

# Rough Heston

Proserpio Lorenzo

As we have seen the skew of the at the money options in the Heston model does not represent well the skew of real-world at the money options. This, among other reasons, made the researcher and practitioner to think about more sophisticated models. As observed by Gatheral, Jaisson and Rosenbaum the log-volatility in the market behaves like a fractional Brownian motion with small Hurst parameter, so the volatility is rough.

## 1 Stylized empirical facts

Heston model reproduces several important features of low frequency price data, provides quite reasonable dynamics for the volatility surface and it can be calibrated efficiently. If we want to surpass that we have to build a model which can reproduce the stylized facts of modern electronic markets in the context of high frequency trading. In practice each market behaves tick-by-tick, indeed we receive an update in price by market-makers whenever there is a trade and the movement is discrete and at least of one tick (usually 1 cent). There are 4 main stylized facts that we can observe in market data:

1. Markets are highly endogenous, as showed by Bouchad. This means that most of the orders have no real economic motivation, but are simply the reaction of algorithms to other orders.
2. Markets at high frequency are much more efficient than at lower frequencies, this means that it is much more difficult to find profitable statistical arbitrage strategies.
3. There is some asymmetry in the liquidity on the bid and the ask side of the order book. Indeed, a market-maker is likely to raise the price by less following a buy order than to lower the price following the same size sell order, as seen by Brunnermeier and Pedersen. This is mostly due to the fact that hedging the first position is easier than the second and that market-makers have usually some inventory.
4. A large proportion of transaction is due to big orders, called metaorders, which are not executed at once, but split in time. Indeed, one of the most challenging part of every trading strategies is to execute it in large volumes without moving changing too much the state of the market.

## 2 Building the model

As in El Euch, Fukasawa & Rosenbaum we will start building a model from Hawkes processes, then slowly including the stylized fact mentioned in the last paragraph and showing that the long-term dynamic of this model will lead to a rough Heston model at the macroscopic scale, in which the leverage effect is still represented.

### 2.1 Hawkes Processes

Hawkes processes are point processes which are said to be self-exciting, in the sense that the instantaneous jump-probability depends on the location of the past events. In particular we will focus on a bivariate Hawkes process,  $(N_t^+, N_t^-)_{t \geq 0}$ , where  $N_t^+$  is the number of upward jumps of one tick and  $N_t^-$  is the number of downward jumps of one tick, both in the interval  $[0, t]$ . The probability to get one-tick upward jump in a time  $dt$  is given by  $\lambda_t^+ dt$ , viceversa by  $\lambda_t^-$ . The array  $(\lambda_t^+, \lambda_t^-)$  is called intensity of the process and it is of the form:

$$\begin{pmatrix} \lambda_t^+ \\ \lambda_t^- \end{pmatrix} = \begin{pmatrix} \mu^+ \\ \mu^- \end{pmatrix} + \int_0^t \begin{pmatrix} \phi_1(t-s) & \phi_3(t-s) \\ \phi_2(t-s) & \phi_4(t-s) \end{pmatrix} \cdot \begin{pmatrix} dN_s^+ \\ dN_s^- \end{pmatrix}$$

where  $\mu^+$  and  $\mu^-$  are positive constants and the components of the matrix are positive and locally integrable functions. The process of the prices  $P_t$  is given by the difference between the number of upward jumps from time 0 and the number of downward jumps, so:

$$P_t = N_t^+ - N_t^-$$

Each component of the intensity can be decomposed into three parts, for example  $\lambda_t^-$  can be decomposed in:

- $\mu_t^-$  which corresponds to the probability that the price will go down because of exogenous reasons;
- $\int_0^t \phi_2(t-s) dN_s^+$  which corresponds to the probability of a downward jump induced by past upward jumps;
- $\int_0^t \phi_4(t-s) dN_s^-$  which corresponds to the probability of a downward jump induced by past downward jumps.

Now we will see that, when the  $\phi_j$  have suitable forms the model can reproduce the stylized effects described in the previous section. Moreover, we want to underlying that, due to how the model is built, the price process assumes discrete values, as in the real world.

### 2.2 Encoding the 2<sup>nd</sup> property

Since the markets at high frequency are expected to be more efficient then this translate in that, over any period of time, we should have on average the same

number of upwards jumps than downwards jumps. This can be translated in:

$$\int_0^t \mathbb{E}[\lambda_s^+] ds = \mathbb{E}[N_t^+] = \mathbb{E}[N_t^-] = \int_0^t \mathbb{E}[\lambda_s^-] ds \quad (1)$$

remembering how we have defined  $\lambda_t^+$  and  $\lambda_t^-$ :

$$\begin{aligned} \mathbb{E}[\lambda_t^+] &= \mu^+ + \int_0^t \phi_1(t-s) \mathbb{E}[\lambda_s^+] ds + \int_0^t \phi_3(t-s) \mathbb{E}[\lambda_s^-] ds \\ \mathbb{E}[\lambda_t^-] &= \mu^- + \int_0^t \phi_4(t-s) \mathbb{E}[\lambda_s^-] ds + \int_0^t \phi_2(t-s) \mathbb{E}[\lambda_s^+] ds \end{aligned}$$

the simplest way to satisfy the equation (1) is to put:

$$\mu^+ = \mu^- \text{ and } \phi_1 + \phi_3 = \phi_2 + \phi_4$$

### 2.3 Encoding the 3<sup>rd</sup> property

Market-makers act as liquidity providers, in practice at the beginning they are long inventory, so the ask side is more liquid than the bid side. This translate into the fact that the conditional probability of an upward jump right after an upward jump is smaller than the conditional probability to observe a downward jump after a downward jump. This means that for  $t \rightarrow 0$  we have:

$$\int_0^t \phi_4(t-s) dN_s^- > \int_0^t \phi_1(t-s) dN_s^+$$

or equivalently

$$\int_0^t \phi_2(t-s) dN_s^+ < \int_0^t \phi_3(t-s) dN_s^-$$

this can be satisfied in several ways, but we make the strong assumption that exists a constant  $\beta > 0$  such that  $\phi_3 = \beta\phi_2$ . Putting all together we have that the structure of our intensity process is:

$$\begin{pmatrix} \lambda_t^+ \\ \lambda_t^- \end{pmatrix} = \mu \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \int_0^t \begin{pmatrix} \phi_1(t-s) & \beta\phi_2(t-s) \\ \phi_2(t-s) & [\phi_1 + (\beta-1)\phi_2](t-s) \end{pmatrix} \cdot \begin{pmatrix} dN_s^+ \\ dN_s^- \end{pmatrix}$$

### 2.4 Encoding the 1<sup>st</sup> property

Markets have an high degree of endogeneity, which means that the proportion of "non-meaningful" orders with respect to the totality of the orders is close to 1. In order to have an intuition on how to include this effect in our model we need to recall dynamical systems. A dynamical system has a stationary point (or equilibrium) if its spectral radius is less than 1, in the same way we have a kernel transition matrix:

$$\int_0^T \begin{pmatrix} \phi_{1,T}(s) & \beta\phi_{2,T}(s) \\ \phi_{2,T}(s) & [\phi_{1,T} + (\beta-1)\phi_{2,T}](s) \end{pmatrix} ds = \int_0^T \Phi_T(s) ds$$

we can extend all the functions on  $[0, \infty)$  with the constant zero. We refer to them with a tilde. The spectral radius in our case is equal to:

$$\sigma\left(\int_0^\infty \tilde{\Phi}_T(s)ds\right) = \|\tilde{\phi}_{1,T}\|_1 + \beta\|\tilde{\phi}_{2,T}\|_1$$

Let  $(\Omega, \mathbb{F} = \{(F_T)_{T \geq 0}\}, \mathbb{P})$  a complete filtered probability space. We can find a sequence  $\{(\tilde{\phi}_{1,T}; \tilde{\phi}_{2,T})\}_T$  of couple of positive functions each one in the respective  $\mathcal{L}^1(F_T)$  such that:

- $\forall T > 0$  we have  $\|\tilde{\phi}_{1,T}\|_1 + \beta\|\tilde{\phi}_{2,T}\|_1 < 1$ ;
- if  $T_2 > T_1$  we have both  $\tilde{\phi}_{1,T_2} \geq \tilde{\phi}_{1,T_1}$  and  $\tilde{\phi}_{2,T_2} \geq \tilde{\phi}_{2,T_1}$ ;
- satisfying:

$$\lim_{T \rightarrow \infty} \left[ \|\tilde{\phi}_{1,T}\|_1 + \beta\|\tilde{\phi}_{2,T}\|_1 \right] = 1$$

then there exist a limit to this sequence and we will call that  $(\tilde{\phi}_1; \tilde{\phi}_2)$  and, due to continuity of the norm, we have that  $\|\tilde{\phi}_1\|_1 + \beta\|\tilde{\phi}_2\|_1 = 1$ . Then we have built our nearly-unstable system for each  $T > 0$  sufficiently large. Moreover notice that, due to the second property of our sequence, we have that the spectral radius, when  $T$  is increasing, is also increasing. We will refer to the matrix obtained with  $(\tilde{\phi}_1; \tilde{\phi}_2)$  as  $\Phi$ . The process of the prices up to time  $T < \infty$  is now indicated as:

$$P_t^T = N_t^{T,+} - N_t^{T,-}$$

with  $N_t^{T,+}$  and  $N_t^{T,-}$  with intensity generated by  $(\tilde{\phi}_{1,T}; \tilde{\phi}_{2,T})$ . From now on, we will denote

$$\sigma\left(\int_0^\infty \tilde{\Phi}_T(s)ds\right) = a_T = \|\tilde{\phi}_{1,T}\|_1 + \beta\|\tilde{\phi}_{2,T}\|_1$$

with  $a_T$  constants and  $a_T \uparrow 1$ . Moreover, we can also construct the  $\Phi_T$  as  $\Phi = a_T \Phi_T$ . We will refer to this as **Assumption 1**.

## 2.5 Encoding the 4<sup>th</sup> property

As showed by Jaisson and Rosenbaum, the effect of metaorders are reflected in the Hawkes framework by considering the condition that the kernel matrix exhibits heavy-tails. In order to encode the metaorders in the framework we need to put some additional assumptions. Let

$$\Psi_T = \sum_{k \geq 1} (\Phi_T)^{*k}$$

where  $(\Phi_T)^{*1} = \Phi_T$  and, for  $k > 1$ ,  $(\Phi_T)^{*k}(t) = \int_0^t \Phi_T(s)(\Phi_T)^{*(k-1)}(t-s)ds$ . The **Assumption 2** is that  $\Psi_T$  is uniformly bounded,  $\Phi$  is differentiable and the derivative of each component of  $\Phi$  is bounded and with finite norm 1. In

order to satisfy this assumption is sufficient, but not necessary, that  $\phi_1$  and  $\phi_2$  are both non-increasing functions in  $\mathcal{L}^1 \cap \mathcal{L}^\infty$  and differentiable. The last assumption, **Assumption 3**, is that there exist  $\alpha \in (1/2, 1)$  and  $C > 0$  such that

$$\alpha t^\alpha \int_t^\infty [\phi_1 + \beta \phi_2](s) ds \xrightarrow[t \rightarrow \infty]{} C$$

and moreover, for some  $\mu > 0$  and  $\lambda^* > 0$ ,

$$T^\alpha(1 - a_T) \xrightarrow[t \rightarrow \infty]{} \lambda^* \text{ and } T^{1-\alpha} \mu_T \xrightarrow[t \rightarrow \infty]{} \mu$$

under this three assumptions also the last stylized effect has been encoded in our model.

