UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI MATEMATICA
"TULLIO LEVI-CIVITA"
Corso di Laurea Magistrale in Mathematics

Two simulation schemes for the rough Heston model: a comparison

Author: **Bertolo Marco**
Student ID: **2087676**
Advisor: **Prof.ssa Callegaro Giorgia**
Co-advisors: **Dott. Pallavicini Andrea, Dott. Daluiso Roberto, Dott. Longoni Riccardo**

---

20 September 2024
Academic year: 2023/2024

Abstract

Since the publication of the pioneering article *Volatility is Rough* [31], it is commonly accepted the fact that the logarithm of the market volatility behaves like a fractional Brownian motion with Hurst index $H < \frac{1}{2}$. Because of this, rough volatility models are very popular nowadays. In particular, we want to focus our attention on the so called rough Heston model: for this type of model, the volatility is described by a one dimensional, affine Volterra process. Even if this choice is coherent with some market features and the characteristic function of the log-price is known in semi-closed form, the drawback is that the variance process is not a Markov process and, as a consequence, some difficulties arises in the simulation. Indeed, for the rough Heston model, the only one simulation method known up to now is the Hybrid Quadratic Exponential (HQE) scheme presented in [28]. On the other hand, one other possibility, is to approximate the rough Heston model with the lifted Heston model which is based on the substitution of the fractional kernel, which appears in the original model, with the sum of an appropriate number of exponential kernels: this choice preserves the Markovian structure of the variance process and simplifies the simulation. Recently, one new simulation scheme which deals with the lifted Heston model was presented in [8]. Thus, the goal of this thesis is to compare this new scheme with the HQE: in particular, we will perform different numerical tests in order to check the level of accuracy of these schemes in reproducing some market features.

Contents

Introduction **1**

1 Stochastic volatility models **5**
1.1 Heston model 6
1.2 Fractional Brownian motion 8
1.3 Volterra processes 11
1.4 Rough Heston model 14
1.4.1 The rough Heston model's characteristic function . . . 14

2 Lifted Heston model **17**
2.1 General results 17
2.2 Convergence to the rough Heston model 19

3 Simulation schemes **23**
3.1 Weak simulation scheme 24
3.1.1 Simulation of the volatility 24
3.1.2 Simulation of the stock price 26
3.1.3 Simulation of the entire process (S, V) 29
3.2 Hybrid Quadratic Exponential scheme 29

4 Numerical tests **37**
4.1 Computational costs 37
4.2 European Call option's smile 38
4.3 Implied volatility surface 40
4.4 The implied volatility Skew 45
4.4.1 European Call 46
4.4.2 Forward Start European Call 49
4.5 Estimate of the Hurst index 55

5 Skew Stickiness Ratio **61**
5.1 Numerical results 67

6 Conclusions and future developments **73**

A General case **75**

B An alternative derivation **77**

C The Quadratic Exponential scheme **81**

Introduction

Rough volatility models are very popular since the publication of the famous article *Volatility is rough* [31], where the authors show that the historical volatility paths are more irregular than those described by standard Brownian motion. This empirical evidence leads to the decision of modeling the variance process upon the fractional extension of standard Brownian motion. This choice produces a dynamics which is not only able to reproduce some statistical property of the volatility, but it also overcome some problems of stochastic volatility models in capturing some market features, for instance the short time explosion of the at the money skew. Nevertheless, rough volatility models are usually less tractable due to the fact that the variance is no more described nor by a Markov process neither by a semimartingale. Among all the models that belong to this class we focus our attention on the rough Heston model. It is presented by O. El Euch and M. Rosenbaum in [25]: they prove that for this model, which can be thought as the rough counterpart of the classical Heston model described in [30], the characteristic function is known in semi-closed form solving numerically a fractional Riccati equation. This allows for pricing options using the Fourier inversion technique, providing the possibility to calibrate the model to European options. Nevertheless, due to the non Markovianity of the variance process, there are problems in finding efficient simulation schemes which are necessary for pricing exotic options: so far, the only one which is available is the Hybrid Quadratic Exponential (HQE) scheme of Gatheral [28].
To overcome this problem, the lifted Heston model was introduced by Abi Jaber [1]. This scheme approximates the variance process of the rough Heston model with a (finite) weighted sum of Markovian process driven by a standard Brownian motion. In addition to the fact that the characteristic function can be computed solving a finite dimensional system of Riccati equations, classical simulation schemes (such as the implicit Euler scheme) are available due to its Markovian structure. The drawback is that using the parameters proposed by Abi Jaber, it is required a high number of approximating factors and this reduces the computational speed. Since recently it has been proposed a new simulation scheme by C. Bayer and S. Breneis [8] which allows to reduce the number of factors necessary for the approximation.

Thus, we want to compare the scheme of Gatheral with that of Bayer and Breneis in order to understand how good they are in reproducing realistic results on the implied volatility surface (for instance the smile and the skew). Moreover, we analyze the regularity of the volatility paths that they generate in order to check the coherence with the rough behaviour of the historical volatility. To conclude, we also analyze their capability in reproducing the asymptotic behaviour of the skew stickiness ratio.
The exposition is organized as follow:

- Chapter 1 starts with an introduction to the classical Heston model and explains its limit in reproducing the short time explosion of the at the money skew. Then, we introduce some results on the fractional Brownian motion and Volterra processes: for the first one, we show that it is neither a Markovian process nor a semimartingale, while for the second we prove a theorem on the existence of continuous weak solution which is useful to conclude that the variance process in the rough Heston model is well defined. Finally, we introduce the rough Heston model.

- Chapter 2 deals with the lifted Heston model. In particular, we dedicate an entire section in showing that, choosing in an appropriate way the parameters, this model represent a good approximation of the rough Heston model. For this purpose, we prove some results on the convergence of the kernel which guarantee the (theoretical) quality of the approximation.

- Chapter 3 is devoted to the description of the simulation schemes that we use in this thesis. Firstly, we describe the scheme of Bayer and Breneis showing in detail how it is derived, then we illustrate also the derivation of the Hybrid Quadratic Exponential scheme of Gatheral.

- Chapter 4 is the core of this thesis and it presents the numerical experiments that aim to compare the schemes described in Chapter 3. First of all, we compare the computational costs of the two schemes showing that the HQE is generally more expensive due to the non Markovian structure of the rough Heston model. At a later time, we perform some tests on the accuracy in reproducing some properties of the implied volatility: we focus our attention on the smile of European Call options and on the skew of European and Forward Start Call options. In the last section, we analyze the roughness of the volatility paths estimating their Hurst index in order to check if they present the rough behaviour described in [31].

- Chapter 5 is dedicated to the skew stickiness ratio. At first, we describe what is its role in linking the variation of the spot price and the variation of the at the money implied volatility. Then, following the idea of Bergomi in [13], we prove an asymptotic result valid for small maturity. We also verify the feasibility of this result for maturity of practical interest.

- Chapter 6 contains the conclusion and a brief discussion on possible future developments of this work.

Chapter 1

Stochastic volatility models

A stochastic volatility model (SVM) is a pricing model in which we assume that the underlying asset has a volatility driven by a stochastic process. In particular these models have the general form

$$\begin{cases} dS_t = rS_t dt + \sqrt{v_t} S_t dW_t, \quad S_0 \geq 0 \\ dv_t = \mu(t, v_t)dt + \sigma(t, v_t)dB_t, \quad v_0 \geq 0 \end{cases} \tag{1.1}$$

where the parameters r, S_0, v_0 represent respectively the risk-free rate, the initial value of the spot and the initial value of the variance.
This kind of models are very useful especially in equity markets since they are able to replicate the so called *leverage effect*, that is the inverse relation observed in the market between the price of a stock and its volatility. [1] It is possible to replicate this phenomenon thanks to the fact that the two Brownian motion are correlated with a correlation factor $\rho dt = d\langle W, B\rangle_t,\ \rho \in [-1, 1]$. The famous Black and Scholes model [17] is just a particular case of this kind of models in which $v_t = v_0$ for all $t \geq 0$. On the other hand a more realistic model is the so called Heston model that we will describe in the next section 1.1.

A fractional volatility model (FVM) is a refined version of a stochastic volatility model in which the volatility is driven by a fBm. The authors of [31] proposed to describe the volatility process with a fBm with Hurst index $H < 1/2$ coherently with some statistical properties that can be observed from historical data. This idea brings to the introduction of the category of rough volatility models. As pointed out by Bayer in [11] we can distinguish two kind of rough models: simple rough models in which the volatility is a function of a fBm and non simple rough models in which the volatility

[1] The idea behind the leverage effect is that financial leverage, such as debt, can amplify the impact of market movements on a company's equity. When the stock price falls, the debt burden becomes relatively larger, leading to increased of financial risk and, consequently, higher volatility in the stock price. (For more details see [16], [19]).

is described by a stochastic Volterra equation with a fractional kernel (For more details on this kind of SDE see [3]).
To this last class of models belongs the rough Heston model which is a more sophisticated development of the classical one.

1.1 Heston model

The Heston model [30] is a one dimensional stochastic volatility model where the underlying asset S and the volatility process v follow the dynamics:

$$\begin{cases} dS_t = S_t\sqrt{v_t}dW_t \\ dv_t = \lambda(\theta - v_t)dt + \sigma\sqrt{v_t}dB_t \end{cases} \tag{1.2}$$

Here the initial value of the underlying S_0 can be observed from the market while the parameters $\lambda, \theta, \sigma, v_0$ are positive and represent respectively the mean reversion rate, the long run variance, the volatility of the variance and the initial value of the variance. Moreover we assume that the risk-free rate $r = 0$ and that the Brownian motions have a correlation factor $\rho dt = d\langle W, B\rangle_t$.
Due to the fact that the volatility follows a CIR process, the fact that $\lambda, \theta, \sigma, v_0$ are positive is not sufficient to guarantee that the process remains positive but it is necessary the so called *Feller condition*: [2]

Proposition 1.1.1. *Consider the CIR process $(v_t)_{t\geq 0}$ in Equation (1.2). If $2\lambda\theta \geq \sigma^2$ then $v_t > 0$ for all $t \in [0, +\infty)$ almost surely. On the contrary if the previous condition is violated the process can reach zero with positive probability.*

The Heston model is very popular due to the fact that it generates reasonable shapes and dynamics for the implied volatility surface and because there is an explicit formula for the characteristic function of the asset log-price. For this reason it is possible to calibrate [3] the model to European options in a very efficient way and then to use the calibrated model in order to price exotic options such as American or Asian-type performing for example a Monte Carlo simulation.

Theorem 1.1.1. *In the Heston model in Equation (1.2), once defined $X_t = \log(S_t)$ and considering the natural filtration $\mathcal{F}_t = \sigma(W_s, B_s | 0 \leq s \leq t \leq T)$ it holds true that*

$$F(t, z, X_t, v_t) = \mathbb{E}(e^{izX_T}|\mathcal{F}_t) = e^{A(t)+B(t)v_t+C(t)X_t}$$

[2] For a rigorous derivation see sections 4,5 and 6 of [26].

[3] The calibration of a pricing model consists in choosing the parameters of the model in such a way that they minimize the distance between the market prices and the corresponding prices given by the model. Typically this procedure is performed using the prices of European options.

where the functions $A(\cdot), B(\cdot), C(\cdot)$ are completely determined solving a Riccati differential equation.

Proof. First of all we apply Itô's Lemma to $X_t = \log(S_t)$, then we get

$$dX_t = -\frac{v_t}{2}dt + \sqrt{v_t}dW_t$$

If we apply the Feynmann-Kac formula for the function $F(t, z, X_t, v_t) = \mathbb{E}(e^{izX_T}|\mathcal{F}_t)$ we find the PDE:

$$\begin{cases} \dfrac{\partial F}{\partial t} - \dfrac{v}{2}\dfrac{\partial F}{\partial x} + \lambda(\theta - v)\dfrac{\partial F}{\partial v} + \dfrac{v}{2}\dfrac{\partial^2 F}{\partial x^2} + \dfrac{\sigma^2 v}{2}\dfrac{\partial^2 F}{\partial v^2} + \sigma\rho v\dfrac{\partial^2 F}{\partial x \partial v} = 0 \\ F(x, v, T) = e^{izx} \end{cases}$$

Supposing that $F(t, z, X_t, v_t) = e^{A(t)+B(t)v_t+C(t)X_t}$ and imposing that it is a solution we have:

$$\begin{cases} F\big[(\dot{A} + \dot{B}v + \dot{C}x) - \dfrac{v}{2}C + \lambda(\theta - v)B + \dfrac{v}{2}C^2 + \dfrac{\sigma^2 v}{2}B^2 + \sigma\rho vBC\big] = 0 \\ A(T) = B(T) = 0 \\ C(T) = iz \end{cases}$$

We can simplify F so this is a Riccati differential equation that can be solved analitically and so the theorem is proved □

One of the main drawbacks of the Heston model is that it is not able to reproduce at the same time the smile [4] for short and long maturity. This is related to the fact that this model has an exponential kernel and so it presents a typical time scale of $\dfrac{1}{\lambda}$. Indeed applying Itô's lemma we derive the equation

$$v_t = e^{-\lambda t}v_0 + \lambda\theta\int_0^t e^{-\lambda(t-s)}ds + \sigma\int_0^t e^{-\lambda(t-s)}\sqrt{v_s}dB_s \tag{1.3}$$

that shows us the fact that the variance in the Heston model is a Volterra process with exponential kernel.
Moreover, as we said above, in [31], the authors show that for a wide range of assets the historical volatility is much rougher than the one reproduced by a standard Brownian motion. Due to this fact it is common since recently to replace the standard Heston model with its rough counterpart that does not present a time scale due to the fact that it present a fractional kernel.
Unfortunately since the volatility in this model is not driven by a standard

[4] The smile is the value of the implied volatility for options of the same kind, written on the same underlying, with the same maturity but with different strike prices. This is called smile because its shape sometimes recalls the one of a smile.

Brownian motion this process is no longer neither a semimartingale nor a Markov process as shown in section 1.2. To overcome this problem it is possible to approximate the rough Heston model with the so called Markovian-lift (see chapter 2) which replaces the fractional kernel with an infinite sum of exponential kernels each one with its typical time scale. [5]

1.2 Fractional Brownian motion

The purpose of this section is to present the main results on fractional Brownian motion in order to show that if the volatility is rough, then it can not be nor a Markov process neither a semimartingale. Moreover we will also prove a result that will be used in the following to study some statistical property of our variance processes. All the definitions and results of this section can be found in the first two chapters of [37].

Definition 1.2.1. *A stochastic process $X = (X_t)_{t\in\mathbb{T}}$ is said to be a Gaussian process if for all $k \in \mathbb{N}$ and for all $t_1 < ... < t_k$ in $\mathbb{T}$ the random vector $(X_{t_1}, ..., X_{t_k})$ is Gaussian.*

Definition 1.2.2. *Let $H \in (0, 1]$, a fractional Brownian motion (FBM) is a centered and continuous Gaussian process $B^H = (B_t^H)_{t\geq 0}$ such that the covariance function is given by:*

$$\mathbb{E}[B_t^H B_s^H] = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}) \tag{1.4}$$

Remark 1. If $H = \frac{1}{2}$ the FBM coincides with the standard Brownian motion, on the other hand it is easy to see that if $H = 1$ then $B_t^H = tB_1^H$ almost surely for all t. Thus in the following we exclude this two trivial cases.

Proposition 1.2.1. *Let B^H be a fractional Brownian motion, then* [6] *:*

- *Self-similarity: for all $a > 0$, $(a^{-H} B_{at}^H)_{t\geq 0} = (B_t^H)_{t\geq 0}$*
- *Stationary increments: for all $h>0$, $(B_{t+h}^H - B_h^H)_{t\geq 0} = (B_t^H)_{t\geq 0}$*
- *Time inversion: $(t^{2H} B_{\frac{1}{t}}^H)_{t\geq 0} = (B_t^H)_{t\geq 0}$*

Proof. Clearly the three processes in the left-hand side of Proposition 1.2.1 are continuous, centered and Gaussian (direct check). So, it remains to prove the condition on the covariance:

[5] The idea is that we can use an infinite number of Heston model, each one with its characteristic time scale, to approximate its rough counterpart.

[6] all the "=" stand for equal by law.

- fix $a > 0$ and $0 \leq s < t$ so

$$\mathbb{E}[a^{-2H}(B^H_{at}B^H_{as})] = \frac{a^{-2H}}{2}((at)^{2H} + (as)^{2H} - (at-as)^{2H}) = \mathbb{E}[B^H_t B^H_s]$$

- fix $h > 0$ and $0 \leq s < t$ so

$$\mathbb{E}[(B^H_{t+h} - B^H_h)(B^H_{s+h} - B^H_h)] = \frac{1}{2}(t^{2H} + s^{2H} - (t-s)^{2H}) = \mathbb{E}[B^H_t B^H_s]$$

- let $0 \leq s < t$ so

$$\mathbb{E}[(ts)^{2H} B^H_{\frac{1}{t}} B^H_{\frac{1}{s}}] = \frac{(ts)^{2H}}{2}\left(\frac{t^{2H} + s^{2H} - (t-s)^{2H}}{(ts)^{2H}}\right) = \mathbb{E}[B^H_t B^H_s]$$

□

The next proposition will be very useful in the estimate of the Hurst index of the volatility process.

Proposition 1.2.2. *Let B^H be a fractional Brownian motion, then for all $t, s, q > 0$ it holds*

$$\mathbb{E}[(B^H_{t+s} - B^H_s)^q] = M_q \; t^{Hq} \tag{1.5}$$

where M_q is the moment of order q of a standard Gaussian.

Proof. From self-similarity and stationariness of increments

$$\mathbb{E}[(B^H_{t+s} - B^H_s)^q] = \mathbb{E}[(B^H_t)^q] = \mathbb{E}[t^{Hq} \; (B^H_1)^q] = M_q \; t^{Hq}$$

□

Theorem 1.2.1 (Stochastic representation)**.** *Let $H \in (0,1)$ with $H \neq \frac{1}{2}$, define*

$$c = \sqrt{\frac{1}{2H} + \int_0^\infty \left((1+u)^{H-\frac{1}{2}} - u^{H-\frac{1}{2}}\right)^2 du} < \infty$$

and let $W = (W_t)_{t\geq 0}$ be a classical Brownian motion. Then the process $B^H = (B^H_t)_{t\geq 0}$ defined as

$$B^H_t = \frac{1}{c}\left(\int_{-\infty}^0 \left((t-u)^{H-\frac{1}{2}} - (-u)^{H-\frac{1}{2}}\right) dW_u + \int_0^t (t-u)^{H-\frac{1}{2}} dW_u\right) \tag{1.6}$$

is a FBM with Hurst parameter H.

The proof of the Theorem 1.2.1 can be found in [37]. Now it is important for us to focus on the non trivial properties of this process, in particular the following theorems show that it is neither a semimartingale nor a Markov process.

Theorem 1.2.2. *Let B^H be a FBM with Hurst index $H \neq \frac{1}{2}$. Then it is not a semimartingale.*

To show the proof it is necessary the following preliminary result:

Lemma 1.2.1. *Let B^H be a FBM with $H \in (0,1)$, let $p \geq 1$ and let $G \sim N(0,1)$. Then in $L^2(\Omega)$ as $n \to \infty$ it holds*

$$\sum_{k=1}^{n} \left(|B^H_{\frac{k}{n}} - B^H_{\frac{k-1}{n}}|\right)^p \longrightarrow \begin{cases} 0 \quad if \quad p > \frac{1}{H} \\ \mathbb{E}(|G|^p) \quad if \quad p = \frac{1}{H} \\ +\infty \quad if \quad p < \frac{1}{H} \end{cases}$$

The proof of this lemma is quite long and can be found in [37]. We assume that it is true and so it is possible to give the proof of Theorem 1.2.2.

Proof. Since in the rough Heston model $H < \frac{1}{2}$ let us focus only on this case. The idea to prove the other case is very similar. By selfsimilarity it is sufficient to consider a time interval [0,1]. Recall that if S is a semimartingale then

$$\sum_{k=1}^{n} \left(S_{\frac{k}{n}} - S_{\frac{k-1}{n}}\right)^2 \longrightarrow \langle S \rangle$$

in probability as $n \to +\infty$ where $\langle S \rangle$ is the quadratic variation of the process S. [7]
But using the previous lemma it holds

$$\sum_{k=1}^{n} \left(B^H_{\frac{k}{n}} - B^H_{\frac{k-1}{n}}\right)^2 \longrightarrow \infty$$

so that B^H can not be a semimartingale. □

Theorem 1.2.3. *Let B^H be a FBM with Hurst index $H \neq \frac{1}{2}$. Then it is not a Markov process.*

[7] The quadratic variation of a stochastic process $(X_t)_{t\geq 0}$ in the interval $[0,T]$ is defined taking a partition $0 = t_1 < ... < t_n = T$ with step $\pi = \max_{1\leq i\leq n} t_i - t_{i-1}$ and considering $\lim_{\pi \to 0^+} \sum_{i=1}^{n} \left(X_{t_i} - X_{t_{i-1}}\right)^2$ .For more details see Chapter 4 of [34].

Proof. Assume by contradiction that it is Markovian, so that since it is Gaussian it must be for all $0 \leq s < t < u$ that [8]

$$\mathbb{E}(B_u^H B_s^H)\mathbb{E}((B_t^H)^2) = \mathbb{E}(B_u^H B_t^H)\mathbb{E}(B_t^H B_s^H) \tag{1.7}$$

Choose $u = 1$ and define for every $s \in (0, 1]$ the function

$$\phi(s) = \mathbb{E}(B_1^H B_s^H) = \frac{1}{2}(1 + s^{2H} - (1-s)^{2H}) \tag{1.8}$$

so by construction it holds for all $0 < s \leq t \leq 1$

$$\frac{\mathbb{E}(B_t^H B_s^H)}{\mathbb{E}((B_t^H)^2)} = \phi(s/t)$$

and

$$\phi(s) = \phi(s/t)\phi(t)$$

Now define $\psi(x) = \log(\phi(e^{-x}))$ for all $x \geq 0$ so that it is easy to observe by the previous constructions that:

- $\psi(0) = 0$
- $\lim_{x \to +\infty} \psi(x) = -\infty$
- $\psi(x + y) = \psi(x) + \psi(y)$

Thus it holds that $\frac{\partial \psi}{\partial x} = -c < 0$ so that $\psi(x) = -cx$ or equivalently $\phi(x) = x^c$ for all $0 \leq x \leq 1$.
Taking the second derivative in (1.8) gives:

$$\phi^{''}(s) = H(2H-1)(s^{2H-2} - (1-s)^{2H-2})$$

and since $H(2H-1) \neq 0$ and $2H - 2 < 0$ it holds $\lim_{s \to 1} |\phi^{''}(s)| = +\infty$.
But it is also true that $\phi^{''}(s) = c(c-1)s^{c-2}$ hence $\lim_{s \to 1} |\phi^{''}(s)| = c|(c-1)| \neq +\infty$.
This leads to a contradiction so that B^H is not a Markov process. □

1.3 Volterra processes

In this section we want to present the so called Volterra processes. This kind of processes belong to the class of the diffusive processes and they can be defined as the solutions of a special class of stochastic equations. In finance it is common to use the so called affine Volterra process. The attractiveness of affine processes arises from the fact that they exhibit a high degree of tractability and from their ability in the description of the rough behaviour that presents the volatility in the market. The other side of the medal is that in general these processes are neither semimartingale nor Markovian.

[8] See chapter 11 of [33].

Definition 1.3.1. *Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. A stochastic Volterra equation is an equation of the form*

$$X_t = X_0 + \int_0^t k(t-s)b(X_s)ds + \int_0^t k(t-s)\sigma(X_s)dW_s \tag{1.9}$$

with $(W_t)_{t\geq 0}$ an m-dimensional Brownian motion, $k \in L^2_{loc}(\mathbb{R}^+, \mathbb{R}^{d\times d})$ and $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ continuous functions. A solution of this equation is said to be a Volterra process.

