Università degli Studi di Padova

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Laurea Triennale in Matematica

Tesi di Laurea

Financial Modeling under Rough and Fractional Stochastic Volatility

Relatrice: **Prof.ssa Giorgia Callegaro**
Candidato: **Antonio Dapporto**
Matricola: **1191151**

Contents

Acknowledgements

First and foremost, I wish to express my deepest gratitude to my supervisor, Giorgia Callegaro, for her continuous support throughout this project. I thank her for having proposed to me the topic of this thesis, for her time, feedbacks and every other aspect of her insights to make this dissertation as thorough as possible.

I am also particularly indebted to all of my family, for their unconditional support and for the confidence they have in me.

Introduction

Fifty years have passed since the publication in 1973 of the seminal article 'The Pricing of Options and Corporate Liabilities' by Fischer Black and Myron Scholes, where the authors presented the Black&Scholes Option Pricing model, bringing a new quantitative approach to pricing options. The outstanding increase in derivatives investing during the 1980's and 1990's can be in part attributed to new sofisticated quantitative techniques for pricing and hedging financial derivatives that followed the Black&Scholes model. The theoretical importance of this model, which is based on the assumption that the underlying follows the dynamic of a geometric Brownian motion, is that it yields an equivalence between the problems of no arbitrage and delta hedging, and the problem of solving a parabolic partial differential equation. More precisely, it can be shown that the Black&Scholes model is complete and arbitrage free, this meaning that (almost) every European derivative[1] is replicable in a unique way, and that the unique replicating strategy is defined by the solution of a parabolic PDE. Moreover, in the case of European Put and Call options, it brings an explicit solution for that PDE. Despite these profound characteristics, the Black&Scholes model lies upon some unrealistic assumptions, and one of these, i.e. the fact that the volatility is considered to be constant, is the starting point of this thesis. By allowing non-constant volatility $\sigma = (\sigma_t)_{t \geq 0}$ in pricing models we are introducing various levels of complexity. This leads to a first distinction between local volatility models, where the volatility is assumed to be a deterministic function of the time and of the price of the underlying, and more advanced models where the volatility itself is a stochastic process, called stochastic volatility (SV) models. In the recent literature there has been an increasing interest towards these SV models, because they are able to better represent many stylized facts about volatility (i.e. empirical observations from the time series of the price of the

---

[1]We call European derivative a derivative whose payoff function depends only on the value of the underlying at the expiration of the contract. We used the adverb 'almost' because some technical assumption (although not very restrictive in the practice) on the payoff function must be required.

underlying asset) and also the so-called implied volatility surface. We will recognize, however, that assuming the (classical) Brownian motion to be the driving stochastic process poses some limitations to the capability of modeling a number of properties of both historical and implied volatility. This will lead us to adopt a different driving stochastic process, called 'fractional Brownian motion'. After investigating some of its properties, we focus on a class of stochastic volatility models recently developed, called 'rough SV models', whose name derives from the irregularity of the sample paths of the volatility process they generate. The study of these models started with the influential article 'Volatility is Rough' by Gatheral, Jaisson and Rosenbaum in 2014, to which we dedicate particular attention, and is currently undergoing great developement. This led to a number of new 'rough' models proposed in the literature. One of these is the 'rough Heston model', based on the well-known classical stochastic volatility model proposed in 1993 by Heston, that we will analyze in detail. Finally, in the last Chapter, we will test the two models just mentioned from the point of view of Option Pricing, discussing their simulation, and a recently developed approach based on Fourier series expansion, called COS methos. Python and Matlab codes for the implementation will be used in parallel with the theoretical results.

Chapter 1

Overview of Volatility modeling

In financial markets, volatility, usually denoted by σ, is a measure of the amount of fluctuation in asset prices and so it is often understood or perceived as a measure of riskiness of the asset or uncertainty about its future value. Since there are various ways one can quantify this risk, a comprehensive and precise study of this quantity is needed. For this reason there are various definitions of volatility, based on the framework and on the purpose of the model.

In this chapter we introduce the notion of volatility. We give at first the definition in the context of the Black and Scholes model, we show the various ways to estimate it based on empirical observations and explain why this model fails to reproduce observed volatility. We then introduce the important concept of implied volatility based on risk-neutral data and the related volatility surface and we give a hint towards the various ways one can model volatility to better understand both its random behaviour and the observed volatility surface. Finally, we give a detailed description of the celebrated Heston model, and show possible alternative models.

We fix once for all a filtered probability space $(\Omega, F, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ satisfying the usual conditions [1].

1.1 Actual volatility

We consider the volatility *actual* when it is given as an input into the model. It is the most fundamental notion of volatility, and for this reason it is also simply called *volatility*.

We start by the simplest circumstance, assuming that the underlying stock price process follows the dynamics of the Black and Scholes model under the

---

[1]$(\Omega, F, \mathbb{P})$ is complete and $(\mathcal{F}_t)_{t \geq 0}$ is complete and right-continuous.

physical measure $\mathbb{P}$:

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \tag{1.1}$$

where $\mu \in \mathbb{R}$ is the average rate or return, $\sigma \in \mathbb{R}_{>0}$ is the (actual) volatility and $B = (B_t)_{t \geq 0}$ is a real Brownian motion on $(\Omega, F, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$. We give various motivation that explain the role of the volatility σ in Equation (1.1):

1. In order to explain the role of σ as a measure of the amount of fluctuation in the stock price, we consider the solution of (1.1), known as *geometric Brownian motion*. Assuming $S_t \geq 0$ for each $t \geq 0$, by applying the Itô formula to $(\log(S_t))_{t \geq 0}$ we have that

$$S_t = S_0 e^{\sigma B_t + \left(\mu - \frac{\sigma^2}{2}\right)t}, \tag{1.2}$$

   which tells us that, the larger σ, the larger the fluctuations (generated by the stochastic term σB_t) of S_t.

2. To clarify the role of σ as a measure of riskiness of the asset, we can at first consider its contribution in the variance function, traditionally considered as a measure of stock's riskiness. Recall that, since Brownian motion is a Gaussian process, we have that:

$$\mathbb{E}[S_t] = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t} \mathbb{E}[e^{\sigma B_t}] = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\frac{1}{2}\sigma^2 t} = S_0 e^{\mu t}. \tag{1.3}$$

   On the other hand, we have

$$\mathbb{E}[S_t^2] = S_0^2 e^{\left(\mu - \frac{\sigma^2}{2}\right)2t} \mathbb{E}[e^{2\sigma B_t}] = S_0^2 e^{\left(\mu - \frac{\sigma^2}{2}\right)2t} \mathbb{E}[e^{\frac{1}{2}4\sigma^2 t}] = S_0^2 e^{2\mu t + \sigma^2 t}. \tag{1.4}$$

   Hence,

$$\mathrm{Var}[S_t] = S_0^2 e^{2\mu t}\left(e^{\sigma^2 t} - 1\right). \tag{1.5}$$

   Therefore, higher values of σ generate higher values of $\mathrm{Var}[S_t]$.

3. To assess the role of σ in uncertainty over future stock value, we proceed heuristically and consider a discretized version of Equation (1.1) with time lag Δt. If Δt is small enough, recalling the purely formal equality $dW_t^2 = dt$ we can consider the following as a good approximation of the stock's rate of return between t and $t + \Delta t$:

$$\frac{S_{t+\Delta t} - S_t}{S_t} \approx \mu \Delta t + \sigma \varepsilon \sqrt{\Delta t}, \tag{1.6}$$

   where $\varepsilon \sim \mathcal{N}(0, 1)$, which shows that $\sigma \sqrt{\Delta t}$ is approximately equal to the standard variation of the stock's rate of return. Note that equation (1.6) holds only approximately for Δt small.

4. Another more precise way to show the role of σ is to consider the stock's logarithmic returns between time 0 and t. Applying Ito's lemma to (1.1), we have that

$$\log(S_t) - \log(S_0) = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t \sim \mathcal{N}\left(\left(\mu - \frac{\sigma^2}{2}\right)t,\ \sigma^2 t\right), \quad (1.7)$$

and so $\sigma_t := \sigma\sqrt{t}$ is the standard deviation of the logarithmic returns which explains why we can interpret volatility as an indicator of the uncertainty on future stock values.

It is common practice to consider one year as the unit of time and so $\sigma = \sigma_1 = \sigma_{annually}$ is the standard deviation of the logarithmic returns over 1 year. Another common assumption is that the number of trading days in one year is $P=252$. Thus, $\sigma = \sigma_{1/252}\sqrt{252} = \sigma_{daily}\sqrt{252}$.

1.2 Historical volatility

We now change point of view and focus on how to derive volatility from empirical observations of the underlying asset S. More precisely, we want to estimate volatility based upon the time series of the logarithm returns of the price of S. A special feature of volatility is that it is not directly observable. In fact, the information that is given to us observing financial markets are bid prices (the maximum price that a buyer is willing to pay for a share of stock or other asset), ask prices (the minimum price that a seller is willing to take for the same security) and trading volumes. A common assumption is to consider the market price as the mid price between the bid price and the ask price. Therefore, to get to know the volatility we must rely on this available data. An estimator of volatility based upon the time series of the underlying over some period in the past is called *Historical* or *Realized* Volatility.

