gamma(\beta+1)}\int_0^t \left(c^2(a,t-s) + d^2(a,t-s)\right) f^{\alpha,\gamma}(s)ds$$

$$d(a,t) = -ia\sqrt{\kappa} - ia\frac{\beta\kappa}{\gamma(\beta+1)}F^{\alpha,\gamma}(t) + \frac{\beta}{2\gamma(\beta+1)}\int_0^t \left(c^2(a,t-s) + d^2(a,t-s)\right) f^{\alpha,\gamma}(s)ds.$$

Proof. Convenient rewriting of $T^\alpha(Y^T - (1,1))$: Using the fact that the complex logarithm[5] is analytic on the set $\mathbb{C}/\mathbb{R}^-$, we can show that there exists $c > 0$ such that for any $x \in \mathbb{C}$ with $|x| < 1/2$,

$$\left|\log(1+x) - x + \frac{1}{2}x^2\right| \leq c|x|^3.$$

Thus, we can write

$$\log\left(Y^T(a,t)\right) = Y^T(a,t) - (1,1) - \frac{1}{2}\left(Y^T(a,t) - (1,1)\right)^2 - \varepsilon^T(a,t),$$

with $|\varepsilon^T(a,t)| \leq c(a)T^{-3\alpha}$. Indeed, for large enough T, we have from Proposition 6.4 that $|Y^{T,+}(a,t) - 1| \leq 1/2$ and $|Y^{T,-}(a,t) - 1| \leq 1/2$, uniformly in t. Now, again from Proposition 6.4, it is easy to see that

$$\left\| i\left(a_T^+, a_T^-\right) + \int_0^t T\left(Y^T(a,t-s) - (1,1)\right).\phi^T(Ts)ds \right\| \leq c(a)T^{-\alpha} \underset{T\to\infty}{\longrightarrow} 0.$$

Hence, for large enough T, the imaginary part of

$$i\left(a_T^+, a_T^-\right) + \int_0^t T(Y^T(a,t-s) - (1,1)).\phi^T(Ts)ds$$

has a norm, which is smaller than π. Therefore,

$$\log\left(\exp\left(i\left(a_T^+, a_T^-\right) + \int_0^t T\left(Y^T(a,t-s) - (1,1)\right).\phi^T(Ts)ds\right)\right)$$

is equal to

$$i\left(a_T^+, a_T^-\right) + \int_0^t T\left(Y^T(a,t-s) - (1,1)\right).\phi^T(Ts)ds.$$

Then, using equation (4.1), we get

$$Y^T(a,t) - (1,1) = \frac{1}{2}\left(Y^T(a,t) - (1,1)\right)^2 + \varepsilon^T(a,t) + ia\sqrt{\kappa}T^{-\alpha}(1,-1)$$
$$- ia\kappa T^{-2\alpha}(1,0) + T\int_0^t \left(Y^T(a,t-s) - (1,1)\right).\phi^T(Ts)ds.$$

Using again the fact that

$$\sum_{k\geq 1}\left(T\phi^T(T.)\right)^{*k} = a_T\frac{T^\alpha}{\gamma}f^{\alpha,\gamma}\chi,$$

together with Lemma A.1, we derive

$$Y^T(a,t) - (1,1) = \frac{1}{2}\left(Y^T(a,t) - (1,1)\right)^2 + \varepsilon^T(a,t) + ia\sqrt{\kappa}T^{-\alpha}(1,-1) - ia\kappa T^{-2\alpha}(1,0)$$
$$+ \frac{a_T}{2}\frac{T^\alpha}{\gamma}\int_0^t \left(Y^T(a,t-s) - (1,1)\right)^2.\chi f^{\alpha,\gamma}(s)ds + \frac{a_T}{\gamma}T^\alpha\int_0^t \varepsilon^T(a,t-s).\chi f^{\alpha,\gamma}(s)ds$$
$$+ ia\sqrt{\kappa}\frac{a_T}{\gamma}(1,-1).\chi F^{\alpha,\gamma}(t) - ia\kappa T^{-\alpha}\frac{a_T}{\gamma}(1,0).\chi F^{\alpha,\gamma}(t).$$