As well as for ordinary stochastic differential equation it is possible to distinguish between strong and weak solution. In particular we said that Equation (1.9) admits a *weak solution* if there exists a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, an $(\mathcal{F}_\sqcup)_{t\geq 0}$ adapted Brownian motion $(W_t)_{t\geq 0}$ and an adapted, continuous process $(X_t)_{t\geq 0}$ such that Equation (1.9) holds $\forall t \geq 0$. Moreover this process is said *strong solution* if in addition it is adapted to the natural filtration of the Brownian motion $(\mathcal{G}_t)_{t\geq 0} = (\sigma(W_s|s \leq t))_{t\geq 0}$. Now we can state an important result on the existence of continuous weak solution under certain condition.

Theorem 1.3.1. *Consider the equation*

$$X_t = X_0 + \int_0^t k(t-s)b(X_s)ds + \int_0^t k(t-s)\sigma(X_s)dW_s \tag{1.10}$$

defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Suppose that the functions $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are continuous and that there exists a positive constant $a > 0$ such that

$$\max\{|b(x)|, |\sigma(x)|\} < a(1 + |x|), \ \forall x \in \mathbb{R}^d.$$

Moreover assume that each component of the kernel $k_{i,j}$, $i, j = 1, ..., d$ satisfies the conditions: $k_{i,j} \in L^2_{loc}(\mathbb{R}^+, \mathbb{R})$, there exists $c \in (0, 2]$ such that

$$\int_0^t k_{i,j}^2(s)ds = O(t^c)$$

and

$$\int_0^T (k_{i,j}(s+t) - k_{i,j}(s))^2 ds = O(t^c)$$

for every $T < \infty$. In addition suppose that there exists a measure $L : \mathbb{R}^+ \to \mathbb{R}^{d\times d}$ of locally bounded variation such that [9]

$$k * L = L * k = \mathbb{I}.$$

Then there exists a continuous weak solution for any initial condition $X_0 \in \mathbb{R}^d$.

[9] The stochastic convolution between a measurable function k and a locally bounded variation measure L is defined by $(k * L)(t) = \int_0^t k(t-s)L(ds)$. In general this operation is not commutative, for instance if you consider non-commuting square matrices.

For the proof see [3] Theorem 3.4. Now we want to use this theorem to derive an important result that will be the core for the derivation of the dynamics in the rough Heston model.

Corollary 1.3.1. *Consider the Volterra equation*

$$X_t = X_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - X_s) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sigma \sqrt{X_s} dW_s \tag{1.11}$$

for $\alpha \in biggl(\frac{1}{2}, 1\Big)$.
Then for any initial value $X_0 \in \mathbb{R}$ *there exists a continuous weak solution.*

Proof. We only need to check that the kernel and the other functions satisfy the hypothesis of Theorem 1.3.1.
The kernel $k(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)}$ is locally square integrable, and we have

$$\int_0^t k^2(s) ds = \frac{t^{2\alpha-1}}{(2\alpha-1)\Gamma^2(\alpha)}$$

as well as

$$\int_0^T (k(s+t) - k(s))^2 ds \leq \frac{1}{(2\alpha-1)\Gamma^2(\alpha)} \Big[(T+t)^{2\alpha-1} - t^{2\alpha-1} + T^{2\alpha-1} \Big]$$

thus we can fix $c = 2\alpha - 1$. Moreover we can consider $L(dx) = \frac{x^{-\alpha}}{\Gamma(1-\alpha)} dx$ and so we have

$$(k * L)(t) = \int_0^t \frac{(t-s)^{\alpha-1} s^{\alpha-1}}{\Gamma(\alpha)\Gamma(1-\alpha)} ds = 1.$$

In addition the functions $b(X) = \lambda(\theta - X)$ and $\sigma(X) = \sigma\sqrt{X}$ are clearly continuous and respectively with linear and square root growth. Thus all the hypothesis are satisfied and then we conclude the existence of a continuous weak solution. □

Remark 2. Under the Lipschitz condition of the functions b and σ it is also possible to prove the existence and uniqueness of continuous strong solution of (1.9). The problem of Equation (1.11) is that the function $\sigma(X) = \sigma\sqrt{X}$ is not lipschitz in zero.

It is possible to prove that Equation (1.11) admits a unique solution in law as demonstrated in Theorem 6.1 of [3]. Nevertheless the proof is very technical and goes beyond the aim of this thesis.
The important fact is that the rough Heston model, that we are going to present in the next section, is based on a variance process of the form given by Equation (1.11) and thanks to these results we have a good definition of the model.

1.4 Rough Heston model

The rough Heston model (RHM) is build up in such a way that the underlying asset S follows the same dynamics of the classical Heston model and the variance v is defined such that

$$v_t = v_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - v_s) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sigma \sqrt{v_s} dB_s \tag{1.12}$$

where the additional parameter $\alpha = H + \frac{1}{2}$ belongs to $(\frac{1}{2}, 1)$.

Remark 3. In the limit when $\alpha \to 1$ the RHM converges to the classical one

Remark 4. As showed by the authors in [31] the rough volatility is coherent with what is observed in the real market and so the dynamics of Equation (1.12) is referred to the real word probability measure $\mathbb{P}$. Nevertheless, for our purpose of pricing options, we can assume that the dynamics given by Equation (1.12) is true also under the risk neutral measure $\mathbb{Q}$ as for semi-martingale models.

The fact that the volatility in the rough Heston model is no more a semi-martingale and a Markov process seems to be a problem for its tractability, however, as showed by the authors in [25], the characteristic function is known even if it is necessary to solve numerically a fractional Riccati differential equation. Nevertheless nowadays there are some available techniques that do it in a quite efficient way (see for instance [18]).

1.4.1 The rough Heston model's characteristic function

Definition 1.4.1. *Define the fractional integral of order $r \in (0, 1]$ of a function f as*

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

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

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

whenever the derivative exists.

Using the above definition it is possible to state the following theorem.

Theorem 1.4.1. *Consider the rough Heston model given by*

$$\begin{cases} S_t = S_0 + \int_0^t \sqrt{v_s} dW_s \\ v_t = v_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - v_s) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sigma \sqrt{v_s} dB_s \end{cases} \tag{1.13}$$

and define $X_t = \log(S_t)$ *and* $L(a,t) = \mathbb{E}(e^{iaX_t})$. *Then it holds*

$$L(a,t) = \exp(\theta \lambda I^1 h(a,t) + v_0 I^{1-\alpha} h(a,t)) \tag{1.14}$$

where $h(a,\cdot)$ *is the solution of the fractional Riccati equation*

$$\begin{cases} D^\alpha h(a,t) = \frac{1}{2}(-a^2 - ia) + (ia\rho\sigma - \lambda) h(a,t) + \frac{\sigma^2}{2} h^2(a,t) \\ I^{1-\alpha} h(a,0) = 0 \end{cases} \tag{1.15}$$

For a complete derivation of this result, we refer to [25], [38], [32]. The idea is to use an appropriate sequence of Hawkes processes that converges in law to the rough Heston model, find their characteristic function and then derive the one of the logarithm of the asset price in the RHM as a limit.
The characteristic function is useful for the calibration of the model but it is not possible to use it for pricing exotic options which require Monte Carlo simulation. Nevertheless due to the lack of Markovianity of the volatility process the simulation is difficult, thus it has been introduce the lifted Heston model which we will present in the next chapter.

Chapter 2

Lifted Heston model

2.1 General results

In order to preserve the Markovian structure of the Heston model and the accuracy of the rough version, it is convenient to consider the lifted Heston model. In section 2.2 we will prove some important results that explain why this model is a good approximation of the rough Heston.
The lifted Heston model was introduced as a multifactor approximation of the rough Heston model in [2]. Then in [1] the author shows one way to reduce the high number of factors that must be determined for an efficient approximation. Here the idea is to consider n factors driven by the same Brownian motion but with different mean reversion. This means that we are working with exponential kernels that determine different time scales so that we overcome the problems arising in the use of the standard Heston model and, at the same time, we preserve the Markovian structure. In particular for a fixed $n \in \mathbb{N}$ the dynamics of the stock price and of the variance is such that:

$$dS_t^n = S_t^n \sqrt{v_t^n} dW_t, \quad S_0^n \geq 0 \tag{2.1}$$

$$v_t^n = g_0^n(t) + \sum_{i=1}^{n} w_i^n U^{n,i} \tag{2.2}$$

$$dU_t^{n,i} = (-x_i^n U_t^{n,i} - \lambda v_t^n)dt + \sigma \sqrt{v_t^n} dB_t, \quad U_0^{n,i} = 0 \quad \forall i = 1, ..., n \tag{2.3}$$

with parameters the function g_0^n, the constants $\lambda, \sigma \in \mathbb{R}^+$, $w_i^n, x_i^n \geq 0 \quad \forall i = 1, ..., n$ and $\rho dt = d\langle B, W \rangle_t$, $\rho \in [-1, 1]$.
In [1] the author gives the rigorous proof of the existence and uniqueness of a continuous strong solution of (2.1)-(2.3).
This approximation of the RHM allows to preserve the quality of the classic Heston model and its rough counterpart, and also to speed up the simulation.

Remark 5. In the following we omit the apex n if there is no ambiguity.

Observing that $\mathbb{E}(U_t^{n,i}) = \int_0^t -\lambda e^{-x_i^n(t-s)}\mathbb{E}(v_s^n)ds$ [1] and taking the expectation on both sides of (2.2) it is possible to find the relation

$$g_0^n(t) = \mathbb{E}(v_t^n) + \lambda \sum_{i=1}^n w_i^n \int_0^t e^{-x_i^n(t-s)}\mathbb{E}(v_s^n)ds \tag{2.4}$$

From now on, we will restrict our attention to the case in which the deterministic function $g_0^n(\cdot)$ is of the form:

$$g_0^n(t) = v_0 + \lambda\theta \sum_{i=1}^n w_i^n \int_0^t e^{-x_i^n(t-s)}ds, \quad v_0, \theta \geq 0 \tag{2.5}$$

Remark 6. It is clear that if $n = 1, w_1^1 = 1$ and $x_1^1 = 0$ we find the classical Heston model.

Theorem 2.1.1. *Consider the evolution described in the lifted Heston model* (2.1)-(2.3) *and define* $X_t^n = \log(S_t^n)$. *Then it holds:*

$$\mathbb{E}(e^{iaX_T^n}|\mathcal{F}_t) = \exp\Big(\phi^n(t,T) + iaX_t^n + \sum_{i=1}^n w_i^n \psi^{n,i}(T-t)U_t^{n,i}\Big), \quad \forall t \leq T \tag{2.6}$$

where $(\psi^{n,i})_{1\leq i\leq n}$ *solves the following n-dimensional system of Riccati ODEs:*

$$(\psi^{n,i})' = -x_i^n\psi^{n,i} + F\big(ia, \sum_{j=1}^n w_j^n\psi^{n,j}\big), \quad \psi^{n,i}(0) = 0, \quad i = 1, ..., n$$

with

$$F(ia, v) = \frac{1}{2}(-a^2 - ia) + (\rho\sigma ia - \lambda)v + \frac{\sigma^2}{2}v^2$$

and

$$\phi^n(t,T) = \int_0^{T-t} F\big(ia, \sum_{i=1}^n w_i^n\psi^{n,i}(s)\big)g_0^n(T-s)ds, \quad t \leq T$$

Here the problem seems to be the high number of parameters that must be determined because we have to consider $2n+5$ parameters: $(v_0, \lambda, \theta, \rho, \sigma)$ and the other $2n$ $(w_i^n, x_i^n)_{i=1,...,n}$. Nevertheless, in [1] the author provides a parameterization of $(w_i^n, x_i^n)_{i=1,...,n}$ in terms of the Hurst index H of a limiting rough volatility model and one additional parameter r_n and so they reduce to 6 once you have chosen the value of H which usually is fixed at 0.1.

[1] Assuming $dU_t = (-xU_t - \lambda v_t)dt + \sigma\sqrt{v_t}dB_t$ then by Itô's lemma we have $d(e^{xt}U_t) = -\lambda e^{xt}v_t dt + e^{xt}\sigma\sqrt{v_t}dB_t$ so that writing the equation in the integral form and taking the expected value on both sides we find our goal

Lemma 2.1.1. *The fractional kernel appearing in Equation* (1.12) *admits the representation*

$$\frac{t^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})} = \int_0^{+\infty} e^{-xt}\mu(dx), \quad \mu(dx) = \frac{x^{-H-\frac{1}{2}}dx}{\Gamma(H+\frac{1}{2})\Gamma(\frac{1}{2}-H)} \tag{2.7}$$

Proof. Observe that

$$\int_0^{+\infty} e^{-xt}x^{-H-\frac{1}{2}}dx = \int_0^{+\infty} e^{-xt}\frac{(xt)^{-H-\frac{1}{2}}}{t^{-H-\frac{1}{2}}}dx. \tag{2.8}$$

Perform the change of variable $tx = z$ so that

$$(2.8) = t^{H-\frac{1}{2}}\int_0^{+\infty} e^{-z}z^{(\frac{1}{2}-H)-1}dz = t^{H-\frac{1}{2}}\Gamma\big(\frac{1}{2}-H\big)$$

□

Using the above lemma and setting $v_0 = 0$, it is possible to rewrite Equation (1.12) as

$$v_t = \int_0^{+\infty} U_t(x)\mu(dx), \quad x > 0 \tag{2.9}$$

where

$$U_t(x) = \int_0^t e^{-x(t-s)}\big(\lambda(\theta - v_s)ds + \sigma\sqrt{v_s}dB_s\big), \quad \forall x > 0 \tag{2.10}$$

This is important because it allows us to look at the dynamics given by Equation (1.12) as a superposition of infinitely many processes $(U(x))_{x>0}$ that have the same dynamics but different mean reversion $x \in (0, +\infty)$.

2.2 Convergence to the rough Heston model

In this section we want to prove some important results about the convergence of the kernel appearing in the LHM to the one of the RHM and so to validate the Markovian approximation that we are doing.
First of all, it is important to observe that thanks to Equation (2.3) we have that Equation (2.2) can be rewritten as

$$v_t^n = g_0^n(t) + \int_0^t K^n(t-s)\big(-\lambda v_s^n ds + \sigma\sqrt{v_s^n}dB_s\big) \tag{2.11}$$

where

$$K^n(t) = \sum_{i=1}^n w_i^n e^{-x_i^n t}, \quad t \geq 0. \tag{2.12}$$

Then, choosing in an appropriate way the values of $(w_i^n, x_i^n)_{1\leq i\leq n}$, we can prove the convergence in $L^2(0,T)$ of K^n to the fractional kernel K.

Theorem 2.2.1. *Let* $\alpha = H + \dfrac{1}{2} \in biggl(\dfrac{1}{2}, 1\Big)$ *and* $(r_n)_{n\geq 1}$ *be a sequence of real numbers such that*

$$r_n \downarrow 1 \quad \textit{and} \quad n\log(r_n) \to +\infty \quad \textit{as } n \to +\infty \tag{2.13}$$

and define $(w_i^n, x_i^n)_{1\leq i\leq n}$ *such that*

$$w_i^n = \frac{(r_n^{1-\alpha}-1)r_n^{(\alpha-1)(1+n/2)}}{\Gamma(\alpha)\Gamma(2-\alpha)} r_n^{(1-\alpha)i} \tag{2.14}$$

$$x_i^n = \left(\frac{1-\alpha}{2-\alpha}\right)\frac{r_n^{2-\alpha}-1}{r_n^{1-\alpha}-1} r_n^{i-1-n/2}. \tag{2.15}$$

Then for all $T > 0$ *it holds*

$$||K^n - K||_{L^2(0,T)} \to 0 \quad n \to +\infty$$

where $K(t) = \displaystyle\int_0^{+\infty} e^{-xt}\mu(dx)$ *and* $\mu(dx) = \dfrac{x^{-H-\frac{1}{2}}dx}{\Gamma(H+\frac{1}{2})\Gamma(-H+\frac{1}{2})}$.

Proof. Set $\eta_i^n = r_n^{i-n/2}$ for $i = 0, ..., n$ and observe that it holds by construction $w_i^n = \displaystyle\int_{\eta_{i-1}^n}^{\eta_i^n} \mu(dx)$. Then it is possible to write:

$$\begin{aligned} K^n(t) - K(t) =& \int_0^{\eta_0^n} e^{-xt}\mu(dx) + \sum_{i=1}^{n}\int_{\eta_{i-1}^n}^{\eta_i^n} \left(e^{-xt} - e^{-x_i^n t}\right)\mu(dx) + \int_{\eta_n^n}^{\infty} e^{-xt}\mu(dx) \\ =& I_1 + I_2 + I_3 \end{aligned}$$

Thus we can estimate the L^2-norm of each integral $(I_j)_{j\in\{1,2,3\}}$. Relying on a second order Taylor expansion, along the lines of the proof of Proposition 7.1 in [20], we get:

$$\Big|\int_{\eta_{i-1}^n}^{\eta_i^n} \left(e^{-xt} - e^{-x_i^n t}\right)\mu(dx)\Big| \leq Ct^2 r_n^{1/2}(r_n - 1)^2 \int_{\eta_{i-1}^n}^{\eta_i^n} (1 \wedge x^{-1/2})\mu(dx), \quad t \leq T$$

for all $i = 1, ..., n$, and with $C > 0$ independent of n, i, t. Thus if we consider the summation over all i we get

$$||I_2||_{L_2} \leq C\frac{T^{5/2}}{\sqrt{5}}\sqrt{r_n}(r_n - 1)^2 \int_0^{\infty} (1 \wedge x^{-1/2})\mu(dx)$$

that gives $||I_2||_{L_2} \to 0$ due to the hypothesis (2.13).
Moreover thanks to the same hypothesis on r_n we have:

$$||I_1||_{L_2} \leq \int_0^{\eta_0^n} \mu(dx) = \frac{(\eta_0^n)^{1-\alpha}}{\Gamma(\alpha)\Gamma(2-\alpha)} = \frac{(r_n)^{-(1-\alpha)n/2}}{\Gamma(\alpha)\Gamma(2-\alpha)} \to 0 \quad as \quad n \to \infty$$

and

$$||I_3||_{L_2} \leq \int_{\eta_n^n}^{\infty} \sqrt{\frac{1-e^{-2xT}}{2x}}\mu(dx) \leq \frac{(r_n)^{(1/2-\alpha)n/2}}{\Gamma(\alpha)\Gamma(2-\alpha)} \to 0 \quad as \quad n \to \infty$$

and so we have proved the theorem. □

This is an important result because it tells us that in the limit the lifted Heston model converges to the rough version which is the one that we want to approximate.
Moreover in [10] the author shows a stronger result:

Theorem 2.2.2. *Let* $L^1(X) = \{f : X \to X : ||f||_{L_1} < \infty\}$ *and* $L^1_{bc}(X) = \{f : X \to X : f \in L^1(X),\ f$ *is bounded and continuous*$\}$. *Assume that* $K^n(t)$ *are chosen such that*

$$e_n = \int_0^T |K(t) - K^n(t)|dt \to 0 \quad as \quad n \to \infty$$

and there exist $C>0$ *such that*

$$K^n(t) < CK(t), \quad (K^n)'(t) < CK'(t) \quad \forall t \leq T$$

Let $f : \mathbb{R} \to \mathbb{R}$ *be a function of the log-stock price and* $\hat{f}$ *its Fourier transform. Define* $f_R(x) = e^{-Rx}f(x)$ *for each* $R \in \mathbb{R}$ *and consider the set* $\mathcal{I} = \{R \in \mathbb{R} :\ f_R \in L^1_{bc}$ *and* $\hat{f}_R \in L^1\}$ *and let* $a \in \mathcal{I} \cap [0,1]$. *Assume that there exist* $C_1, \delta > 0$ *such that*

$$|\hat{f}(b+ia)| \leq C_1(1+|b|)^{-(1+\delta)}$$

Then there exist $\bar{C}, \epsilon > 0$ *independent of* C_1, δ, f *such that if* $e_n < \epsilon$ *then for* $\delta \neq 6$ *it holds*

$$|\mathbb{E}(f(X_T)) - \mathbb{E}(f(X_T^n))| \leq \hat{C}C_1\left(\frac{1}{|6-\delta|} + \frac{1}{\delta}\right)e_n^{\frac{\delta}{6}\wedge 1}. \tag{2.16}$$

Moreover for $\delta = 6$ *it holds*

$$|\mathbb{E}(f(X_T)) - \mathbb{E}(f(X_T^n))| \leq \hat{C}C_1 \log(e_n^{-1})e_n \tag{2.17}$$

The rigorous proof can be found in [10]. For us it is important to notice that this result is crucial since it tells us that we can work with the lifted Heston model instead of using the rough one without losing too much precision if we choose n such that $e_n < \epsilon$ [2]. Now we want to show that it is possible to choose $K^n(t)$ s.t. $e_n \to 0$ because, up to now, we only know that $||K^n - K||_{L_2} \to 0$.

[2] The constants C_1, δ are fixed so thanks to the hypothesis on the convergence to zero of e_n we can increase the number of factors in the Markovian kernel in order to move down the error under a certain tolerance

Theorem 2.2.3. *Under the same assumption of Theorem 2.2.1 it holds*

$$||K^n - K||_{L^1(0,T)} \to 0 \quad as\ n \to \infty$$

Proof. The idea is to use the same approximation of Theorem 2.2.1 so we give only a sketch of the proof. Define $\eta_i^n = r_n^{i-n/2}$ for $i = 0, ..., n$ and $I_1,\ I_2$ and I_3 as we did in the proof of Theorem 2.2.1. First all we have

$$\Big|\int_{\eta_{i-1}^n}^{\eta_i^n} \big(e^{-xt} - e^{-x_i^n t}\big)\mu(dx)\Big| \leq Ct^2 r_n^{1/2}(r_n - 1)^2 \int_{\eta_{i-1}^n}^{\eta_i^n} (1 \wedge x^{-1/2})\mu(dx), \quad t \leq T$$

Thus taking the sum over all $i = 1, ..., n$ and thanks to the hypothesis (2.13) on the asymptotic behaviour of r_n we get $||I_2||_{L_1} \to 0$.
Moreover

$$||I_1||_{L_1} \leq T\tilde{C}(\alpha) r_n^{-n/2(1-\alpha)} \to 0$$

and also

$$||I_3||_{L_1} \leq \tilde{C}_1(\alpha) \int_{\eta_n^n}^{\infty} \frac{1 - e^{-xT}}{x} x^{-\alpha} dx \to 0$$

where $\tilde{C}(\alpha), \tilde{C}_1(\alpha)$ are two positive constants depending on $\alpha = \dfrac{1}{2} + H$. □

Remark 7. In [10] the authors provide a more sophisticated estimate of the error, in particular they prove that using Gaussian quadrature rules they can achieve the rate of convergence

$$||K^n - K||_{L^1(0,T)} \leq Ce^{-2.38\sqrt{\alpha n}}.$$

This means that if we choose in a good way the nodes and the weights, the error goes to zero exponentially fast for all $H > -\dfrac{1}{2}$ and so we need a lower number of factors n to approximate the rough kernel.