### 2.5.1 Wiener-Hopf equations

**Assumption 2** and the definition of the  $\Psi_T$  seems a little obscure. This assumption is there only because it allows us to use the following result on integral equations:

**Lemma.** *Let  $g$  be a measurable locally bounded function from  $\mathbb{R}$  to  $\mathbb{R}^2$  and  $\Phi : \mathbb{R}^+ \rightarrow \mathcal{M}^2(\mathbb{R})$  be a matrix-value function with integrable components such that the spectral radius of  $\int_0^\infty \phi(s) ds < 1$ . Then there exists a unique locally bounded function  $f$  from  $\mathbb{R}$  to  $\mathbb{R}^2$  solution of*

$$f(t) = g(t) + \int_0^t \langle \Phi(t-s), f(s) \rangle ds, \quad t \geq 0$$

given by

$$f(t) = g(t) + \int_0^t \langle \Psi(t-s), g(s) \rangle ds, \quad t \geq 0$$

where  $\Psi = \sum_{k \geq 1} \Phi^{*k}$ .

## 2.6 From microstructure to macrostructure

If **Assumptions 1,2 and 3** hold then it happens that the asymptotic behaviour of the microstructural model that we have built behaves like an Heston model, more precisely a rough version of it. Indeed, using the same notation as in **Assumption 3**, let:

$$\lambda = \frac{\alpha \lambda^*}{CT(1 - \alpha)}$$

then it holds true the following theorem:

**Theorem.** *As  $T \rightarrow \infty$  then the rescaled microscopic price*

$$\sqrt{\frac{1 - a_T}{\mu T^\alpha}} P_t^T$$

converges in the sense of finite dimensional laws to the following rough Heston model:

$$P_t = \frac{1}{1 - (\|\phi_1\|_1 - \|\phi_2\|_1)} \sqrt{\frac{2}{\beta + 1}} \int_0^t \sqrt{v_s} dW_s$$

where  $v_t$  is the solution to the following rough SDE:

$$v_t = \frac{1}{\Gamma(\alpha)} \left[ \int_0^t (t-s)^{\alpha-1} \lambda (1+\beta-v_s) ds + \int_0^t (t-s)^{\alpha-1} \lambda \sqrt{\frac{1+\beta^2}{\lambda^* \mu (1+\beta)^2}} \sqrt{v_s} d\tilde{W}_s \right]$$

where  $(W, \tilde{W})$  are two correlated Brownian motions with:

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

furthermore, the process  $v_t$  has Hölder regularity  $\alpha - 1/2 - \varepsilon$  for each  $\varepsilon > 0$ .

The proof is really technical and out of our scope. It can be found in El Euch, Fukasawa and Rosenbaum.

## 2.7 Mittag-Leffler functions

As we can see in the previous theorem we are implicitly assuming that  $v_0 = 0$ . Well, for a practitioner this is a severe limitation. However, it is not immediate to obtain the same result if  $v_0$  is not zero. In this subsection we will give some definitions which we will use to obtain a more general result. First of all, we define the Mittag-Leffler functions. Let  $\alpha, \beta \in \mathbb{R}^+$ , then the Mittag-Leffler function  $E_{\alpha, \beta}$  is defined for  $t \in \mathbb{C}$  and  $C \in \mathbb{C}$  as

$$E_{\alpha, \beta}(Ct^\alpha) = \sum_{j=0}^{\infty} \frac{C^j t^{j\alpha}}{\Gamma(\alpha j + \beta)}$$

if  $\alpha \in (0, 1)$  and  $\lambda \in \mathbb{R}^+$ , we can also define

$$\begin{aligned} f^{\alpha, \lambda}(t) &= \lambda t^{\alpha-1} E_{\alpha, \alpha}(-\lambda t^\alpha) \mathbf{1}_{t \geq 0} \\ F^{\alpha, \lambda}(t) &= \int_0^t f^{\alpha, \lambda}(s) ds \end{aligned}$$

the first one can be proven to be a density function on  $\mathbb{R}^+$  and it is called Mittag-Leffler density function. The Mittag-Leffler density function has many nice properties, among the others, we are interested in the followings:

$$\begin{aligned} f^{\alpha, \lambda}(t) &\sim \frac{\alpha}{\lambda \Gamma(1-\alpha)} t^{-\alpha-1} \quad \text{if } t \rightarrow \infty \\ F^{\alpha, \lambda}(t) &\sim \frac{\lambda}{\Gamma(1+\alpha)} t^\alpha \quad \text{if } t \rightarrow 0^+ \\ F^{\alpha, \lambda}(t) &\sim 1 + \frac{1}{\lambda \Gamma(1-\alpha)} t^{-\alpha} \quad \text{if } t \rightarrow \infty \\ D^\alpha[E_\alpha(Ct^\alpha) - 1] &= CE_\alpha(Ct^\alpha) \end{aligned}$$

when not specified  $\beta$  is assumed to be 1. Moreover we will use this lemma:

**Lemma.** *Let  $\alpha \in (0, 1]$ ,  $u \in \mathbb{C}$  with  $u = a + ib$  with  $a \in \mathbb{R}^+$  and  $b \in [-1/(1 - \rho^2), 0]$ . Define as  $C = \sqrt{u(u + i) - \rho^2 u^2}$ . Then for any positive integer  $p$  and  $t \in \mathbb{R}^+$  this expansion holds*

$$E_\alpha(-Ct^\alpha) = \sum_{j=1}^p \frac{(-1)^{j-1} t^{-j\alpha}}{C^j \Gamma(1 - j\alpha)} + \mathcal{O}(|Ct^\alpha|^{-1-p}) \quad \text{if } t \rightarrow \infty$$

## 2.8 Adjusting the initial volatility

Luckily, we will see that in order to adjust the initial volatility is sufficient to consider an appropriate inhomogeneous intensity for our bi-dimensional Hawkes process and an appropriate kernel matrix. Indeed, using the same notation as in the previous session, we can make a very particular choice for the  $\Phi_T$ . We will choose the following: suppose that exist  $\beta \geq 0$ ,  $\alpha \in (1/2, 1)$  and  $\lambda > 0$  such that

$$a_T = 1 - \lambda T^{-\alpha}, \quad \Phi_T(s) = a_T f^{\alpha,1}(s) \chi$$

where

$$\chi = \frac{1}{\beta + 1} \begin{pmatrix} 1 & \beta \\ 1 & \beta \end{pmatrix}$$

with this particular choice we see that **Assumption 1, 2** and **3** are satisfied. Moreover, all the first three properties are still well encoded into the model. We need also to notice that  $\chi$  is an idempotent matrix, this will be useful later when deriving the characteristic function. Now to identify the appropriate inhomogeneous intensity  $\hat{\mu}_T(\cdot)$  is far more complicated and not so useful, so we will give only the final candidate:

$$\hat{\mu}_T(t) = \mu T^{\alpha-1} + \varepsilon \mu T^{\alpha-1} \left[ \frac{1 - \int_0^t a_T f^{\alpha,1}(s) ds}{1 - a_T} - \int_0^t a_T f^{\alpha,1}(s) ds \right]$$

with  $\varepsilon, \mu$  positive constants. In the process to obtain this appropriate candidate, written in *The characteristic function of rough Heston models* of Euch and Rosenbaum, they obtained also an explicit form for  $\Psi_T$  as

$$\Psi_T(Tt) = \frac{a_T f^{\alpha,\lambda}(t)}{T(1 - a_T)}$$

which will be useful in the derivation of the characteristic function. We now need to define the microscopic process converging to the log-price of the rough Heston model, for a positive function  $\gamma(\cdot)$ , as

$$P_t^T = \sqrt{\frac{\gamma(t)}{2}} \sqrt{\frac{1 - a_T}{T^\alpha \mu}} (N_{tT}^{T,+} - N_{tT}^{T,-}) - \frac{\gamma(t)}{2} \frac{1 - a_T}{T^\alpha \mu} N_{tT}^{T,+}$$

and finally we have the following

**Theorem 1.** *As  $T \rightarrow \infty$ , under the assumptions made in this section, the sequence of processes  $(P_t^T)_{t \in (0,1)}$  converges in law for the Skorokhod topology to*

$$P_t = \int_0^t \sqrt{v_s} d\tilde{W}_s - \frac{1}{2} \int_0^t v_s ds$$

where  $v$  is the unique solution of the rough stochastic differential equation

$$v_t = \gamma(t)\varepsilon + \frac{1}{\Gamma(\alpha)} \left[ \int_0^t (t-s)^{\alpha-1} \lambda(\gamma(s) - v_s) ds + \lambda \sqrt{\frac{\gamma(s)(1+\beta^2)}{\lambda\mu(1+\beta)^2}} \int_0^t (t-s)^{\alpha-1} \sqrt{v_s} dW_s \right]$$

with  $(\tilde{W}, W)$  correlated Brownian motions with

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

### 3 The rough Heston model

We have seen that the rough Heston model is what arise from taking the limit of Hawkes processes. Here we add also the drift term to the equation and substitute  $\alpha$  with  $H + 1/2$ . Given a stock price process  $S = (S_t)_{t \geq 0}$  the rough Heston model, under risk-free probability measure  $\mathbb{Q}$ , is the following :

$$\begin{cases} dS_t = (r - q)S_t dt + S_t \sqrt{v_t} d\tilde{W}_t \\ v_t = v_0 + \frac{\lambda}{\Gamma(H + \frac{1}{2})} \int_0^t \frac{\gamma(s) - v_s}{(t-s)^{\frac{1}{2}-H}} ds + \frac{\theta}{\Gamma(H + \frac{1}{2})} \int_0^t \frac{\sqrt{v_s}}{(t-s)^{\frac{1}{2}-H}} dW_s \end{cases}$$

where:

- $r$  is the risk-free rate;
- $q$  is the yield of the underlying;
- $W$  and  $\tilde{W}$  are correlated Brownian motions with  $d\langle \tilde{W}, W \rangle_t = \rho dt$ ;
- $H \in (0, 1/2)$  is the Hurst exponent of the fractional Brownian motion;
- $\theta$  is the volatility of the volatility;
- $\lambda \geq 0$  is a constant representing the "speed" of the mean reversion;
- $\gamma(\cdot)$  is a positive  $F_0$ -measurable function representing the mean reversion level for the volatility.