Let

$$\varepsilon_1^T(a,t) = \frac{1}{2}\left(Y^T(a,t) - (1,1)\right)^2 + \varepsilon^T(a,t) - ia\kappa T^{-2\alpha}(1,0) + \frac{a_T}{\gamma}T^\alpha\int_0^t \varepsilon^T(a,t-s).\chi f^{\alpha,\gamma}(s)ds.$$

We have

$$Y^T(a,t) - (1,1) = \varepsilon_1^T(a,t) + ia\sqrt{\kappa}T^{-\alpha}(1,-1) + \frac{a_T}{2}\frac{T^\alpha}{\gamma}\int_0^t \left(Y^T(a,t-s) - (1,1)\right)^2.\chi f^{\alpha,\gamma}(s)ds$$
$$-ia_T a\frac{\kappa}{\gamma(\beta+1)}T^{-\alpha}F^{\alpha,\gamma}(t)(1,\beta).$$

Let now

$$\varepsilon_2^T(a,t) = -\frac{1}{2}\int_0^t \left(Y^T(a,t-s) - (1,1)\right)^2.\chi f^{\alpha,\gamma}(s)ds + ia\frac{\kappa}{(\beta+1)}T^{-2\alpha}F^{\alpha,\gamma}(t)(1,\beta).$$

We obtain

$$Y^T(a,t) - (1,1) = \varepsilon_1^T(a,t) + \varepsilon_2^T(a,t) + ia\sqrt{\kappa}T^{-\alpha}(1,-1) + \frac{1}{2\gamma}T^\alpha\int_0^t \left(Y^T(a,t-s) - (1,1)\right)^2.\chi f^{\alpha,\gamma}(s)ds$$
$$-ia\frac{\kappa}{\gamma(\beta+1)}T^{-\alpha}F^{\alpha,\gamma}(t)(1,\beta).$$

Using Proposition 6.4, we easily see that $T^{2\alpha}\varepsilon_1^T$ and $T^{2\alpha}\varepsilon_2^T$ are uniformly bounded in t and T. We now set

$$\theta^T(a,t) = (\theta^{T,+}(a,t), \theta^{T,-}(a,t)) = T^\alpha(Y^T(a,t) - (1,1))$$

and

$$r^T(a,t) = T^\alpha(\varepsilon_1^T(a,t) + \varepsilon_2^T(a,t)).$$

We have that $T^\alpha r^T$ is uniformly bounded in t and T and

$$\theta^T(a,t) = r^T(a,t) + ia\sqrt{\kappa}(1,-1) - ia\frac{\kappa}{\gamma(\beta+1)}F^{\alpha,\gamma}(t)(1,\beta) + \frac{1}{2\gamma}\int_0^t \left(\theta^T(a,t-s)\right)^2 .\chi f^{\alpha,\gamma}(s)ds.$$

Convergence of θ^T: For fixed a, we now show that $t \to \theta^T(a,t)$ is a Cauchy sequence in the space of continuous functions $C([0,1],\mathbb{R}^2)$ equipped with the sup-norm. Let $\delta > 0$ and $T_0 > 1$ such that for $T > T_0$, $\|r^T(a,t)\|_\infty \leq \frac{\delta}{2}$ for any $t \in [0,1]$. Then for $T > T_0$, $T' > T_0$ and $t \in [0,1]$,

$$\|\theta^T(a,t) - \theta^{T'}(a,t)\| \leq \delta + \frac{1}{2\gamma}\int_0^t \|(\theta^T(a,t-s))^2.\chi - (\theta^{T'}(a,t-s))^2.\chi\| f^{\alpha,\gamma}(s)ds.$$

Because θ^T is uniformly bounded in t and T, we get

$$\|\theta^T(a,t) - \theta^{T'}(a,t)\| \leq \delta + C(a)\int_0^t \|\theta^T(a,t-s) - \theta^{T'}(a,t-s)\| f^{\alpha,\gamma}(s)ds.$$

Using Lemma A.3 in the Appendix, this enables us to show that θ^T is a Cauchy sequence. Consequently, $\theta^T(a,t)$ converges uniformly in t to $(c(a,t), d(a,t))$, where (c,d) is solution to the following equation:

$$c(a,t) = ia\sqrt{\kappa} - ia\frac{\kappa}{\gamma(\beta+1)}F^{\alpha,\gamma}(t) + \frac{1}{2\gamma(\beta+1)}\int_0^t \left(c^2(a,t-s) + d^2(a,t-s)\right) f^{\alpha,\gamma}(s)ds$$

$$d(a,t) = -ia\sqrt{\kappa} - ia\frac{\beta\kappa}{\gamma(\beta+1)}F^{\alpha,\gamma}(t) + \frac{\beta}{2\gamma(\beta+1)}\int_0^t \left(c^2(a,t-s) + d^2(a,t-s)\right) f^{\alpha,\gamma}(s)ds.$$