Chapter 3

Simulation schemes

In this chapter we want to describe efficient simulation schemes both for the rough Heston model and for its lifted version. Thanks to Theorem 2.2.3 and Remark 7 we expect a good approximation from the simulation of the lifted model and, once again, we underline that due to its Markovian structure the simulation is expected to be computationally easier. As we said in the previous chapter, starting from the dynamics of the variance process [1]

$$v_t = v_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1}(\theta - \lambda v_s)ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1}\sigma\sqrt{v_s}dW_s \tag{3.1}$$

and approximating the fractional kernel $K(t) = \dfrac{t^{\alpha-1}}{\Gamma(\alpha)}$ with

$$K^n(t) = \sum_{j=1}^{n} w_i e^{-x_i t} \tag{3.2}$$

we get the Markovian approximation:

$$dS_t = \sqrt{V_t}S_t(\rho dW_t + \sqrt{1-\rho^2}dB_t), \quad S_0 = s_0 \geq 0 \tag{3.3}$$

$$dV_t^i = -x_i(V_t^i - v_0^i)dt + (\theta - \lambda V_t)dt + \sigma\sqrt{V_t}dW_t, \quad V_0^i = v_0^i \tag{3.4}$$

for $i = 1, ..., n$ where $V_t = \sum_{i=1}^{n} w_i V_t^i = \mathbf{w} \cdot \mathbf{V}_t$ and v_0^i are chosen such that $\sum_{i=1}^{n} w_i v_0^i = v_0$. Now we want to compare two possible simulation schemes: the first has been provided by the authors of [8], it deals with the lifted model and it is an extension of the Moment Matching scheme of [35], while the second has been provided in [28] and works directly on the rough model.

[1] We adapt the dynamics of Equation (1.12) according to what is presented in [8]. In particular the only difference is that we bring λ inside the parenthesis and so one should think to θ as $\theta\lambda$ in Equation (1.12).

3.1 Weak simulation scheme

Looking at Equations (3.3)-(3.4) it is clear that the volatility process does not depend on the stock price so this suggests to provide a simulation scheme for the volatility and then use it to simulate the stock price preserving the correlation ρ between the two processes.

3.1.1 Simulation of the volatility

Consider Equation (3.4) and split it into the deterministic and the stochastic part.
Define $D(\mathbf{z}, h) := \mathbf{Z}_h$ the solution at time h of the ODE

$$dZ_t^i = -x_i(Z_t^i - v_0^i)dt + (\theta - \lambda Z_t)dt, \quad Z_0^i = z^i, \quad Z_t = \sum_{i=1}^n w_i Z_t^i \tag{3.5}$$

and $S(\mathbf{y}, h) := \mathbf{Y}_h$ the solution at time h of the SDE [2]

$$dY_t^i = \sigma\sqrt{Y_t}dW_t, \quad Y_0^i = y^i, \quad Y_t = \sum_{i=1}^n w_i Y_t^i \tag{3.6}$$

Now observe that the first one can be solved exactly and the solution is given by

$$D(\mathbf{z}, h) = \mathbf{Z}_h = e^{Ah}\mathbf{z} + A^{-1}(e^{Ah} - Id)b$$

where $A = -\lambda \mathbf{1}\mathbf{w}^T - \text{diag}(\mathbf{x})$ and $b = \theta\mathbf{1} + \text{diag}(\mathbf{x})\mathbf{v}_0$.
On the other hand, we have

$$dY_t = \mathbf{w} \cdot d\mathbf{Y}_t = \sum_{i=1}^n w_i \sigma\sqrt{Y_t}dW_t = \bar{w}\sigma\sqrt{Y_t}dW_t, \quad Y_0 = \mathbf{w} \cdot \mathbf{y} \tag{3.7}$$

where $\bar{w} = \sum_{i=1}^n w_i$. This is an SDE which is studied in [35] where the authors provide an approximated solution using a scheme which is based on the idea that we can simulate a discrete random variable $\hat{Y}_h$ which matches the first five moment of our original variable $Y_h = \sum_{j=1}^n w_j Y_h^j$ for which the dynamics is given by Equation (3.7). In order to do this we need to define the values:

$$x = \mathbf{w} \cdot \mathbf{y}, \quad z = \sigma^2 \bar{w}^2 h$$

and use them to construct the quantities

[2] The solutions $D(\mathbf{z}, h)$ and $S(\mathbf{y}, h)$ are vectors since they contain the components that must be used to construct the variance process $(V_t)_{t\geq 0}$.

- $m_1 = x, \quad m_2 = x^2 + xz, \quad m_3 = x^3 + 3x^2 z + \frac{3}{2} x z^2$
- $p_1 = \dfrac{m_1 x_2 x_3 - m_2(x_2 + x_3) + m_3}{x_1(x_3 - x_1)(x_2 - x_1)}$
- $p_2 = \dfrac{m_1 x_1 x_3 - m_2(x_1 + x_3) + m_3}{x_2(x_3 - x_2)(x_1 - x_2)}$
- $p_3 = \dfrac{m_1 x_1 x_2 - m_2(x_1 + x_2) + m_3}{x_3(x_1 - x_3)(x_2 - x_3)}$
- $x_1 = x + Cz - \sqrt{(3x + C^2 z)z}$
- $x_2 = z + (C - \frac{3}{4})z$
- $x_3 = x + Cz + \sqrt{(3x + C^2 z)z}$
- $C = \dfrac{6 + \sqrt{3}}{4}$

Then define the discrete, one dimensional, random variable $\hat{Y}_h$ which takes value x_i with probability p_i for $i = 1, 2, 3$.

Remark 8. It is important to observe that it holds $\mathbb{E}(Y_h^i) = m_i$ for $i = 1, 2, 3$ and the values $(x_i, p_i)_{i=1,2,3}$ are chosen such that $\mathbb{E}(\hat{Y}_h^k) = \mathbb{E}(Y_h^k)$ for $k = 1, ..., 5$. For a rigorous derivation of this result we refer to [35] and [8]. Thus, since this method simulates a random variable which matches the first five moments of Y_h, in the in the following we refer to this scheme as Moment Matching scheme without ambiguity.

Now we want to simulate the entire process $\mathbf{Y}$. Thus proceed in this way:

1. Since the right hand side of (3.6) is the same for all $i = 1, ..., n$ we deduce that the solution must be of the form $Y_h^i = y^i + Q$ where Q is a scalar random variable
2. Observe that $\displaystyle\sum_{i=1}^{n} w_i Y_h^i = \mathbf{w} \cdot \mathbf{y} + \bar{w} Q \Rightarrow Q = \frac{Y_h - \mathbf{w} \cdot \mathbf{y}}{\bar{w}}$
3. Define the vector $\hat{\mathbf{Q}} = \left(\dfrac{\hat{Y}_h - \mathbf{w} \cdot \mathbf{y}}{\bar{w}} \right) \mathbf{1}$.
4. Set the approximated solution

$$\hat{S}(\mathbf{y}, h) = \hat{\mathbf{Y}}_h = \mathbf{y} + \hat{\mathbf{Q}}$$

Finally we combine the solution of the ODE and the approximated solution of the SDE using the Strang splitting and get

$$A^{CIR}(\mathbf{v}, h) = D\Big(\hat{S}\Big(D\Big(\mathbf{v}, \frac{h}{2}\Big), h\Big), \frac{h}{2}\Big) \tag{3.8}$$

for approximating $\mathbf{V}_h$ given $\mathbf{v}$.

Remark 9. The Strang splitting method is a numerical scheme for the solution of differential equation. The general scheme can be described as follow:

- Consider the differential equation $\frac{dy}{dt} = f_0(y) + f_1(y), \quad y(0) = y_0$ and suppose that you computed the value $y(s)$ at a certain time s. Now we want to consider the interval $[s, s + \Delta s]$.

- Make half a step with the operator f_0 such that $\frac{dy_1}{dt} = f_0(y_1),\ y_1(s) = y(s),\ t \in [s, s + \frac{\Delta s}{2}]$.

- Make a full step with the operator f_1 such that $\frac{dy_2}{dt} = f_1(y_2),\ y_2(s) = y_1(s + \frac{\Delta s}{2}),\ t \in [s, s + \Delta s]$.

- Make another half a step with the operator f_0 such that $\frac{dy_3}{dt} = f_0(y_3),\ y_3(s + \frac{\Delta s}{2}) = y_2(s + \frac{\Delta s}{2}),\ t \in \Big[s + \frac{\Delta s}{2}, s + \Delta s\Big]$.

- The global solution at time $s + \Delta s$ is given by $y(s + \Delta s) = y_3(s + \Delta s)$.

For more details on this method consult [36].

3.1.2 Simulation of the stock price

Thanks to the fact that we have a simulation scheme for the volatility, now we are able to simulate the entire process $(S, \mathbf{V})$ where $\mathbf{V}$ is the the vector of variances in $\mathbb{R}^n$ defined by (3.4).
First of all, it is easy to check that the solution of Equation (3.3) is given by

$$S_t = S_0 \exp\Big(\int_0^t \sqrt{V_s}(\rho dW_s + \sqrt{1-\rho^2} dB_s) - \frac{1}{2}\int_0^t V_s ds\Big) \tag{3.9}$$

thus it will be useful to define the process $\mathbf{Y} = Y^i$ for $i = 1, ..., n$ where

$$Y_t^i = \int_0^t V_s^i ds. \tag{3.10}$$

So if we now consider Equations (3.3)-(3.4) and the one that we have introduced $dY_t^i = V_t^i dt, \quad Y_0^i = y^i$, we can split the system into two parts:

1. Consider the first system of SDE $S^W((s, \mathbf{v}, \mathbf{y}), h)$ given by

$$dS_t = \sqrt{V_t} S_t \rho dW_t, \quad S_0 = s$$
$$dV_t^i = -x_i(V_t^i - v_0^i)dt + (\theta - \lambda V_t)dt + \sigma\sqrt{V_t}dW_t, \quad V_0^i = v_0^i$$
$$dY_t^i = V_t^i dt, \quad Y_0^i = y^i.$$

2. Let the second one be $S^B((s, \mathbf{v}, \mathbf{y}), h)$ such that

$$dS_t = \sqrt{V_t} S_t \sqrt{1-\rho^2} dB_t, \quad S_0 = s$$
$$dV_t^i = 0, \quad V_0^i = v_0^i$$
$$dY_t^i = 0, \quad Y_0^i = y^i.$$

The second can be solved explicitly since the volatility is constant and it gives us

$$S_h = s \exp\left(\sqrt{v(1-\rho^2)} B_h - \frac{1}{2} v(1-\rho^2)h \right)$$
$$\mathbf{V}_h = \mathbf{v}$$
$$\mathbf{Y}_h = \mathbf{y}$$

For the solution of S^W we have to use a numerical approximation, so first of all we compute $\hat{\mathbf{V}}_h = A^{CIR}(\mathbf{v}, h)$ and use it to compute with the trapezoidal rule $\hat{\mathbf{Y}}_h = \mathbf{y} + \frac{h}{2}(\mathbf{v} + \hat{\mathbf{V}}_h)$ using the relation between the processes $\mathbf{V}$ and $\mathbf{Y}$ given by Equation (3.10).
Now suppose that the solution has the form

$$S_t = s \exp\left(at + \sum_{i=1}^n b_i (Y_t^i - Y_0^i) + \sum_{i=1}^n c_i (V_t^i - V_0^i) \right). \tag{3.11}$$

Thus taking the Itô's differential we have

$$\begin{aligned} dS_t &= S_t \left(a dt + \sum_{i=1}^n b_i dY_t^i + \sum_{i=1}^n c_i dV_t^i + \frac{1}{2} \sum_{i,j=1}^n c_i c_j d\langle V^i, V^j \rangle_t \right) \\ &= S_t \Bigg[\left(a + \sum_{i=1}^n c_i x_i v_0^i + \theta \sum_{i=1}^n c_i \right) dt \\ &+ \sum_{i=1}^n \left(b_i - c_i x_i - \lambda w_i \sum_{j=1}^n c_j + \frac{1}{2}\sigma^2 w_i \sum_{j,k=1}^n c_j c_k \right) V_t^i dt + \sum_{i=1}^n \sigma c_i \sqrt{V_t} dW_t \Bigg] \end{aligned}$$

thus, in order to have a solution of the system of SDE S^W, we have to impose

$$a + \sum_{i=1}^{n} c_i x_i v_0^i + \theta \sum_{i=1}^{n} c_i = 0$$

$$b_i - c_i x_i - \lambda w_i \sum_{j=1}^{n} c_j + \frac{1}{2}\sigma^2 w_i \sum_{j,k=1}^{n} c_j c_k = 0, \quad i = 1, ..., n$$

$$\sum_{i=1}^{n} \sigma c_i = \rho$$

It is important to observe that the last equation has not a unique solution and once we have determined the coefficients $(c_i)_{i=1,...,n}$ we have also fixed the others by virtue of the first two equations.
In [8] the authors suggest to choose

$$c_i = \frac{\rho}{\sigma}\delta_{i,1}$$

where $\delta_{i,1} = 1$ if $i = 1$ and 0 for any other $j \neq 1$. It seems to be the choice that makes the numerical algorithm as stable as possible, especially after we have sorted $0 \leq x_1 < ... < x_n$ [3]. Fixing the coefficients c_i in this way we have that

$$a = -(x_1 v_0^i + \theta)\frac{\rho}{\sigma}, \quad b_i = \frac{\rho}{\sigma} x_1 \delta_{i,1} + \lambda w_i \frac{\rho}{\sigma} - \frac{1}{2} w_i \rho^2$$

and therefore

$$S_t = s \exp\left(\frac{\rho}{\sigma}\left(-(x_1 v_0^1 + \theta)t + x_1(Y_t^1 - Y_0^1) + \left(\lambda - \frac{\rho\sigma}{2}\right)(Y_t - Y_0) + (V_t^1 - V_0^1)\right)\right)$$

Hence we have an approximate solution of S^W given by

$$\hat{\mathbf{V}}_h = A^{CIR}(\mathbf{v}, h)$$

$$\hat{\mathbf{Y}}_h = \mathbf{y} + \frac{h}{2}(\mathbf{v} + \hat{\mathbf{V}}_h)$$

$$\hat{S}_h = s \exp\left(\frac{\rho}{\sigma}\left(-(x_1 v_0^1 + \theta)h + x_1(\hat{Y}_h^1 - Y_0^1) + \left(\lambda - \frac{\rho\sigma}{2}\right)(\hat{Y}_h - Y_0) + (\hat{V}_h^1 - V_0^1)\right)\right)$$

Now we have to combine the solutions S^B and S^W: in [8] the authors assert that numerical evidence suggest to use the randomized Leapfrog splitting. In order to do this consider a random variable $U \sim Unif(0, 1)$ independent of everything else in the model, then set

$$\hat{S}^{rH}((s, \mathbf{v}, \mathbf{y}), h) = \begin{cases} \hat{S}^W(S^B((s, \mathbf{v}, \mathbf{y}), h), h) & if\ U \leq \frac{1}{2} \\ S^B(\hat{S}^W((s, \mathbf{v}, \mathbf{y}), h), h) & otherwise \end{cases} \tag{3.12}$$

[3] This is just an empirical evidence, the authors do not prove it in a rigorous way

Remark 10. The Leapfrog method is a scheme for the numerical solution of differential equations. The idea is that given a differential equation of the form

$$\frac{dy}{dt} = F(y)$$

the time derivative at time $n\Delta t$ can be approximated by

$$\frac{U((n+1)\Delta t) - U((n-1)\Delta t)}{2\Delta t}$$

thus

$$U((n+1)\Delta t) = U((n-1)\Delta t) + 2\Delta t F(U(n\Delta t)).$$

For more details and a rigorous description we refer to [15].

3.1.3 Simulation of the entire process (S, V)

Now that we have a simulation scheme for the variance and for the stock price we can combine then to simulate the entire process (S, V). The general description can be given in the following way:

- For a fixed discretization step Δt, given the vector $\mathbf{V}_{n-1}$ compute $\mathbf{V}_n$ i.e. the variances at time t_n using the scheme given in Equation (3.8) .
- Compute the vector $\mathbf{Y}_n = \mathbf{Y}_{n-1} + \frac{\Delta t}{2}(\mathbf{V}_n + \mathbf{V}_{n-1})$.
- Use the scheme described in (3.12) to compute S_n given S_{n-1}.
- Compute the variance at time t_n using $V_n = \mathbf{w} \cdot \mathbf{V}_n$.

Remark 11. Looking at the simulation scheme described above it is evident the Markovian structure of the model. Indeed for the simulation at time t_n we only need the value of the processes at time t_{n-1}.

3.2 Hybrid Quadratic Exponential scheme

The next simulation scheme that we want to study is the Hybrid Quatratic Exponential scheme which was presented by Gatheral in [28]. This scheme works directly on the rough version of the model. The idea is to combine the Quadratic Exponential[4] (QE) scheme of [5] and the Hybrid scheme of [12]. To do this we deal with rough processes of the form

$$S_t = S_0 + \int_0^t S_s \sqrt{v_s} dW_s \tag{3.13}$$

[4]See appendix B

$$v_t = \mathbb{E}(v_t) + \int_0^t K(t-s)\sqrt{v_s}dB_s = \xi(t) + \int_0^t K(t-s)\sqrt{\xi_s(s)}dB_s \quad (3.14)$$

with B, W two correlated Brownian motions with correlation ρ, forward variance $\xi_u(t) = \mathbb{E}_u(v_t)$ (if $u = 0$ we write $\xi(t)$ instead of $\xi_u(t)$) and $K(\cdot)$ an appropriate kernel.
In particular we will assume that $K(t) = t^{H-1/2}L_k(t),\ t \in (0, +\infty]$ with $H \in \left(0, \frac{1}{2}\right)$ and $L_k : (0, +\infty] \to [0, +\infty)$ of class C^∞, slowly varying at 0 [5] and bounded away from 0.

For a given time horizon t we fix a time step $\Delta = t/N$. Since there is no ambiguity we write n instead of $n\Delta$, so that for every $n \leq N$ we get

$$\begin{aligned} v_n = v_{n\Delta} &= \mathbb{E}(v_{n\Delta}) + \sum_{j=1}^{n} \int_{(j-1)\Delta}^{j\Delta} K(n\Delta - s)\sqrt{v_s}dB_s \\ &= \xi(n\Delta) + \sum_{j=1}^{n} \int_{(j-1)\Delta}^{j\Delta} K(n\Delta - s)\sqrt{v_s}dB_s \\ &= \hat{\xi}_n + u_n \end{aligned}$$

where

$$\hat{\xi}_n = \mathbb{E}(v_n|\mathcal{F}_{n-1}) = \xi_n + \sum_{j=1}^{n-1} \int_{(j-1)\Delta}^{j\Delta} K(n\Delta - s)\sqrt{v_s}dB_s \quad (3.15)$$

is an $\mathcal{F}_{n-1}$ adapted variable, and

$$u_n = \int_{(n-1)\Delta}^{n\Delta} K(n\Delta - s)\sqrt{v_s}dB_s \quad (3.16)$$

is the martingale part. In addition to this consider the n-th increment of $X_t = \log(S_t)$ parallel to the volatility process and define it as

$$\chi_n = \int_{(n-1)\Delta}^{n\Delta} \sqrt{v_s}dB_s \quad (3.17)$$

Thus our goal is to find a way to simulate the random variables $\hat{\xi}_n,\ u_n,\ \chi_n$ in order to be able to reconstruct the value of v_n and S_n. Moreover we observe that at time t_{n-1} the value of $\hat{\xi}_n$ is known since it is $\mathcal{F}_{n-1}$ measurable. Thus when we are at time t_{n-1} and we want to perform the simulation for time t_n we need to take into account the fact that $\hat{\xi}_{n+1},\ u_n,\ \chi_n$ are correlated, and then we need to compute the correlation matrix. For this purpose it is useful to introduce the following quantities:

[5] A function $F : (0, +\infty) \to (0, +\infty)$ is said to be slowly varying at 0 if for any positive constant $c > 0$ it holds $\lim_{t\to 0^+} \frac{F(ct)}{F(t)} = 1$.

Definition 3.2.1. *For every $p, q \geq 0$ define*

$$\zeta_p(\Delta) = \int_0^\Delta K(s + p\Delta)ds, \quad \zeta_{p,q}(\Delta) = \int_0^\Delta K(s + p\Delta)K(s + q\Delta)ds$$

Using the Definition 3.2.1, we are able to prove the following result that is crucial for our goal of computing the correlation matrix.

Lemma 3.2.1. *Let*

$$\zeta_{0,0}(\Delta) = \int_0^\Delta K^2(s)ds$$

and assume that $K(t) = t^{H-1/2}L_k(t)$, $t \in (0, \Delta]$ with $H \in \left(0, \frac{1}{2}\right)$ and $L_k : (0, \Delta] \to [0, +\infty)$ of class C^∞, slowly varying at 0 and bounded away from 0.
Then assuming that the simulated forward variance curve $\mathbb{E}(v_s|\mathcal{F}_{n-1})$ is twice differentiable

$$\mathrm{var}(u_n|\mathcal{F}_{n-1}) = \frac{\zeta_{0,0}(\Delta)}{2H + 1}(\hat{\xi}_n + 2Hv_{n-1}) + O(\Delta^{2+2H})$$

Proof. Using the definition of u_n and applying Itô's isometry we get

$$\mathrm{var}(u_n|\mathcal{F}_{n-1}) = \int_{(n-1)\Delta}^{n\Delta} K^2(n\Delta - s)\mathbb{E}(v_s|\mathcal{F}_{n-1})ds$$

Perform a change of variable $r = s - (n-1)\Delta$ so that

$$\begin{aligned}\mathbb{E}(v_s|\mathcal{F}_{n-1}) &= v_{n-1} + \left(\frac{\mathbb{E}(v_n|\mathcal{F}_{n-1}) - v_{n-1}}{\Delta}\right)r + O(r^2) \\ &= v_{n-1} + \left(\frac{\hat{\xi}_n - v_{n-1}}{\Delta}\right)r + O(r^2)\end{aligned}$$

Moreover it holds

$$\int_{(n-1)\Delta}^{n\Delta} r\ K^2(n\Delta - s)ds = \int_0^\Delta r\ K^2(\Delta - r)dr = \frac{\zeta_{0,0}(\Delta)}{2H + 1}\Delta + O(\Delta^{2+2H})$$

Then, using $\beta = \dfrac{1}{2H + 1}$, we can directly compute the error

$$\begin{aligned}&\mathrm{var}(u_n|\mathcal{F}_{n-1}) - \zeta_{0,0}(\Delta)(\beta\hat{\xi}_n + (1-\beta)v_{n-1}) \\ &= \int_{(n-1)\Delta}^{n\Delta}\left(v_{n-1} + \frac{\hat{\xi}_n - v_{n-1}}{\Delta}r - \beta\hat{\xi}_n - (1-\beta)v_{n-1} + O(r)\right)K^2(n\Delta - r)dr \\ &= \frac{\hat{\xi}_n - v_{n-1}}{\Delta}\int_0^\Delta (r - \beta\Delta)K^2(\Delta - r)dr + O(\Delta^{2+2H}) \\ &= O(\Delta^{2+2H})\end{aligned}$$

□

Remark 12. From now on $\bar{v}_n = \dfrac{1}{2H+1}(\hat{\xi}_n + 2Hv_{n-1})$

Thanks to Lemma 3.2.1 and making the piecewise constant approximation $v_s \approx \bar{v}_n,\ s \in ((n-1)\Delta, n\Delta]$ we can compute the covariances:

$$\begin{aligned}
\mathrm{var}(u_n|\mathcal{F}_{n-1}) &\approx \bar{v}_n\zeta_{0,0}(\Delta)\\
\mathrm{var}(\hat{\xi}_{n+1}|\mathcal{F}_{n-1}) &\approx \bar{v}_n\zeta_{1,1}(\Delta)\\
\mathrm{var}(\chi_n|\mathcal{F}_{n-1}) &\approx \bar{v}_n\Delta\\
\mathrm{cov}(u_n, \hat{\xi}_{n+1}|\mathcal{F}_{n-1}) &\approx \bar{v}_n\zeta_{0,1}(\Delta)\\
\mathrm{cov}(u_n, \chi_n|\mathcal{F}_{n-1}) &\approx \bar{v}_n\zeta_0(\Delta)\\
\mathrm{cov}(\chi_n, \hat{\xi}_{n+1}|\mathcal{F}_{n-1}) &\approx \bar{v}_n\zeta_1(\Delta)
\end{aligned} \tag{3.18}$$

and thus plugging them into the correlation matrix we get

$$R = \begin{pmatrix} 1 & \rho_{u,\chi} & \rho_{u,\xi} \\ \rho_{u,\chi} & 1 & \rho_{\chi,\xi} \\ \rho_{u,\xi} & \rho_{\chi,\xi} & 1 \end{pmatrix} \tag{3.19}$$

where all the entries are independent of the time step n and take values

$$\rho_{u,\chi} = \frac{\zeta_0(\Delta)}{\sqrt{\Delta\zeta_{0,0}(\Delta)}},\ \rho_{u,\xi} = \frac{\zeta_{0,1}(\Delta)}{\sqrt{\zeta_{1,1}(\Delta)\zeta_{0,0}(\Delta)}},\ \rho_{\chi,\xi} = \frac{\zeta_1(\Delta)}{\sqrt{\Delta\zeta_{1,1}(\Delta)}}$$

Now we want to use a specific kernel in order to match the dynamics described in Equation (2.2). To do this we introduce two kernels.