In order to proceed with this chapter we need to define the fractional integral and the fractional derivative. We define the fractional integral of order  $\alpha \in (0, 1]$  of a function  $f$  as

$$I^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s) ds$$

whenever the integral exists. We define the fractional derivative of order  $\alpha \in (0, 1]$  of a function  $f$  as

$$D^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t (t-s)^{-\alpha} f(s) ds$$

whenever it exists.

### 3.1 Inference of $\lambda\gamma(\cdot)$ from the forward variance curve

A variance swap with maturity  $T$  is a contract which pays out the realized variance of a financial underlying, computed as the sum of the squares of daily log-returns, in exchange for a fixed strike called the variance swap variance  $V_0^T$  that is determined in such a way that the initial value of the contract is zero. In practice can be seen as a way to trade implied volatility with future realized volatility. We can define the volatility of a variance swap at time zero as:

$$\hat{\sigma}_0^T := \sqrt{\frac{V_0^T}{T}}$$

Using the definition the volatility of the swap at time  $t$  is

$$(\hat{\sigma}_t^T)^2 = \frac{1}{T-t} \mathbb{E} \left[ \int_t^T v_s ds \middle| F_t \right] = \frac{1}{T-t} \int_t^T \xi_t(s) ds$$

where  $\xi_t(s) := \mathbb{E}[v_s | F_t]$  (with  $s > t$ ) and it is called the forward variance curve. So equivalently we can derive  $\xi_0(\cdot)$  as

$$\xi_0(t) = (\hat{\sigma}_0^t)^2 + t \frac{d}{dt} [(\hat{\sigma}_0^t)^2]$$

In *Perfect hedging in rough Heston models* El Euch and Rosenbaum showed that there is a link between  $\lambda\gamma(\cdot)$  and the forward variance curve. Assume that the curve admits a fractional derivative of order  $\alpha$ , then  $\lambda\gamma(\cdot)$  can be chosen so that the model is consistent with the market observed forward variance curve by taking

$$\lambda\gamma(t) = D^\alpha [\xi_0(\cdot) - v_0](t) - \lambda\xi_0(t) \quad (2)$$

Equation (2) can be obtained by calculating the conditional expected value of  $v_t$  using the second equation of the rough Heston model, showing that the forward variance curve is locally integrable (so the expected value of the stochastic integral is zero), then fractionally differentiate the LHS and RHS and reorder.

Using this fact and assuming that  $\lambda$  is sufficiently small, we can rewrite the dynamic in a compact way as:

$$(\star) \quad \begin{cases} dS_t = (r - q)S_t dt + S_t \sqrt{v_t} \{ \rho dW_t + \sqrt{1 - \rho^2} dW_t^\perp \} \\ v_t = \xi_0(t) + \frac{\theta}{\Gamma(H + \frac{1}{2})} \int_0^t \frac{\sqrt{v_s}}{(t - s)^{\frac{1}{2} - H}} dW_s \end{cases}$$

the hypothesis that  $\lambda$  must be sufficiently small is sensible since the volatility is *slowly* mean reverting.

## 4 The Characteristic function

Let  $x_t$  be the log-spot price, we wish to obtain the characteristic function of terminal log-spot  $x_T$  conditional on the initial log-price  $x_0$  and the initial forward variance curve  $\xi_0(\cdot)$ . In mathematical terms:

$$\phi_{rH}(u, T; 0) = \mathbb{E}_{\mathbb{Q}}[e^{iu x_T} | x_0, \xi_0(\cdot)]$$

### 4.1 The Characteristic function of an Hawkes process

We need to do a step back and restart from the 2-dimensional Hawkes process presented in the second section of this chapter, called  $N$ . Using the same notation as before, let  $L(u, t)$  be the characteristic function of the 2-dimensional Hawkes process  $N$  conditional at time  $t$ :

$$L(u, t) = \mathbb{E}[e^{i\langle u, N_t \rangle}], \quad u \in \mathbb{R}^2$$

then

**Theorem 2.** *We have*

$$L(u, t) = \exp \left\{ \int_0^t \langle C(u, t - s) - (1, 1), \mu(s) \rangle ds \right\}$$

where  $C : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{C}^2$  is the solution of this integral equation:

$$C(u, t) = \exp \left\{ iu + \int_0^t \Phi_T^\top(s) [C(u, t - s) - (1, 1)] ds \right\}$$

In order to prove this statement we define two auxiliary independent 2 dimensional point processes  $(\tilde{N}_j)_{j=1,2}$ . We will refer to  $\tilde{N}_1^2$  for the second component of the first auxiliary process, in the same way to the others. Let  $(\tilde{N}_j^k)_{j=1,2}$  be two bivariate Hawkes processes with kernel matrix  $\Phi : \mathbb{R}^+ \rightarrow \mathcal{L}^2(\mathbb{R}^+) \cap \mathcal{L}^1(\mathbb{R}^+)$ . We will denote their characteristic function at time  $t$  as  $L_j(u, t)$ . Now, for each  $j$  and each  $t \geq 0$ , let  $N_t^{0,1}$  be the number of upwards jumps occurred up to time  $t$  and  $N_t^{0,2}$  be the number of downwards jumps occurred up to time  $t$ , each one is a Poisson process with rates, respectively,  $\mu_1(t)$  and  $\mu_2(t)$ . We also define as

$\tau_1^k < \dots < \tau_{N_t^{0,k}}^k \in [0, t]$  the arrival times of jumps of type  $k$  (where  $k = 1$  for up,  $k = 2$  for down) of  $N$  up to time  $t$ . The number of jumps of type  $k$  arrived at time  $\tau_u^k$  has the same law as  $(\tilde{N}_{j,t-\tau_u^k}^k)_{j=1,2}$  where  $\tilde{N}$  is taken independent from  $N$ . So we can write the following equality, in law:

$$N_t^k = N_t^{0,k} + \sum_{j=1}^2 \sum_{l=1}^{N_t^{0,j}} \tilde{N}_{j,t-\tau_l^j}^{k,l}$$

where  $(\tilde{N}_j^{k,l})_{j=1,2}$  are  $l$  independent copies of  $(\tilde{N}_j^k)_{j=1,2}$ , also independent of  $(N^{0,k})$ . Using these considerations we can obtain:

$$\mathbb{E}[e^{i\langle u, N_t \rangle} | N_t^0] = e^{i\langle u, N_t^0 \rangle} \prod_{j=1}^2 \prod_{l=1}^{N_t^{0,j}} L_j(u, t - \tau_l^j)$$

Now, fixing  $k$ , remember that, conditional on  $N^{0,k}, t$ , we have that the vector of the arrival times has the same law as the order statistics  $(X_{(1)}, \dots, X_{(N^{0,k}, t)})$  built from iid variables  $(X_1, \dots, X_{N_t^{0,k}})$  with density with support in  $[0, t]$ . We will refer to that density as  $\frac{\mu_k(s) \mathbf{1}_{s \leq t}}{\int_0^t \mu_k(s) ds}$ . So, if we use the fact that the  $X_i$  are iid we obtain:

$$\mathbb{E}[e^{i\langle u, N_t \rangle} | N_t^0] = e^{i\langle u, N_t^0 \rangle} \prod_{j=1}^2 \left[ \left( \int_0^t L_j(u, t-s) \frac{\mu_j(s)}{\int_0^t \mu_j(r) dr} ds \right)^{N_t^{0,j}} \right]$$

using again independence and rearranging we obtain:

$$L(u, t) = \exp \left\{ \sum_{j=1}^2 \int_0^t (e^{iu_j} L_j(u, t-s) - 1) \mu_j(s) ds \right\}$$

using the same trick and remembering that also  $(\tilde{N}_j^k)_{j=1,2}$  are bivariate Hawkes processes with kernel matrix  $\Phi$  we can write:

$$L_k(u, t) = \exp \left\{ \sum_{j=1}^2 \int_0^t (e^{iu_j} L_j(u, t-s) - 1) \Phi_{j,k}(s) ds \right\}$$

now it is enough to define

$$C(u, t) = \begin{pmatrix} e^{iu_1} L_1(u, t) \\ e^{iu_2} L_2(u, t) \end{pmatrix}$$

and do the substitution in the previous equations to obtain the conclusion of the proof.

## 4.2 Intuition about the result

Using what we obtained in **section 2.8**, **3.1** and **4.1** we wish to give an intuition for the following result:

**Theorem 3.** *Consider the rough Heston model as  $(\star)$  with  $\rho \in (-1/\sqrt{2}, 1/\sqrt{2}]$ . Then the characteristic function of the terminal log-spot  $x_T$  conditional on the initial state  $(x_0, \xi_0)$  is*

$$\phi_{rH}(u, T; 0) = \exp \left\{ iux_0 + iu(r - q)T + \int_0^T D^{H+1/2} h(u, T-s) \xi_0(s) ds \right\}$$

where  $h(u, \cdot)$  is the unique continuous solution of the fractional Riccati Cauchy problem

$$\begin{cases} D^{H+1/2} h(u, \cdot) = -\frac{u^2 + iu}{2} + iu\theta\rho h(u, \cdot) + \frac{\theta^2}{2} h^2(u, \cdot) \\ I^{1/2-H} h(u, 0) = 0 \end{cases}$$

**NOTE:** the assumption that  $\rho \in (-1/\sqrt{2}, 1/\sqrt{2}]$  is simply because we have changed parameters

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

with  $\beta > 0$  and we need to use **Theorem 1**.