□

6.2.3 | End of the proof of Theorem 4.1

Deriving the characteristic function

Let $a \in \mathbb{R}$. Recall that from Section 4.1, we have

$$L^T\left(a_T^+, a_T^-, tT\right) = \exp\left(\int_0^t \left(T^\alpha(Y^{T,+}(a,t-s) - 1) + T^\alpha(Y^{T,-}(a,t-s) - 1)\right)\left(T^{1-\alpha}\hat{\mu}(sT)\right)ds\right)$$

and furthermore, from Proposition 6.5,

$$T^\alpha(Y^{T,+}(a,t) - 1) + T^\alpha(Y^{T,-}(a,t) - 1)$$

converges uniformly in t to $c(a,t) + d(a,t)$. Also, using Remark 2.4, we have

$$T^{1-\alpha}\hat{\mu}(tT) = \mu + \mu\xi\left(\frac{t^{-\alpha}}{\gamma}(Tt)^\alpha\int_{tT}^\infty \varphi(s)ds + \gamma T^{-\alpha}\int_0^{tT}\varphi(s)ds\right)$$

and therefore $T^{1-\alpha}\hat{\mu}(tT)$ converges toward

$$\mu\left(1 + \xi\frac{t^{-\alpha}}{\gamma\Gamma(1-\alpha)}\right).$$

In addition, using Proposition 6.4, we get that for given $t \in [0,1]$ and for any $s \in [0,t]$

$$|T^\alpha(Y^{T,+}(a,t-s) - 1) + T^\alpha(Y^{T,-}(a,t-s) - 1)|(T^{1-\alpha}\hat{\mu}(sT)) \leq c(a)(1 + s^{-\alpha}).$$

The right-hand side of the last inequality is integrable over $[0,t]$. Therefore, using the convergence of $L^T(a_T^+, a_T^-, tT)$ toward $L_p(a,t)$ and applying the dominated convergence theorem, we obtain

$$L_p(a,t) = \exp\left(\int_0^t g(a,s)\left(1+\xi\frac{(t-s)^{-\alpha}}{\gamma\Gamma(1-\alpha)}\right)ds\right),$$

where $g(a,t) = \mu(c(a,t)+d(a,t))$. Thus, we have shown that

$$L_p(a,t) = \exp\left(\int_0^t g(a,s)ds + \frac{V_0}{\theta\gamma}I^{1-\alpha}g(a,t)\right).$$

Integral equation for g: We now prove that g is solution of an integral equation. First remark that

$$d(a,t) = \beta c(a,t) - ia(1+\beta)\sqrt{\kappa}.$$

Hence, $g(a,t) = \mu(\beta+1)(c(a,t)-ia\sqrt{\kappa})$, which can be written

$$-ia\frac{\mu\kappa}{\gamma}F^{\alpha,\gamma}(t) + \frac{\mu}{2\gamma}\int_0^t ((c(a,s)-ia\sqrt{\kappa}+ia\sqrt{\kappa})^2 + (\beta(c(a,s)-ia\sqrt{\kappa})-ia\sqrt{\kappa})^2)f^{\alpha,\gamma}(t-s)ds.$$

Thus,

$$\begin{aligned} g(a,t) = &-ia\frac{\mu\kappa}{\gamma}F^{\alpha,\gamma}(t) + \frac{1+\beta^2}{2\mu\gamma(1+\beta)^2}\int_0^t (g(a,s))^2 f^{\alpha,\gamma}(t-s)ds - a^2\frac{\mu\kappa}{\gamma}F^{\alpha,\gamma}(t) \\ &+ia\frac{\sqrt{\kappa}(1-\beta)}{\gamma(\beta+1)}\int_0^t g(a,s)f^{\alpha,\gamma}(t-s)ds. \end{aligned}$$