Definition 3.2.2. *The power-law kernel is defined by*

$$k(t) = \frac{\sigma}{\Gamma(H+1/2)}t^{H-\frac{1}{2}} = c\ t^{H-\frac{1}{2}}.$$

The Mittag-Leffler kernel is given by

$$K(t) = \sigma t^{H-\frac{1}{2}}\sum_{n=0}^{\infty}\frac{(-\lambda t^{H-\frac{1}{2}})^n}{\Gamma((n+1)(H+1/2))}.$$

where $H \in \left(0, \dfrac{1}{2}\right)$ *is the Hurst index.*

Remark 13. As proved in [27], in the rough Heston model the kernel which appears in Equation (3.14) is a Mittag-Leffler kernel. If $\lambda = 0$ then it reduces to a power-law kernel.

Now we need to compute the quantities $\zeta_{i,j}(\Delta)$ for different values of $i, j \geq 0$. As we will see in the next Lemma this is possible only if we consider a power-law kernel and only for some values of i, j.

Lemma 3.2.2. *For the power-law kernel*

$$\zeta_i(\Delta) = \frac{c\Delta^{H+\frac{1}{2}}}{H+\frac{1}{2}}\left((i+1)^{H+\frac{1}{2}} - i^{H+\frac{1}{2}}\right)$$

and

$$\zeta_{i,i}(\Delta) = \frac{c^2}{2H}\Delta^{2H}\big((i+1)^{2H} - i^{2H}\big).$$

The proof consists in making a direct computation of the integrals. Unfortunately when we use Mittag-Leffler kernel it is not possible to derive a closed form formula due to the fact that we are integrating a series. Thus it is necessary to perform a numerical integration as pointed out in the thesis *Monte Carlo simulation schemes for rough Heston model* of Scaramelli.
Even for the power-law kernel we still have the problem of the computation of $\zeta_{i,j}(\Delta)$ if $i \neq j$. However in [28] Gatheral says that numerical experiments suggest that for $i \geq 0$ and $j \geq 1$ it is possible to compute $\zeta_{i,j}(\Delta)$ using the approximation

$$\zeta_{i,j}(\Delta)\Delta \approx \zeta_i(\Delta)\zeta_j(\Delta). \tag{3.20}$$

Thus, using (3.20), we find out the approximated correlation matrix

$$\tilde{R} = \begin{pmatrix} 1 & \tilde{\rho} & \tilde{\rho} \\ \tilde{\rho} & 1 & 1 \\ \tilde{\rho} & 1 & 1 \end{pmatrix} \tag{3.21}$$

where $\tilde{\rho} \approx \dfrac{\zeta_0(\Delta)}{\sqrt{\zeta_{0,0}(\Delta)\Delta}} = \rho_{u,\chi}$.

Remark 14. The approximation given by Equation (3.20) was derived for the power-law kernel. Nevertheless in the thesis *Monte Carlo simulation schemes for rough Heston model* of Scaramelli they say that this approximation is still true also for the Mittag-Leffler kernel and so we do the same.

Using the approximation $\tilde{R}$ we observe that χ_n and $\hat{\xi}_{n+1}$ are perfectly correlated. Then it comes out that in practice, it is sufficient to simulate two random variables consistent with the correlation matrix $\tilde{R}$. To do this we use a bivariate version of the Quadratic Exponential scheme, following what is made in [28].
Consider the random variables u_n, χ_n defined by (3.16) and (3.17). Using the linear regression we get

$$u_n \approx \beta_{u\chi}\chi_n + \epsilon_n$$

where $\beta_{u\chi} = \dfrac{\zeta_0(\Delta)}{\Delta}$ and $\operatorname{cov}(\chi_n, \epsilon_n) = 0$. Moreover, the variance has to be positive, then we have to be ensure that

$$v_n = u_n + \hat{\xi}_n = \beta_{u\chi}\chi_n + \epsilon_n + \hat{\xi}_n \geq 0.$$

Thus we give the following lemma which describes the procedure that have to be applied.

Lemma 3.2.3. *Let* $\beta_{u\chi}\hat{\chi}_n = \beta_{u\chi}\chi_n + \frac{1}{2}\hat{\xi}_n$ *and* $\hat{\epsilon}_n = \epsilon_n + \frac{1}{2}\hat{\xi}_n$ *be generated independently using the QE scheme with the following condition:*

$$\mathbb{E}(\beta_{u\chi}\hat{\chi}_n|\mathcal{F}_{n-1}) = \frac{1}{2}\hat{\xi}_n, \quad \mathbb{E}(\hat{\epsilon}_n|\mathcal{F}_{n-1}) = \frac{1}{2}\hat{\xi}_n$$

$$\text{var}(\hat{\chi}_n|\mathcal{F}_{n-1}) = \bar{v}_n\Delta, \quad \text{var}(\hat{\epsilon}_n|\mathcal{F}_{n-1}) = \bar{v}_n\big(\zeta_{0,0}(\Delta) - \frac{1}{\Delta}\zeta_0(\Delta)\big)$$

Then $v_n = \beta_{u\chi}\chi_n + \epsilon_n + \hat{\xi}_n \geq 0$ *and* $u_n = \beta_{u\chi}\chi_n + \epsilon_n$ *is such that*

$$\text{var}(u_n|\mathcal{F}_{n-1}) = \bar{v}_n\zeta_{0,0}(\Delta), \quad \text{cov}(u_n, \chi_n|\mathcal{F}_{n-1}) = \bar{v}_n\zeta_0(\Delta)$$

The proof can be found in [28]. Using Lemma 3.2.3 we give the following steps to describe the Hybrid Quadratic Exponential scheme:

1. Given χ_k for $k < n$, fix $0 < \delta << 1$ and compute

$$\hat{\xi}_n = \max\left\{\delta, \xi_n + \sum_{k=1}^{n-1}\sqrt{\frac{\zeta_{n-k+1,n-k+1}(\Delta)}{\Delta}}\,\chi_k\right\}.$$

2. Compute the ratios

$$\psi_n^{\beta_{u\chi}\hat{\chi}_n} = \frac{4\text{var}[\beta_{u\chi}\hat{\chi}_n|\mathcal{F}_{n-1}]}{\hat{\xi}_n^2} = \frac{4\zeta_0(\Delta)^2\bar{v}_n}{\Delta\hat{\xi}_n^2}$$

and

$$\psi_n^{\hat{\epsilon}_n} = \frac{4\text{var}[\hat{\epsilon}_n|\mathcal{F}_{n-1}]}{\hat{\xi}_n^2} = \frac{4\bar{v}_n}{\Delta\hat{\xi}_n^2}\left(\zeta_{0,0}(\Delta) - \frac{\zeta_0(\Delta)^2}{\Delta}\right).$$

3. Simulate $\beta_{u\chi}\hat{\chi}_n$ and $\hat{\epsilon}_n$ using the bivariate QE scheme based on the ratios at the previous point. Then using Lemma 3.2.3 compute

$$\chi_n = \frac{(\beta_{u\chi}\hat{\chi}_n - \hat{\xi}_n/2)}{\beta_{u\chi}}$$

$$\epsilon_n = \hat{\epsilon}_n - \hat{\xi}_n/2.$$

4. Compute $v_n = u_n + \hat{\xi}_n = \beta_{u\chi}\chi_n + \epsilon_n + \hat{\xi}_n = \frac{\zeta_0(\Delta)}{\Delta}\hat{\chi}_n + \hat{\epsilon}_n$.

5. $X_n = X_{n-1} - \frac{1}{4}(v_n + v_{n-1})\Delta + \sqrt{1-\rho^2}\sqrt{\frac{1}{2}(v_n + v_{n-1})\Delta}Z_n^{\perp} + \rho\chi_n$ where $Z_n^{\perp}$ is an independent standard Gaussian.

6. Compute S_n using the fact that $X_n = \log(S_n)$ and so $S_n = e^{X_n}$.

Remark 15. Looking at the HQE scheme we see that u_n and χ_n are generated at each step. In addition these two random variables have zero expected value so we get

$$\hat{\xi}_{n+1} = \mathbb{E}(v_{n+1}|\mathcal{F}_n) = \xi_{n+1} + \sum_{k=1}^{n} \sqrt{\frac{\zeta_{n-k+2,n-k+2}(\Delta)}{\Delta}} \chi_k.$$

This is an evidence of the non-Markovianity of the two processes ξ and v since to compute $\hat{\xi}_{n+1}$ (which is $\mathcal{F}_n$ measurable) we need the entire history of χ_k, $k \leq n$.

Chapter 4

Numerical tests

In this chapter we want to study some numerical properties of the two simulation schemes presented in chapter 3. In order to do this, we fix the same parameters presented in [1]:

$$\lambda = 0.3,\ \sigma = 0.3,\ \theta = 0.02,\ V_0 = 0.02,\ \rho = -0.7,\ S_0 = 1.$$

In addition to this we fix also the Hurst index H to 0.1.
The chapter is organized in the following way: in section 4.1 we compare the computational costs of the two schemes, then we focus our attention on European Call options: firstly we analyze the smile for a single maturity in section 4.2, and subsequently the whole volatility surface in section 4.3. At a later time, in section 4.4 we will focus on the skew both for standard European both for Forward Start Call options. Finally we give an estimate of the Hurst index produced by the two schemes in section 4.5.

4.1 Computational costs

First of all, we want to investigate on the computational cost of the two simulation schemes. In [8] the authors point out that the cost of the Hybrid Quadratic Exponential scheme is $O(M^2)$ while for the Moment Matching simulation scheme it is $O(N^2M)$, where N is the number of factors in the approximating kernel given by Equation (3.2) and M is the number of time steps. This implies that since usually we use $N = 2$ or $N = 3$ the computational cost for the Moment Matching scheme is approximately linear.
In order to verify it with an example we fix $N = 2$, a maturity $T = 1$ year and 11 linearly spaced values in $[-0.5, 0.5]$ for the log-moneyness. [1] Then we compute the smile with the two methods performing 1 million simulations and choosing as discretization steps $M \in \{1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024\}$.

[1] For a Call option written on an underlying with current stock price S_0 and strike K we define its log-moneyness as $\log(\frac{K}{S_0})$.

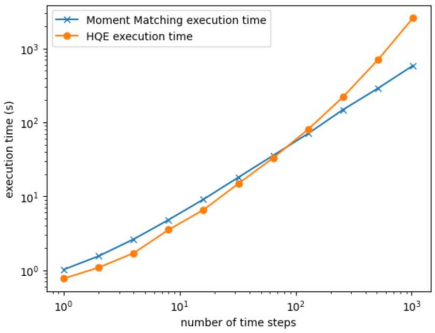


Figure 4.1: Computational costs (in seconds) of HQE and Moment Matching scheme for the computation of the smile at time $T = 1$, considering 11 linearly spaced values in $[-0.5, 0.5]$ for the log-moneyness, using different time steps and performing 1 million samples.

The implied volatility is computed with the bisection method once we have computed the price with the Monte Carlo simulation. As we can observe in Figure 4.1 the cost for the HQE scheme seems to be similar to the one of the Moment Matching scheme for $M \leq 128$, while for higher number of time steps it is much expensive. The rate of growth of the computational cost of the HQE scheme seems to be greater than $O(M^2)$: indeed, in that case, the graph should be a straight line with slope 2.

4.2 European Call option's smile

Now we consider Plain Vanilla Call options with different strikes and we investigate on the quality of the smile approximation, trying to understand what is the optimal choice for the number of time steps M and the optimal number of factors N that must be considered in the Moment Matching scheme. Due to the fact that the characteristic function of the logarithm of the stock price is known in semi-closed form both in the rough Heston model and for its Markovian approximation, we take as reference point the smile given by the the Fourier inversion technique in the RHM and look at the maximal relative error over all the strikes. [2]
Thus, we fix the maturity $T = 1$ year and 11 linearly spaced values of the log-moneyness in $[-0.5, 0.5]$. Of course for the Moment Matching scheme it is necessary to fix also the nodes and the weights: we did it according to what is presented in [10] and the values are reported in Table 4.1.
The value of the relative errors are reported in Table 4.2 and looking at them

[2] To compute the characteristic function in the rough Hestom model it has been used Adam's scheme (see e.g. [22]) while for the characteristic function of the Markovian lift it has been used a predictor corrector scheme (see e.g. [2], [9]).

	N=2		N=3		
Nodes	0.05	8.7171	0.03333	2.2416	46.831
Weights	0.7673	3.2294	0.5554	1.1111	6.0858

Table 4.1: Nodes and weights computed according to [10] for $T = 1,\ H = 0.1$, and $N = 2, 3$.

we can observe three important features:

- The errors, when $N = 3$, are bigger than for $N = 2$ for a number of time steps $M \leq 128$, this is related to the fact that the maximal mean reversion is greater for $N = 3$ and so we need a thinner discretization step. Indeed, for greater values of M, the errors are under 1%.

- There is a clear trade-off between a more accurate Markovian approximation and a less accurate simulation. In other words if we fix a number of time steps M and we increase N, we know by the theory that the Markovian approximation is closer to the original model but due to the higher mean reversion we loose precision in the simulation. In particular for a fixed N we have to fix a number of steps $M > x_N$, where x_N is the maximal mean reversion, to be able to appreciate the contribution of this mean reversion. [3]

- The error of the HQE scheme decreases more slowly than the one of the Moment Matching. This is a problem because to achieve a greater precision, we will need a higher number of time steps and this increases noticeably the computational time as we saw in Figure 4.1.

To conclude this section we report an example of smiles in Figure 4.2 together with the corresponding (relative) errors in Figure 4.3. As we can observe the Moment Matching smile and the Markovian smile are perfectly covered by the rough one. On the other hand we can see that there is a little difference with the HQE smile for strikes out of the money, indeed looking at Table 4.2 we see that the Moment Matching error is of order 0.5% while the HQE one is above 1%.

[3] It is well known that a process of the form $dY_t = k(\theta - Y_t)dt + \sigma\sqrt{Y_t}dB_t$ has a mean reversion time of $1/k$ so to appreciate its mean reverting behaviour we expect that a thinner discretization step is needed.

M	HQE	Mom. matching N=2	Mom. matching N=3
1	0.3002	0.2532	0.3851
2	0.2959	0.1609	0.2620
4	0.1912	0.0696	0.1651
8	0.1048	0.0193	0.0884
16	0.0429	0.0069	0.0326
32	0.0185	0.0059	0.0122
64	0.0175	0.0115	0.0223
128	0.0097	0.0081	0.0092
256	0.0286	0.0143	0.0085
512	0.0089	0.0218	0.0033
1024	0.0117	0.0297	0.0082

Table 4.2: Maximal relative errors of the implied volatility smile with 1 million paths. The Monte Carlo error is always below 0.03%. The error of the Markovian approximation is 0.0025 for $N = 2$ and 0.0006 for $N = 3$.

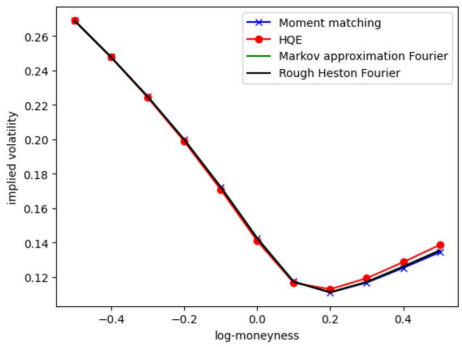


Figure 4.2: Implied volatility smiles using 1 million paths and $M = 32$ time steps and nodes and weights reported in Table 4.1 for $N = 2$. The rough Heston and the Markovian smiles are obtained using the Fourier inversion technique.

4.3 Implied volatility surface

Now we are going to study the whole implied volatility surface. This means that we have to consider different maturities and also different strikes. Thus we fix:

$$T^i = \frac{i}{16}, \quad K^i = e^{x\sqrt{T^i}} \quad i = 1, ..., 16, \quad x = -0.1, -0.09, ..., 0.1 \tag{4.1}$$

We rescale the log-strikes according to $\sqrt{T}$ because we want to consider only liquid options. [4]

[4] An option is said to be liquid if we can buy or sell it quickly and without modifying its market value. The intuition is the following: if you fix a strike K and you have two

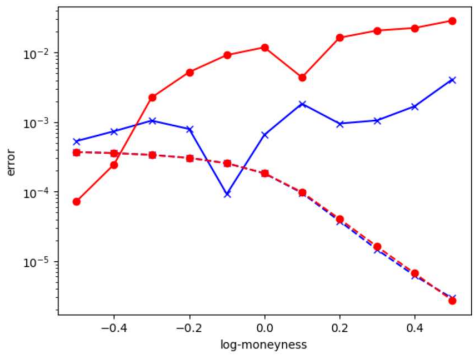


Figure 4.3: The relative errors with respect to the rough smile for the Moment Matching scheme (blue line) and the HQE scheme (red line) using $M = 32$ time steps and nodes and weights reported in Table 4.1 for $N = 2$. The dashed lines represent the Monte Carlo errors.

Since now we are working on an entire surface we could proceed changing the nodes and the weights for every maturity and performing different Monte Carlo simulations for options with different maturities in order to replicate perfectly the smile. However, this is not a good choice because if we need to simulate an entire portfolio with many different assets, we need to fix the nodes and the weights once for all, in order to perform a Monte Carlo simulation that can be used for pricing all the options in the portfolio. Thus, we need to choose the interval $[0, T_0]$ that will be used to determine the nodes and the weights. The authors of [10] tell us that the best choice supported by empirical evidence is to choose

$$T_0 = T_{min}^{\alpha(N)} T_{max}^{1-\alpha(N)}$$

where T_{min} and T_{max} are respectively the first and the last maturity that we consider and the value of the exponent are given by:

$$\alpha(1) = \frac{3}{5},\ \alpha(2) = \frac{1}{2},\ \alpha(3) = \frac{1}{3}$$

$$\alpha(4) = \frac{1}{4},\ \alpha(5) = \frac{1}{6},\ \alpha(6) = \frac{1}{10}$$

$$\alpha(N) = 0, N \geq 7$$

This choice produces the nodes and the weights reported in Table 4.3.
The relative errors and the computational times are reported in Table 4.4 and Table 4.5. In this case, we can observe that to achieve a relative error

maturities $T_1 < T_2$ then the option with the longer maturity has much time to reach the level in/out of the money, thus to compare options with different maturities we rescale their strikes.

	N=2		N=3		
Nodes	0.2	34.8683	0.084	5.6485	118.0062
Weights	1.336	5.6628	0.8039	1.6079	8.8078

Table 4.3: Nodes and weights computed according to [10] for $T = \frac{i}{16},\ i = 1, ..., 16,\ H = 0.1$ and $N = 2, 3$.

M	HQE	Mom. matching N=2	Mom. matching N=3
16	0.0865	0.1046	0.1659
32	0.0787	0.0458	0.0854
64	0.0544	0.0225	0.0407
128	0.0309	0.0171	0.0161
256	0.0155	0.0183	0.0054
512	0.0097	0.0184	0.0026

Table 4.4: Maximal relative errors of the implied volatility surface with 1 million paths. The error of the Markovian approximation is 0.0177 for $N = 2$ and 0.0018 for $N = 3$. The Monte Carlo error is always under 0.03%.

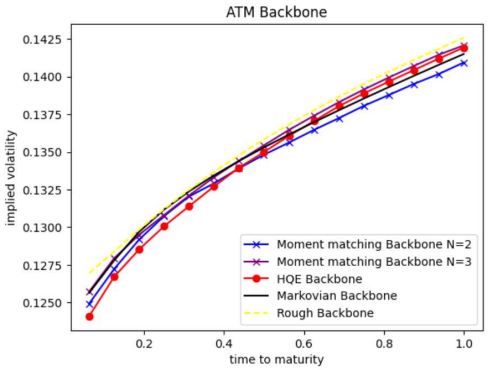
M	HQE	Mom. matching N=2	Mom. matching N=3
16	27.38	19.82	22.44
32	46.01	29.19	29.38
64	93.02	45.68	47.82
128	222.21	80.74	88.52
256	613.13	151.53	160.42
512	1666.54	295.95	310.35

Table 4.5: Computational time (in seconds) for the implied volatility surface with 1 million samples. The computational times using the Fourier inversion technique are respectively 182.96 for the rough Heston model, 78.48, 83.26 for the Markovian approximation with $N = 2, 3$.

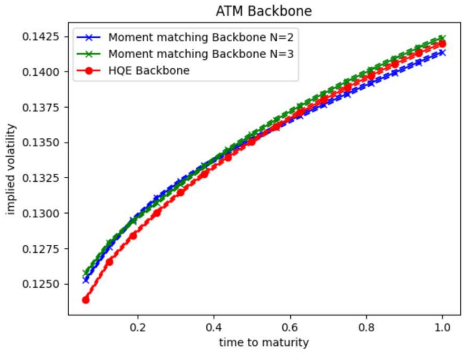
under 1% with the Moment Matching scheme fixing $N = 2$ we need more than 512 time steps so it is better to choose $N = 3$ which needs more or less 200 time steps to achieve an error of 1% while the HQE scheme needs 500 time steps.[5] On the other hand, the data reported in Table 4.5 suggest that the best possible choice is the Moment Matching scheme with $N = 3$ due to the explosion of the computational time of the HQE scheme. Moreover, as we can observe, the computational time of the volatility surface of the rough Heston model is comparable to the one of the Monte Carlo simulation, even if we have to compute numerically the characteristic function.

On the left side in Figure 4.4 it is possible to observe the at the money

[5]The estimate of the number of steps is made using linear regression.



(a) ATM backbone using 1 million samples, $M = 128$ time steps, nodes and weights reported in Table 4.3. The Markovian and the rough backbones are obtained using the Fourier inversion technique.



(b) ATM backbone for HQE scheme and Moment Matching scheme performing 10 million samples, $N = 2, 3$ and $M = 128$. The dashed lines represent the 99% confidence interval for each backbone.