As written in **section 2.8** and remembering the definition of  $a_T$  we have that

$$P_t^T = \sqrt{\frac{\lambda\gamma(t)}{2\mu}} T^{-\alpha} (N_{tT}^{T,+} - N_{tT}^{T,-}) - \frac{\lambda\gamma(t)}{2\mu} T^{-2\alpha} N_{tT}^{T,+}$$

and we know that, if  $T \rightarrow \infty$ , this sequence of processes converges in law to  $P$  where  $P_t = \log(S_t/S_0)$ . Now let  $N^T = (N^{T,+}, N^{T,-})$  be a sequence of two dimensional Hawkes processes (varying  $T$ ) satisfying all the assumptions in **section 2.8** and denote with  $L^T(u, t)$  the characteristic function of the process  $N^T$  at point  $u = (u^+, u^-) \in \mathbb{C}^2$  and time  $t$ . If we fix a scalar  $\bar{u} \in \mathbb{C}$  and let it be

$$\bar{u}_+^T = \bar{u} \sqrt{\frac{\lambda\gamma(tT)}{2\mu}} T^{-\alpha} - \bar{u} \frac{\lambda\gamma(tT)}{2\mu} T^{-2\alpha}, \quad \bar{u}_-^T = -\bar{u} \sqrt{\frac{\lambda\gamma(tT)}{2\mu}} T^{-\alpha}$$

then we have that, since convergence in law implies pointwise convergence of the characteristics we have that

$$L^T((\bar{u}_+^T, \bar{u}_-^T), tT) \rightarrow L(\bar{u}, t) \quad \text{if } T \rightarrow \infty \quad (3)$$

where  $L(\cdot, t)$  is the characteristic function of  $P$  at time  $t$ . Now, we have to notice the following fact, which will be useful later: taking the definition of  $\hat{\mu}(\cdot)$

given in **section 2.8** and the asymptotic properties of the Mittag-Leffler density function given in **section 2.7** we may write, for each  $t \in (0, 1]$ , that

$$\begin{aligned} T^{1-\alpha} \hat{\mu}(tT) &= T^{1-\alpha} \mu_T + \varepsilon T^{1-\alpha} \mu_T \left[ \frac{T^\alpha}{\lambda} \int_{tT}^\infty f^{\alpha,1}(s) ds + \lambda T^{-\alpha} \int_0^{tT} f^{\alpha,1}(s) ds \right] \\ &= \mu_T \left[ 1 + \frac{\varepsilon t^{-\alpha}}{\lambda} \cdot (tT)^\alpha \int_{tT}^\infty f^{\alpha,1}(s) ds \right] + \mu_T \varepsilon \lambda T^{-\alpha} \int_0^{tT} f^{\alpha,1}(s) ds \\ &\xrightarrow{T \rightarrow \infty} \mu + \frac{\mu \varepsilon t^{-\alpha}}{\lambda \Gamma(1-\alpha)} \end{aligned}$$

Thanks to **Theorem 2.** we can write the characteristic function of our Hawkes process as

$$\begin{aligned} L^T((\bar{u}_+^T, \bar{u}_-^T), tT) &= \exp \left\{ \int_0^{tT} \hat{\mu}_T(s) [C^{T,+}((\bar{u}_+^T, \bar{u}_-^T), tT - s) - 1] \right. \\ &\quad \left. + C^{T,-}((\bar{u}_+^T, \bar{u}_-^T), tT - s) - 1] ds \right\} \end{aligned}$$

where  $C^T((\bar{u}_+^T, \bar{u}_-^T), t) = (C^{T,+}((\bar{u}_+^T, \bar{u}_-^T), t), C^{T,-}((\bar{u}_+^T, \bar{u}_-^T), t))$  is the solution to its respective integral equation written in **Theorem 2.** Now we define

$$Y^T(\bar{u}, t) = (Y^{T,+}(\bar{u}, t), Y^{T,-}(\bar{u}, t)) = C^T((\bar{u}_+^T, \bar{u}_-^T), tT)$$

and we can rewrite the characteristic function as

$$\begin{aligned} L^T((\bar{u}_+^T, \bar{u}_-^T), tT) &= \exp \left\{ \int_0^t T^\alpha [(Y^{T,+}(\bar{u}, t-s) - 1) + \right. \\ &\quad \left. (Y^{T,-}(\bar{u}, t-s) - 1)] \cdot [T^{1-\alpha} \hat{\mu}(sT)] ds \right\} \end{aligned}$$

Since we have (3) we can expect that as  $T \rightarrow \infty$  then  $T^\alpha(Y^T(\bar{u}, t) - (1, 1))$  converges to some functions  $(c(\bar{u}, t), d(\bar{u}, t))$ , this can be shown as in the last section of *The characteristic function of rough Heston models* of El Euch and Rosenbaum. Using the fact that  $(Y^T(\bar{u}, t) - (1, 1)) = \mathcal{O}(T^{-\alpha})$  (in the sense component-wise) we can expand  $\log(Y^T(\bar{u}, t))$  around  $(1, 1)$  (where  $\log(\cdot)$  has been applied on each component) and obtain

$$\log(Y^T(\bar{u}, t)) = Y^T(\bar{u}, t) - (1, 1) - \frac{1}{2}(Y^T(\bar{u}, t) - (1, 1))^2 + o(T^{-2\alpha})(t)$$

and using the characteristic function above and solving for  $Y^T(\bar{u}, t) - (1, 1)$  we

obtain

$$\begin{aligned}
Y^T(\bar{u}, t) - (1, 1) &= i\bar{u}\sqrt{\frac{\lambda\gamma(t)}{2\mu}}(1, -1)T^{-\alpha} - i\bar{u}\frac{\lambda\gamma(t)}{2\mu}(1, 0)T^{-2\alpha} \\
&+ T \int_0^t \phi_T^\top(Ts)(Y^T(\bar{u}, t-s) - (1, 1))ds \\
&+ \frac{1}{2}(Y^T(\bar{u}, t) - (1, 1))^2 + o(T^{-2\alpha})(t)
\end{aligned}$$

where with  $(Y^T(\bar{u}, t) - (1, 1))^2$  we intend  $\langle (Y^T(\bar{u}, t) - (1, 1)), (Y^T(\bar{u}, t) - (1, 1)) \rangle$ . Now, we wish to use the lemma in **section 2.5.1** to solve the integral part and then using the explicit form for  $\Psi_T(T\cdot)$  written in **section 2.8**. So we need to notice

$$\begin{aligned}
\sum_{k \geq 1} (T\Phi_T(T\cdot))^{*k} &= \sum_{k \geq 1} T\Phi_T^{*k}(T\cdot) \\
&= T \sum_{k \geq 1} (a_T f^{\alpha, \lambda})^{*k}(T\cdot) \chi^k \\
&= T\Psi_T(T\cdot) \chi \\
&= \frac{a_T T^\alpha}{\lambda(\beta+1)} f^{\alpha, \lambda}(\cdot) \begin{pmatrix} 1 & \beta \\ 1 & \beta \end{pmatrix}
\end{aligned}$$

and then, applying the lemma, we obtain

$$\begin{aligned}
Y^T(\bar{u}, t) - (1, 1) &= i\bar{u}\sqrt{\frac{\lambda\gamma(t)}{2\mu}}(1, -1)T^{-\alpha} - i\bar{u}\frac{a_T\gamma(t)}{2\mu(1+\beta)}(1, \beta)T^{-\alpha}F^{\alpha, \lambda}(t) \\
&+ \frac{a_T T^\alpha}{2\lambda(\beta+1)} \int_0^t f^{\alpha, \lambda}(s) \begin{pmatrix} 1 & 1 \\ \beta & \beta \end{pmatrix} (Y^T(\bar{u}, t-s) - (1, 1))^2 ds \\
&+ o(T^{-\alpha})(t)
\end{aligned}$$

so we have that  $(c(\bar{u}, t), d(\bar{u}, t))$  must satisfy the integral equations

$$\begin{aligned}
c(\bar{u}, t) &= i\bar{u}\sqrt{\frac{\lambda\gamma(t)}{2\mu}} - i\bar{u}\frac{\gamma(t)}{2\mu(1+\beta)}F^{\alpha, \lambda}(t) \\
&+ \frac{1}{2\lambda(\beta+1)} \int_0^t f^{\alpha, \lambda}(s)(c^2(\bar{u}, t-s) + d^2(\bar{u}, t-s))ds \\
d(\bar{u}, t) &= -i\bar{u}\sqrt{\frac{\lambda\gamma(t)}{2\mu}} - i\bar{u}\frac{\beta\gamma(t)}{2\mu(1+\beta)}F^{\alpha, \lambda}(t) \\
&+ \frac{\beta}{2\lambda(\beta+1)} \int_0^t f^{\alpha, \lambda}(s)(c^2(\bar{u}, t-s) + d^2(\bar{u}, t-s))ds
\end{aligned}$$

and defining  $h(\bar{u}, t) := \mu[c(\bar{u}, t) + d(\bar{u}, t)]$  then we have that

$$\begin{aligned} L(\bar{u}, t) &= \exp \left\{ \lambda \gamma(t) I^1 h(\bar{u}, t) + \varepsilon \gamma(t) I^{1-\alpha} h(\bar{u}, t) \right\} \\ &= \exp \left\{ \lambda \gamma(t) I^1 h(\bar{u}, t) + v_0 I^{1-\alpha} h(\bar{u}, t) \right\} \end{aligned}$$

where  $h$  is the solution of the fractional Riccati Cauchy problem

$$\begin{cases} D^\alpha h(u, t) = -\frac{u^2 + iu}{2} + \lambda(iu\theta\rho - 1)h(u, t) + \frac{\lambda^2\theta^2}{2}h^2(u, t) \\ I^{1-\alpha}h(u, 0) = 0 \end{cases}$$

now if we reformulate in terms of the forward variance curve, add a drift term and suppose that the starting log-spot price is not 1 we obtain the result that we wanted to prove at the beginning of this section.

### 4.3 Rational approximation of the solution

We have a quasi-closed form for the characteristic function and we wish to obtain the solution to the fractional equation. Unfortunately, this solution is not known, so we will use the Padé approximants to obtain a fast and reliable approximation. In this section we will follow the work of Gatheral and Radoicic in *Rational approximation of the rough Heston solution*. In particular we will derive an expansion for small times for the characteristic function, an expansion for long times and then we will derive a Padé rational expansion to match the two formulae.

#### 4.3.1 Small times expansion

For the small times expansion we will follow the work of Alòs *et al.*. Only for this paragraph suppose that interest rates, borrow costs and yields are zero. Let  $H(x_t, w_t(T))$  be a solution of this equation

$$-\frac{\partial H}{\partial w}(x_t, w_t(T)) + \frac{1}{2} \frac{\partial^2 H}{\partial x^2}(x_t, w_t(T)) - \frac{1}{2} \frac{\partial H}{\partial x}(x_t, w_t(T)) = 0$$

where

$$w_t(T) = \mathbb{E} \left[ \int_t^T v_s ds \middle| F_t \right] = \int_0^T \xi_0(s) ds - \int_0^t v_s ds =: M_t - \int_0^t v_s ds$$

and  $x_t$  is the log price as before. Now we need these definition

**Definition 1.** Let  $A_t$  and  $B_t$  two stochastic processes. Then

$$(A \diamond B)_t(T) := \mathbb{E} \left[ \int_t^T d\langle A_t, B_t \rangle_s \middle| F_t \right]$$

provided that the expectation is finite.