Using the definition of κ in Section 6.2, we deduce

$$\begin{aligned} g(a,t) = &\frac{\theta}{2}(-a^2-ia)F^{\alpha,\gamma}(t) + ia\frac{\sqrt{\theta}(1-\beta)}{\sqrt{2\gamma\mu}(\beta+1)}\int_0^t g(a,s)f^{\alpha,\gamma}(t-s)ds \\ &+\frac{1+\beta^2}{2\mu\gamma(1+\beta)^2}\int_0^t g^2(a,s)f^{\alpha,\gamma}(t-s)ds, \end{aligned}$$

and from those of ρ and ν in Section 4.1, we finally obtain that $g(a,t)$ is equal to

$$\frac{\theta}{2}(-a^2-ia)F^{\alpha,\gamma}(t) + ia\rho\nu\int_0^t g(a,s)f^{\alpha,\gamma}(t-s)ds + \frac{\nu^2}{2\theta}\int_0^t (g(a,s))^2 f^{\alpha,\gamma}(t-s)ds.$$

Thus,

$$L_p(a,t) = \exp\left(\int_0^t g(a,s)\left(1+\xi\frac{(t-s)^{-\alpha}}{\gamma\Gamma(1-\alpha)}\right)ds\right),$$

with

$$g(a,t) = \int_0^t \left(\frac{\theta}{2}(-a^2-ia) + ia\rho\nu g(a,s) + \frac{\nu^2}{2\theta}(g(a,s))^2\right)f^{\alpha,\gamma}(t-s)ds.$$

Let us now set $h = g/(\theta\gamma)$. Then

$$L_p(a,t) = \exp\left(\int_0^t h(a,s)\left(\theta\gamma + V_0\frac{(t-s)^{-\alpha}}{\Gamma(1-\alpha)}\right)ds\right),$$

with

$$h(a,t) = \int_0^t \left(\frac{1}{2}(-a^2 - ia) + ia\gamma\rho\nu h(a,s) + \frac{(\gamma\nu)^2}{2}(h(a,s))^2\right)\frac{1}{\gamma}f^{\alpha,\gamma}(t-s)ds. \tag{6.5}$$

Using Lemma A.2, we have that equation (6.5) can also be written under the following form:

$$D^\alpha h(a,t) = \frac{1}{2}(-a^2 - ia) + \gamma(ia\rho\nu - 1)h(a,s) + \frac{(\gamma\nu)^2}{2}(h(a,s))^2, \quad I^{1-\alpha}h(a,0) = 0.$$

6.2.4 | Uniqueness of the solution of (4.6)

For a given $a \in \mathbb{R}$, consider two continuous solutions $h_1(a,.)$ and $h_2(a,.)$ of (4.6) or equivalently of (6.5). We have that $|h_1(a,t) - h_2(a,t)|$ is smaller than

$$\int_0^t \left(|a\rho\nu||h_1(a,s) - h_2(a,s)| + \frac{\gamma\nu^2}{2}|\left(h_1(a,s)\right)^2 - \left(h_2(a,s)\right)^2|\right)f^{\alpha,\gamma}(t-s)ds.$$

Using the continuity of $h_1(a,.)$ and $h_2(a,.)$, this is also smaller than

$$c(a)\int_0^t |h_1(a,s) - h_2(a,s)|f^{\alpha,\gamma}(t-s)ds.$$

Thanks to Lemma A.3, this gives $h_1(a,.) = h_2(a,.)$.

ACKNOWLEDGMENTS

We thank Masaaki Fukasawa, Jim Gatheral, and Antoine Jacquier for many interesting discussions and Christa Cuchiero and Josef Teichmann for very relevant comments about the affine nature of the processes considered in this work.

ENDNOTES

[1] Of course by T we implicitly mean T_n with $n \in \mathbb{N}$ tending to infinity.

[2] From now on we write a dot between quantities to emphasize matrix product.

[3] Note that we call this equation rough because of the presence of the kernel $(t-s)^{\alpha-1}$. However, it is not directly related to rough paths theory.

[4] Actually, using a more complex framework of Hawkes processes, one can show that the results still hold for any $\rho \in [-1,1]$; see El Euch and Rosenbaum (2018).

[5] The complex logarithm is defined on $\mathbb{C}/\mathbb{R}^-$ by $\log(z) = \log(|z|) + i\arg(z)$, with $\arg(z) \in (-\pi,\pi]$.