Figure 4.4

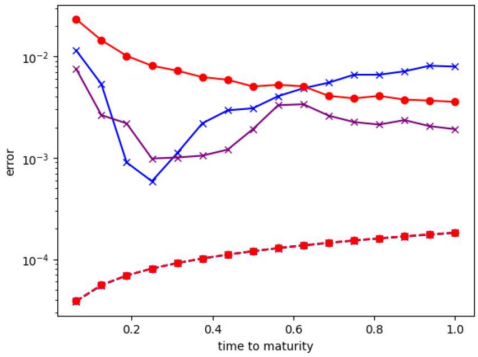
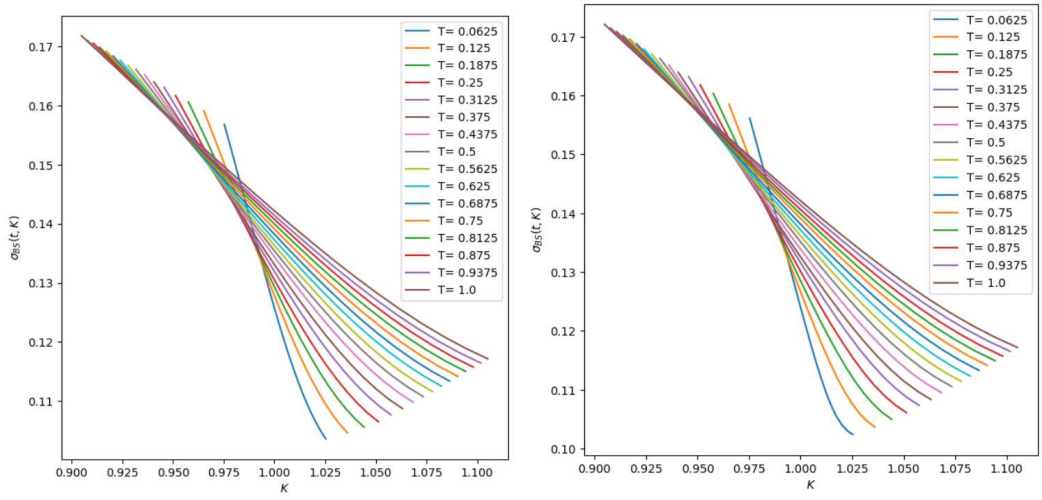


Figure 4.5: The relative errors with respect to the rough backbone for the Moment matching scheme with $N = 2$ (blue line) and $N = 3$ (purple line) as well as for the HQE scheme (red line) using $M = 128$ time steps. The dashed lines represent the Monte Carlo errors.

backbones [6] computed with the the Fourier inversion technique as well as with the simulation schemes with $M = 128$ and performing 1 million samples. As we can observe, the general behaviour is the same, but the backbone generated by the Moment Matching scheme with $N = 3$ is very close to our reference curve generated by the Fourier inversion technique: this confirms the fact that increasing N the lifted model will approximate the rough counterpart. On the right side we repeat the experiment keeping fixed the

[6]The backbone is obtained looking at the implied volatility of options with the same strike price K but different maturities.

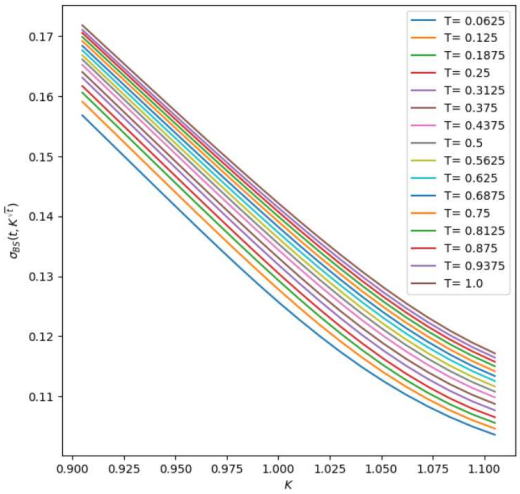
discretization step but increasing the number of samples to 10 million and adding a confidence interval of 99%.
The relative errors, as well as the Monte Carlo errors can be seen in Figure 4.5: as we have already observed, in this case, the best possible choice is the Moment Matching scheme using $N = 3$ factors, since it gives a smaller error and it is computationally cheaper.
To conclude this section, we also add Figure 4.6 which represents the value of the implied volatility surface for the maturities reported in (4.1): as we can observe, the slope of the surfaces decreases as soon as the maturity increases and this is coherent with the theory. Moreover, due to our choice of rescaling the strikes, we can observe that for shorter maturities, the strikes are closer to the at the money level . In Figure 4.7, we plot the same slices of the implied volatility surfaces of Figure 4.6 but in a different system of coordinates given by the transformation $(t, K) \to (t, K^{\sqrt{t}})$.
As we can observe, in this system of coordinates the smiles present very similar behaviour. In particular they seem to have the same slope: this observation confirms what we said above on the fact that to have liquid options we need to rescale the log-strikes according to $\sqrt{t}$.



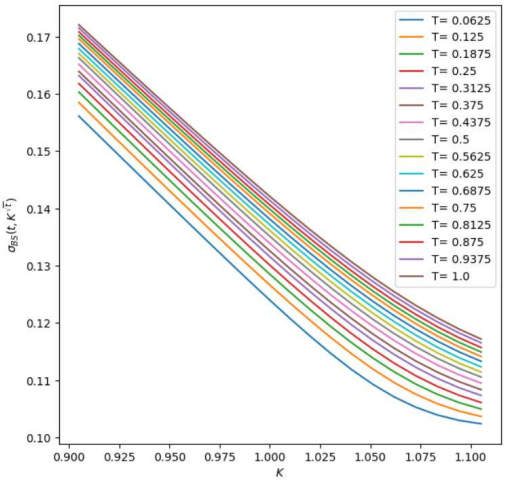
(a) Slices of the implied volatility surface generated by Moment Matching scheme performing 1 million samples, using $M = 128$ time steps, $N = 3$.

(b) Slices of the implied volatility surface generated by HQE scheme performing 1 million samples, using $M = 128$ time steps.

Figure 4.6



(a) Slices of the implied volatility surface generated by Moment Matching scheme performing 1 million samples, using $M = 128$ time steps, $N = 3$.



(b) Slices of the implied volatility surface generated by HQE scheme performing 1 million samples, using $M = 128$ time steps.

Figure 4.7

4.4 The implied volatility Skew

Given a pricing model where the stock price is described by a stochastic process $(S_t)_{t\geq 0}$ and a European Call with strike K and maturity T we know that the implied volatility is a map

$$\sigma_{BS} : (T, K) \to \sigma_{BS}(T, K)$$

defined such that

$$C(T, K) = C_{BS}(T, K, \sigma_{BS}(T, K))$$

where $C_{BS}(\cdot, \sigma_{BS}(T, K))$ denotes the price in the Black-Scholes model with volatility $\sigma_{BS}(T, K)$. The same relation can be given considering the relative strike $k = K/S_0$ instead of K using the trivial fact that

$$(S_T - K)^+ = (S_T - k\ S_0)^+.$$

Then we define the at the money implied volatility skew at time T as

$$\Psi(T) = \frac{\partial}{\partial k}\sigma_{BS}(T, k)\bigg|_{k=1} \tag{4.2}$$

Remark 16. In the literature it is common to find the definition

$$\Psi(T) = \frac{\partial}{\partial k}\sigma_{BS}(T, k)\bigg|_{k=0}, \quad k = \log(K/S_0).$$

These two definitions are perfectly compatible once you decide to consider the implied volatility as a function of the relative strike or of the relative log-strike.

Our goal now is to study the skew that our simulation schemes produce. It is well known that the at the money skew of implied volatility explodes for short maturities as we can see for instance from Figure 4.8 taken from [1]. The incapability of the classic Heston model (as well as the classic stochastic volatility models) to reproduce this behaviour is one of the main motivations that brings to the introduction of the fractional volatility models [7]. Thus, we want to investigate if our simulation schemes are able to match this behaviour especially in the limit as $t \to 0^+$.

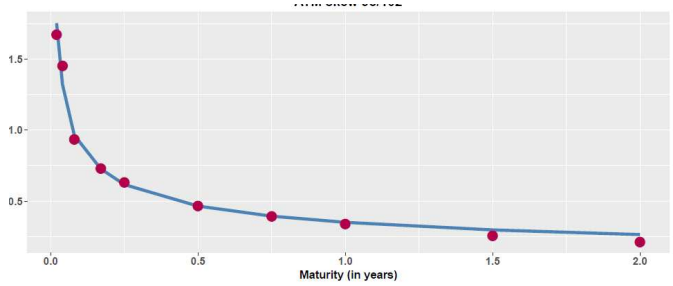


Figure 4.8: Term structure of the at the money skew of the implied volatility for the S&P index on June 20, 2018 (red dots) and the exponential fit $f(t) = 0.35\ t^{-0.41}$.

4.4.1 European Call

The first step is to investigate the behaviour of the at the money skew for European Call options. Of course we have to compute it numerically and so we are going to follow these steps:

- Fix 25 linearly spaced maturities $T_i = \dfrac{i}{25}, \quad i = 1, ..., 25.$

- Fix 4 relative strike prices

$$k_1 = 0.999,\ k_2 = 0.9995,\ k_3 = 1.0005,\ k_4 = 1.001$$

close to the at the money level. As we said before these strikes will be rescaled according to $k^*_{j,i} = k_j^{\sqrt{T_i}}, \quad j = 1, ..., 4, \quad i = 1, ..., 25.$

[7] For more details we refer to Chapter 7 of [4] and the pioneering paper [31].

- Compute the implied volatility surface and then for each maturity T_i compute the incremental ratio

$$\Psi_1(T_i) = \frac{|\sigma_{BS}(T_i, k^*_{3,i}) - \sigma_{BS}(T_i, k^*_{2,i})|}{k^*_{3,i} - k^*_{2,i}}$$

and

$$\Psi_2(T_i) = \frac{|\sigma_{BS}(T_i, k^*_{4,i}) - \sigma_{BS}(T_i, k^*_{1,i})|}{k^*_{4,i} - k^*_{1,i}}$$

- Fix a tolerance $e = 10^{-3}$ so that if

$$\frac{|\Psi_1(T_i) - \Psi_2(T_i)|}{\Psi_1(T_i)} < e$$

we accept $\Psi_1(T_i)$ as skew otherwise we tighten our strikes and repeat the procedure.

The methodology described above applies both to simulation schemes and close form solutions with the only difference that the prices that we need to compute the implied volatility are obtained performing a Monte Carlo simulation or using the Fourier inversion technique.

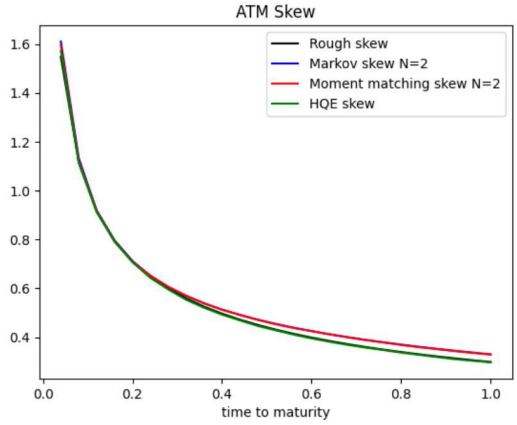
	N=2		N=3		
Nodes	0.25	43.5854	0.09746	6.5545	136.9341
Weights	1.4607	6.1477	0.8531	1.7064	9.3478

Table 4.6: Nodes and weights for the maturities $T = 0.04, ..., 1,\ H = 0.1,$ and $N = 2, 3$.

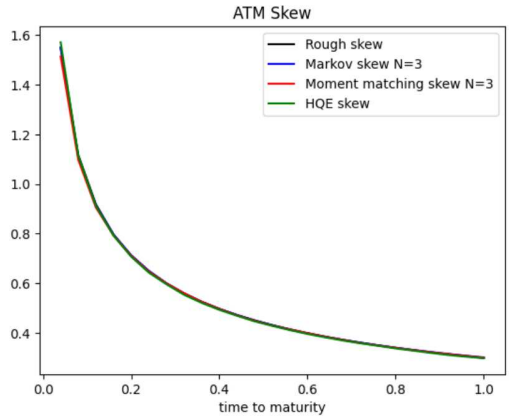
	N=2		N=3		
Nodes	0.7906	137.8292	0.21	14.1212	295.0156
Weights	2.3151	9.7435	1.1597	2.3197	12.7069

Table 4.7: Nodes and weights for the maturities $T = 0.004, ..., 1,\ H = 0.1,$ and $N = 2, 3$.

In Figure 4.9 we report the graphs of the skews: on the left hand side, we can observe that choosing $N = 2$ we have not a good fit for longer maturity while this problem disappears fixing $N = 3$ as we can see from the graph on the right. In Figure 4.10 we plot the relative errors with respect to our benchmark which is the rough skew: as we can see, choosing $N = 2$ we reach an error of 10% while the maximal error is around 2% choosing $N = 3$ or using the HQE scheme. In addition, the choice of the Moment Matching scheme with 3 approximating factors seems to behave in a better way due to the fact that generally the error is below 1%. Moreover, we can see that



(a) ATM skew performing 1 million samples for maturities $T = 0.04, 0.08, ..., 1$, $H = 0.1$, $M = 250$ time steps and $N = 2$ approximating factors for the Markovian and the Moment Matching skew. The rough and the Markovian skew are obtained by the Fourier inversion technique.



(b) ATM skew performing 1 million samples for maturities $T = 0.04, 0.08, ..., 1$, $H = 0.1$, $M = 250$ time steps and $N = 2$ approximating factors for the Markovian and the Moment Matching skew. The rough and the Markovian skew are obtained by the Fourier inversion technique.

Figure 4.9

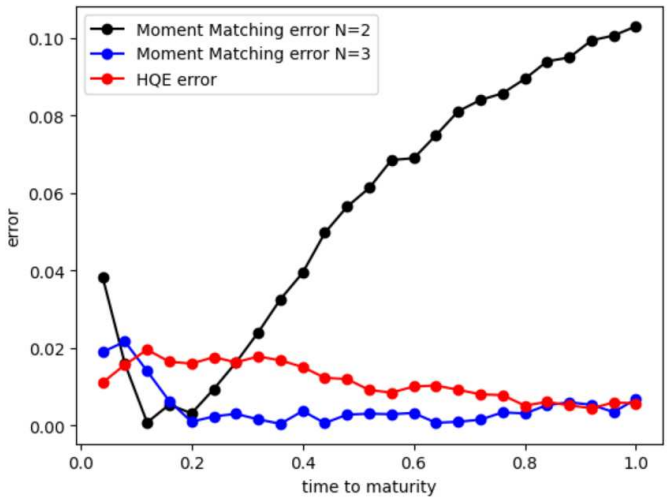


Figure 4.10: Relative errors of ATM skew with maturities $T = 0.04, 0.08, ..., 1$, $H = 0.1$, $M = 250$ time steps for the HQE scheme and Moment Matching scheme with $N = 2, 3$ approximating factors.

the error for the first maturity is greater for the Moment Matching scheme with $N = 2$: this is related to the fact that the maximal mean reversion for $N = 2$ that we can observe in Table 4.6 is too small to mimic the explosion

for short maturities. For completeness, we report also the maximal errors for the Markovian approximation which are respectively 10.5% and 0.8% for $N = 2, 3$. This is coherent with the behaviour that we see in Figure 4.9: in effect the graph of the Markovian approximation seems to be very close to the Moment Matching one generated with the same number of approximating factors.

Now, we want to study the behaviour also for shorter maturities, therefore we apply the same methodology that we used before but with maturities $T_i = \frac{i}{250}, \quad i = 1, ..., 250$. This means that we consider options with daily maturities starting from tomorrow until 1 year with the usual convention of 250 days per year. Unfortunately, it is not possible to compute the skew for all these maturities with Fourier technique in an acceptable time, so we fix 50 linearly spaced maturities starting from $T_1 = 0.004$ till $T_{50} = 0.984$ in order to match the maturities that we consider in the Monte Carlo simulation. The results are plotted in Figure 4.11 and, looking at these graphs, we can clearly see that the best choice to fit the rough skew is to fix $N = 3$ once again. Indeed, using $N = 2$ we have some problems for longer maturity (as before) but also in the short time limit. Thus, even if we have not computed the errors, we can conclude that the best choice to fit the rough skew using the Moment Matching scheme is to use (at least) 3 approximating factors. To conclude this section, we want to underline the fact that this choice is independent of the maturities that we consider, since it approximates in a better way the rough skew both for shorter and longer times to maturity.

4.4.2 Forward Start European Call

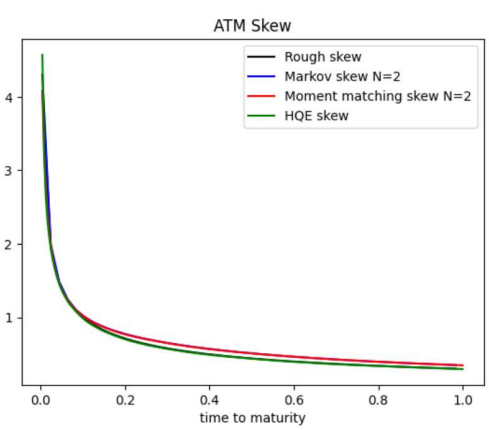
The next step is to study the implied volatility skew for the Forward Start European (FSE) options. To do this, we need to slightly change what we did in the previous section.

Definition 4.4.1. *A Forward Start European Call option written on S with maturity T, forward-start date $t \in (0, T)$ and relative strike k is defined by the payoff*

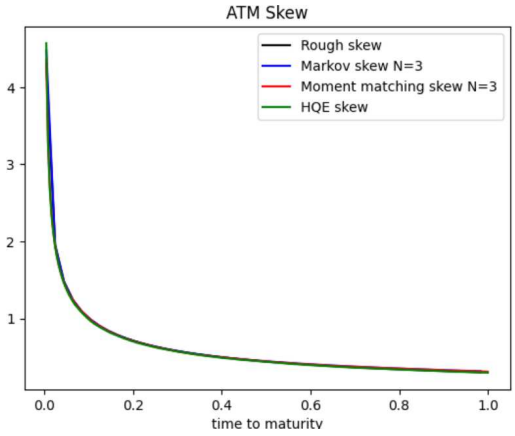
$$(S_T - kS_t)^+ \tag{4.3}$$

Remark 17. It is clear that if we want to determine the price of a FSE option with payoff given by Definition 4.4.1 at a time $t_1 > t$ this reduces to the price of a standard European Call with strike kS_t.

Clearly, it is necessary to adapt the definition of the implied volatility skew since now our strike $K = kS_t$ is random. To compute the implied volatility skew we need a formula to determine the price of Forward Start options under the Black-Scholes model. Thus, we need the following theorem which can be found in [4].



(a) ATM skew performing 1 million samples for maturities $T = 0.004, 0.008, ..., 1$, $H = 0.1$, $M = 1000$ time steps and $N = 2$ approximating factors for the Markovian and the Moment Matching skew. The rough and the Markovian skew are obtained by the Fourier inversion technique choosing 50 linearly spaced maturities $T = 0.004, 0.024, ..., 0.984$.



(b) ATM skew performing 1 million samples for maturities $T = 0.004, 0.008, ..., 1$, $H = 0.1$, $M = 1000$ time steps and $N = 3$ approximating factors for the Markovian and the Moment Matching skew. The rough and the Markovian skew are obtained by the Fourier inversion technique choosing 50 linearly spaced maturities $T = 0.004, 0.024, ..., 0.984$.

Figure 4.11

Theorem 4.4.1. *Consider the Black-Scholes model with volatility σ, interest rate r and an option whose payoff is given by Equation (4.3). Then for all $s < t < T$ it holds:*

$$e^{-r(T-s)}\mathbb{E}_s\big[(S_T - kS_t)^+\big] = S_s e^{-r(T-t)}\mathbb{E}_t\big[(\hat{S}_T - k)^+\big] \tag{4.4}$$

where $\hat{S}$ is the spot process with current value at time t given by $S_t = 1$.

Proof. By definition of no arbitrage price $\forall s \leq t$ we have

$$\begin{aligned}
\mathrm{C}\big(s, T, kS_t\big) &= e^{-r(T-s)}\mathbb{E}_s\big[(S_T - kS_t)^+\big] \\
&= e^{-r(T-s)}\mathbb{E}_s\big[\mathbb{E}_t\big[(S_T - kS_t)^+\big]\big] \\
&= e^{-r(T-s)}\mathbb{E}_s\big[e^{r(T-t)}\mathrm{C}(t, T, kS_t)\big] \\
&= e^{-r(t-s)}\mathbb{E}_s\big[S_t\big(N(d_1) - ke^{-r(T-t)}N(d_2)\big)\big] \\
&= S_s\hat{\mathrm{C}}(t, T, k)
\end{aligned}$$

where $d_1 = \dfrac{\log(\frac{S_t}{kS_t}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}$, $d_2 = d_1 - \sigma\sqrt{T-t}$ and $\hat{\mathrm{C}}(t, T, k)$ is the Black-Scholes price at time t of a Call with maturity T, $S_t = 1$ and strike k. □

Thanks to Theorem 4.4.1 we can define the implied volatility and consequently the skew for Forward Start options.

Definition 4.4.2. *Consider a Forward Start European Call option with starting date t, maturity T and relative strike k. Suppose that its market price is given by $C(t,T,k)$. Then we can define the implied volatility $\sigma_{BS}(t,T,k)$ as the value of volatility such that*

$$C(t,T,k) = C^{BS}(t,T,k,\sigma_{BS}(t,T,k))$$

where $C^{BS}(t,T,k,\sigma_{BS}(t,T,k))$ is the price of the option in the Black-Scholes model with volatility $\sigma_{BS}(t,T,k)$ given by Theorem 4.4.1.

Now we can define the implied volatility skew similarly to what we did in the case of European Call options.

Definition 4.4.3. *Consider a Forward Start options written on S with starting date t, maturity T and relative strike k. Let $\sigma_{BS}(t,T,k)$ be its implied volatility, thus we define the at the money forward skew at time T*

$$\Psi_t(T) = \frac{\partial}{\partial k}\sigma_{BS}(t,T,k)\Big|_{k=1} \tag{4.5}$$

Remark 18. In Chapter 9 of [4] the authors demonstrate that rough volatility models can describe a blow-up of the at the money skew as time to maturity decreases. Therefore, we expect that also the lifted model present this behaviour since, in the limit, it converges to the rough version. Unfortunately, we are not able to determine neither the characteristic function that must be used to price Forward Start options in the rough Heston model, nor for the lifted version. Due to this lack, we base our analysis only on Monte Carlo simulations.

We keep fixed the maximal maturity $T = 1$ year and we choose to split our interval in two different ways: in the first case we take $T_0 = 0.02 < T_1 = 0.04 < ... < T = 1$ while in the second case we fix $T_0 = 0.004 < T_1 = 0.008 < ... < T = 1$ [8]. Moreover, for each choice of the discretization step, we decide to take three possible starting dates $t_1 = 0.3$ years, $t_2 = 0.5$ years and $t_3 = 0.9$ years.
From numerical experiments we observe that the method that we use to compute the skew for standard European options does not behave well in the case of Forward Start options, since it requires a high computational time to achieve the desired level of precision. This is principally related to the fact that the computation of a derivative using the finite difference method is ill conditioned by the Monte Carlo error. The problem is that pushing

[8] These choices correspond respectively to 50 and 250 linearly spaced maturities in the interval $[0,1]$ neglecting the first.

forward the starting date, the standard deviations increase: for instance, if we choose a discretization step of 0.004 we have that the standard deviations of the asset price performing 1 million simulations are

$$\sigma_{0.3} = 8\%, \quad \sigma_{0.5} = 10.6\%, \quad \sigma_{0.9} = 14.5\%.$$

So, this implies that as soon as we push forward the starting date, we need to tighten our strikes, to perform a high number of Monte Carlo simulations and this slows down the computations.
Nevertheless we overcome this problem adopting a scheme based on the use of Chebyshev interpolant. This method is very useful because it allows us to approximate the smile with a polynomial for which the derivative can be computed in closed form without numerical approximations. The choice of Chebyshev interpolant is motivated by the fact that it presents very good property of convergence with respect to the uniform norm as we will see.