Close-to-Close

There are many estimators of volatility, each with their benefits and drawbacks. The simplest one is the so called Close-to-Close estimator (some authors call it Historical Volatility tout-court). Let's suppose we observe the stock price process S at equidistant discrete points $t_0, t_1, ..., t_n$ representing the closing times of $n + 1$ consecutive trading sessions.

S_{t_k} is the closing price[2] at the closing time t_k of k-th trading session, for $k = 0, 1, ..., n$. We thus observe the closing prices $S_{t_0}, ..., S_{t_n}$. Let $\Delta t = t_i - t_{i-1}$.
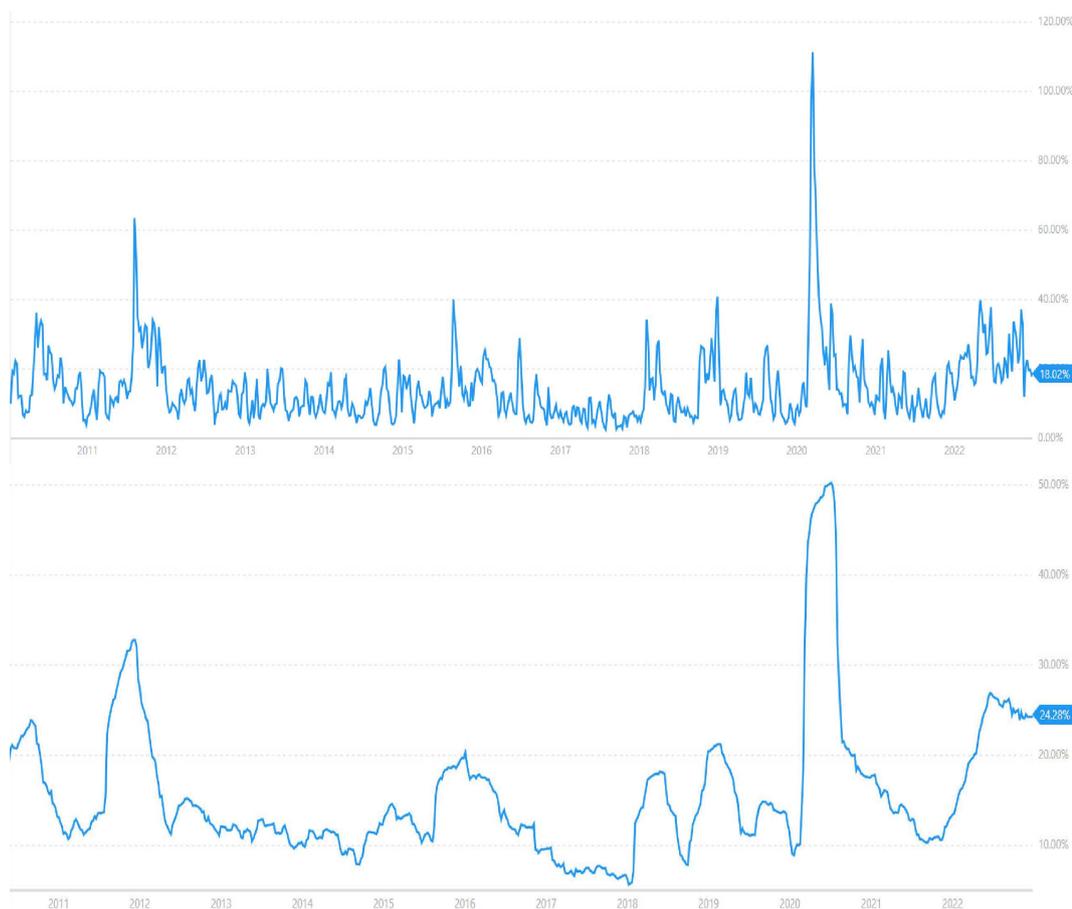
Figure 1.1: Close-to-Close Historical Volatility for S&P500 with n=10 (above) and n=100 (below). Since all returns are equally weighted, while they are in the estimate of volatility, any large return will stay in the estimate of volatility until the 10 (or 100) days have passed. This gives rise to a plateauing of volatility, clearly visible for n=100, and is totally spurious. It has been calculated using the software developed by https://portfolioslab.com/.

In order to estimate σ we use Equation (1.7) from which it's clear that the

---

[2]Notice that the closing price of a trading session does not necessarily coincide with the opening price of the next trading session. This is due to various activities occurring during the time of the day when stock market remains closed, one of them being After-Hours Trading.

log-increments

$$\rho_i = \log(S_{t_i}) - \log(S_{t_{i-1}}), \qquad i = 1, ..., n$$

are independent (since so are the increments of B) and normally distributed with $\text{Var}[\rho_i] = \sigma^2 \Delta t$. An unbiased estimator for $\text{Var}[\rho_i]$ is the corrected sample variance[3]

$$V^* = \frac{1}{n-1} \sum_{i=1}^{n} (\rho_i - \bar{\rho})^2,$$

where $\bar{\rho} = \frac{1}{n} \sum_{i=1}^{n} \rho_i$ is the sample mean. Therefore the Close-to-Close estimator for σ is defined as

$$\sigma_{CC} = \sqrt{\frac{1}{\Delta t \cdot (n-1)} \sum_{i=1}^{n} (\rho_i - \bar{\rho})^2}.$$

Figure 1.1 above picture shows an important feature of the (Historical) volatility: it is not constant. This contradiction with the unvarying σ that appears in the B-S model is the motivation for more complex models discussed next.

Parkinson

Close-to-Close historical volatility is calculated using only stock's closing prices, and in many cases it is not precise enough. Stock prices could jump considerably during a trading session, and return to the opening value at the end. That means that a big amount of price information is not taken into account.
Parkinson's Historical Volatility estimator, formulated in 1980 by Parkinson in [Par80], uses the stock's high and low price of the trading session. That is useful as Close-to-Close prices could show little difference between opening and closing price, while large price movements could have happened during the session. It is defined as:

$$\sigma_{Parkinson} = \sqrt{\frac{1}{4n \log(2)} \sum_{i=1}^{n} \log\left(\frac{h_i}{l_i}\right)^2},$$

---

[3]This means that $\mathbb{E}[V^*] = \text{Var}[\rho_i]$.

where h_i and l_i are the high and low price on i-th session respectively. However, it is common to observe jumps in the stock's opening and closing price as markets are most active during the opening and closing of a trading session and these jumps are not captured by Parkinson volatility estimator (unless they coincide with the high or low price measured) and thus it tends to underestimate the volatility.

Garman-Klass

Garman-Klass's Historical Volatility, proposed by Garman and Klass in 1980 in [GK80], improves Parkinson's by incorporating all commonly available prices of the current trading session: opening (o_i), high (h_i), low (l_i) and closing (c_i) prices (OHLC).

$$\sigma_{GK} = \sqrt{\frac{1}{2n} \sum_{i=1}^{n} \log\left(\frac{h_i}{l_i}\right)^2 - \frac{2\log(2) - 1}{n} \log\left(\frac{c_i}{o_i}\right)^2}.$$

Despite being more effective than both the Close-to-Close and the Parkinson estimators, this method is not robust for opening jumps in price and trend movements. For this reason, more efficient method for assessing historical volatility that takes into account price trends have been subsequently developed, like the one proposed by Rogers and Satchell

Rogers-Satchell

The three estimators we have seen so far, are not efficient for estimating volatility when the asset price is trending, i.e. when the average rate of return is particularly high. Rogers and Satchell proposed in 1994 [RSY94] a volatility estimator that is particularly tailored for this circumstance and that, like Garman-Klass, takes into account all of the prices that characterize a trading session (OHLC). As a result, it provides generally a more precise estimation of volatility (i.e. the variance of the estimator is decreased), particularly when the underlying is trending strongly.

$$\sigma_{RS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(log\left(\frac{h_t}{c_t}\right) log\left(\frac{h_t}{o_t}\right) + log\left(\frac{l_t}{c_t}\right) log\left(\frac{l_t}{o_t}\right) \right)}$$

The main disadvantage of this estimator is that it does not take into account price movements between trading sessions. It means that it tends to underestimate volatility, since price jumps periodically occur in the market precisely at the moments between sessions.

Yang-Zhang

A more comprehensive estimator that also considers the gaps between sessions was developed, based on the Rogers-Satchell, by Yang and Zhang in 2000. It handles, like the Rogers-Satchell estimator, both opening jumps and trending prices but ,unlike the former, it is a minimum-variance and unbiased estimator.

We avoid the formula, which is a little intricate. It can be found in [YZ00].