REFERENCES

Albrecher, H., Mayer, P., Schoutens, W., & Tistaert, J. (2007, January). The little Heston trap. *Wilmott Magazine*, pp. 83–92.

Bacry, E., Delattre, S., Hoffmann, M., & Muzy, J.-F. (2013). Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance*, *13*(1), 65–77.

Bacry, E., Delattre, S., Hoffmann, M., & Muzy, J.-F. (2013). Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Processes and Their Applications*, *123*(7), 2475–2499.

Bacry, E., Jaisson, T., & Muzy, J.-F. (2016). Estimation of slowly decreasing Hawkes kernels: Application to high frequency order book modelling. *Quantitative Finance*, *16*(8), 1179–1201.

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

Bennedsen, M., Lunde, A., & Pakkanen, M. S. (2017a). *Decoupling the short-and long-term behavior of stochastic volatility*. Preprint arXiv:1610.00332.

Bennedsen, M., Lunde, A., & Pakkanen, M. S. (2017b). Hybrid scheme for Brownian semistationary processes. *Finance and Stochastics*, *21*(4), 931–965.

Bouchaud, J.-P., & Potters, M. (2003). *Theory of financial risk and derivative pricing: From statistical physics to risk management*. Cambridge, UK: Cambridge University Press.

Carr, P., & Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, *2*(4), 61–73.

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics*, *10*(4), 407–432.

Diethelm, K., Ford, N. J., & Freed, A. D. (2002). A predictor-corrector approach for the numerical solution of fractional differential equations. *Nonlinear Dynamics*, *29*(1-4), 3–22.

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

Diethelm, K., & Freed, A. D. (1999). The fracpece subroutine for the numerical solution of differential equations of fractional order. In *Forschung und Wissenschaftliches Rechnen 1998* (pp. 57–71). Gottingen, Germany: Gesellschaft für Wisseschaftliche Datenverarbeitung Gottingen.

Dragulescu, A. A., & Yakovenko, V. M. (2002). Probability distribution of returns in the Heston model with stochastic volatility. *Quantitative Finance*, *2*(6), 443–453.

El Euch, O., Fukasawa, M., & Rosenbaum, M. (2016). *The microstructural foundations of leverage effect and rough volatility.* Working Paper, École Polytechnique.

El Euch, O., & Rosenbaum, M. (2018). Perfect hedging in rough Heston models. *Annals of Applied Probability*, *28*(6), 3813–3856. https://doi.org/10.1214/18-AAP1408.

Forde, M., Jacquier, A., & Lee, R. (2012). The small-time smile and term structure of implied volatility under the Heston model. *SIAM Journal on Financial Mathematics*, *3*(1), 690–708.

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

Gatheral, J. (2011). *The volatility surface: A practitioner's guide* (Vol. 357). Hoboken, NJ: John Wiley & Sons.

Gatheral, J., Jaisson, T., & Rosenbaum, M. (2018). Volatility is rough. *Quantitative Finance*, *18*(6), 933–949. https://doi.org/10.1080/14697688.2017.1393551.

Guennoun, H., Jacquier, A., Roome, P., & Shi, F. (2018). Asymptotic behavior of the fractional Heston model. *SIAM Journal on Financial Mathematics*, *9*(3), 1017–1045. https://doi.org/10.1137/17M1142892.

Hardiman, S. J., Bercot, N., & Bouchaud, J.-P. (2013). Critical reflexivity in financial markets: A Hawkes process analysis. *European Physical Journal B*, *86*(10), 1–9.

Haubold, H. J., Mathai, A. M., & Saxena, R. K. (2011). Mittag-Leffler functions and their applications. *Journal of Applied Mathematics*, Article 298628.

Hawkes, A. G., & Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, *11*(3), 493–503.

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

Itkin, A. (2005). *Pricing options with VG model using FFT*. Preprint arXiv physics/0503137.

Jacod, J., & Shiryaev, A. (2013). *Limit theorems for stochastic processes* (Vol. 288). Berlin, Germany: Springer Science & Business Media.

Jacquier, A., & Roome, P. (2013). The small-maturity Heston forward smile. *SIAM Journal on Financial Mathematics*, *4*(1), 831–856.

Jacquier, A., & Roome, P. (2016). Large-maturity regimes of the Heston forward smile. *Stochastic Processes and Their Applications*, *126*(4), 1087–1123.