Definition 4.4.4. *Given a function $f : [-1, 1] \to \mathbb{R}$ and a set of distinct points $\{x_1, ..., x_n\}$ then the Lagrange interpolant is the polynomial of degree at most $n - 1$ defined by*

$$p_{n-1}(x) = \sum_{k=1}^{n} f(x_k) L_k(x)$$

where

$$L_k(x) = \prod_{j \neq k} \frac{x - x_j}{x_k - x_j}$$

is the k-th Lagrange polynomial.
Choosing the point $x_i = cos(\frac{i\pi}{n})$, $i = 1, ..., n$ and computing the Lagrange interpolant we will find the Chebyshev interpolant.

Remark 19. The choice of the interval $[-1, 1]$ for the domain of the function can be easily extended to a more general case of a close interval $[a, b]$ using the affine transformation:

$$y \in [-1, 1] \to \frac{(b - a)}{2}(y - a) + a.$$

The Chebyshev interpolant and its derivatives have very good property of convergence which can be summerized in the following theorem.

Theorem 4.4.2. *Let $f : [-1, 1] \to \mathbb{R}$ be an analytic function which is analitically continuable to the closed Bernstein ellipse E_ρ of radius $\rho > 1$. Then for each integer $m > 0$ there exists a positive constant c such that*

$$||f^{(m)} - p_{n-1}^{(m)}||_\infty < c\rho^{-n}$$

For the proof and for more details on Chebyshev interpolant we suggest to consult [6] and [7]. For us, it is important since if we suppose that the smile function is "smooth enough" we can compute the skew for Forward Start options applying the following procedure:

- Fix the partition $T_1 < ... < T_n$ of the interval $[0, 1]$ as indicated above and choose the starting date t.

- Simulate the paths of the underlying asset and choose the value corresponding to the starting date.

- Fix 7 relative strike prices in the interval $[0.997, 1.003]$ according to the rule
$$k_i = 1 + \frac{3}{1000} \cos\left(\frac{(2i-1)\pi}{14}\right), \ i = 1, ..., 7.$$
These strikes correspond to the Chebyshev nodes that we need to compute the interpolant, moreover, as we said before, they will be rescaled according to $k^*_{j,i} = k_j^{\sqrt{T_i}}, \quad j = 1, ..., 7, \quad i = 1, ..., n$ and then we multiply them by the values of S_t. [9]

- Compute the implied volatility surface using Theorem 4.4.1 and then for each maturity T_i compute the Chebyshev polynomial of degree at most six $p_6(k)$ that fits the smile.

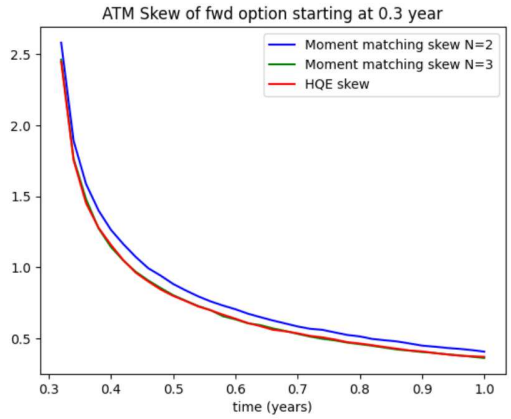
- Thanks to the property on convergence of Chebyshev interpolant its derivative $p'_6(k)$ is a good approximation for the derivative of the smile that is the skew. Thus, we compute the derivative of $p_6(k)$ and take as skew $p'_6(1)$.

We underline the two great advantages given by this procedure:

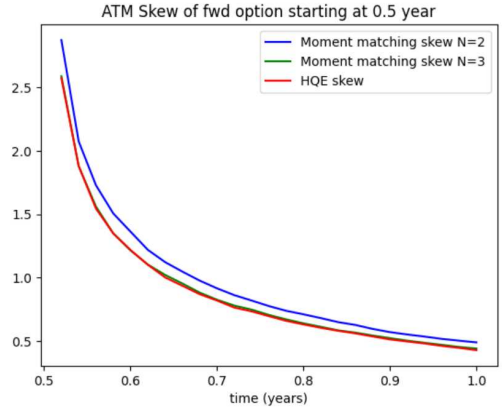
- It allows us to reduce the number of Monte Carlo simulations because we need only 1 million samples and subsequently we use only the interpolants for all the computations. In particular, the number of interpolants is equal to the number of maturities (one for each smile) and they can be computed in a very efficient way, since we have a closed form formula given in Definition 4.4.4.

[9] Plotting the true smile vs the interpolant it seems that the quality of the approximation does not deteriorate even if the rescaled nodes do not correspond exactly to the Chebyshev nodes.

- We are substituting a numerical derivative with the exact one because the derivative of a polynomial is well known and we compute it in $k = 1$.

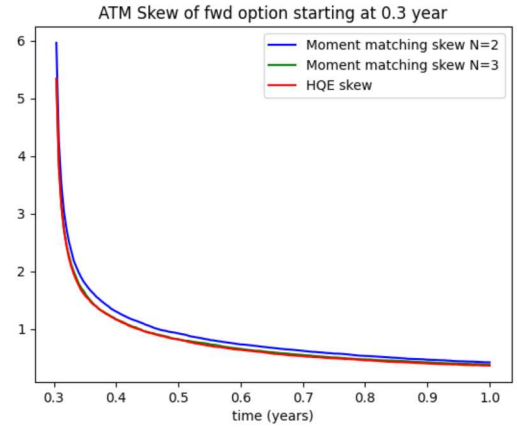


(a) ATM forward implied volatility skew for a Call option with starting date $t = 0.3$, maturities $T = 0.32, 0.34, ..., 1$ performing 1 million samples with $M = 500$ time steps.

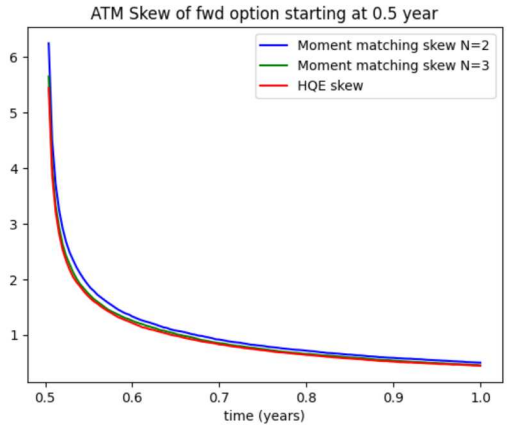


(b) ATM forward implied volatility skew for a Call option with starting date $t = 0.5$, maturities $T = 0.52, 0.54, ..., 1$ performing 1 million samples with $M = 500$ time steps.

Figure 4.12



(a) ATM forward implied volatility skew for a Call option with starting date $t = 0.3$, maturities $T = 0.304, 0.308, ..., 1$ performing 1 million samples with $M = 1000$ time steps.



(b) ATM forward implied volatility skew for a Call option with starting date $t = 0.5$, maturities $T = 0.504, 0.508, ..., 1$ performing 1 million samples with $M = 1000$ time steps.

Figure 4.13

What we can observe in Figure 4.12 and Figure 4.13 puts in evidence some important features:

- The Moment Matching scheme is able to reproduce the blow-up of the at the money skew for short maturities. Indeed, as for standard European options, we have an explosion. This is related to the higher mean reversion which dominates the others for short maturities.

- The best choice for the number of factors N in the Markovian approximation seems to be $N = 3$. Indeed the Moment Matching skew is (quite) perfectly overlapping the HQE skew in both cases. This fact is coherent with what we observed in the case of standard European options for which we had a reference skew computed via the Fourier inversion technique.

The last thing that we want to verify is that the quality of the skew does not deteriorate if we push forward the starting date. To do this, we fix the last starting date $t = 0.9$ and in order to have significant results we fix the daily discretization step as we did above. The results can be seen in Figure 4.14 and they confirm that the quality is maintained also for a very long starting date.

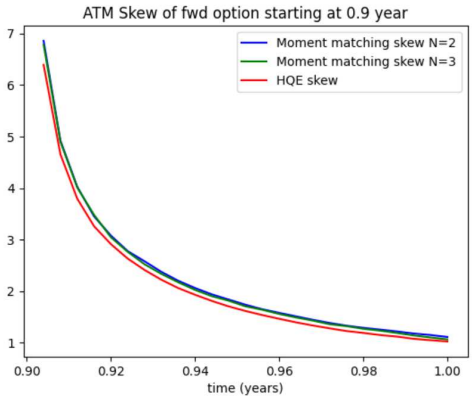


Figure 4.14: ATM forward implied volatility skew for a Call option with starting date $t = 0.9$, maturities $T = 0.904, 0.908, ..., 1$ performing 1 million samples with $M = 1000$ time steps using Moment Matching scheme with $N = 2, 3$.

4.5 Estimate of the Hurst index

In the spirit of [31] we want to estimate the Hurst index produced by the sample path of the volatility process $(\sqrt{V_t})_{t\geq 0}$. In particular, our goal is to check that the volatility behaves like a fractional Brownian motion with

Hurst index $H < \frac{1}{2}$.
Then we define $\sigma_t = \sqrt{V_t}$ and we estimate the (discrete) q-variation

$$m(q, \Delta) = \frac{1}{N} \sum_{i=1}^{N} |\sigma_{i\Delta} - \sigma_{(i-1)\Delta}|^q \tag{4.6}$$

for different values of $q > 0$ and $\Delta = \frac{T}{N}$ where T is the temporal horizon that we want consider. Thus, we need a discretization of the interval $[0, T]$ which is determined for different values of N. Of course N can not be too small if we want a good approximation of the q-variation of the volatility.

Remark 20. In the original paper [31] the authors show that the process $(\log(\sigma_t))_{t\geq 0}$ behaves like a fractional Brownian motion. Nevertheless, in [1] Abi Jaber estimates the q-variation of the volatility which has a rough behaviour with $H < \frac{1}{2}$. Our choice is also motivated by the fact that for the HQE scheme the volatility can reach the zero level so we can not consider the logarithm.

Thanks to Proposition 1.2.2 our guess is that

$$m(q, \Delta) = C_q \Delta^{Hq}$$

and so this enables us to recover the Hurst index H. We want to underline the fact that this test is different from the others because it wants to determine a path-wise property of the variance while the previous experiments deal with its law. Due to this reason, we consider also the drift implicit Euler scheme which is briefly described in the following.

- Fix a time step $\Delta t = \frac{T}{M}$ and a number of factors N.

- Approximate the variance V^i using

  $V^i_{t_{m+1}} = V^i_{t_m} - x_i(V^i_{t_{m+1}} - v^i_0)\Delta t + (\theta - \lambda V_{t_{m+1}})\Delta t + \sigma\sqrt{(V_{t_m})^+}\Delta W_{t_m}$

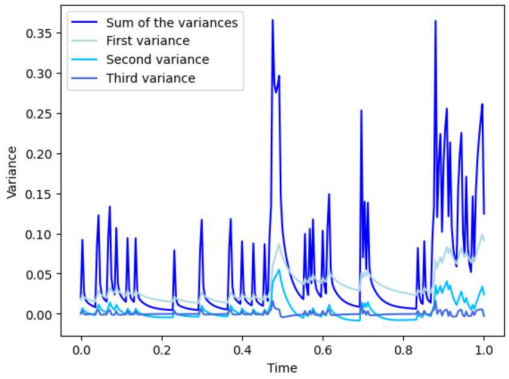
  for $m = 0, ..., M-1$ and $\Delta W_{t_m} = W_{t_{m+1}} - W_{t_m}$.

- Take $V_{t_{m+1}} = \big(\sum_{i=1}^{N} w_i V^i_{t_{m+1}}\big)^+$ [10]

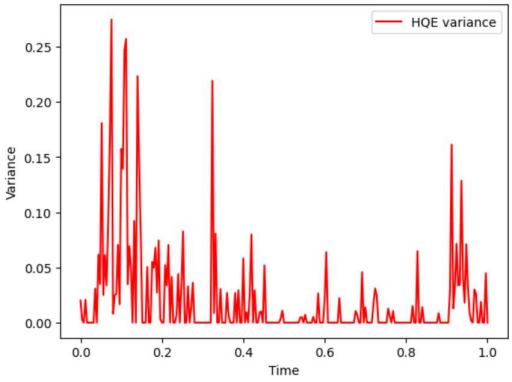
[10] This choice is not the only one that guarantees that the variance process remains non negative, indeed we could also decide to define $V_{t_{m+1}} = |\sum_{i=1}^{N} w_i V^i_{t_{m+1}}|$

We choose to add also this simulation scheme because the theory guarantees that as soon as the time steps $\Delta t \to 0$ the simulated dynamics converges to the original one.
First of all, we want to check graphically what is the behaviour of the simulated variance with the different methods in order to understand if they are able to trick the human eye. [11] Thus, we fix the maturity $T = 1$ year and we consider the daily value of the variance which corresponds to a discretization of the interval $[0, 1]$ into 250 sub-intervals of equal length. In addition, we tighten our time step splitting $[0, 1]$ into 2500 sub-intervals.

As we can observe in Figure 4.15, Figure 4.16 the HQE scheme seems to



(a) Variance samples of the three components V^i, $i = 1, 2, 3$ and of the total variance $V = \mathbf{w} \cdot \mathbf{V}$ for the Moment Matching scheme using $M = 250$ time steps.

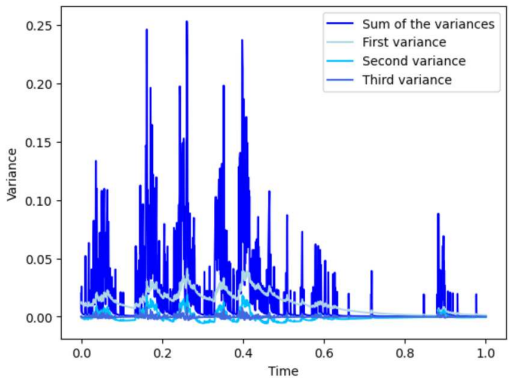


(b) Variance sample for the Hybrid Quadratic Exponential scheme using $M = 250$ time steps.

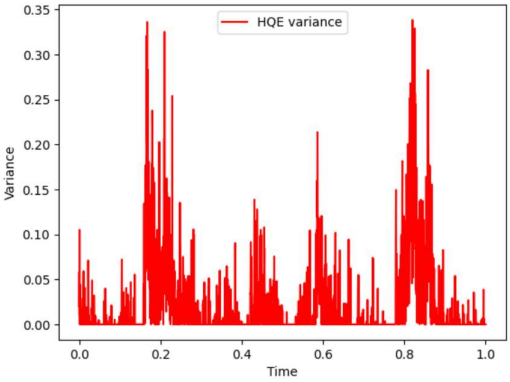
Figure 4.15

replicate in a good way the rough behaviour of the market variance even if the simulation reaches the level zero very often (in this case more that 60 % of the values are zero). On the contrary, the Moment Matching scheme seems not to be able to trick the human eye mimicking the rough behaviour of the market variance. To understand if this is a problem related to the simulation scheme or to the lifted Heston model, we repeat the experiment using the drift implicit Euler scheme keeping fixed the number of factors $N = 3$ and all the other parameters. As we can observe from Figure 4.17, this scheme seems to replicate the rough behaviour of the variance process, thus we conclude that the problem was related to the choice of the simulation scheme. Indeed, it is not clear at all why, even tightening the time steps, there are some points in which the exponential behaviour seems to overcome the noise.

[11] We want to understand if the simulated paths are similar to what we observe from historical market variance.

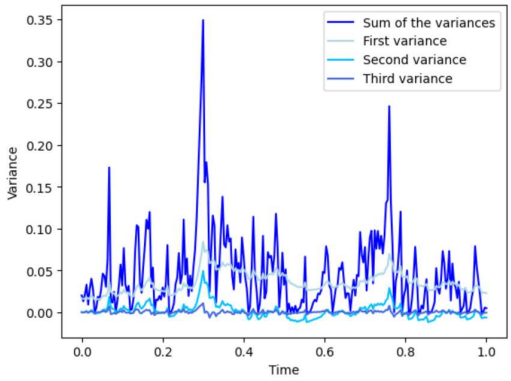


(a) Variance samples of the three components V^i, $i = 1, 2, 3$ and of the total variance $V = \mathbf{w} \cdot \mathbf{V}$ for the Moment Matching scheme using $M = 2500$ time steps.

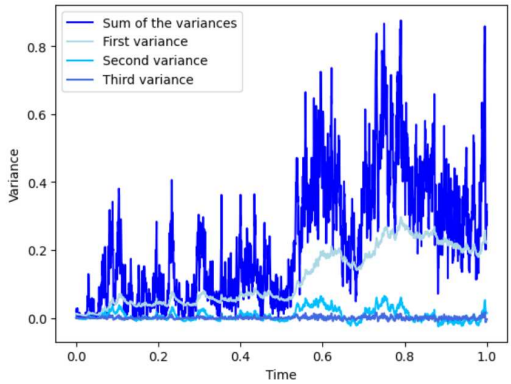


(b) Variance sample for the Hybrid Quadratic Exponential scheme using $M = 2500$ time steps.

Figure 4.16



(a) Variance samples of the three components V^i, $i = 1, 2, 3$ and of the total variance $V = \mathbf{w} \cdot \mathbf{V}$ for the drift implicit Euler scheme using $M = 250$ time steps.



(b) Variance samples of the three components V^i, $i = 1, 2, 3$ and of the total variance $V = \mathbf{w} \cdot \mathbf{V}$ for the drift implicit Euler scheme using $M = 2500$ time steps.

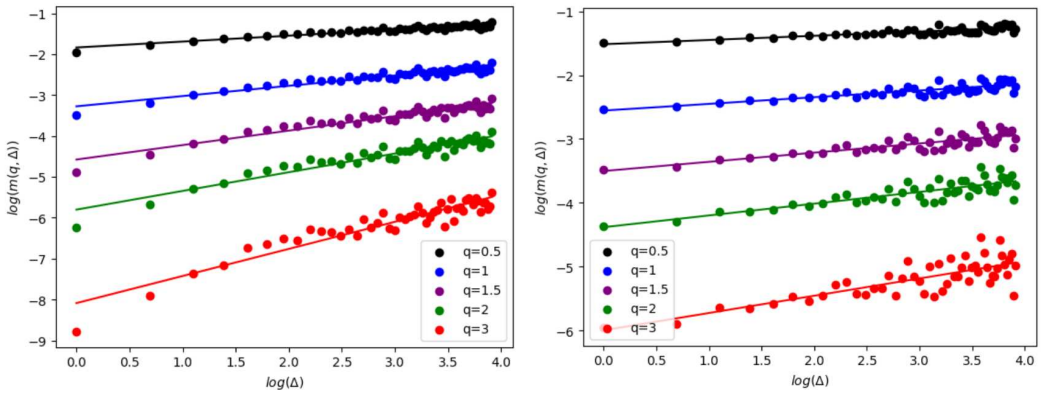
Figure 4.17

Now that we have an idea of the behaviour of the sample paths we want to estimate the Hurst index. To have a good sample for our estimate we decide to fix $\Delta = 1, ..., 50$, then we face the problem of the choice of the discretization step. We decide to fix it with the aim of having a good trade off between the accuracy of the estimate and the noise that arise in the computation of the q-th moment.
From empirical evidence the choice of $M = 2500$ for the Moment Matching

scheme and the HQE scheme seems to perform in an efficient way. We consider a time window of 1 year so, with the convention of 250 trading days per year and 6.5 trading hours per day, this corresponds to a choice of

$$\Delta = 0.65 \ hours, ..., 5 \ days$$

As we can see in Figure 4.18 the moments seem to lie on a straight line in both cases. Now we plot the slope of the regression straight line against q



(a) Logarithm of the moments of order $q = 0.5, 1, 1.5, 2, 3$ for the Moment Matching scheme using $M = 2500$ time steps and $\Delta = 1, ..., 50$.

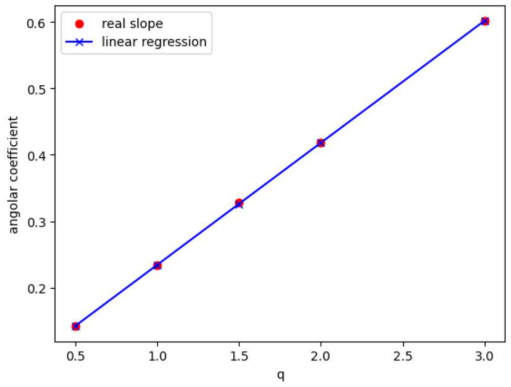
(b) Logarithm of the moments of order $q = 0.5, 1, 1.5, 2, 3$ for the HQE scheme using $M = 2500$ time steps and $\Delta = 1, ..., 50$.

Figure 4.18

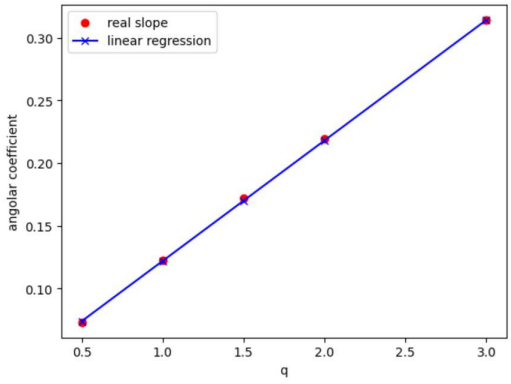
and we observe that there is a linear relation between them as we can see from Figure 4.19.
In particular, we find that the estimated Hurst index is 0.19 for the Moment Matching scheme and 0.09 for the HQE scheme. This confirms our guess on the rough behaviour of the volatility.

Remark 21. We want to remark the fact that we repeat the test also doubling the number of the values for the Delta and keeping fixed the number of time steps. We obtain very similar values for the Hurst index, however the moments are affected by a greater noise and so this makes the estimate less accurate.

We also repeat the experiment for the estimate of the Hurst index in the case of the drift implicit Euler scheme: the results are very similar to those of the Moment Matching scheme. As a matter of fact, we find an Hurst index $H = 0.17$. As pointed out by Abi Jaber in [1] for the drift implicit Euler scheme, we have seen that also the Moment Matching scheme presents a rough behaviour (at least for a sufficiently thin time step) with an Hurst



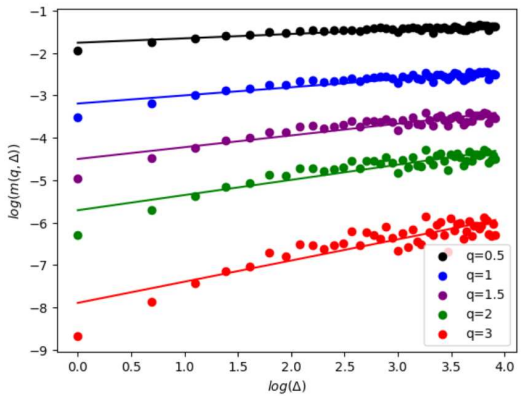
(a) q against $\zeta = Hq$ for the Moment Matching scheme using $M = 2500$ time steps.



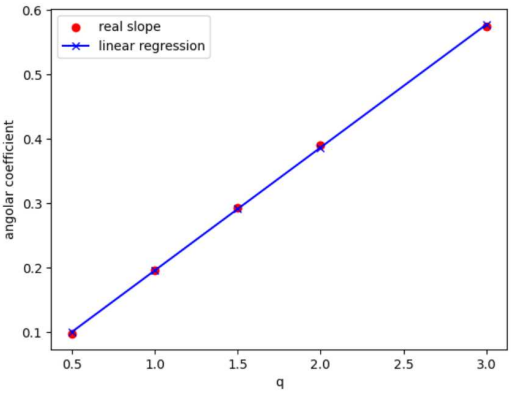
(b) q against $\zeta = Hq$ for the Hybrid Quadratic Exponential scheme using $M = 2500$ time steps.

Figure 4.19

index way below 0.5 which is what we expected since the volatility process, in this case, is driven by a Brownian motion.