1.3 Implied volatility

Recall that the Black and Scholes model is complete and there exist a unique risk-neutral measure $\mathbb{Q}$, under which the price of a European Call option at time $t = 0$ with expiration date T is defined by:

$$C_{BS}(\sigma, S_0, K, T, r) = S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2), \qquad (1.8)$$

where Φ is the standard normal cumulative distribution function,

$$d_1 = \frac{1}{\sigma\sqrt{T}}\Big[\log\Big(\frac{S_0}{K}\Big) + \Big(r + \frac{\sigma^2}{2}\Big)T\Big], \qquad d_2 = d_1 - \sigma\sqrt{T}.$$

Actually, the price can also be expressed in the form

$$C_{BS}(\sigma, S_0, K, T, r) = S_0 C_{BS}^\star(\sigma, k, T, r), \qquad (1.9)$$

where $k = \log\Big(\frac{S_0}{K}\Big)$, and $C_{BS}^\star$ is a function whose expression can be easily deduced from formula (1.8). The number $m = \frac{S_0}{K}$ is called "moneyness" of the option and we say that the Call option is "in the money" if $m > 1$, since we are in a position of potential profit; otherwise we say it is "out of the money". We recall that $\nu = \partial_\sigma C_{BS}$ is called "Vega" and measures the sensitivity of the Call option with respect to the volatility. We have

$$\begin{aligned}
\nu = \partial_\sigma C_{BS} &= g^{'}(d_1)\partial_\sigma d_1 + Ke^{-rT}\Phi^{'}(d_1 - \sigma\sqrt{T})\sqrt{T} \\
&= Ke^{-rT}\Phi^{'}(d_1 - \sigma\sqrt{T})\sqrt{T} \qquad (1.10) \\
&= S_t\sqrt{T}\Phi^{'}(d_1) > 0
\end{aligned}$$

where we used the fact that

$$g(d) := S_t\Phi(d) - Ke^{-rT}\Phi(d - \sigma\sqrt{T}) \quad \text{is such that} \quad g^{'}(d_1) = 0 \qquad (1.11)$$

and that

$$S_t \Phi^{'}(d_1) = K e^{-rT} \Phi^{'}(d_1 - \sigma\sqrt{T}) \tag{1.12}$$

Equation (1.10) shows that the Vega is always strictly positive. Intuitively this is due to the fact that an option is a contract giving a right, but not an obligation, therefore one takes advantage of the greater riskiness of the underlying asset. It follows that the price of a Call option[4] is a strictly increasing function of the volatility and C_{BS} is an invertible function of the volatility. Thus, having fixed all other parameters, there exist a unique value $\hat{\sigma}^{impl}$ of the volatility that, plugged into (1.8), produces an observed option price. We call this value "implied volatility".

Having fixed the parameters S_0 and r, let us suppose that, for each value (k, T), we observe in the real market the prices $C^m(k, T)$ of european Call options. Then we can define, as explained above, the "implied volatility function" $(k, T) \mapsto \hat{\sigma}^{impl}(k, T)$ whose graph is called "implied volatility surface". Every section, with T fixed, of the implied volatility surface is usually called "smile". Generally we can say that market quotation tends to give more value (greater implied volatility) to the extreme cases "in" or "out of the money". This reflects the fact that some situations in the market are perceived as more risky.
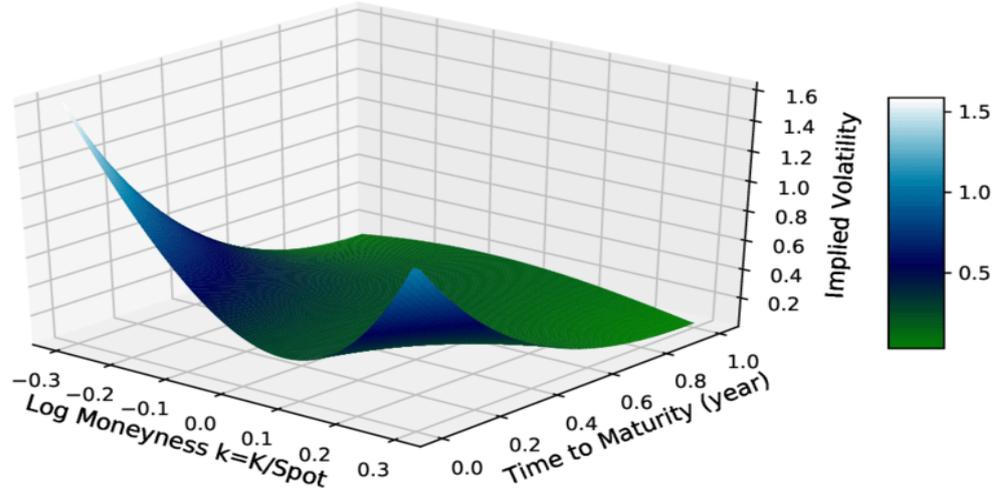

Figure 1.2: Implied Volatility surface of the SPX index on 17/06/2020, see [Yos20].

Clearly, since S_0 is fixed and thanks to relation (1.9) we can also consider

---

[4]The payoff function of a Call Option is given by $f(x) := [x - K]^+$, where K is the strike price.

$\hat{\sigma}^{impl}$ as a function of (K, T). In this case every section, for T fixed, of the implied volatility surface is called "skew" because we tipically observe an appreciable slant to that curve, with higher implied volatility for lower strikes.

In a study by Szakmary et al. of 2003 [Sza+03], using data from 35 futures options from 8 separate exchanges, the authors conclude that implied volatility outperforms historical volatility as a predictor of the subsequently observed volatility in the underlying futures prices over the remaining life of the option.

However, given one smile for a fixed expiration, little can be said about the process generating it. An important feature of the volatility surface, more sensitive to the choice of volatility dynamics between models, is the term structure of at-the-money (ATM) volatility skew, defined at time to expiration T as

$$\psi(T) = \left| \frac{\partial}{\partial k} \hat{\sigma}^{impl}(k, T) \right|_{k=0}.$$

Empirically, as shown in figure below, we tipically observe that $\psi(T)$ is proportional to $T^{-\alpha}$ for some $0 < \alpha < 1/2$.
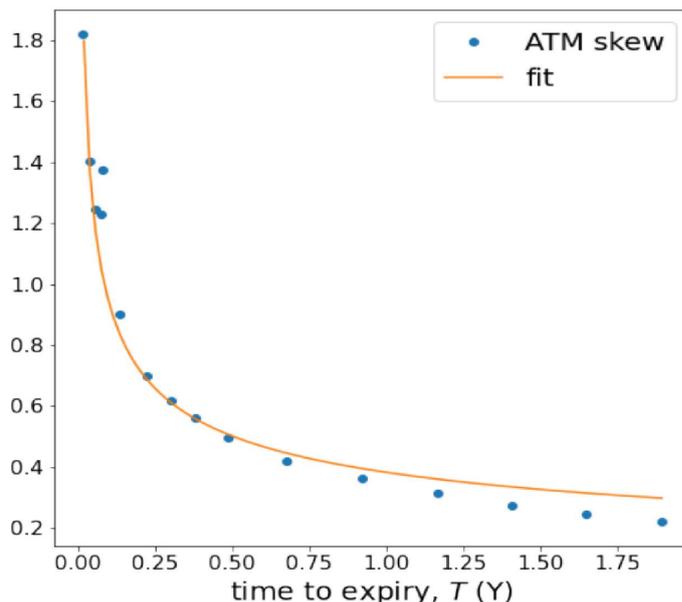

Figure 1.3: The blue dots are the estimates of the SPX ATM volatility skew as of 17/06/2020, see [Yos20]. Approximation of the ATM skew is done using a numerical symmetric derivative; the orange curve is the power-law $\psi(\tau) = 0.4\tau^{-0.4}$.

Clearly, if there exist a unique correct Black and Scholes model and if people

priced option correctly according to formula (1.8) then the implied volatility function should be constant. However, a stylized fact, derived from years of observations of financial markets, suggest that these assumptions are unrealistic, and so we expect the surface to be curved, as in Figure 1.2.. Therefore, as Riccardo Rebonato famously stated, implied volatility is "the wrong number to put in the wrong formula to get the right price"(Rebonato, 1999). This motivates the need to go beyond a Black-Scholes world.

1.4 Beyond Black-Scholes

What we have seen so far from Section 1.2 and 1.3 about Historical and Implied volatility is that Black and Scholes assumption about volatility being constant is wrong, because it leads to systematic differences between market prices and Black-Scholes prices. In order to improve this model it is necessary to introduce different volatility models. Generally speaking, the models with non-constant volatility can be divided in two groups:

- Volatility is endogenous, i.e. it is described by a deterministic function that depends on the same risk factors of the underlying asset. In this case, the completeness of the market is generally preserved. Local and path-dependent volatility models belong to this category the. We will focus only on local models.

- Volatility is exogenous, i.e. , it is described by a process that is driven by some additional risk factors (e.g. Brownian motions or even more complicated ones like the fractional stochastic processes). In this case the corresponding market model is generally incomplete since we have more risk factors than risky assets (remember that volatility is not considered to be directly observable or tradeable in the market). These are called stochastic volatility models and they are the main topic of this thesis.