**Definition 2.** Let  $H_t := H(x_t, w_t(T))$ , defined as before, then

$$(x \diamond M)_t(T) \cdot H_t := \mathbb{E} \left[ \int_t^T d\langle x_t, M_t \rangle_s \middle| F_t \right] \frac{\partial^2 H_t}{\partial x \partial w}$$

**Definition 3.** Let  $\mathbb{F}_0 = M$ . Then the forest of order  $k \in \mathbb{N}$  is defined recursively as:

$$\mathbb{F}_k = \frac{1}{2} \sum_{l=0}^{k-2} \sum_{j=0}^{k-2} \mathbf{1}_{l+j=k-2} \mathbb{F}_l \diamond \mathbb{F}_j + x \diamond \mathbb{F}_{k-1}$$

In the last definition we dropped the subscript and the point of evaluation for  $\diamond$ , from now on, unless specified, it is always  $T$ . Using this notation Alòs, Gatheral and Radoicic obtained this powerful result

**Theorem 4.** If  $H_t$  is a solution of the differential equation presented at the beginning of this section,  $\mathbb{E}[H_T | F_t]$  is finite and for each  $j \geq 0$  the integrals in each forest  $\mathbb{F}_j$  exist. Then

$$\mathbb{E}[H_T | F_t] = e^{\sum_{j=1}^{\infty} \mathbb{F}_j} \cdot H_t$$

where the exponential is to be understood as a formal power series and  $\cdot$  is the operator in **Definition 2**.

this theorem gives us an exact representation of the conditional expectation for every model that can be written in the forward variance form without assuming Markovianity. We can now rewrite the rough Heston model in the forward variance form as:

$$\begin{cases} dS_t = S_t \sqrt{v_t} \{ \rho dW_t + \sqrt{1 - \rho^2} dW_t^\perp \} =: S_t \sqrt{v_t} d\tilde{W}_t \\ d\xi_t(u) = \frac{\theta}{\Gamma(H + \frac{1}{2})} \frac{\sqrt{v_t}}{(u - t)^{\frac{1}{2} - H}} dW_t \end{cases}$$

and, excluding the term which are  $F_t$ -measurable and do not contribute in the tree computations

$$\begin{aligned} dx_t &= \sqrt{v_t} d\tilde{W}_t + F_t\text{-measurable terms} \\ dM_t &= \frac{\theta(T - t)^\alpha}{\Gamma(\alpha + 1)} \sqrt{v_t} dW_t \end{aligned}$$

and proceeding with the computations

$$\mathbb{F}_1 = x \diamond M = \frac{\rho\theta}{\Gamma(\alpha + 1)} \int_t^T \xi_t(s) (T - s)^\alpha ds$$

if we define for  $j \in \mathbb{N}$

$$I_t^{(j)}(T) := \int_t^T \xi_t(s) (T - s)^{j\alpha} ds$$



then we have

$$\begin{aligned} dI_t^{(j)}(T) &= \int_t^T (T-s)^{j\alpha} d\xi_t(s) ds \\ &= \frac{\theta\Gamma(1+j\alpha)}{\Gamma(1+j\alpha+\alpha)} \sqrt{v_t}(T-t)^{(j+1)\alpha} dW_t + \text{drift terms} \end{aligned}$$

and in the computation of  $\diamond$  the drift terms do not contribute. With this notation we have:

$$x \diamond M = \frac{\rho\theta}{\Gamma(\alpha+1)} I_t^{(1)}(T)$$

in  $\mathbb{F}_2$  we have two trees:

$$\begin{aligned} M \diamond M &= \frac{\theta^2}{\Gamma(\alpha+1)^2} I_t^{(2)}(T) \\ x \diamond (x \diamond M) &= \frac{\rho\theta}{\Gamma(\alpha+1)} \mathbb{E} \left[ \int_t^T d\langle x, I^{(1)} \rangle_s ds \middle| F_t \right] \\ &= \frac{\rho^2\theta^2}{\Gamma(2\alpha+1)} I_t^{(2)}(T) \end{aligned}$$

it can be proven by induction that each tree in the forest  $\mathbb{F}_j$  is equal to  $\theta^j I_t^{(j)}(T)$  multiplied by a constant. Now, let's consider this characteristic function

$$H_t(u) = \phi(u, T; t) := \exp \left\{ iux_t - \frac{u^2 + ui}{2} w_t(T) \right\}$$

this clearly satisfy the PDE at the beginning of this section and, moreover, it holds true, through differentiation, that

$$e^{\sum_{j=1}^{\infty} \mathbb{F}_j} \cdot \phi(u, T; t) = e^{\sum_{j=1}^{\infty} \tilde{\mathbb{F}}_j(u)} \phi(u, T; t)$$

where  $\tilde{\mathbb{F}}_j(u)$  is defined as  $\mathbb{F}_j$  but with each occurrence of  $\partial/\partial w$  replaced with  $-(u^2 + ui)/2$  and each occurrence of  $\partial/\partial x$  replaced with  $iu$ . Using **Theorem 4.** we have that

$$\phi_{rH}(u, T; t) = \mathbb{E}_{\mathbb{Q}} [e^{iux_T} | F_t] = e^{\sum_{j=1}^{\infty} \tilde{\mathbb{F}}_j(u)} \phi(u, T; t)$$

since we have used **Theorem 4.** we have to be sure that each tree in each forest  $\mathbb{F}_j$  exists, this is true if the forward variance curve is bounded on finite intervals, in that case each tree is of order  $(T-t)^{j\alpha+1}$ . From the calculations that we made before on  $\mathbb{F}_k$  we have that

$$\tilde{\mathbb{F}}_k(u) = \beta_k(u) \theta^k I_t^{(k)}(T)$$

with  $\beta_k(u)$  a coefficient dependent on  $k$  and  $u$ . Define also  $\tilde{X}(u) = iux_t$ . Firstly we compute for  $l < j$

$$\begin{aligned}\tilde{\mathbb{F}}_l(u) \diamond \tilde{\mathbb{F}}_j(u) &= \theta^{l+j+2} \beta_l(u) \beta_j(u) \frac{\Gamma(l\alpha+1)\Gamma(j\alpha+1)}{\Gamma(l\alpha+\alpha+1)\Gamma(j\alpha+\alpha+1)} I_t^{(l+j+2)}(T) \\ \tilde{X}(u) \diamond \tilde{\mathbb{F}}_k(u) &= iu\rho\theta^k \frac{\Gamma(k\alpha-\alpha+1)}{\Gamma(k\alpha+1)} \beta_{k-1}(u)\end{aligned}$$

Using the recursion formula in **Definition 3.** with the correct modifications (substituting  $\mathbb{F}_j$  with  $\tilde{\mathbb{F}}_j(u)$  and  $x$  with  $\tilde{X}(u)$ ) we obtain a recursion formula to express the coefficients  $\beta_j(u)$  as

$$\begin{aligned}\beta_0(u) &= -\frac{u^2 + iu}{2} \\ \beta_k(u) &= \frac{1}{2} \sum_{l=0}^{k-2} \sum_{j=0}^{k-2} \mathbf{1}_{l+j=k-2} \beta_l(u) \beta_j(u) \frac{\Gamma(l\alpha+1)\Gamma(j\alpha+1)}{\Gamma(l\alpha+\alpha+1)\Gamma(j\alpha+\alpha+1)} \\ &\quad + iu\rho \frac{\Gamma(k\alpha-\alpha+1)}{\Gamma(k\alpha+1)} \beta_{k-1}(u)\end{aligned}$$

and now defining  $h(u, t)$  as the formal power series

$$h(u, t) = \sum_{j=0}^{\infty} \frac{\Gamma(\alpha j + 1)}{\Gamma(\alpha j + \alpha + 1)} \beta_j(u) \theta^j t^{(j+1)\alpha}$$

we have that for small times is converging, satisfy the fractional Riccati equation, the boundary condition and, moreover, thanks to what we noticed before

$$\phi_{rH}(u, T; t) = e^{\sum_{j=1}^{\infty} \tilde{\mathbb{F}}_j(u)} \phi(u, T; t) = \exp \left\{ iux_t + \int_t^T D^\alpha h(u, T-s) \xi_t(s) ds \right\}$$

which was exactly what we obtained in the previous sections (minus the drift term). Notice that the function  $\theta h(u, t)$  depends only on the quantity  $\theta t^\alpha$ , so if we do the following change of variable  $\tilde{t}^\alpha = \theta t^\alpha$  we can rewrite our fractional differential equation as

$$\begin{aligned}D^\alpha h(u, \tilde{t}) &= -\frac{u^2 + ui}{2} + iu\rho h(u, \tilde{t}) + \frac{1}{2} h(u, \tilde{t})^2 \\ &= \frac{1}{2} [h(u, \tilde{t}) - r_-] [h(u, \tilde{t}) - r_+]\end{aligned}$$

with  $r_\pm = -iu\rho \pm \sqrt{u^2 + iu - \rho^2 u^2}$ .

### 4.3.2 Long times expansion

If we define  $C := (r_+ - r_-)/2$ . Then this proposition holds:

**Proposition 5.** Let  $h_\infty(u, \tilde{t}) := r_-[1 - E_\alpha(-C\tilde{t}^\alpha)]$ . For  $u \in \mathbb{C}$  with  $\Re(u) \geq 0$  and  $\tilde{t} \in \mathbb{R}^+$ , if  $\tilde{t} \rightarrow \infty$  then  $h_\infty(u, \tilde{t})$  solves the fractional equation in **Theorem 3.** up to an error term of  $\mathcal{O}(|C\tilde{t}^\alpha|^{-2})$ .