Jaisson, T., & Rosenbaum, M. (2015). Limit theorems for nearly unstable Hawkes processes. *The Annals of Applied Probability*, *25*(2), 600–631 .

Jaisson, T., & Rosenbaum, M. (2016). Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *The Annals of Applied Probability*, *26*(5), 2860–2882.

Janek, A., Kluge, T., Weron, R., & Wystup, U. (2011). FX smile in the Heston model. In P. Cizek, W. K. Härdle, & R. Weron (Eds.), *Statistical tools for finance and insurance* (2nd ed., pp. 133–162). New York, NY: Springer.

Kahl, C., & Jäckel, P. (2005, September). Not-so-complex logarithms in the Heston model. *Wilmott Magazine*, pp. 94–103.

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

Li, C., & Tao, C. (2009). On the fractional Adams method. *Computers & Mathematics with Applications*, *58*(8), 1573–1588.

Livieri, G., Mouti, S., Pallavicini, A., & Rosenbaum, M. (2018). Rough volatility: Evidence from option prices. *IISE Transactions*, *50*(9), 767–776. https://doi.org/10.1080/24725854.2018.1444297.

Mainardi, F. (2014). On some properties of the Mittag-Leffler function $\mathrm{E}_\alpha(-\mathrm{t}^\alpha)$, completely monotone for $\mathrm{t} > 0$ with $0 < \alpha < 1$. *Discrete & Continuous Dynamical Systems - B*, *19*(7), 2267–2278. https://doi.org/10.3934/dcdsb.2014.19.2267.

Mandelbrot, B. B. (1997). The variation of certain speculative prices. In *Fractals and scaling in finance* (pp. 371–418). New York, NY: Springer.

Mathai, A. M., & Haubold, H. J. (2008). *Special functions for applied scientists*. New York, NY: Springer.

Pascucci, A., & Mazzon, A. (2017). The forward smile in local-stochastic volatility models. *Journal of Computational Finance*, *20*(3), 1–29. https://doi.org/10.21314/JCF.2016.319.

Poon, S.-H. (2009). *The Heston option pricing model*. Unpublished manuscript.

Revuz, D., & Yor, M. (1999). *Continuous martingales and Brownian motion* (Vol. 293). Berlin, Germany: Springer Science & Business Media.

Samko, S. G., Kilbas, A. A., & Marichev, O. I. (1993). *Fractional integrals and derivatives*: *Vol. 1993. Theory and applications*. Newark, NJ: Gordon and Breach.

Schmelzle, M. (2010). *Option pricing formulae using Fourier transform: Theory and application*. Preprint, http://pfadintegral.com.

How to cite this article: El Euch O, Rosenbaum M. The characteristic function of rough Heston models. *Mathematical Finance*. 2019;29:3–38. https://doi.org/10.1111/mafi.12173

APPENDIX

We gather in this section some useful technical results.

A.1 Mittag–Leffler functions

Let $(\alpha, \beta) \in (\mathbb{R}_+^*)^2$. The Mittag–Leffler function $E_{\alpha,\beta}$ is defined for $z \in \mathbb{C}$ by

$$E_{\alpha,\beta}(z) = \sum_{n\geq 0} \frac{z^n}{\Gamma(\alpha n + \beta)}.$$

For $(\alpha,\gamma) \in (0,1)\times\mathbb{R}_+$, we also define

$$f^{\alpha,\gamma}(t) = \gamma t^{\alpha-1}E_{\alpha,\alpha}(-\gamma t^{\alpha}), \quad t>0,$$

$$F^{\alpha,\gamma} = \int_0^t f^{\alpha,\gamma}(s)ds, \quad t \ge 0.$$

The function $f^{\alpha,\gamma}$ is a density function on $\mathbb{R}_+$ called Mittag–Leffler density function. The following properties of $f^{\alpha,\gamma}$ and $F^{\alpha,\gamma}$ can be found in Haubold, Mathai, and Saxena (2011), Mainardi (2014), and Mathai and Haubold (2008). We have

$$f^{\alpha,\gamma}(t) \underset{t\to 0^+}{\sim} \frac{\gamma}{\Gamma(\alpha)}t^{\alpha-1}, \qquad f^{\alpha,\gamma}(t) \underset{t\to\infty}{\sim} \frac{\alpha}{\gamma\Gamma(1-\alpha)}t^{-(\alpha+1)}$$

and

$$F^{\alpha,\gamma}(t) = 1 - E_{\alpha,1}(-\gamma t^{\alpha}), \qquad F^{\alpha,\gamma}(t) \underset{t\to 0^+}{\sim} \frac{\gamma}{\Gamma(\alpha+1)}t^{\alpha}, \qquad 1-F^{\alpha,\gamma}(t) \underset{t\to\infty}{\sim} \frac{1}{\gamma\Gamma(1-\alpha)}t^{-\alpha}.$$