In what follows, we suppose $(\mathcal{F}_t)_{t \geq 0}$ to be a Brownian filtration, that is $\mathcal{F}_t = \sigma\{(B_s)_{s \leq t}\}$, $t \geq 0$, is the sigma-algebra generated by $(B_s)_{s \leq t}$, where B is a Brownian motion on the probability space $(\Omega, F, \mathbb{P})$.

1.4.1 Local volatility

In a local volatility model the dynamics of the underlying asset under the physical measure $\mathbb{P}$ is given by:

$$dS_t = \mu(S_t, t)S_t dt + \sigma(S_t, t)S_t dB_t \qquad (1.13)$$

where $\mu : \mathbb{R}^2_{>0} \to \mathbb{R}$ and $\sigma : \mathbb{R}^2_{>0} \to \mathbb{R}_{>0}$ are deterministic functions.

We fix a maturity T and we focus on the interval $[0, T]$.

We use $\mathbb{M}^2_{loc}[0, T]$ to indicate the vector space of progressively measurable processes $X = (X_t)_{t \in [0,T]}$ on $(\Omega, F, \mathbb{P}, (\mathcal{F}^B_t)_{t \in [0,T]})$ such that $\int_0^T X_T^2(\omega)dt < \infty$ for almost every $\omega \in \Omega$.

Since we have the same number of risk factors and risky asset, the model is complete and there exist a unique risk-neutral measure $\mathbb{Q}$, under which we have:

$$dS_t = r_t S_t dt + \sigma(S_t, t) S_t dB_t^{\mathbb{Q}} \tag{1.14}$$

where $B^{\mathbb{Q}}$ is the Brownian motion on $(\Omega, F, \mathbb{Q}, (\mathcal{F}^B_t)_{t \in [0,T]})$ defined by $dB_t^{\mathbb{Q}} = dB_t + \lambda_t dt$, where $\lambda_t \in \mathbb{M}^2_{loc}[0, T]$, called *market price of risk*, is defined by

$$dZ_t = -Z_t \lambda_t dB_t, \qquad Z_0 = 1 \tag{1.15}$$

with $Z_t := \frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}^B_t} = \mathbb{E}^{\mathbb{P}}\Big[\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|\mathcal{F}^B_t\Big], \quad t \in [0, T]$.

Since we are only interested in volatility, we assume for simplicity $r_t = r \in \mathbb{R}_{>0}$ constant.

We now want to extend the concept of implied volatility from the black and Scholes world to the local volatility model. Again, having fidex S_0 and r, we look for a deterministic function $\sigma(K, T)$ such that, plugged as a coefficient into the SDE (1.14) (with $\sigma(S_t, t) = \sigma(K, T)|_{(K=S_t, T=t)}$), the model provides exactly the observed market prices C^m of European Call options. More formally, given a function $C^m(K, T)$ with sufficient regularity conditions that will be discussed specified later, we look for a function $\sigma : \mathbb{R}^2_{>0} \to \mathbb{R}_{>0}$ such that

$$C^m(K, T) = B_0 \mathbb{E}^{\mathbb{Q}}[B_T^{-1}(S_T - K)^+] \tag{1.16}$$

where $B_T = e^{rT}$.

Starting from Breeden and Litzenberger [BL78], Dupire [Dup93] solved the problem in 1993, at least theoretically, assuming that $C^m(K, T) : \mathbb{R}^2_{>0} \to \mathbb{R}_{>0}$ is of class $C^{2,1}$. He proved that under this condition there exist a unique function $\sigma(K, T)$ of class $C^{2,1}$ that solves Equation (1.16). This function is given by

$$\sigma(K, T) = \sqrt{\frac{2\left(\frac{\partial}{\partial T}C^m(K, T) + rK\frac{\partial}{\partial K}C^m(K, T)\right)}{K^2 \frac{\partial^2}{\partial K^2}C^m(K, T)}} \tag{1.17}$$

Therefore, defining $\sigma(S_t, t) = \sigma(K, T)|_{(K=S_t, T=t)}$, we have that, under the assumptions on C^m above, there exists a unique local volatility model defined by the SDE (1.13) that is able to reproduce market prices of European Call options.

Unfortunately, formula (1.17) cannot be used in practice since the market prices are known only at a finite number of strikes and maturities. Furthermore, we only used European Call options to define the function σ, and thus our model could fail to represent the market prices of non-Vanilla contingent claims. This explains why there have been serious concerns about the effectiveness and the validity of the model (1.13) and why, when used, they have been unable to predict future movements of volatilities, leading to unrealistic future dynamics of the volatility surface.

1.4.2 Stochastic volatility

Tipically, a stochastic volatility (SV) model under the physical measure $\mathbb{P}$ takes the form

$$dS_t = \mu(S_t, t)S_t dt + \sigma_t S_t dB_t^1 \tag{1.18}$$

$$d\sigma_t = \alpha(\sigma_t, t)dt + \beta(\sigma_t, t)dB_t^2 \tag{1.19}$$

where $B_t = (B_t^1, B_t^2)$ is a 2-dimensional correlated brownian motion, that is:

$$B_t = A\overline{B_t}, \qquad A = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}$$

with $\overline{B_t}$ a standard 2-dimensional Brownian motion[5] and $\rho \in]-1, 1[$. Again, we fix a maturity T and we focus on the interval $[0, T]$.

We recall that, given a market model with d risk factors (the correlated Brownian motions) and $N \leq d$ risky assets, a *market price of risk* is a d-dimensional stochastic process $\lambda \in \mathbb{M}_{loc}^2[0, T]$ such that:

(i) $\lambda_t^i = \frac{\mu_t^i - r_t}{\sigma_t^i}, i = 1, ..., N$

(ii) the solution Z of the SDE

$$dZ_t = -Z_t \bar{A}^{-1} \lambda_t dB_t, \qquad Z_0 = 1 \tag{1.20}$$

is a $\mathbb{P}$-martingale. Here $\bar{A} := \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

It can be shown using Girsanov theorem that there exists a one-to-one correspondence between *risk-neutral measures* $\mathbb{Q}$ and *market prices of risk* λ. Given a market price of risk λ, the associated risk-neutral measure $\mathbb{Q}$ is given

---

[5]This means that B is a 2-dimensional a.s. continuous Gaussian process with zero mean and $\text{Cov}(B_s^1, B_t^2) = \delta_{ij} \min\{s, t\}$, with δ_{ij} the delta of Kronecker.

by $\dfrac{d\mathbb{Q}}{d\mathbb{P}} = Z_T$.

Under suitable conditions (see Karatzas-Shreve, chapter 5) on the real deterministic functions μ, α, β, for fixed initial conditions the theory of SDEs guarantees that a unique solution to the system (1.18)-(1.19) exists. Moreover there exists a (non unique) 2-dimensional market price of risk process

$$\lambda_t = (\lambda_t^1, \lambda_t^2), \qquad \lambda_t^1 = \frac{\mu(S_t, t) - r}{\sigma_t},$$

with associated risk-neautral measure $\mathbb{Q}$. Then we have the following risk-neutral dynamics under $\mathbb{Q}$:

$$dS_t = rS_t dt + \sigma_t S_t dB_t^{1,\mathbb{Q}} \tag{1.21}$$

$$d\sigma_t = \hat{\alpha}(\sigma_t, t)dt + \beta(\sigma_t, t)dB_t^{2,\mathbb{Q}}, \tag{1.22}$$

where $\hat{\alpha}(\sigma_t, t) = \alpha(\sigma_t, t) - \lambda_t \beta(\sigma_t, t)$ and $B_t^{i,\mathbb{Q}}$ is defined by $dB_t^{i,\mathbb{Q}} = dB_t^i + \lambda_t dt$.

It is also possible to derive, under the hypothesis of the Feynman-Kac Theorem, the partial differential equation satisfied by the value function of a European contingent claim of the form $F(S_T, \sigma_T)$, where F is the deterministic payoff function. The risk-neutral price under $\mathbb{Q}$ is equal to $H_t^{\mathbb{Q}} = f(t, S_t, \sigma_t)$, with

$$f(t, s, \sigma) = \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)} F(S_T^{t,s,\sigma}, \sigma_T^{t,s,\sigma})],$$

where $(S_T^{t,s,\sigma}, \sigma_T^{t,s,\sigma})$ is the solution at time T of the system (1.21)-(1.22) with initial values (s, σ) at time t. Then, f is solution of the Cauchy problem:

$$\begin{cases} \frac{\partial f}{\partial t} + \mathcal{L}^\lambda f - rf = 0 & \text{on} \quad [0, T[\times \mathbb{R}_{>0}^2 \\ f(T, s, \sigma) = F(s, \sigma) & \text{on} \quad \mathbb{R}_{>0}^2 \end{cases}$$

where

$$\mathcal{L}^\lambda = \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 f}{\partial s^2} + \rho s \sigma \beta \frac{\partial^2 f}{\partial s \partial \sigma} + \frac{1}{2}\beta^2 \frac{\partial^2 f}{\partial \sigma^2} + rs \frac{\partial f}{\partial s} + \hat{\alpha} \frac{\partial f}{\partial \sigma} \tag{1.23}$$

We emphasize that the dependence on λ of the differential operator (1.23) (given by the presence of λ in the term $\hat{\alpha}$) reflects the fact that we have different pricing PDEs for different risk-neutral measures.