To prove this proposition is sufficient to apply the last property shown in **section 2.7** and use the lemma. Notice also that the definition of  $C$  in the lemma is coherent with the definition of  $C$  in this section. Looking at **Proposition 4.** it raises a natural *ansatz* for  $h(u, \tilde{t})$  when  $\tilde{t} \rightarrow \infty$  and it is

$$h(u, \tilde{t}) = r_- \sum_{j=0}^{\infty} \gamma_j \frac{\tilde{t}^{-j\alpha}}{C^j \Gamma(1-j\alpha)}$$

for some coefficients  $(\gamma_j)_{j=0}^{\infty}$ . Using now the last property of Mittag-Leffler functions we, after changing the index, obtain

$$\frac{1}{r_-} D^\alpha h(u, \tilde{t}) = C \sum_{j=1}^{\infty} \gamma_{j-1} \frac{\tilde{t}^{-j\alpha}}{C^j \Gamma(1-j\alpha)}$$

then, assuming that our *ansatz* is the solution to the fractional differential equation we have also the following

$$\begin{aligned} \frac{1}{r_-} D^\alpha h(u, \tilde{t}) &= \frac{1}{r_-} \frac{1}{2} (h(u, \tilde{t}) - r_-)(h(u, \tilde{t}) - r_+) \\ &= \sum_{j=1}^{\infty} \gamma_j \frac{\tilde{t}^{-j\alpha}}{C^j \Gamma(1-j\alpha)} \left( -C + \frac{r_-}{2} \sum_{j=1}^{\infty} \gamma_j \frac{\tilde{t}^{-j\alpha}}{C^j \Gamma(1-j\alpha)} \right) \end{aligned}$$

using the identity principle for power series we obtain

$$\begin{aligned} \gamma_0 &= 1 \\ \gamma_1 &= -1 \\ \gamma_2 &= 1 + \frac{r_-}{2C} \frac{\Gamma(1-2\alpha)}{\Gamma(1-\alpha)^2} \\ &\vdots \\ \gamma_j &= -\gamma_{j-1} + \frac{r_-}{2C} \sum_{i=1}^{\infty} \sum_{l=1}^{\infty} \mathbf{1}_{i+l=j} \gamma_i \gamma_l \frac{\Gamma(1-j\alpha)}{\Gamma(1-i\alpha)\Gamma(1-l\alpha)} \end{aligned}$$

### 4.3.3 Padé approximation

We define the rational approximation of  $h(u, \tilde{t})$  with  $m$  terms at the numerator and  $n$  terms at the denominator as

$$h^{(m,n)}(u, \tilde{t}) = \frac{\sum_{j=1}^m p_j \tilde{t}^{j\alpha}}{\sum_{j=1}^n q_j \tilde{t}^{j\alpha}}$$

such that

$$h^{(m,n)}(u, \tilde{t}) \sum_{j=1}^n q_j \tilde{t}^{j\alpha} - \sum_{j=1}^m p_j \tilde{t}^{j\alpha} = \mathcal{O}(\tilde{t}^{\alpha(m+n+1)})$$

Notice that for  $t \rightarrow \infty$ , thanks to the expansion for long times, we have that

$$h^{(m,n)}(u, \tilde{t}) \sim \frac{p_m \tilde{t}^{m\alpha}}{q_n \tilde{t}^{n\alpha}}$$

and we have seen that for long times it must be finite, so this can happen if and only if  $m = n$ . So the only admissible approximation of  $h$  are the ones of type  $h^{(m,m)}$ . As denoted in the paper we will choose  $m = 3$ , although there is no theoretical motivation for this choice, in practice happens to be both fast to compute and sufficiently accurate for our scope. Then, setting WLOG the constant at the denominator equal to 1:

$$h^{(3,3)}(u, \tilde{t}) = \frac{p_1 \tilde{t}^\alpha + p_2 \tilde{t}^{2\alpha} + p_3 \tilde{t}^{3\alpha}}{1 + q_1 \tilde{t}^\alpha + q_2 \tilde{t}^{2\alpha} + q_3 \tilde{t}^{3\alpha}}$$

thanks to the small time expansion for small  $x$  we have, for some coefficients  $b_1, b_2, b_3$ :

$$h(u, \tilde{t}) = b_1 \tilde{t}^\alpha + b_2 \tilde{t}^{2\alpha} + b_3 \tilde{t}^{3\alpha} + \mathcal{O}(\tilde{t}^{4\alpha})$$

meanwhile for long times and coefficients  $c_0, c_1, c_2$ :

$$h(u, \tilde{t}) = c_0 + \frac{c_1}{\tilde{t}^\alpha} + \frac{c_2}{\tilde{t}^{2\alpha}} + \mathcal{O}(\tilde{t}^{-3\alpha})$$

using the definition at the beginning of this section we obtain the following equations

$$\begin{aligned} p_1 &= b_1 \\ p_2 - p_1 q_1 &= b_2 \\ p_1 q_1^2 - p_1 q_2 - p_2 q_1 + p_3 &= b_3 \\ p_3 &= c_0 q_3 \\ p_2 q_3 - p_3 q_2 &= c_1 q_3^2 \\ p_1 q_3^2 - p_2 q_2 q_3 - p_3 q_1 q_3 + p_3 q_2^2 &= c_2 q_3^3 \end{aligned}$$

this linear system can be solved and the solution is

$$\begin{aligned}
p_1 &= b_1 \\
p_2 &= \frac{b_1^3 c_1 + b_1^2 c_0^2 + b_1 b_2 c_0 c_1 - b_1 b_3 c_0 c_2 + b_1 b_3 c_1^2 + b_2^2 c_0 c_2 - b_2^2 c_1^2 + b_2 c_0^3}{b_1^2 c_2 + 2b_1 c_0 c_1 + b_2 c_0 c_2 - b_2 c_1^2 + c_0^3} \\
p_3 &= c_0 q_3 \\
q_1 &= \frac{b_1^2 c_1 - b_1 b_2 c_2 + b_1 c_0^2 - b_2 c_0 c_1 - b_3 c_0 c_2 + b_3 c_1^2}{b_1^2 c_2 + 2b_1 c_0 c_1 + b_2 c_0 c_2 - b_2^2 c_1^2 + c_0^3} \\
q_2 &= \frac{b_1^2 c_0 - b_1 b_2 c_1 - b_1 b_3 c_2 + b_2^2 c_2 + b_2 c_0^2 - b_3 c_0 c_1}{b_1^2 c_2 + 2b_1 c_0 c_1 + b_2 c_0 c_2 - b_2^2 c_1^2 + c_0^3} \\
q_3 &= \frac{b_1^3 + 2b_1 b_2 c_0 + b_1 b_3 c_1 - b_2^2 c_1 + b_3 c_0^2}{b_1^2 c_2 + 2b_1 c_0 c_1 + b_2 c_0 c_2 - b_2^2 c_1^2 + c_0^3}
\end{aligned}$$

notice that  $I^{1-\alpha} h^{(3,3)}(u, 0) = 0$ .

## 5 Pricing

Using the facts that  $h$  is a solution of the fractional Riccati equation and that  $h^{(3,3)}$  approximates it well then we have:

$$\begin{aligned}
\phi_{rH}(u, T; 0) &= \exp \left\{ iux_0 + iu(r - q)T - \int_0^T \frac{u^2 + iu}{2} \xi_0(s) ds \right. \\
&\quad + \int_0^T iu\theta \rho h(u, T - s) \xi_0(s) ds \\
&\quad \left. + \int_0^T \frac{\theta^2}{2} h^2(u, T - s) \xi_0(s) ds \right\}
\end{aligned}$$

and then

$$\begin{aligned}
\hat{\phi}_{rH}(u, T; 0) &= \exp \left\{ iux_0 + iu(r - q)T - \int_0^T \frac{u^2 + iu}{2} \xi_0(s) ds \right. \\
&\quad + \int_0^T iu\theta \rho h^{(3,3)}(u, T - s) \xi_0(s) ds \\
&\quad \left. + \int_0^T \frac{\theta^2}{2} [h^{(3,3)}(u, T - s)]^2 \xi_0(s) ds \right\}
\end{aligned}$$

where  $\hat{\phi}_{rH}(u, T; 0) \approx \phi_{rH}(u, T; 0)$ . Now if we denote the volatility of a variance swap at time 0 with tenor  $T$  with  $\hat{\sigma}_0^T$  and remember the definition of the forward variance curve we have:

$$T(\hat{\sigma}_0^T)^2 = \int_0^T \xi_0(s) ds$$

where we have to notice also that  $\hat{\sigma}_0^T$  is the strike of a variance swap, which, for some tenors, can be observed in the market. The parametrization that we chose to use for the volatility of a variance swap is the Gompertz function, which is the following

$$\hat{\sigma}_0^T = z_1 e^{-z_2 e^{-z_3 T}}$$

where  $z_1, z_2$  and  $z_3$  are positive constants. The effects of the parameters are:  $z_1$  is the asymptote (i.e. the long time implied future volatility),  $z_2$  sets the displacement along the  $x$ -axis (i.e. time to maturity) and  $z_3$  sets the growth rate. We used the least squared method to fit the parameters and, on our dataset, we obtained:

$$z_1 = 0.2393444554 \quad z_2 = 0.2355916752 \quad z_3 = 0.1927188249$$

using these facts the formula becomes:

$$\begin{aligned} \hat{\phi}_{rH}(u, T; 0) = \exp \Big\{ & iux_0 + iuT \left[ r - q + (iu - 1) \frac{(\hat{\sigma}_0^T)^2}{2} \right] \\ & + iu\theta\rho \int_0^T h^{(3,3)}(u, T-s) \xi_0(s) ds \\ & + \frac{\theta^2}{2} \int_0^T [h^{(3,3)}(u, T-s)]^2 \xi_0(s) ds \Big\} \end{aligned}$$

Now that we have a good approximation of the characteristic formula we can apply the Lewis's formula to evaluate the price of an european call option as

$$C_{0,K} = S_0 e^{-qT} - \frac{\sqrt{S_0 K} e^{-(r+q)T/2}}{\pi} \int_0^\infty \frac{\Re\{\phi_{rH}(u - i/2, T; 0)\}}{u^2 + 1/4} du$$

where the integral is performed numerically using standard integration techniques. Due to the intrinsic structure of the scripting language that we chose to use (Python) we have implemented this formula. However, we wish also to present a viable alternative in the next section.

## 5.1 Fourier transform technique for Vanilla Options

An alternative method has been developed by Carr and Madan to compute vanilla options prices leveraging the Fast Fourier Transform algorithm. As in the Heston model, we have a closed (but approximated) formula for the characteristic function. If we denote with  $f_{rH}$  the risk-neutral density function (conditional to time  $T$ ) then

$$\phi_{rH}(u, T; 0) = \int_{\mathbb{R}} e^{iux} f_{rH}(x) dx$$

so, in other words, the characteristic function is the Fourier transform of the transition density function. Consequently, if we denote with  $k$  the log of the

strike, the initial value of a call with strike  $K$  and tenor  $T$  is

$$C_{0,K} = \int_k^\infty e^{-rT} (e^s - e^k) f_{rH}(s) ds$$

and if  $k \rightarrow -\infty$ , or equivalently  $K \rightarrow 0$ , the value of a call with such a strike is equal to the discounted value of the spot at that time, hence the function is not in  $\mathcal{L}^2(\mathbb{R})$ . For the next considerations this property is needed, so we will dampen the call price with an exponential kernel. The modified price is defined as

$$c_{0,K} = e^{\beta k} C_{0,K}$$

with  $\beta > 0$ . The Fourier transform of  $c_{0,K}$  is

$$\begin{aligned} \psi(u, T; 0) &= \int_{\mathbb{R}} e^{iuk} \int_k^\infty e^{\beta k} e^{-rT} (e^s - e^k) f_{rH}(s) ds dk \\ &= \frac{e^{-rT} \phi_{rH}(u - (\beta + 1)i, T; 0)}{\beta^2 + \beta - u^2 + iu(2\beta + 1)} \end{aligned}$$

so the price of the call can be obtained using the inverse transform as

$$\begin{aligned} C_{0,K} &= \frac{e^{-\beta k}}{2\pi} \int_{\mathbb{R}} e^{-iuk} \psi(u, T; 0) du \\ &= \frac{e^{-\beta k}}{\pi} \int_0^\infty e^{-iuk} \psi(u, T; 0) du \\ &= \frac{e^{-\beta k}}{\pi} \int_0^\infty e^{-iuk} \frac{e^{-rT} \phi_{rH}(u - (\beta + 1)i, T; 0)}{\beta^2 + \beta - u^2 + iu(2\beta + 1)} du \end{aligned} \tag{4}$$

where in the second step we used the fact that  $C_{0,K} \in \mathbb{R}$  so the imaginary part of the integral must be odd and even in its real part. For the modified call value to be integrable in the positive log strike direction, and hence for it to be square-integrable as well, a sufficient condition is provided by  $\psi(0, T; 0)$  being finite, which is equivalent to  $\phi_{rH}(-(\beta + 1)i, T; 0)$  being finite and this happens if and only if

$$\mathbb{E}[S_T^{\beta+1}] < \infty$$

if we find a suitable  $\beta$  then we have that  $\psi \in \mathbb{R}^+$ , indeed  $\phi_{rH}$  is bounded by  $\mathbb{E}[S_T^{\beta+1}]$ , which is independent from  $u$  and then

$$|\psi(u, T; 0)|^2 \leq \frac{\text{cost.}}{u^4}$$

this also give us an estimation for the truncation error in the right tail, indeed

$$\int_a^\infty |\psi(u, T; 0)| du < \frac{\sqrt{\text{cost.}}}{a}$$

so we can choose a suitable  $a$  such that the truncation error in computing the transform is below a certain tolerance.