(a) Logarithm of the moments of order $q = 0.5, 1, 1.5, 2, 3$ for the drift implicit Euler scheme using $M = 2500$ time steps and $\Delta = 1, ..., 50$.



(b) q against $\zeta = Hq$ for the drift implicit Euler scheme using $M = 2500$ time steps.

Figure 4.20

Chapter 5

Skew Stickiness Ratio

The concept of skew stickiness ratio (SSR) was introduced for the first time by Bergomi in [13]. Therein he demonstrates the asymptotic behaviour of the skew stickiness ratio in the case of stochastic volatility models. Since the classic Heston model, described by Equation (1.2), is a special case of the more general rough version, the aim of this chapter is to study the SSR in the rough Heston model.
The skew stickiness ratio was introduced to derive a relation between the implied volatility and the stock price. Since they are both observable in the market, it makes sense to ask whether there exists a relation that links them. In particular, the goal is to understand how a variation of the spot impacts the implied volatility. The importance of being able in linking the movements of the spot and the movements of the implied volatility is that in this way traders can make Delta hedging instead of the more expensive Vega hedging.

Let us focus on a European Call option with current value of the spot S_0, maturity T and strike K. We know that its market price can be expressed using the Black and Scholes model by the relation

$$C(T, K, S_0) = C_{BS}(T, K, S_0, \sigma_{BS}(T, k))$$

where $k = \log(K/S_0)$ is the log-moneyness and $\sigma_{BS}(\cdot)$ is the implied volatility.
Thus, since the price is affected by the spot and by the implied volatility, to hedge options using the Black-Scholes formula, traders need to hedge two effects:

- The effect of a change of the spot price (Delta hedging)

- The effect of a change of the implied volatility conditional on a change of the spot (Vega hedging).

Let us focus on at the money options, i.e., with $k = 0$. Market makers typically assume that the variation of the at the money implied volatility and the variation of the stock price are linked by the linear relation

$$\delta\sigma_{BS}(T,0) = \alpha(T)\frac{\delta S}{S_0} + \epsilon = \alpha(T)\delta X + \epsilon \tag{5.1}$$

where ϵ represents the Gaussian noise independent of the rest and with zero mean, and $\delta X = \dfrac{\delta S}{S_0} = \dfrac{S_T - S_0}{S_0}$.
Then, taking the conditional expectation w.r.t. δX on both sides of Equation (5.1), we observe that

$$\alpha(T) = \frac{\mathbb{E}\big[\delta\sigma_{BS}(T,0)|\delta X\big]}{\delta X}.$$

If we consider the at the money skew

$$\psi(T) = \frac{\partial}{\partial k}\sigma_{BS}(T,k)\bigg|_{k=0}$$

then we can finally compute the skew stickiness ratio as

$$\mathcal{R}(T) = \frac{\alpha(T)}{\psi(T)}. \tag{5.2}$$

Looking at Equation (5.2), we observe that the SSR can be thought as a measure of how much the at the money implied volatility moves conditional on a move of S in units of at the money skew. In the market it is common to distinguish two regimes according to the value of the SSR:

- Sticky-strike: The implied volatility is fixed for a given strike independent of the stock price.

- Sticky-delta: The at the money implied volatility is fixed independently on the current stock price.

Remark 22. The sticky-strike regime implies that if we consider the at the money implied volatility and a stock price increase from S_0 to $S_0 + \delta S$ then, if we change the system coordinates using the transformation $(t,k) \to (t,K)$ [1] where $K = S_0 e^k$, and using a first order Taylor expansion for the implied volatility we have

$$\delta\sigma_{BS}(T,0) = \sigma_{BS}(T, S_0 + \delta S) - \sigma_{BS}(T,S_0) \approx \psi(T)\frac{\delta S}{S_0}$$

that implies $R(T) = 1$. On the other hand, the sticky-delta regime implies that $\delta\sigma_{BS}(T,0) = 0$ because the at the money implied volatility remains fixed independently on the value of the spot and thus $R(T) = 0$.

[1] Observe that this change of coordinates is such that $\dfrac{\partial}{\partial K} = \dfrac{1}{K}\dfrac{\partial}{\partial k}$.

It has been demonstrated in [13] that for stochastic volatility models $\mathcal{R}(T) \to 2$ for $T \to 0$. Thus, as demonstrated in [29], we want to show in a different way that for the rough Heston model $R(T) \to \left(\frac{3}{2} + H\right)$ for $T \to 0$. To this end, we need to formalize in a rigorous way the concept of skew stickiness ratio.

Definition 5.0.1. *Define $\sigma_t(T,k)$ the implied volatility observed at time t, for a Call option with maturity T and strike price $K = S_0 e^k$. Then considering $X_t := log(S_t/S_0)$ we define the skew stickiness ratio*

$$\mathcal{R}_t(T) := \frac{\mathbb{E}[(\sigma_t(T, X_t) - \sigma_0(T,0))X_t]}{\psi(T)\mathbb{E}[(X_t)^2]}$$

where $\psi(T) = \dfrac{\partial}{\partial k}\sigma_{BS}(T,k)\Big|_{k=0}$.

Remark 23. Definition The skew stickiness ratio is the linear regression coefficient of the at the money implied volatility level in the time interval $(0,t)$ against the movement X_t in unit of at the money skew. This is exactly the definition given by Bergomi in [13]: we can observe that it involves the maturity T but also the time step t. Then, our aim is to compute the limit

$$\mathcal{R} = \lim_{T\to 0}\lim_{t\to 0}\mathcal{R}_t(T). \tag{5.3}$$

Moreover, it is important to observe that Definition 5.0.1 involves only expected values, this implies that in general the short time limit does not hold pathwise.

To follow the idea of Bergomi, it is convenient to introduce two other processes: the forward variance and the the implied variance swap volatility.

Definition 5.0.2. *Given the dynamics a variance process $(V_t)_{t\geq 0}$ we define:*

- *The forward variance for date T observed a time $t \in [0,T]$ $\xi_t(T) = \mathbb{E}[V_T|\mathcal{F}_t]$, where $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration.*

- *The implied variance swap volatility for date T observed at time $t < T$ $\hat{\sigma}_{t,T}$ such that $\xi_t(T) = \dfrac{d}{dT}\big((T-t)\hat{\sigma}^2_{t,T}\big)$.*

We recall that the dynamics in our model is given by

$$dX_t = -\frac{1}{2}V_t dt + \sqrt{V_t}dB_t, \quad X_0 = 0 \tag{5.4}$$

with

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\lambda(\theta - V_s)ds + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\sigma\sqrt{V_s}dW_s$$

where $\alpha = \left(\frac{1}{2} + H\right) \in \left(\frac{1}{2}, 1\right)$, $\theta, \sigma > 0$, $\lambda \geq 0$ and $\rho dt = d\langle W, B\rangle_t$ with $\rho \in [-1, 1]$.
Now we want to follow the same procedure applied by Bergomi in [13], thus we need to understand what is the dynamics of the forward variance curve.

Proposition 5.0.1. *Assume that the dynamics of the model is given by Equation* (5.4) *with* $\lambda > 0$. *Then the forward variance curve evolves according to*

$$d\xi_t(T) = \sigma(T-t)^{\alpha-1}\left(\sum_{n\geq 0}\frac{(-\lambda(T-t)^\alpha)^n}{\Gamma((n+1)\alpha)}\right)\sqrt{\xi_t(t)}dW_t. \tag{5.5}$$

In particular for $(T-t) \to 0$ *it holds*

$$\sigma(T-t)^{\alpha-1}\left(\sum_{n\geq 0}\frac{(-\lambda(T-t)^\alpha)^n}{\Gamma((n+1)\alpha)}\right) \sim \frac{1}{\Gamma(\alpha)}\sigma(T-t)^{\alpha-1}. \tag{5.6}$$

For the proof we refer to [27]. The important thing is that thanks to Proposition 5.0.1, applying the Equation (5.6) we can assume that $\lambda = 0$ because we are interested in the short time limit. Indeed, if we suppose that $\lambda = 0$, then

$$V_T = V_0 + \frac{1}{\Gamma(\alpha)}\int_0^T (T-s)^{\alpha-1}\sigma\sqrt{V_s}dW_s. \tag{5.7}$$

Thus, taking the conditional expectation at time $t \leq T$ on both sides of (5.7), and using the properties of the stochastic integral, we find

$$\xi_t(T) = V_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (T-s)^{\alpha-1}\sigma\sqrt{V_s}dW_s \tag{5.8}$$

which can be written as

$$d\xi_t(T) = \frac{\sigma}{\Gamma(\alpha)}(T-t)^{\alpha-1}\sqrt{\xi_t(t)}dW_t. \tag{5.9}$$

Essentially, using Proposition 5.0.1 and Equation (5.9) we observe that when $(T-t) \to 0$ the mean reversion λ has no effect.

Remark 24. The same result can be obtained working with $\lambda > 0$ and then passing to the limit for $T \to 0$ but it requires more computations and so we prefer to proceed assuming $\lambda = 0$ from the beginning. The general case can be found in Appendix A.

Moreover, using Definition 5.0.2, we can derive an approximation for the implied variance swap volatility: integrating on both sides we find

$$\hat{\sigma}^2_{t,T} = \frac{1}{(T-t)}\int_t^T \xi_t(s)ds.$$

Now, applying the chain rule and taking only the lowest order term in dt we obtain

$$d\hat{\sigma}_{t,T} = \frac{1}{2\hat{\sigma}_{t,T}(T-t)} \int_t^T d\xi_t(s)ds. \tag{5.10}$$

After that, we want to use Equations (5.9)-(5.10) to compute the skew stickiness ratio using the approximation given by Bergomi in [13] where he replaces the implied volatility with the volatility of the variance swap. Hence, we have to determine the ratio

$$\mathcal{R}_t(T) = \frac{\mathbb{E}[d\hat{\sigma}_{t,T} dX_t]}{\psi(T)\mathbb{E}[(dX_t)^2]}. \tag{5.11}$$

Remark 25. To have an intuition we can think at the quantity in Equation (5.11) as the ratio between the infinitesimal covariation of the processes $(\hat{\sigma}_{t,T})_{t\geq 0}$ and $(X_t)_{t\geq 0}$, and the infinitesimal quadratic variation of $(X_t)_{t\geq 0}$. This ratio is measured in unit of at the money skew.

The next step is to compute the at the money skew. In order to do this we use the well known asymptotic approximation [2]

$$\frac{\partial}{\partial k}\sigma_{BS}(T,k)\Big|_{k=0} \approx \frac{1}{6\sqrt{T}} \frac{\mathbb{E}\big[(X_T - \mathbb{E}(X_T))^3\big]}{\Big(\mathbb{E}\big[(X_T - \mathbb{E}(X_T))^2\big]\Big)^{3/2}}. \tag{5.12}$$

Once again, we underline the fact that we neglect higher order factors in the computations since we are interested in short time limit with respect to T. Now assume that $(X_t)_{t\geq 0}$ evolves according to Equation (5.4). Applying Itô's lemma we derive:

$$X_T^2 = \int_0^T V_t(1 - X_t)dt + 2\int_0^T X_t\sqrt{V_t}dB_t$$

and

$$X_T^3 = 3\int_0^T X_t V_t\left(1 - \frac{X_t}{2}\right)dt + 3\int_0^T X_t^2\sqrt{V_t}dB_t.$$

Thus, we can compute the first three moments which we need to give the asymptotic estimate of the skew. In doing this, we obtain:

- $\mathbb{E}(X_T) = -\frac{1}{2}\int_0^T \xi(t)dt$
- $\mathbb{E}(X_T^2) = \int_0^T \xi(t)dt - \mathbb{E}\left(\int_0^T X_t V_t dt\right)$
- $\mathbb{E}(X_T^3) = \mathbb{E}\left(-\frac{3}{2}\int_0^T X_t^2 V_t dt + 3\int_0^T X_t V_t dt\right)$.

[2] For a complete derivation of this result see Chapter 9 of [14].

In particular we can go forward and, neglecting the drift term of X_t which will be of higher order in the limit for $T \to 0$ and by means of the Itô isometry, we can compute

$$\begin{aligned}
\mathbb{E}\Big(\int_0^T X_t V_t dt\Big) &= \int_0^T \Big(\int_0^t \mathbb{E}\big[\sqrt{V_s} V_t dB_s\big]\Big) dt \\
&= \int_0^T \Big(\int_0^t \mathbb{E}\big[\sqrt{V_s}\big(V_0 + \int_0^t \frac{(t-u)^{\alpha-1}}{\Gamma(\alpha)} \sigma \sqrt{V_u} dW_u\big) dB_s\big]\Big) dt \\
&= \int_0^T \Big(\int_0^t \frac{\rho\sigma}{\Gamma(\alpha)} (t-s)^{\alpha-1} \mathbb{E}(V_s) ds\Big) dt \\
&= \frac{V_0 \rho \sigma T^{1+\alpha}}{\Gamma(\alpha+2)}.
\end{aligned}$$

Then, at first order, we obtain the approximations:

$$\Big(\mathbb{E}\big[(X_T - \mathbb{E}(X_T))^2\big]\Big)^{3/2} = \Big(\int_0^T \xi(s) ds\Big)^{3/2}$$

and

$$\mathbb{E}\big[(X_T - \mathbb{E}(X_T))^3\big] = \frac{3 V_0 \rho \sigma T^{1+\alpha}}{\Gamma(\alpha+2)}.$$

Now we compute the ratio

$$\frac{\mathbb{E}[d\hat{\sigma}_{t,T} dX_t]}{\mathbb{E}[(dX_t)^2]}.$$

For the numerator we obtain

$$\begin{aligned}
\mathbb{E}[d\hat{\sigma}_{t,T} dX_t] &= \frac{1}{2\hat{\sigma}_{t,T}(T-t)} \int_t^T \mathbb{E}[d\xi_t(s) dX_t] ds \\
&= \frac{1}{2\hat{\sigma}_{t,T}(T-t)} \int_t^T \frac{\rho\sigma}{\Gamma(\alpha)} \mathbb{E}(V_t) dt (s-t)^{\alpha-1} ds \\
&= \frac{\rho\sigma (T-t)^{\alpha-1} V_0 dt}{2\hat{\sigma}_{t,T} \Gamma(\alpha+1)}
\end{aligned}$$

while

$$\mathbb{E}[(dX_t)^2] = V_0 dt.$$

Thus, computing everything for $t = 0$ and using the properties of the Gamma function, we find that the skew stickiness ratio becomes

$$\mathcal{R}(T) = \frac{(\alpha+1)\big(\int_0^T \xi(s) ds\big)^{3/2}}{V_0 \hat{\sigma}_{0,T} T^{3/2}}. \tag{5.13}$$

To conclude we need to apply Lebesgue differentiation theorem:

Theorem 5.0.1. *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Consider $B(x_0, r)$ the ball with center in $x_0 \in \mathbb{R}^n$ and radius $r > 0$. Then if $|B(x_0, r)|$ is the Lebesgue measure of the ball*

$$\lim_{r \to 0} \frac{1}{|B(x_0, r)|} \int_{B(x_0, r)} f(x)dx = f(x_0).$$

Applying Theorem 5.0.1 we deduce that $\hat{\sigma}^2_{0,T} = \frac{1}{T}\int_0^T \xi(s)ds \to V_0$ as $T \to 0$ and then we can conclude that

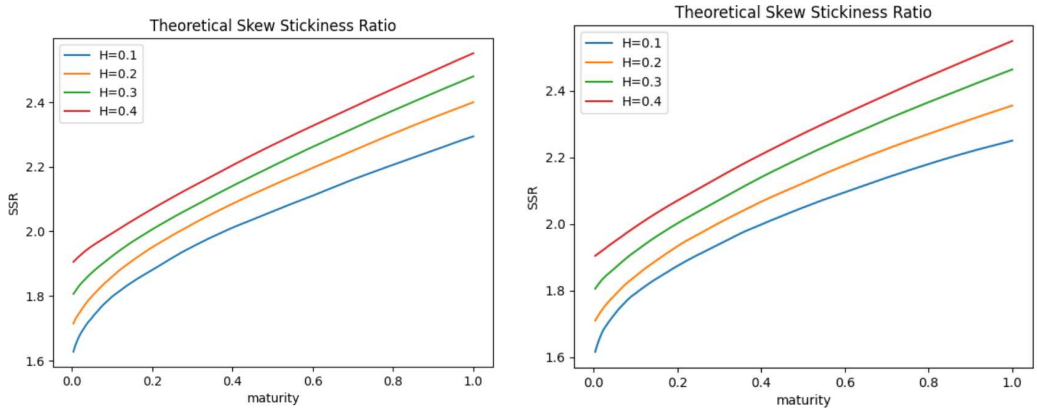
$$\mathcal{R}(T) \to \alpha + 1 = H + \frac{3}{2} \quad as \quad T \to 0.$$

Remark 26. We propose also an alternative derivation of this result in Appendix B. The idea is to use the asymptotic relation, valid for short maturities, between the skews of local and implied volatility described in [21]. Nevertheless, we are not able to prove it in whole generality and we have to make some restrictive assumptions.

5.1 Numerical results

Now that we have derived the asymptotic behaviour of the skew stickiness ratio in the rough Heston model, we want to investigate on the performance of the simulation schemes of our interest.
First of all, we need to understand how good is the approximation given by Equation (5.13) for maturities of practical use, thus we fix the time interval $[0, 1]$, we split it into 250 linearly spaced maturity (so ranging from 1 day till 1 year) and we plot the theoretical skew stickiness ratio in Figure 5.1 for different values of the Hurst index H. The integral in (5.13) is computed using the trapezoidal rule, while the expected value of the variance is computed performing 1 million Monte Carlo simulations using the Moment Matching and the HQE scheme described in chapter 3.
In such a way, we find that the relative error with respect to our benchmark of 1.6 is around 1.5% for the Moment Matching scheme while is around 0.9% for the HQE scheme. Thus, in order to reduce the error it is necessary to consider shorter maturities. Nevertheless, it is useless from a practical point of view since the smallest maturity that we can find in the market is 1 day and moreover, it is also expensive from a computational point of view, since we need to tighten the discretization step.

Now we want to compute the skew stickiness ratio using Equation (5.2), which is consistent with what we would do if we have had market data. In order to do this, we need to simulate sample path of the spot, and then we can use it to compute the at the money skew and the at the money implied volatility. Therefore, we follow these steps:



(a) Theoretical skew stickiness ratio for maturities ranging from 1 day till 1 year computed with Moment Matching scheme with $N = 3$.

(b) Theoretical skew stickiness ratio for maturities ranging from 1 day till 1 year computed with Hybrid Quadratic Exponential scheme.

Figure 5.1

- Fix a time step $\Delta t = 1$ day, define $t_i = i\Delta t,\ i = 0, ..., 250$ and simulate a sample path of the spot over a time window of 1 year.

- Relying on Equation (5.2) we need to compute the at the money skew: we do this using Chebyshev interpolant instead of finite difference method, in order to speed up the computations. Then, we fix seven strikes
$$k_i = 1 + \frac{3}{1000}\cos\left(\frac{(2i-1)\pi}{14}\right),\ i = 1, ..., 7$$
close to the at the money level.

- Fix 4 maturities for which we compute the skew stickiness ratio. In particular, since we are interested on the short time limit we decide to choose $T \in \{1 \text{ day},\ 1 \text{ week},\ 2 \text{ weeks},\ 1 \text{ month}\}$.

- For each maturity T compute
$$\mathbb{E}[(S_{t_i+T} - k_j^{\sqrt{T}} S_{t_i})^+ | \mathcal{F}_{t_i}] \quad i = 0, ..., 250 \text{ and } j = 1, ..., 7.$$
.

- Compute $\delta X_i = \dfrac{S_{t_i+T} - S_{t_i}}{S_{t_i}}$, the at the money implied volatility $\sigma_i(T, 0)$ and the at the money skew $\psi_i(T)$ for $i = 0, ..., 250$.

- Compute the linear regression coefficient between the $(\delta X_i \psi_i(T))_{i=0,...,249}$ and $(\sigma_{i+1}(T,0) - \sigma_i(T,0))_{i=0,...,249}$.

The first problem that we encounter is how to determine the forward price which means how to compute

$$\mathbb{E}[(S_{t_i+T} - S_{t_i})^+ | \mathcal{F}_{t_i}] \quad i = 0, ..., 250.$$

Clearly to do this with the Fourier inversion technique, it necessary to use the conditional characteristic function, but Theorem 1.4.1 gives us the characteristic function of the log-price only at $t = 0$. This problem may be overcome (at least theoretically) thanks to Corollary 2 at page 119 of [24]. Nevertheless, this result would be completely useless due to the huge computational time that it requires.
Even if we decide to use the Markovian lift, where the conditional characteristic function is given by Theorem 2.1.1, we face some problems in the computations.
Then we work with Monte Carlo method, relying on the Moment Matching scheme with 3 approximating factors since, as we saw in section 4.1, it is efficient from a computational point of view and it is also able to fit the rough skew. [3] In this case, we face the problem of the choice of the nodes $(x_i)_{i=1,2,3}$ and the weights $(w_i)_{i=1,2,3}$ that must be used to compute the prices: to be consistent, we fix the same that we use to simulate the spot path, they are reported in Table 5.1. Moreover, we simulate also the sample path of the volatility, so that at time each time t_i we can use the values $V_{t_1}^1, V_{t_i}^2, V_{t_i}^3$ to compute the variance $V_{t_i} = \sum_{j=1}^{3} w_i V_{t_i}^j$ that is necessary to compute the implied volatility.

Nodes	0.2099	14.1212	295.0156
Weights	1.1597	2.3197	12.7069

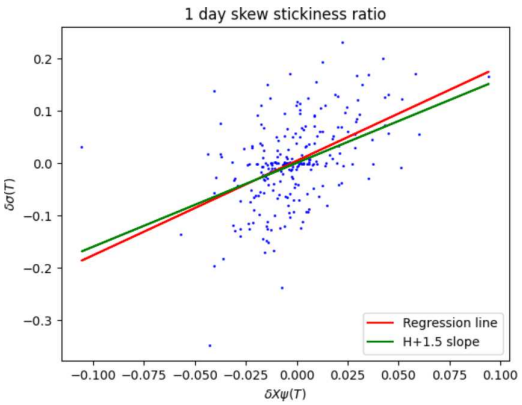
Table 5.1: Nodes and weights used to compute the skew stickiness ratio.

Then, applying the steps described above we give the estimates of the skew stickiness ratio reported in Table 5.2.

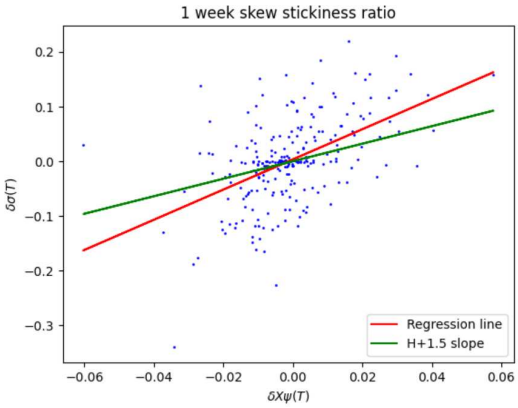
	1 day	1 week	2 weeks	1 month
SSR	1.804	2.757	3.723	4.979

Table 5.2: Estimated values of the skew stickiness ratio for maturity $T \in \{1 \text{ day}, 1 \text{ week}, 2 \text{ weeks}, 1 \text{ month}\}$.

[3] Unfortunately it is not possible to reproduce this experiment for HQE scheme because it would require too much time.