Stochastic volatility models are useful because they model volatility trough time in a relistic way. In fact, since options with different strikes and maturities have different Black-Scholes implied volatilities, the Black&Scholes

model does not give a realistic representation of the dynamic of asset prices. On the other hand, in SV models, the volatility itself is modeled as a continuous Brownian semi-martingale and, as we will see in chapter 3, this is consistent with the observed dynamic of asset prices. However, while stochastic volatility dynamics are more realistic than local volatility dynamics, generated option prices are not consistent with observed European Call option prices and lead to volatility surfaces whose shapes can differ substantially from that of the empirically observed volatility surfaces. Hence, SV models are usually unable to capture the smile (or skew) of the implied volatility surface, and neither are able to fit the term structure of ATM skew $\psi(T)$. Adding jumps to these models (e.g. Merton or Bates model) can help to better fit the volatility surface, but we will not deepen this topic.

1.4.3   Heston model

It is a stylized fact that, at least in equity markets, although the level and orientation of the volatility surface changes over time, the general overall shape of the volatility surface does not change, at least to a first approximation. This suggests that it is desirable to model volatility as a time-homogenous process, that is, a process whose parameters α and β in Equation (1.19) are independent of asset price and time. The model that we now present exhibits this feature.

In the classical Heston stochastic voaltility model formulated by Heston in 1993 [Hes93], the underlying asset and the volatility $\sigma_t = \sqrt{\nu_t}$ have the following dynamics under the physical measure $\mathbb{P}$:

$$dS_t = \mu S_t dt + \sqrt{\nu_t} S_t dB_t^1 \tag{1.24}$$

$$d\nu_t = k(\hat{\nu} - \nu_t)dt + \eta\sqrt{\nu_t}dB_t^2 \tag{1.25}$$

where $\mu \in \mathbb{R}$ and $k, \hat{\nu}, \eta \in \mathbb{R}_{>0}$ are constant parameters, and $B = (B^1, B^2)$ is a 2-dimensional correlated Brownian motion.

In analogy to what we have seen in Section 1.1 for the Black-Scholes model, considering a discretized version of (4.37) as a good approximation, here the process ν can be interpreted (up to the term dt) as the instantaneous variance of the asset's rate of return, and with a small term abuse is often called the variance of the asset. The interest rate r is supposed to be constant. Equation (4.38) was previously suggested by Cox, Ingersoll, Ross (1985) as a model for the short interest rate (the CIR model), and its solution is called "mean-reverting square root process".

Remark 1.1. *For $k > 0$, the drift is positive if $\nu_t < \hat{\nu}$ and it is negative if $\nu_t > \hat{\nu}$, and so the process is "pushed" towards the value $\hat{\nu}$ that can be inter-*

preted as a long-term mean of the variance. The other parameters represent: the drift term μ is the mean rate of return of the asset, k the speed of mean reversion and η the volatility of the variance.

What about existence and uniqueness for the solutions of (4.37)-(4.38)? For what concerns the asset, by Itô's formula we have that the explicit solution of (4.37) is

$$S_t = S_0 \exp\Big(\int_0^t \sqrt{\nu_s} dB_s^1 + \int_0^t \big(\mu - \frac{\nu_s}{2} ds\big)\Big)$$

For the variance process, we have the following remarkable result (see Ikeda-Watanabe [IS89], p. 168):

Theorem 1.1. *For any $\nu_0 \geq 0$, there exists a unique non-negative strong solution to (4.38) starting from ν_0.*

Since, in general, the solution ν can reach the origin, we denote by τ the stopping time defined by

$$\tau = \inf\{t \geq 0 | \nu_t = 0\},$$

with $\tau(\omega) = \infty$ if $\nu_t(\omega) > 0$ for all $t \geq 0$.
We have the so-called Feller condition (see Proposition 6.2.3 in [LL07]):

Proposition 1.1. *For $\nu_0 \geq 0$ we have that:*

- *if $\eta^2 \leq 2k\hat{\nu}$, then $\tau = \infty$ $\quad$ $\mathbb{P}$-a.s.*

- *if $\eta^2 > 2k\hat{\nu} \geq 0$, then $\tau < \infty$ $\quad$ $\mathbb{P}$-a.s.*

In order to define a market price of risk, we need to be sure that $\sqrt{\nu_t} > 0$. To this end we exploit Proposition above and assume $\eta^2 \leq 2k\hat{\nu}$. As already mentioned in the previous Section 1.4.2, a market price of risk is a 2-dimensional process $\lambda_t = (\lambda_t^1, \lambda_t^2)$ with

$$\lambda_t^1 = \frac{\mu - r}{\sqrt{\nu_t}}.$$

Since in general λ^2 is not uniquely determined, λ is not unique and so we do not have a uniquely associated risk-neutral measure. However, there exists a natural choice for λ^2 that simplifies computations and that gives rise to a well defined market price of risk. If we define

$$\lambda_t^2 = \frac{a\nu_t + b}{\sqrt{\nu_t}} \tag{1.26}$$

with $a, b \in \mathbb{R}$, under the risk-neutral measure $\mathbb{Q}$ associated to λ we have:

$$dS_t = rS_t dt + \sqrt{\nu_t} S_t dB_t^{1,\mathbb{Q}}$$

$$d\nu_t = \tilde{k}(\tilde{\nu} - \nu_t)dt + \eta\sqrt{\nu_t}dB_t^{2,\mathbb{Q}},$$

where

$$\tilde{k} = k + a\eta, \qquad \tilde{\nu} = \frac{k\hat{\nu} - b\eta}{\tilde{k}},$$

and therefore ν is a mean-reverting square root procees also under $\mathbb{Q}$. We see that the dynamics of the instantaneous variance ν has exactly the same form (up to different constants) under $\mathbb{P}$ and $\mathbb{Q}$, which explains why the Heston model is usually specified directly under a risk-neutral world. The choice of parameters a, b in Equation (1.26) is obtained by calibrating the model to the available market data.

The risk-neutral price of a derivative $F(S_T, \nu_T)$ is equal to $H_t^{\mathbb{Q}} = f(t, S_t, \nu_t)$, where $f = f(t, s, \nu)$ is solution to the Cauchy problem:

$$\begin{cases} \frac{\partial f}{\partial t} + \mathcal{L}^\lambda f - rf = 0 & \text{on} \quad [0, T[\times \mathbb{R}_{>0}^2 \\ f(T, s, \nu) = F(s, \nu) & \text{on} \quad \mathbb{R}_{>0}^2 \end{cases} \qquad (1.27)$$

where now

$$\mathcal{L}^\lambda = \frac{1}{2}\nu s^2 \frac{\partial^2 f}{\partial s^2} + \rho s \nu \eta \frac{\partial^2 f}{\partial s \partial \nu} + \frac{1}{2}\eta^2 \nu \frac{\partial^2 f}{\partial \nu^2} + rs\frac{\partial f}{\partial s} + \tilde{k}(\tilde{\nu} - \nu)\frac{\partial f}{\partial \nu} \qquad (1.28)$$

To solve numerically the PDE (1.27) one can rely on general numerical methods, like e.g. finite difference methods. However, since the log-characteristic function of S

$$\mathbb{E}^{\mathbb{Q}}\left[e^{i\xi log S_T}\right]$$

can be computed explicitly, analytical approximations of the price of European options are available by Fourier techniques: this will be discussed in Chapter 4, where we will present two different approaches. These formulas are generally preferable to standard numerical techniques because of their precision and computational efficiency in the valuation of European options, which becomes critical when calibrating the model to known option prices.

In conclusion, the existence of a fast and easily implemented quasi-closed form soluton for European options is one of the main advantages of the Heston model over other (potentially more realistic) stochastic volatility models and the reason for its great popularity.

Finally, recall that, since it is a stochastic volatility model, the Heston model does not fit the observed implied volatility surface. To fill this gap and generate consistent volatility surface, more advanced stochastic volatility models based on jumps (e.g. Merton model, Bates model) or on fractional Brownian motion have been proposed. In the next Chapter 2 we will discuss the latter in detail.

Chapter 2

Fractional Brownian motion and long memory

In this chapter we review the main properties of the fractional Brownian motion and discuss its role in volatility modelling. We introduce some mathematical and statistical concepts, like self-similarity, long-range dependence and regularity of paths, that will help us to build more advanced volatility models in the next Chapter 3. Finally, based on these properties, we will study the fractional Ornstein-Uhlenbeck process.