## 5.2 Numerical implementation

Let consider the first step in (4), if we truncate the integral till  $a$  and choose  $N$  and  $\Delta$  such that  $a = N\Delta$  we can discretize the integral, using the Simpson's rule, on a grid  $[0, \Delta, \dots, N\Delta]$  and obtain

$$C_{0,K} \approx \frac{e^{-\beta k}}{\pi} \sum_{j=1}^N e^{-i\Delta(j-1)k} \psi(\Delta(j-1), T; 0) \frac{\Delta}{3} [3 + (-1)^j - \delta_{j-1}]$$

where  $\delta_j$  is the Kronecker delta. Suppose now that we are interested mainly in at-the-money call values, so with  $k$  near 0. The FFT returns  $N$  values of  $k$  and we use a regular spacing size  $\lambda$ , so our grid for the strikes is  $[-b := -N\lambda/2, \dots, N\lambda/2 - \lambda] = [k_1, \dots, k_N]$  so the value at time 0 of a call option with strike  $K_v$  with  $k_v$  in the grid becomes

$$C_{0,K_v} \approx \frac{e^{-\beta k_v}}{\pi} \sum_{j=1}^N e^{-i\lambda\Delta(j-1)(v-1)} e^{ib\Delta(j-1)} \psi(\Delta(j-1), T; 0) \frac{\Delta}{3} [3 + (-1)^j - \delta_{j-1}]$$

and in order to use the FFT algorithm we need

$$\lambda\Delta = \frac{2\pi}{N}$$

for the strikes that are not in the grid we will use spline interpolation.

**NOTE:** Carr and Madan developed also a modification to price far OTM options, here we will not show the process, but only the final formula

$$C_{0,K_v} \approx \frac{\Delta/3\pi}{\sinh(-\beta k_v)} \sum_{j=1}^N e^{-i\lambda\Delta(j-1)(v-1)} e^{ib\Delta(j-1)} \gamma(\Delta j - \Delta) [3 + (-1)^j - \delta_{j-1}]$$

where

$$\zeta(u) = e^{-rT} \left( \frac{1}{1+iu} - \frac{e^{rT}}{iv} - \frac{\phi_{rH}(u-i, T; 0)}{u^2 - iu} \right)$$

$$\gamma(u) = \frac{\zeta(u-i\beta) - \zeta(u+i\beta)}{2}$$

## 6 Calibration

The calibration is done in two steps. Firstly, we use a least square algorithm to fit the prices of the variance swaps to the ones observed in the market. In this way we fix  $z_1, z_2$  and  $z_3$ . After that, let  $\Xi := [H, \rho, \theta]^\top$  be our vector of parameters. We denote with  $\sigma^*(K_i, T_i)$  the market implied volatility for calls with strike  $K_i$  and maturity  $T_i$  and with  $\sigma(\Xi; K_i, T_i)$  the implied volatility for calls under the rough Heston model with parameters  $\Xi$ . Given  $n$  call options we define:

$$r_i(\Xi) := \sigma(\Xi; K_i, T_i) - \sigma^*(K_i, T_i) \quad i = 1, \dots, n$$



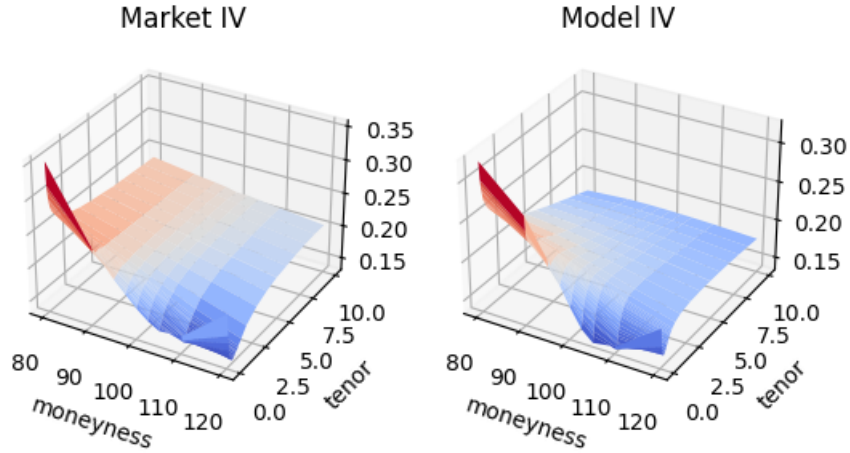
and the residual vector  $r(\Xi) = [r_1(\Xi), \dots, r_n(\Xi)]^\top$ . Then we have to remember that the parameters are subjected to certain constraints:  $H \in (0, 1/2)$ ,  $\theta > 0$  and  $\rho \in [-1, 1]$ . In **Theorem 3.** we used the additional hypothesis  $\rho \in (-1/\sqrt{2}, 1/\sqrt{2})$ , however, this is not strictly necessary and using the forests approach is possible to obtain the same result with  $\rho \in [-1, 1]$ . With this notation the calibration of the Heston model is an inverse problem in the nonlinear least square form as:

$$\min_{\substack{\Xi \\ H \in (0, 1/2) \\ \theta > 0 \\ \rho \in [-1, 1]}} \frac{1}{2} \|r(\Xi)\|^2$$

since we suppose to have  $n \gg 5$  (where 5 is the number of parameters that we have to determine) it is an overdetermined problem. To tackle this kind of problem we will use the Trust Region Reflective algorithm. For the interested reader see Branch, Coleman, Li in *A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems*.

## 6.1 Numerical results

Here we report the implied volatility surface calibrated to all the strikes and tenors together:



where the parameters are:

$$\theta = 0.2817 \quad H = 0.0010 \quad \rho = -0.6995$$

the mean relative percentage error obtained is 6.4480% and it took around 3mins on a standard laptop. As we can see we were able to obtain pretty

similar surface especially for shorter tenors, meanwhile for longer tenors seems that we are underestimating the implied volatility. It is interesting to notice that we have obtained a similar  $\rho$  for both the Heston and the rough Heston model.

## 7 Simulation

In this section we will explain the HQE scheme proposed by Gatheral in *Efficient simulation of affine forward variance models* to simulate price paths under the rough Heston model. Firstly, we will rewrite the rough Heston model in the forward variance form and under the log-spot dynamics as:

$$\begin{cases} dx_t = \left( r - q - \frac{v_t}{2} \right) dt + \sqrt{v_t} d\tilde{W}_t \\ d\xi_t(u) = \frac{\theta}{\Gamma(H + \frac{1}{2})} \frac{\sqrt{v_t}}{(u-t)^{\frac{1}{2}-H}} dW_t =: \kappa(u-t) \sqrt{v_t} dW_t \end{cases}$$

If we discretize the variance process with timestep  $\Delta = T/N$  (where  $N$  is the number of steps) we obtain:

$$v_{n\Delta} = \xi_0(n\Delta) + \sum_{k=1}^n \int_{(k-1)\Delta}^{k\Delta} \kappa(n\Delta - s) \sqrt{v_s} dW_s$$

we can now define the  $F_{n-1}$ -adapted part as

$$\hat{\xi}_n := \xi_0(n\Delta) + \sum_{k=1}^{n-1} \int_{(k-1)\Delta}^{k\Delta} \kappa(n\Delta - s) \sqrt{v_s} dW_s$$

and the martingale increment as

$$u_n = \int_{(n-1)\Delta}^{n\Delta} \kappa(n\Delta - s) \sqrt{v_s} dW_s$$

so

$$v_{n\Delta} = \hat{\xi}_n + u_n$$

moreover let's denote also for  $i, j \geq 0$

$$\mathcal{K}_i = \int_0^\Delta \kappa(s + i\Delta) ds \quad \text{and} \quad \mathcal{K}_{i,j} = \int_0^\Delta \kappa(s + i\Delta) \kappa(s + j\Delta) ds$$

and, doing the calculations, we obtain

$$\begin{aligned} \mathcal{K}_i &= \frac{\theta}{\alpha \Gamma(\alpha)} \Delta^\alpha [(i+1)^\alpha - i^\alpha] \\ \mathcal{K}_{i,i} &= \frac{\theta^2}{2H [\Gamma(\alpha)]^2} \Delta^{2H} [(i+1)^{2H} - i^{2H}] \end{aligned}$$

to conclude this introductory part we will state the following result, which can be obtained using the Itô's isometry:

**Lemma 6.** *Assuming that the forward variance curve is twice differentiable then*

$$\text{Var}[u_n|F_{n-1}] = \frac{\mathcal{K}_{0,0}}{2H+1} [\hat{\xi}_n + 2Hv_{(n-1)\Delta}] + \mathcal{O}(\Delta^{2+2H}) =: \bar{v}_n \mathcal{K}_{0,0} + \mathcal{O}(\Delta^{2+2H})$$

the last definition that we need is

$$\chi_n = \int_{(n-1)\Delta}^{n\Delta} \sqrt{v_s} dW_s$$

## 7.1 The HQE scheme

Using **Lemma 6.** we have that for  $s \in ((n-1)\Delta, n\Delta]$  it is true that  $v_s \approx \bar{v}_n$ , from there we get the corresponding approximations

$$\begin{aligned} \text{Var}[\hat{\xi}_{n+1}|F_{n-1}] &\approx \bar{v}_n \mathcal{K}_{1,1} \\ \text{Var}[\chi_n|F_{n-1}] &\approx \bar{v}_n \Delta \\ \text{Cov}[u_n, \hat{\xi}_{n+1}|F_{n-1}] &\approx \bar{v}_n \mathcal{K}_{0,1} \\ \text{Cov}[u_n, \chi_n|F_{n-1}] &\approx \bar{v}_n \mathcal{K}_0 \\ \text{Cov}[\chi_n, \hat{\xi}_{n+1}|F_{n-1}] &\approx \bar{v}_n \mathcal{K}_1 \end{aligned}$$

then the correlation matrix takes this form

$$\begin{pmatrix} 1 & \frac{\mathcal{K}_0}{\sqrt{\Delta}\sqrt{\mathcal{K}_{0,0}}} & \frac{\mathcal{K}_{0,1}}{\sqrt{\mathcal{K}_{1,1}}\sqrt{\mathcal{K}_{0,0}}} \\ \frac{\mathcal{K}_0}{\sqrt{\Delta}\sqrt{\mathcal{K}_{0,0}}} & 1 & \frac{\mathcal{K}_1}{\sqrt{\Delta}\sqrt{\mathcal{K}_{1,1}}} \\ \frac{\mathcal{K}_{0,1}}{\sqrt{\mathcal{K}_{1,1}}\sqrt{\mathcal{K}_{0,0}}} & \frac{\mathcal{K}_1}{\sqrt{\Delta}\sqrt{\mathcal{K}_{1,1}}} & 1 \end{pmatrix}$$

where all the entries are independent of the time step  $n$  and are all functions of only  $H$ . Moreover we will leverage the following approximations for a sufficiently small  $\Delta$

$$\mathcal{K}_{0,1} \approx \frac{\mathcal{K}_1 \mathcal{K}_0}{\Delta} \quad \text{and} \quad \mathcal{K}_{1,1} \approx \frac{\mathcal{K}_1^2}{\Delta}$$

with that our correlation matrix becomes

$$R = \begin{pmatrix} 1 & \frac{\sqrt{2H}}{H+1/2} & \frac{\sqrt{2H}}{H+1/2} \\ \frac{\sqrt{2H}}{H+1/2} & 1 & 1 \\ \frac{\sqrt{2H}}{H+1/2} & 1 & 1 \end{pmatrix}$$

### 7.1.1 The Andersen Quadratic Exponential scheme

Andersen developed the Quadratic Exponential scheme to integrate the Heston model which matches the means and the variances. Moreover this scheme also guarantees that  $v_{\Delta n}$  stays positive in every point. We need the following definitions:

$$\begin{aligned}\psi_n &= \frac{\text{Var}[u_n|F_{n-1}]}{\hat{\xi}_n^2} \\ \beta_n^2 &= \frac{2}{\psi_n} - 1 + \sqrt{\frac{2}{\psi_n}} \sqrt{\frac{2}{\psi_n} - 1} \\ Z_n &\sim \mathcal{N}(0, 1) \\ U_n &\sim \mathcal{U}(0, 1) \\ p_n &= \frac{2}{1 + \psi_n} \\ \gamma_n &= \frac{1}{2} \hat{\xi}_n (1 + \psi_n)\end{aligned}$$

Then the algorithm works in this way:

---

#### Algorithm 1 QE Scheme

---

- 1: Evaluate  $\psi_n$ .
- 2: **if**  $\psi_n \geq 3/2$  **then**
- 3:     Simulate  $v_n$  as

$$v_n = -\mathbf{1}_{U_n < p_n} \gamma_n \log \frac{U_n}{p_n}$$

- 4: **else**
- 5:     Simulate  $v_n$  as

$$v_n = \frac{\hat{\xi}_n}{1 + \beta_n^2} (\beta_n + Z_n)^2$$

- 6: **end if**
- 

here at each step we simulate only one random variable, Gatheral developed an extension of this for the rough Heston model.

### 7.1.2 Hybrid simulation step

Linear regression gives

$$u_n \approx \frac{\mathcal{K}_0}{\Delta} \chi_n + \varepsilon_n$$

where the three quantities on the RHS are uncorrelated. The next lemma guarantees the positivity of the variance process and that the covariance matrix presented in the previous session is correctly matched.

**Lemma 7.** Let  $\frac{\kappa_0}{\Delta}\tilde{\chi}_n = \frac{\kappa_0}{\Delta}\chi_n + \frac{\hat{\xi}_n}{2}$  and  $\tilde{\varepsilon}_n = \varepsilon + \frac{\hat{\xi}_n}{2}$  be generated using the QE scheme presented in the previous section with the following conditional mean and variances:

$$\mathbb{E}\left[\frac{\kappa_0}{\Delta}\tilde{\chi}_n|F_{n-1}\right] = \frac{\hat{\xi}_n}{2}; \quad \mathbb{E}[\tilde{\varepsilon}_n|F_{n-1}] = \frac{\hat{\xi}_n}{2}$$

$$\text{Var}[\tilde{\chi}_n|F_{n-1}] = \bar{v}_n\Delta; \quad \text{Var}[\tilde{\varepsilon}_n|F_{n-1}] = \bar{v}_n\left(\kappa_{0,0} - \frac{\kappa_0^2}{\Delta}\right)$$

Then  $v_{n\Delta} = \frac{\kappa_0}{\Delta}\chi_n + \varepsilon_n + \frac{\hat{\xi}_n}{2} \geq 0$ . Moreover, with  $u_n = \frac{\kappa_0}{\Delta}\chi_n + \varepsilon_n$

$$\text{Var}[u_n|F_{n-1}] = \bar{v}_n\kappa_{0,0} \quad \text{and} \quad \text{Cov}[u_n, \chi_n|F_{n-1}] = \bar{v}_n\kappa_0$$

Now we have everything we need to present the HQE scheme, but before doing that we need to recognize that there is nothing which guarantees that  $\hat{\xi}_n$  is always positive. Indeed, we have to impose this additional condition by choosing an  $\epsilon > 0$  which will be the lower bound of the variance process.

---

**Algorithm 2** HQE Scheme

---

1: Given  $\chi_k$ , for  $k < n$ , with  $\epsilon$  very small, compute

$$\hat{\xi}_n = \max\left[\epsilon, \xi_0(n\Delta) + \sum_{k=1}^{n-1} \sqrt{\frac{\kappa_{n-k+2, n-k+2}}{\Delta}} \chi_k\right]$$

2: Simulate  $\tilde{\chi}_n$  and  $\tilde{\varepsilon}_n$  accordingly to **Lemma 7**.

3: Compute  $v_{n\Delta} = \frac{\kappa_0}{\Delta}\tilde{\chi}_n + \tilde{\varepsilon}_n$ .

4: Update the log-price, with  $\tilde{v}_n = \frac{v_{n\Delta} + v_{(n+1)\Delta}}{2}$  and  $Z_n^\perp \sim \mathcal{N}(0, 1)$  as

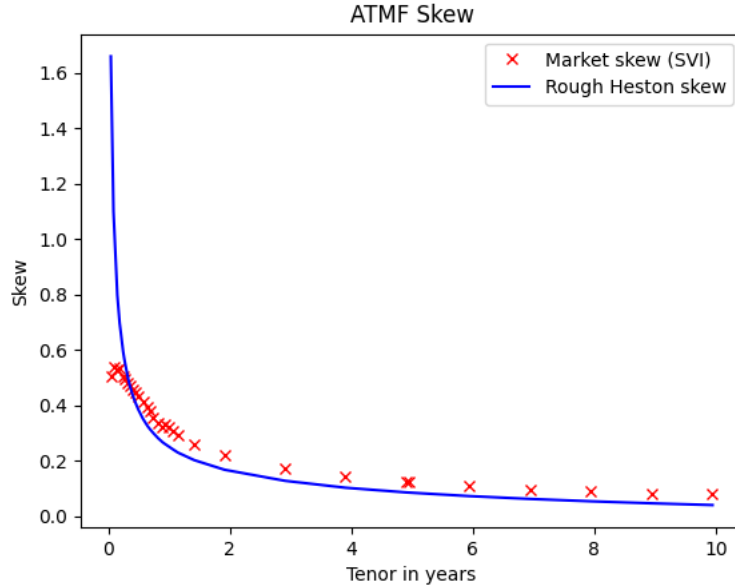
$$X_{n\Delta} = X_{(n-1)\Delta} + \left(r - q - \frac{\tilde{v}_n}{2}\right)\Delta + \sqrt{1 - \rho^2}\sqrt{\Delta\tilde{v}_n}Z_n^\perp + \rho\chi_n$$


---

## 8 Is Heston's ATMF skew problem solved?

In the previous chapter we have seen that the Heston model does not fit well the at-the-money forward skew, mainly because the ATMF decays as  $1/T$  for  $T \rightarrow \infty$ . If we repeat the same process as in that paragraph with the calibrated rough Heston we will obtain:

a much better fit! This is due to the fact that we are using a power-law kernel for the fractional Brownian motion, and for  $T \rightarrow \infty$  behaves like  $T^{-\alpha} \approx T^{-0.5}$ .



## References

1. Gatheral, Jaisson, Rosenbaum. Volatility is rough. ArXiv, 2014.
2. Bouchaud, The endogenous dynamics of markets: price impact and feedback loops. ArXiv, 2010.
3. Brunnermeier, Pedersen. Market Liquidity and Funding Liquidity. The Review of Financial Studies, 2009.
4. El Euch, Fukasawa, Rosenbaum. The microstructural foundations of leverage effect and rough volatility. ArXiv, 2016.
5. Jaisson, Rosenbaum. Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. ArXiv, 2015.
6. El Euch, Gatheral, Rosenbaum. Roughening Heston. Risk, 2019.
7. El Euch, Gatheral, Rosenbaum. Perfect hedging in rough Heston models. ArXiv, 2017.
8. El Euch, Rosenbaum. The characteristic function of rough Heston models. ArXiv, 2016.
9. Gatheral, Radoicic. Rational approximation of the rough Heston solution. International Journal of Theoretical and Applied Finance, 2019.

10. Alòs, Gathera, Radoicic. Exponentiation of conditional expectations under stochastic volatility. *Quantitative Finance*, 2020.
11. Branch, Coleman, Li. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM Journal of scientific computing*, 1999.
12. Gatheral. Efficient simulation of affine forward variance models. *Risk*, 2022.
13. Leif Andersen. Simple and efficient simulation of the Heston stochastic volatility model. *The Journal of Computational Finance*, 2008.
14. Bennedsen, Lunde and Pakkanen. Hybrid scheme for Brownian semistationary processes. *ArXiv*, 2017.
15. Lewis. A simple option formula for general jump-diffusion and other exponential Lévy processes. *OptionCity.net*, 2001.
16. Carr, Madan. Option valuation using the fast Fourier transform. *Journal of computational finance*, 2001.