Finally, for $\alpha \in (1/2,1)$, $f^{\alpha,\gamma}$ is square-integrable and its Laplace transform is given for $z \ge 0$ by

$$\hat{f}^{\alpha,\gamma}(z) = \int_0^{\infty} f_{\alpha,\gamma}(s)e^{-zs}ds = \frac{\gamma}{\gamma+z^{\alpha}}.$$

A.2 Wiener–Hopf equations

The following result is used extensively in this work to solve Wiener–Hopf type equations; see, for example, Bacry et al. (2013).

Lemma A.1. *Let g be a measurable locally bounded function from $\mathbb{R}$ to $\mathbb{R}^d$ and $\phi : \mathbb{R}_+ \to \mathcal{M}^{\mathbf{d}}(\mathbb{R})$ be a matrix-valued function with integrable components such that $\mathcal{S}(\int_0^{\infty}\phi(s)ds) < 1$. Then there exists a unique locally bounded function f from $\mathbb{R}$ to $\mathbb{R}^d$ solution of*

$$f(t) = g(t) + \int_0^t \phi(t-s).f(s)ds, \quad t \ge 0$$

given by

$$f(t) = g(t) + \int_0^t \psi(t-s).g(s)ds, \quad t \ge 0,$$

where $\psi = \sum_{k\ge 1}\phi^{*k}$.

A.3 Fractional differential equations

We end this appendix with some useful results about fractional differential equations. The next lemma can be found in Samko, Kilbas, and Marichev (1993).

Lemma A.2. *Let h be a continuous function from $[0,1]$ to $\mathbb{R}$, $\alpha \in (0,1]$ and $\gamma \in \mathbb{R}$. There is a unique continuous solution to the equation*

$$D^{\alpha}y(t) = \gamma y(t) + h(t), \quad I^{1-\alpha}y(0) = 0$$

given by

$$y(t) = \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}\left(\gamma(t-s)^{\alpha}\right) h(s)ds.$$

We also have the following useful result.

Lemma A.3. *Let h be a nonnegative continuous function from $[0,1]$ to $\mathbb{R}$ such that for any $t \in [0,1]$,*

$$h(t) \leq \varepsilon + C \int_0^t f^{\alpha,\gamma}(t-s)h(s)ds,$$

for some $\varepsilon \geq 0$ and $C \geq 0$. Then for any $t \in [0,1]$,

$$h(t) \leq C'\varepsilon,$$

with

$$C' = 1 + C\gamma \int_0^1 s^{\alpha-1} E_{\alpha,\alpha}\left(\gamma(C-1)s^{\alpha}\right) ds > 0.$$

In particular, if $\varepsilon = 0$ then $h = 0$.

Proof. Let

$$f(t) = h(t) - C \int_0^t f^{\alpha,\gamma}(t-s)h(s)ds,$$

and $g = h - f$. The function g is solution of

$$g(t) = C \int_0^t f^{\alpha,\gamma}(t-s)\left(g(s) + f(s)\right) ds.$$

Thus, from Lemma A.2, g is the unique solution of

$$D^{\alpha} g(t) = \gamma(C-1)g(t) + C\gamma f(t), \quad I^{1-\alpha} g(0) = 0.$$

Hence using again Lemma A.2, we deduce that

$$g(t) = C\gamma \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}\left(\gamma(C-1)(t-s)^{\alpha}\right) f(s)ds.$$

Therefore,

$$g(t) \leq C\gamma\varepsilon \int_0^t s^{\alpha-1} E_{\alpha,\alpha}\left(\gamma(C-1)s^{\alpha}\right) ds.$$

Using that $h = f + g$ together with the fact that $E_{\alpha,\alpha}$ is nonnegative, we get the result. □