2.1   Fractional Brownian motion

Fractional Brownian motion (fBm) was first introduced in 1940 within a Hilbert space framework by Kolmogorov in [Kol40], where it was called "Wiener Helix". The name fractional Brownian motion is due to Mandelbrot and Van Ness, who in 1968 provided in [MV68] a stochastic integral representation of this process in terms of a standard Brownian motion. As we will see, the fBm is of interest in many applications in various fields, including financial mathematics, because of its capability of modelling short-range and long-range dependent phenomena.

Definition 2.1. *A fractional Brownian motion (fBm) with Hurst paramter $H \in (0,1)$ is a centered Gaussian process $(B_t^H)_{t \geq 0}$ with covariance function*

$$\mathbb{E}[B_t^H B_s^H] = \frac{1}{2}\big(t^{2H} + s^{2H} - |t-s|^{2H}\big). \tag{2.1}$$

It follows directly from (2.1) that $B_0^H = 0$ a.s.

For $H = 1/2$, the covariance function is $\mathbb{E}[B_t^H B_s^H] = t \wedge s$ and thus we obtain a standard Brownian motion. This justifies the name: B^H is a generalization of Brownian motion obtained by allowing the Hurst parameter to differ from $1/2$. Later on we will uncover the meaning of the Hurst parameter.

We recall that the distribution of a Gaussian process is uniquely determined by its mean and covariance functions. Therefore, for each value of $H \in (0, 1)$, the distribution of B^H is uniquely determined by Definiton 2.1. Nevertheless, this definition does not guarantee the existence of a fBm. To show that a fBm actually exists, it would be sufficient to check that the covariance function (2.1) is symmetric and positive semi-definite[1]. We will show the existence in Section 2.1.6, giving an explicit integral representation with respect to a standard Brownian motion which satisfies the properties of Definition 2.1.

2.1.1   Basic properties

Self-similarity and stationarity of increments

We summarize the basic properties of the fBm in the following proposition, whose proof can be found in Appendix.

Proposition 2.1. *Let B^H be a fBm with Hurst parameter $H \in (0, 1)$. Then:*

(1)[Self-similarity] $(\alpha^H B_t^H)_{t \geq 0} \overset{law}{=} (B_{\alpha t}^H)_{t \geq 0}$, for every $\alpha > 0$.

(2)[Stationarity of increments] $(B_{t+h}^H - B_h^H)_{t \geq 0} \overset{law}{=} (B_t^H)_{t \geq 0}$, for all $h > 0$.

(3)[Time inversion] $(t^{2H} B_{1/t}^H)_{t > 0} \overset{law}{=} (B_t^H)_{t > 0}$.

Conversly, any continuous Gaussian process $B^H = (B_t^H)_{t \geq 0}$ with $B_0^H = 0$, $Var(B_1^H) = 1$ and such that (1) and (2) hold, is a fractional Brownian motion of Hurst parameter H.

Remark:

(1): roughly speaking, self-similarity means that the patterns of a time-scaled sample path in any time interval have a similar shape to those of the original process, when properly space-rescaled.

(2): actually, more can be said about the increments than simply stationarity. In fact since the fBm is Gaussian, its increments are normally distributed: for all s, t such that $0 \leq s < t$,

$$B_t^H - B_s^H \sim \mathcal{N}\big(0, (t-s)^{2H}\big). \tag{2.2}$$

---

[1]Given $I \in \mathbb{R}$, a function $K : I \times I \to \mathbb{R}$ is positive semi-definite if for any $n \in \mathbb{N}$, $t_1, \dots t_n \in I$ and $u \in \mathbb{R}^n$, $\sum_{i,j=1}^{n} K(t_i, t_j) u_i u_j \geq 0$.

Correlation between two increments

It follows immediatly from Definition 2.1 that, for $r < u \leq s < t$,

$$\mathbb{E}[(B_t^H - B_s^H)(B_u^H - B_r^H)] = \frac{1}{2}\left(|s-u|^{2H} + |t-r|^{2H} - |t-u|^{2H} - |s-r|^{2H}\right). \quad (2.3)$$

For $H \neq 1/2$ and $r < u < s < t$, we have the following integral representation of (2.3):

$$\mathbb{E}[(B_t^H - B_s^H)(B_u^H - B_r^H)] = H(2H-1)\int_r^u \left(\int_s^t (x-y)^{2H-2}dx\right)dy. \quad (2.4)$$

It follows from (2.2) and (2.4) that the fBm has negatively correlated increments if $H < 1/2$ (we say that the fBm is "antipersistent"), and positively correlated increments if $H > 1/2$ (we say that the fBm is "persistent"). It is also easy to see that the correlation increases with H. Intuitively, when $H < 1/2$ this means that if the last increment has been negative, the next one is more likely to be positive (and vice-versa), while when $H > 1/2$ the signs of the increments tend to be preserved. In the first case, the process can be used to model random systems with *intermittent* time series data, while in the second case the process presents an aggregation behavior and this property can be used to describe *cluster* phenomena (systems with *memory*).
These properties are closely linked to long-range dependence, a crucial feature of fBm that we are going to discuss.

2.1.2 Long-range dependence

Definition 2.2. *A stationary discrete-time stochastic process $(X_n)_{n\in\mathbb{N}}$ exhibits* long-range dependence (or long memory) *if the autocovariance function $\rho(n) := \mathrm{Cov}(X_k, X_{k+n})$ satisfies*[2]

$$\lim_{n\to\infty}\frac{\rho(n)}{cn^{-\alpha}} = 1, \quad (2.5)$$

for some constants $c \in \mathbb{R}$ and $\alpha \in (0,1)$.
In this case the correlation between X_k and X_{k+n} decays slowly as $n \to \infty$ and

$$\sum_{n=1}^{\infty}|\rho(n)| = \infty \quad (2.6)$$

---

[2]Notice that, since X is stationary, $\mathrm{Cov}(X_k, X_{k+n})$ does not depend on k.

In contrast, processes with summable covariance functions are said to exhibit *short-range dependence* or *short memory*. Given a fBm B^H, let us define the increment process with time lag 1 by $X_n = B_n^H - B_{n-1}^H$. From Proposition 2.1 (2), we have that $(X_n)_{n\in\mathbb{N}}$ is stationary and the autocovariance function

$$\rho_H(n) = \text{Cov}(X_k, X_{k+n}) = \frac{1}{2}\big[(n+1)^{2H} + (n-1)^{2H} - 2n^{2H}\big]$$

satisfies

$$\lim_{n\to\infty} \frac{\rho_H(n)}{H(2H-1)n^{-(2-2H)}} = 1. \tag{2.7}$$

Summarizing, we obtain:

- For $H \in (1/2, 1)$, $\sum_{n=1}^{\infty} \rho_H(n) = \infty$ and the increment process exhibits long-range dependence.

- For $H \in (0, 1/2)$, $\sum_{n=1}^{\infty} |\rho_H(n)| < \infty$ and the increment process exhibits short-range dependence.

Alternative definitions of *long-range dependence* (or *long memory*) can be found in the literature, all of them capturing the same essential feature: a slowly decaying autocorrelation function[3]. One of the most well established definitions uses the notion of *slowly varying* function.

Definition 2.3. *A function $L : (0, +\infty) \to (0, +\infty)$ is called slowly varying at infinity if, for all $a > 0$:*

$$\lim_{x\to\infty} \frac{L(ax)}{L(x)} = 1 \tag{2.8}$$

Examples of slowly varying functions at infinity are logarithms, iterated logarithms and functions that converge to positive constants. Polynomials and the positive functions $2 + \sin x$, e^{-x}, e^x are not slowly varying at infinity.

Definition 2.4. *A stationary process $(X_t)_{t\geq 0}$ exhibits long-range dependence (or long memory) if the autocovariance function $\rho(t) := \text{Cov}(X_h, X_{h+t})$ satisfies:*

$$\rho(t) = t^{-\beta}L(t), \qquad t \geq 0, \tag{2.9}$$

where L is a slowly varying function at infinity and $\beta \in (0, 1)$.

---

[3]For a stationary process $(X_n)_{n\in\mathbb{N}}$ with finite variance, the autocorrelation function is $\text{Corr}(X_k, X_{k+n}) = \frac{\text{Cov}(X_k, X_{k+n})}{\sigma_k \sigma_{k+n}} = \frac{\rho(n)}{\sigma_k^2}$, where σ_k is the standard deviation of X_k, and thus exhibits long-term properties analogous to those of the autocovariance function $\rho(n)$.

For $X_t = B_t^H - B_{t-1}^H$, the two definitions are equivalent. In fact, recalling Equation (2.7), also valid in the continuous case, and taking $L(t) = t^\beta \rho_H(t), \beta = 2 - 2H \in (0,1)$, we have that, for $H \in (1/2, 1)$, L converges to a positive constant, and thus it is slowly varying at infinity. Hence, we retrieve the long range-dependence property for $H \in (1/2, 1)$, and short-range dependence for for $H \in (0, 1/2)$.