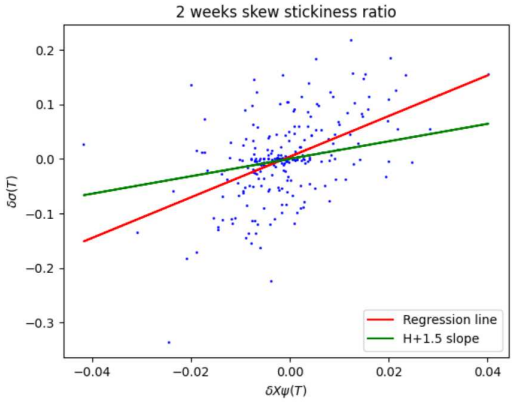


(a) 1 day skew stickiness ratio computed with the Moment Matching scheme.

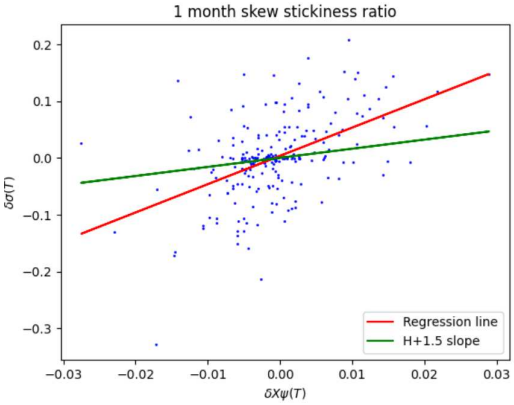


(b) 1 week skew stickiness ratio computed with the Moment Matching scheme.

Figure 5.2



(a) 2 weeks skew stickiness ratio computed with the Moment Matching scheme.



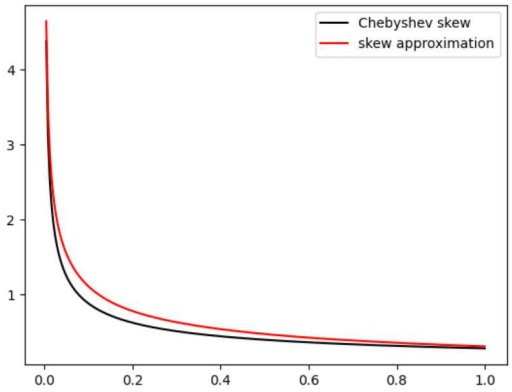
(b) 1 month skew stickiness ratio computed with the Moment Matching scheme.

Figure 5.3

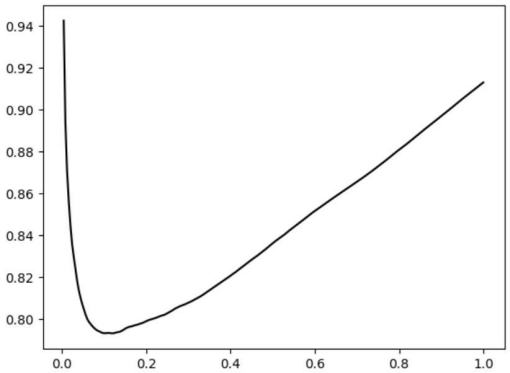
In Figure 5.2 and Figure 5.3 we report the diagrams with the theoretical level of $H + \frac{3}{2}$ and the estimated one.
As we can observe, the estimated values are far above the asymptotic level of 1.6 even for short maturity.
This can be caused by the first order approximation of the skew

$$\frac{\partial}{\partial k}\sigma_{BS}(T,k)\Big|_{k=0} \approx \frac{\rho\sigma V_0}{2\Gamma(H+5/2)}\frac{T^{H+1}}{(\int_0^T \xi(s)ds)^{\frac{3}{2}}}. \tag{5.14}$$

Indeed, it is true for small maturity T, so we need to understand how good



(a) The rough Heston skew computed with the Moment Matching scheme using Chebyshev interpolant and the first order approximation.



(b) Ratio between the rough Heston skew computed with the Moment Matching scheme using Chebyshev interpolan and the first order approximation.

Figure 5.4

is this approximation for the maturity of our interest. Looking at Figure 5.4 we observe that the first order approximation overestimates the skew: this implies that if we replace the rough skew with the approximation given by Equation (5.14) we would find a lower value for the skew stickiness ratio.
In addiction, we also report the values of the residual sum of squares (SSE) and the coefficient of determination (R^2). They are both statistical indicator of how good our data are described by a linear regression model. In particular we define:

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{5.15}$$

and

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{5.16}$$

where y is the vector of the observed values of $\delta\sigma$, $\hat{y}$ is the vector of the values of $\delta\sigma$ predicted by the model, and $\bar{y}$ is the mean value.
The values of these statistical indicator are reported in Table 5.3.

	1 day	1 week	2 weeks	1 month
SSE	0.0049	0.0044	0.0042	0.0037
R^2	0.2294	0.2522	0.2497	0.2500

Table 5.3: Values of the residual sum of squares and the R^2 coefficient for maturity $T \in \{1 \text{ day}, \ 1 \text{ week}, \ 2 \text{ weeks}, \ 1 \text{ month}\}$.

If we define $b_1 := \dfrac{\text{cov}(x,y)}{\text{var}(x)}$ and $b_0 := \bar{y} - b_1\bar{x}$ then

$$\hat{y} = b_0 + b_1 x. \tag{5.17}$$

Thus, it is possible to write Equation (5.16) as

$$R^2 = \frac{\text{cov}^2(x,y)}{\text{var}(x)\text{var}(y)} \tag{5.18}$$

that allows us to interpret the R^2 coefficient as a statistical measure that indicates the proportion of the variance in the dependent variable y that is predictable from the independent variable x. In this case we observe that only 25% (more or less) is generally captured by model, then the stickiness ratio can predict only a fraction of the variation of the implied volatility knowing the variation of the spot.

Chapter 6

Conclusions and future developments

The rough volatility models are famous for their capability in reproducing short time explosion of the at the money skew of Plain Vanilla options. Among these models, we focused on the rough Heston model. It is a challenging task to find an efficient simulation scheme for this model due to its non Markovian structure and the singularity of the kernel, indeed, at the moment, the only one available scheme is the Hybrid Quadratic Exponential (HQE) of Gatheral. One other possibility is to work with the Markovian lift which is a multi-factor stochastic volatility model which converges to the rough Heston when the number of approximating factors n goes to infinity. In this setting we analyzed the Moment Matching scheme proposed by Bayer and Breneis.
Thanks to the fact that the characteristic function of the log-price is known in semi-closed form we were able to price Plain Vanilla options using the Adam's scheme and therefore we used these prices as benchmark for the simulation schemes of our interest.
Our analysis showed that both schemes are able to fit the implied volatility surface and the skew explosion typical of rough models. In particular, the HQE scheme and the Moment Matching scheme with 3 approximating factors seems to share the same behaviour in terms of the skew both for Plain Vanilla and for Forward Start options.
We investigated also the Hurst index produced by the sample path of the variance process and we verified that both present a rough behaviour with an Hurst index way below 0.5. Nevertheless, the variance process generated by the HQE scheme reaches very often the zero level while the Moment Matching always produces a positive variance.
To understand the cause of this behavior, in the future, it will be interesting to investigate if there exists a condition on the parameters of the model which guarantees that the variance always remains positive both for the rough He-

ston model and for the Markovian lift.
The other advantage that we noticed in using the Moment Matching is the lower computational time that requires to achieve a satisfactory level of accuracy. This is a clear consequence of the Markovian structure of the underlying model.
The last object that we analyzed was the skew stickiness ratio. We faced some difficulties in the empirical validation of the theoretical result on the short time limit that we obtained. In particular, our empirical analysis was based on the Moment Matching scheme which is built on an approximation of the rough Heston model. Further research could study a fast algorithm to obtain the conditional characteristic function in order to compute the real stickiness ratio generated by the original model. Finally, we also suggest to deepen the relation between the local volatility and the skew stickiness ratio that we found in the sketched proof in Appendix B.

Appendix A

General case

Consider the dynamics

$$dX_t = -\frac{1}{2}V_t dt + \sqrt{V_t}dB_t$$

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\lambda(\theta - V_s)ds + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\sigma\sqrt{V_s}dW_s$$

where $\alpha = \frac{1}{2} + H \in \left(\frac{1}{2}, 1\right)$, $\theta, \sigma, \lambda > 0$ and $\rho dt = d\langle W, B\rangle_t$ with $\rho \in [-1, 1]$. Since $\lambda > 0$ the kernel which appears in the dynamics of the forward variance curve (see Proposition 5.0.1) is

$$k(x) = \sigma x^{\alpha-1}\sum_{n\geq 0}\frac{(-\lambda x^\alpha)^n}{\Gamma((n+1)\alpha)}.$$

Thus, doing the same computations that we did for $\lambda = 0$ we have that in the small time limit:

- $\mathbb{E}\big[(X_T - \mathbb{E}(X_T))^2\big] = \int_0^T \xi(s)ds$
- $\mathbb{E}\big[(X_T - \mathbb{E}(X_T))^3\big] = 3\rho\int_0^T\int_0^t k(t-s)\xi(s)dsdt$
- $E[(dX_t)^2] = \xi(t)dt$
- $\mathbb{E}[d\hat{\sigma}_{t,T}dX_t] = \frac{\rho\xi(t)dt}{2\hat{\sigma}_{t,T}(T-t)}\int_t^T k(s-t)ds.$

Looking at Proposition 2.1 in [27] it holds

$$\xi(t) = V_0 - (V_0 - \theta)\int_0^t \frac{\lambda}{\sigma}k(s)ds.$$

Then we can use it to deduce the estimate

$$\int_0^t k(t-s)\xi(s)ds \sim \frac{V_0\sigma t^\alpha}{\alpha\Gamma(\alpha)} + O(t^{\alpha+1}) =: \tilde{k}(t).$$

So thanks to the previous approximation the SSR can be computed by

$$\mathcal{R}(T) = \frac{1}{T^{1/2}\hat{\sigma}_{0,T}} \frac{\int_0^T k(t)dt \big(\int_0^T \xi(t)dt\big)^{3/2}}{\int_0^T \tilde{k}(t)dt}.$$

Now observe that:

- $\displaystyle\int_0^T k(t)dt \sim \frac{\sigma T^\alpha}{\alpha\Gamma(\alpha)} + O(T^{\alpha+1})$
- $\displaystyle\int_0^T \tilde{k}(t)dt \sim \frac{V_0\sigma T^{\alpha+1}}{\Gamma(\alpha+2)} + O(T^{\alpha+2}).$

Passing to the limit for $T \to 0$ and using Lebesgue differentiation theorem we conclude that

$$\mathcal{R}(T) \to \alpha + 1 = H + \frac{3}{2}.$$

Appendix B

An alternative derivation

Let us define the local volatility $\eta(t,k)$ as the function of time t and strike k that solves the Dupire equation (see [23])

$$\eta^2(t,k) := \frac{2\partial_t C(t,k)}{k^2\partial_k^2 C(t,k)} \tag{B.1}$$

where $C(t,k)$ is the price of a Call option with maturity t and strike k. As usual we indicate with $\sigma(t,k)$ the implied volatility of a Call option with maturity t and strike k.
If we indicate with $B(t,k,\sigma)$ the price in the Black model with volatility σ, then using the relation

$$B(t,k,\sigma(t,k)) = C(t,k) \tag{B.2}$$

we can write the Dupire equation in terms of the implied volatility. In particular we obtain

$$\eta^2 = \frac{\sigma^2 + 2\sigma t\partial_t\sigma}{1 + 2k\sqrt{t}(y+\sigma\sqrt{t})\partial_k\sigma + k^2\sigma t\partial_k^2\sigma + k^2 ty(y+\sigma\sqrt{t})(\partial_k\sigma)^2} \tag{B.3}$$

where $y = -\left(\frac{1}{\sigma\sqrt{t}}\log(k) + \frac{1}{2}\sigma\sqrt{t}\right)$. In [21] the authors show that writing the Equation (B.3) in the coordinate system

$$\theta(t) = t, \quad \zeta(t,k) = t^{H-\frac{1}{2}}\log(k) \tag{B.4}$$

they are able to derive the asymptotic relation between the skew of the local and the implied volatility

$$\partial_\zeta\eta(0,0) = \left(H + \frac{3}{2}\right)\partial_\zeta\sigma(0,0). \tag{B.5}$$

The idea is to use this result to prove the asymptotic behaviour of the skew stickiness ratio.
Starting from the dynamics of the log-price $X_t := \log(S_t)$

$$dX_t = -\frac{1}{2}V_t dt + \sqrt{V_t}dB_t, \quad X_0 = 0 \tag{B.6}$$

where B_t is a Brownian motion under the risk neutral measure we can define the skew stickiness ratio observed at time t for a maturity T as

$$\mathcal{R}_t(T) := \frac{\mathbb{E}[(\sigma_t(T, S_t) - \sigma_0(T, S_0))|S_t - S_0]}{\partial_k \sigma_0(T, S_0)(S_t - S_0)} \tag{B.7}$$

where $\sigma_t(T, k)$ is the implied volatility observed at time t of a Call option with maturity T and strike k.
Then the quantity of our interest is given by

$$\mathcal{R} := \lim_{t\to 0}\lim_{T\to t} \mathcal{R}_t(T). \tag{B.8}$$

Remark 27. It is not clear at all if it is possible to invert the two limits. Indeed in [13] Bergomi defines $\mathcal{R} := \lim_{T\to 0}\lim_{t\to 0} \mathcal{R}_t(T)$.

First of all we pass into the coordinate system (θ, ζ) given by Equation (B.4). In this system we can observe that

$$\partial_k = \theta^{H-\frac{1}{2}} e^{-\zeta\theta^{\frac{1}{2}-H}} \partial_\zeta. \tag{B.9}$$

Thus, letting $T \to t$ we obtain

$$\lim_{t\to 0} \mathcal{R}_t(t) = \lim_{t\to 0} \frac{\mathbb{E}[\sigma_t(t, \zeta(t, S_t)) - \sigma_0(t, 0)|S_t]}{t^{H-\frac{1}{2}} \partial_\zeta \sigma_0(t, 0)(S_t - S_0)}. \tag{B.10}$$

Recall that the local volatility can be computed as the conditional expectation of the stochastic volatility and this holds also in the case of rough Heston model as shown in [21], thus

$$\eta_t^2(\theta(T), \zeta(T, k)) = \mathbb{E}_t[V_T|S_T = k]. \tag{B.11}$$

Furthermore, by the chain rule, we obtain

$$\partial_k \sigma_0^2(\theta, \zeta) = 2\sigma_0(\theta, \zeta)\partial_k \sigma_0(\theta, \zeta) \tag{B.12}$$

so that

$$\lim_{t\to 0} \mathcal{R}_t(t) = \lim_{t\to 0} \frac{\mathbb{E}[\sigma_t^2(t, \zeta(t, S_t)) - \sigma_0^2(t, 0)|S_t]}{2\sigma_0(t, 0)t^{H-\frac{1}{2}} \partial_\zeta \sigma_0(t, 0)(S_t - S_0)}. \tag{B.13}$$

To continue we have to show that $\lim_{t\to 0}\zeta(t,S_t)=0$.
By definition

$$\zeta(t,S_t) = t^{H-\frac{1}{2}}\left(\frac{-1}{2}\int_0^t V_s ds + \int_0^t \sqrt{V_s}dB_s\right) \tag{B.14}$$

$$= t^{H+\frac{1}{2}}\left(\frac{c}{t}\int_0^t V_s ds + \frac{1}{t}\int_0^t \sqrt{V_s}dB_s\right) \tag{B.15}$$

$$= t^{H+\frac{1}{2}}(\mathbf{I}_1+\mathbf{I}_2) \tag{B.16}$$

It is straightforward to prove that $t^{H+\frac{1}{2}}\mathbf{I}_1 \to 0$ when $t\to 0$ (Use Lebesgue differentiation theorem and $H+\frac{1}{2}>0$).
For small time it is true the estimate $\int_0^t \sqrt{V_s}dB_s \approx \sqrt{V_0}B_t = \sqrt{V_0}\sqrt{t}\mathcal{N}$ where $\mathcal{N}$ is a standard Gaussian.
Then we obtain

$$t^{H+\frac{1}{2}}\mathbf{I}_2 \approx t^H\sqrt{V_0}\mathcal{N}$$

and since $H>0$ we can conclude that

$$\lim_{t\to 0}\zeta(t,S_t)=0 \quad a.s.$$

Then we obtain

$$\begin{aligned}
\mathcal{R} &= \lim_{t\to 0}\frac{\mathbb{E}[\sigma_t^2(t,\zeta(t,S_t))-\sigma_0^2(t,0)|S_t]}{2\sigma_0(t,0)t^{H-\frac{1}{2}}\partial_\zeta\sigma_0(t,0)(S_t-S_0)}\\
&= \lim_{t\to 0}\frac{\eta_0^2(t,\zeta(t,S_t))-\eta_0^2(t,0)}{2\sigma_0(t,0)t^{H-\frac{1}{2}}\partial_\zeta\sigma_0(t,0)(S_t-S_0)}\\
&= \frac{2\eta_0(0,0)\partial_\zeta\eta_0(0,0)}{2\sigma_0(0,0)\partial_\zeta\sigma_0(0,0)}\\
&= H+\frac{3}{2}
\end{aligned}$$

where we use the fact that $\eta_t^2(t,S_t)=\sigma_t^2(t,S_t)=V_t$ for all $t\geq 0$.

Remark 28. The strong assumption we made in this derivation is the applicability of the chain rule. This is likely related to certain regularity conditions on the volatility process.

Appendix C

The Quadratic Exponential scheme

The Quadratic Exponential scheme, is a simulation scheme that was developed by Andersen in [5] to guarantee that the variance process

$$dv_t = \lambda(\theta - v_t)dt + \sigma\sqrt{V_t}dW_t$$

remains non-negative at each step t_n. First of all, we consider the ratio

$$\psi_n = \frac{\text{var}[v_n|\mathcal{F}_{n-1}]}{(\mathbb{E}[v_n|\mathcal{F}_{n-1}])^2}$$

then we distinguish two cases:
If $\psi_n \leq 2$ then

$$v_n = \alpha_n(\beta_n + Z_n)^2$$

where Z_n is a standard Gaussian and

$$\beta_n^2 = \frac{2}{\psi_n} - 1 + \sqrt{\frac{2}{\psi_n}}\sqrt{\frac{2}{\psi_n} - 1}; \quad \alpha_n = \frac{\mathbb{E}[v_n|\mathcal{F}_{n-1}]}{1+\beta_n^2}.$$

On the contrary, if $\psi_n \geq 1$ then

$$v_n = -\mathbb{I}_{(U_n<p_n)}\gamma_n \log\left(\frac{U_n}{p_n}\right)$$

with $U_n \sim \text{Unif}(0,1)$ and

$$p_n = \frac{2}{1+\psi_n}; \quad \gamma_n = \frac{1}{2}\mathbb{E}[v_n|\mathcal{F}_{n-1}](1+\psi_n).$$

It is straightforward to check that means and variances are correctly matched. Moreover, since the two regions overlap Andersen suggests to use the value $\frac{3}{2}$ as watershed.
The values of $\alpha_n,\ \beta_n,\ \gamma_n$ and p_n are obtained applying a moment matching procedure.

Bibliography

[1] Eduardo Abi Jaber. Lifting the Heston model. *Quantitative finance*, 19(12):1995–2013, 2019.

[2] Eduardo Abi Jaber and Omar El Euch. Multifactor approximation of rough volatility models. *SIAM journal on financial mathematics*, 10(2):309–349, 2019.

[3] Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine Volterra processes. 2019.

[4] Elisa Alòs and David Garcia Lorite. *Malliavin calculus in finance: Theory and practice*. Chapman and Hall/CRC, 2021.

[5] Leif BG Andersen. Efficient simulation of the Heston stochastic volatility model. *Available at SSRN 946405*, 2007.

[6] Kendall Atkinson. *An introduction to numerical analysis*. John wiley & sons, 1991.

[7] Kendall Atkinson and Weimin Han. *Theoretical numerical analysis*, volume 39. Springer, 2005.

[8] Christian Bayer and Simon Breneis. Efficient option pricing in the rough Heston model using weak simulation schemes. *arXiv preprint arXiv:2310.04146*, 2023.

[9] Christian Bayer and Simon Breneis. Markovian approximations of stochastic Volterra equations with the fractional kernel. *Quantitative Finance*, 23(1):53–70, 2023.

[10] Christian Bayer and Simon Breneis. Weak Markovian approximations of rough Heston. *arXiv preprint arXiv:2309.07023*, 2023.

[11] Christian Bayer, Peter K Friz, Paul Gassiat, Jorg Martin, and Benjamin Stemper. A regularity structure for rough volatility. *Mathematical Finance*, 30(3):782–832, 2020.

[12] Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. Hybrid scheme for Brownian semistationary processes. *Finance and Stochastics*, 21:931–965, 2017.

[13] Lorenzo Bergomi. Smile dynamics IV. *Available at SSRN 1520443*, 2009.

[14] Lorenzo Bergomi. *Stochastic volatility modeling*. CRC press, 2015.

[15] Charles K Birdsall and A Bruce Langdon. *Plasma physics via computer simulation*. CRC press, 2018.

[16] Fischer Black. Proceedings of the American Statistical Association. Business and Economic Statistics Section, 1976.

[17] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

[18] Giorgia Callegaro, Martino Grasselli, and Gilles Pagès. Fast hybrid schemes for fractional Riccati equations (rough is not so tough). *Mathematics of Operations Research*, 46(1):221–254, 2021.

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

[20] Laure Coutin and Monique Pontier. Approximation of the fractional brownian sheet via Ornstein-Uhlenbeck sheet. *ESAIM: Probability and Statistics*, 11:115–146, 2007.

[21] Enrico Dall'Acqua, Riccardo Longoni, Andrea Pallavicini, et al. Rough-heston local-volatility model. *International Journal of Theoretical and Applied Finance (IJTAF)*, 26(06n07):1–18, 2023.

[22] Kai Diethelm, Neville J Ford, and Alan D Freed. Detailed error analysis for a fractional Adams method. *Numerical algorithms*, 36:31–52, 2004.

[23] Bruno Dupire et al. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[24] Omar El Euch. *Quantitative Finance under rough volatility*. PhD thesis, Sorbonne université, 2018.

[25] Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. *Mathematical Finance*, 29(1):3–38, 2019.

[26] William Feller. Two singular diffusion problems. *Annals of mathematics*, pages 173–182, 1951.

[27] Martin Forde, Stefan Gerhold, and Benjamin Smith. Small-time, large-time, and $H \to 0$ asymptotics for the rough Heston model. *Mathematical Finance*, 31(1):203–241, 2021.

[28] Jim Gatheral. Efficient simulation of affine forward variance models. *Risk, February*, 2022.

[29] Jim Gatheral and London Quantminds. Skew-stickiness under rough volatility. 2023.

[30] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.

[31] Gatheral J, Jaisson Th, and Rosenbaum M. Volatility is rough (2018). *Quantitative finance*.

[32] Thibault Jaisson and Mathieu Rosenbaum. Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. 2016.

[33] Olav Kallenberg and Olav Kallenberg. *Foundations of modern probability*, volume 2. Springer, 1997.

[34] Jean-François Le Gall. *Brownian motion, martingales, and stochastic calculus*. Springer, 2016.

[35] Gytenis Lileika and Vigirdas Mackevičius. Second-order weak approximations of CKLS and CEV processes by discrete random variables. *Mathematics*, 9(12):1337, 2021.

[36] Robert I McLachlan and G Reinout W Quispel. Splitting methods. *Acta Numerica*, 11:341–434, 2002.

[37] Ivan Nourdin et al. *Selected aspects of fractional Brownian motion*, volume 4. Springer, 2012.

[38] El Euch Omar, Fukasawa Masaaki, and Rosenbaum Mathieu. The microstructural foundations of leverage effect and rough volatility. *arXiv preprint arXiv:1609.05177*, 2016.