2.1.3 Regularity of paths

Hölder continuity

There are several ways to establish the continuity of fBm. All of them are based on the formula

$$\mathbb{E}[(B_t^H - B_s^H)^2] = |t - s|^{2H},$$

which follows from Equation (2.2). The regularity of fBm is dictated by its Hurst parameter H: the larger H is, the smoother fBm becomes, as stated in the next proposition, which follows immediately from Kolmogorov-Chentsov continuity theorem.

Proposition 2.2. *The fractional Brownian motion B^H has a continuous modification. Moreover, for any $\gamma \in (0, H)$, this modification is γ-Hölder continuous on each finite interval.*

Assumption: To avoid speaking about a continuous modification each time, in the rest of this thesis we will assume the continuity of fBm itself.

Finally, we mention that, by using well-known facts about Gaussian processes, it is possible to show that the exact modulus of continuity of fBm is $\omega(\delta) = \delta^H |log\delta|^{1/2}$ (see e.g. [Lif95], p.103). Consequently, it is only Holder continuous of order up to H, but not H-Holder continuous. Figure 2.1 illustrates the dependence of fBm on H. We can see that he sample paths of fBm exhibit higher regularity for greater values of Hurst parameter, in accordance with Proposition 2.2.

Path differentiability

Despite being γ-Holder continuous for $\gamma \in (0, H)$, the fBm B^H has almost surely nowhere differentiable sample paths. This follows from the law of iterated logarithm for fBm, which generalizes the well known theorem for classical Brownian motion (see [LS01]):

Theorem 2.1. *If $(B_t^H)_{t \geq 0}$ is a fBm with Hurst parameter $H \in (0, 1)$, then, for all $t_0 \geq 0$, almost surely*

$$\limsup_{t \downarrow 0} \frac{B_{t_0+t}^H - B_{t_0}^H}{t^H |2 \log \log t|^{1/2}} = 1, \qquad \liminf_{t \downarrow 0} \frac{B_{t_0+t}^H - B_{t_0}^H}{t^H |2 \log \log t|^{1/2}} = -1. \qquad (2.10)$$

It follows immediately that

$$\limsup_{t \downarrow 0} \left| \frac{B_{t_0+t}^H - B_{t_0}^H}{t} \right| = \infty \qquad \mathbb{P} - a.s.$$

and therefore, for each $t_0 \geq 0$, $\mathbb{P}$-almost surely the function $t \mapsto B_t^H$ is not differentiable in $t = t_0$. Actually, this result can be improved by showing that, $\mathbb{P}$-almost surely, the function $t \mapsto B_t^H$ is not differentiable in any $t \geq 0$.
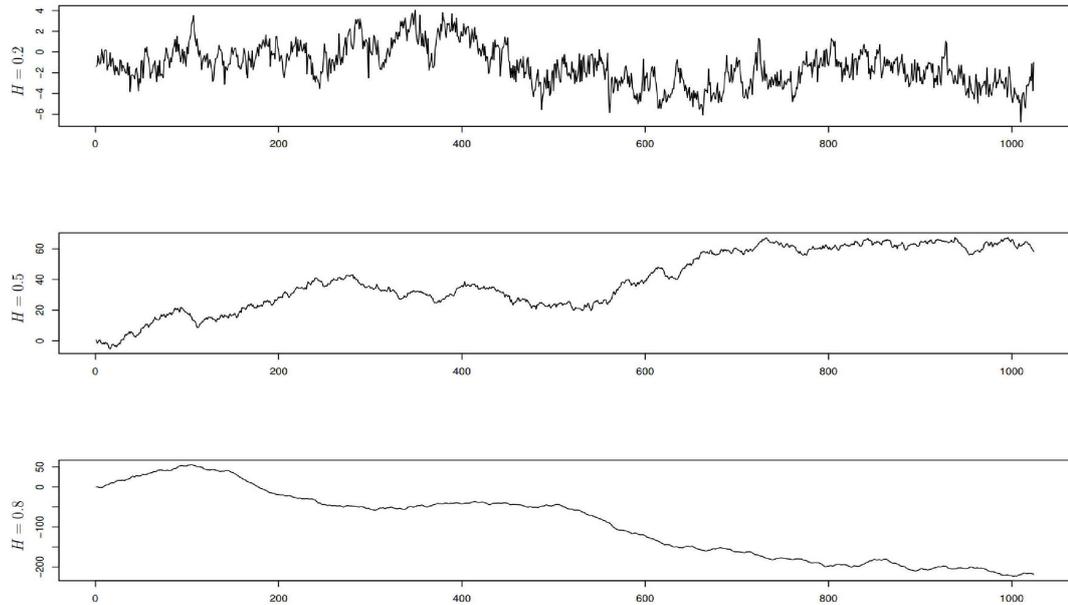

Figure 2.1: Samples of fractional Brownian motion for H = 0.2, H = 0.5 and H = 0.8. Taken from [Nou12].

2.1.4   Markov and Semimartingale property

Fractional Brownian motion with Hurst parameter $H \neq 1/2$ has properties that deviate significantly from standard Brownian motion, semimartingales

and others processes classically used in probability theory. In fact, in order to gain interesting modelling properties, like short-range or long-range dependence, we must 'sacrifice' some good properties of both Brownian motion and semimartingales, and this is the reason why fBm is often quite difficult to handle.

Markov property

We recall the definition of Markov process.

Definition 2.5. *Let X be a real-valued stochastic process. X is called Markov process if it satisfies, for all Borel set $A \subset \mathbb{R}$ and all real numbers $t > s > 0$,*

$$\mathbb{P}(X_t \in A | X_u, u \leq s) = \mathbb{P}(X_t \in A | X_s). \qquad (2.11)$$

For a fBm we have the following result (see [Huy03] for a proof):

Theorem 2.2. *Let B^H be a fractional Brownian motion of Hurst parameter $H \in (0,1), H \neq 1/2$. Then B^H is not a Markov process.*

The proof of this result, as presented in [Huy03], is based on a property of the covariance function of a centered Gaussian Markov process, call it $(X_t)_{t \geq 0}$. For such a process, the author proves that the autocovariance function is given by:

$$R(t,s) = \frac{R(t,s)R(t_0,t_0)}{R(s,s)}, \qquad t > s > t_0 \geq 0. \qquad (2.12)$$

From here, the proof of Theorem 2.2 is rather straightforward, (see [Huy03] for the details).

Semimartingale property

Now we study the asymptotic behavior of the p-variation of the fBm. As a byproduct, we will see that fBm is never a semimartingale except, of course, when it is the standard Brownian motion.
Recall that a real-valued process is called a semimartingale if it can be decomposed as the sum of a local martingale and a càdlàg adapted process of locally bounded variation. We start with a general preliminary result, which may be viewed as a law of large numbers for fBm (see [Nou12], p.19 for a proof).

Theorem 2.3. *Let B^H be a fractional Brownian motion of Hurst parameter $H \in (0,1)$. Let $G \sim \mathcal{N}(0,1)$ and let $f : \mathbb{R} \to \mathbb{R}$ a measurable function such that $\mathbb{E}[f^2(G)] < \infty$. Then, as $n \to \infty$,*

$$\frac{1}{n} \sum_{k=1}^{n} f(B_k^H - B_{k-1}^H) \xrightarrow{L^2} \mathbb{E}[f(G)]$$

With a direct application of Theorem 2.3 with $f(x) = |x|^p$, we deduce the following result about the p-variation of fBm.

Corollary 2.1. *Let B^H be a fractional Brownian motion of Hurst parameter $H \in (0,1)$, and let $p \in [1, \infty[$. Then, as $n \to \infty$, one has*

$$\sum_{k=1}^{n} |B_{k/n}^H - B_{(k-1)/n}^H|^p \xrightarrow{L^2} \begin{cases} 0 & \text{if } p > 1/H \\ \mathbb{E}[|G|^P] & \text{if } p = 1/H, \text{ with } G \sim \mathcal{N}(0,1) \\ +\infty & \text{if } p < 1/H \end{cases}$$

We are now ready for the main result (see, e.g., [Rog97] for the proof).

Theorem 2.4. *Let B^H be a fractional Brownian motion of Hurst index $H \in (0,1), H \neq 1/2$. Then B^H is not a semimartingale.*

2.1.5 Integration with respect to fBm

Theorem 2.4 explains why integrating with respect to fBm is a non-trivial problem. In fact semimartingales form the largest class of "good integrators" for which a well defined notion of stochastic integral exists. Moreover, for a stochastic process, being a semimartingale is the weakest condition to require in order to apply Itô's lemma and consequently all the properties from of Itô's calculus. Nevertheless, it is still possible to develop a stochastic calculus with a fractional Brownian motion as integrator.
There are essentially two different approaches:

   (i) *Pathwise approach*;

   (ii) *Malliavin calculus.*

We focus in this thesis on the first approach. We refer to Nualart [Nua02] for a complete treatment of Malliavin calculus, and to Biagini et al. [Bia+08] for its applications to fBm and to finance in general.

Pathwise approach

As a consequence of Corollary 2.1 we have that, for every $H \in (0, 1)$, B^H is not of bounded variation. This makes Riemann-Stieltjes pathwise integration problematic, since one of the usually required assumptions for the integrator function in the Riemann-Stieltjes integral is having bounded variation. Fortunately, there exists another way to define the integral pathwise.

If $(u_t)_{t \in [0,T]}$ is a stochastic process with γ-Holder continuous trajectories, where $\gamma > 1 - H$, then the (Young-)integral $\int_0^T u_t dB_t^H$ exists pathwise (i.e. for every fixed $\omega \in \Omega$, the (Young-)integral $\int_0^T u_t(\omega) dB_t^H(\omega)$ exists). The Young integral is an extension of Riemann-Stieltjes integral, in the sense that if both integrals are well defined for given integrand and integrator functions, then they coincide. This follows from the next theorem, based on [You36], which can be proved using the Hahn-Banach theorem.

For any $\alpha \in [0, 1]$, we denote by Λ^α the set of α-Holder continuous functions, that is, the set of functions $f : [0, T] \to \mathbb{R}$ satisfying

$$|f|_\alpha := \sup_{0 \leq s < t \leq T} \frac{|f(t) - f(s)|}{(t - s)^\alpha} \quad < \infty.$$

We equip Λ^α with the norm $||f||_\alpha := |f|_\alpha + |f|_\infty$, where $|f|_\infty = \sup_{t \in [0,T]} |f(t)|$.

Theorem 2.5. *Let $f \in \Lambda^\alpha$ with $\alpha \in (0, 1)$, and let $\beta \in (0, 1)$ be such that $\alpha + \beta > 1$. The linear operator $T_f : C^1 \subset \Lambda^\beta \to \Lambda^\beta$ defined by $T_f(g) = \int_0^\cdot f(u)g'(u)du$ is continuous with respect to the norm $|| \cdot ||_\beta$. By density, it extends in a unique way to an operator $\bar{T}_f : \Lambda^\beta \to \Lambda^\beta$.*

As a consequence, we are now able to define a so-called Young integral (see [Nua02]).

Definition 2.6. *Let $f \in \Lambda^\alpha$ and $g \in \Lambda^\beta$ with $\alpha, \beta \in (0, 1)$, $\alpha + \beta > 1$. The Young integral is (well-)defined as $\int_0^\cdot f(u)dg(u) := \bar{T}_f(g)$.*

From Proposition 2.2 we know that B^H is γ-Hölder continuous for any $\gamma \in (0, H)$. Thus, taking $\alpha > 1 - \gamma$, $\beta = \gamma$, $f = (u_t)_{t \in [0,T]}$ an α-Hölder continuous process and $g = B^H$, we have the well-defined pathwise integral $\int_0^T u_t dB_t^H$ (we omit the dipendence on $\omega \in \Omega$). The applicability of this method leads to a first distinction:

- For $H > 1/2$, the method above is particularly useful, because it includes processes of the form $u_t = F(B_t^H)$, where F is a continuously differentiable function. In fact in this case u_t has the same Holder index as B_t^H, and we can take $\alpha = \beta = \gamma > 1/2$.

- In the case $H < 1/2$ things are not so simple. A powerful approach known as *rough path theory* may be used to give sense to $\int_0^T u_t dB_t^H$, at least when H is not too small. However, we do not need to develop a systematic theory of integration with respect to fBm when $H \in (0, 1/2)$ because the integrands we will encounter will be constants and therefore, thanks to Theorem 2.5 and Definiton 2.6, we have that the integral $\int_0^T c dB_t^H$ is well defined for every $H \in (0, 1)$ (here c is a constant, hence α-Holder continuous for every $\alpha \in (0, 1)$).

2.1.6   Stochastic integral representations

In this section we show that the fractional Brownian motion can be represented as a stochastic integral with respect to Brownian motion in (at least) two different ways. We will make use of the two-sided classical Brownian motion $B = (B_t)_{t \in \mathbb{R}}$ defined as

$$B_t = \left\{ \begin{array}{ll} B_t^1 & \text{if } t \geq 0 \\ B_{-t}^2 & \text{if } t < 0 \end{array} \right. ,$$

where B^1 and B^2 are independent (one-sided) classical Brownian motions.

Mandelbrot-van Ness

The first representation of fBm was given by Mandelbrot and Van Ness in [MV68], and is also called *moving average* representation. It can also be used as a proof of the existence of fBm.

Proposition 2.3. *Let $H \in (0, 1), H \neq 1/2$, set*

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

and let $B = (B_t)_{t \in \mathbb{R}}$ be a two-sided classical Brownian motion. Then the process $B^H = (B_t^H)_{t \geq 0}$ defined as

$$B_t^H = \frac{1}{c_H} \Big(\int_{-\infty}^0 \left((t-u)^{H-1/2} - (-u)^{H-1/2} \right) dB_u + \int_0^t \left(t-u \right)^{H-1/2} dB_u \Big)$$

is a fractional Brownian motion of Hurst parameter H.

Notice that the Mandelbrot-van Ness representation uses the entire history of the Brownian motion $(B_s)_{s \in [0,t]}$, which has an impact on the memory of the process. This reflects the *memory* property of fBm, as we have seen in Section 2.1.2.

fBm as a Volterra process

Our second representation allows us to write fBm in the form of a so-called Volterra process, that is, in the form $B_t^H = \int_0^t K_H(t,s)dB_s$, where $B = (B_t)_{t \geq 0}$ is a classical Brownian motion and K_H is an explicit square integrable kernel, whose expression is a bit convoluted. The advantage is that the kernel in this representation has compact support. We give the expression only for $H \in (0, 1/2)$, since we will only analyze this case in the next chapter. However, an analogous formula is valid for $H \in (1/2, 1)$ (see [Nou12], p. 16).

Proposition 2.4. *Let $H \in (0, 1/2)$ and, for $t > s > 0$, set*

$$K_H(t,s) = d_H \left[\left(\frac{t(t-s)}{s} \right)^{H-1/2} - (H-1/2)s^{1/2-H} \int_s^t u^{H-3/2}(u-s)^{H-1/2}du \right],$$

where

$$d_H = \sqrt{\frac{2H}{(1-2H) \int_0^1 (1-x)^{-2H} x^{H-1/2} dx}} \quad .$$

Let $B = (B_t)_{t \geq 0}$ be a classical Brownian motion, and define $B^H = (B_t^H)_{t \geq 0}$ by

$$B_t^H = \int_0^t K_H(t,s)dB_s.$$

Then B^H is a fractional Brownian motion of Hurst parameter H.

The above explicit representation is also extensively used when performing numerical simulations of fBm, given the fact that, generally speaking, it is the representation to use when solving some fractional differential equations (see Section 4.3.3 for a brief explanation), i.e. differential equations involving fractional derivatives, like the fractional Riccati equation.

2.2 Fractional Ornstein-Uhlenbeck processes

We now present a generalization of the classical Ornstein-Uhlenbeck process, the so-called fractional Ornstein-Uhlenbeck process (fOU), which will be fundamental in the next Chapter 3 on fractional stochastic volatility models. We recall that the classical Ornstein-Uhlenbeck process with parameter $\lambda > 0$ and $\sigma > 0$ starting at $x \in \mathbb{R}$, is the unique strong solution (see, e.g. [Pas11] p. 327) of the Langevin equation

$$\begin{cases} dX_t = -\lambda X_t dt + \sigma dB_t & t \geq 0 \\ X_0 = x \end{cases}, \tag{2.13}$$

where B is a classical Brownian motion. It is given by

$$X_t^x = e^{-\lambda t}\left(x + \sigma \int_0^t e^{\lambda u} dB_u\right) \quad t \geq 0. \tag{2.14}$$

This model is popular in interest rate modelling, where it is used in the Vasicek model.

Actually, Langevin equation can be defined in a more generale setting, where the noise term σdB_t is replaced by a general Gaussian noise term dN_t and the constant $x \in \mathbb{R}$ is replaced by a random variable $\xi \in L^0(\Omega)^4$. Thus, it becomes

$$\begin{cases} X_t = \xi - \lambda \int_0^t X_s ds + N_t & t \geq 0 \\ X_0 = \xi \end{cases} \tag{2.15}$$

Notice that, since the above Equation (2.15) only involves an integral with respect to ds and not a stochastic integral, it can be solved path-wise for much more general noise processes $(N_t)_{t\geq0}$ than Brownian motion. This leads us to the question whether for $H \in (0,1), H \neq 1/2$ and for fixed initial conditions, the Langevin equation with fractional Brownian noise $(\sigma B_t^H)_{t\geq0}$ has a solution, if it is unique and eventually what are its main features. In this case we would call the solution a *fractional Ornstein-Uhlenbeck* process with Hurst parameter H.

Existence and uniqueness

Cheridito et al. have shown in [CKM03] that, for each $H \in (0,1)$ and $\xi \in L^0(\Omega)$,

$$X_t^{H,\xi} := e^{-\lambda t}\left(\xi + \sigma \int_0^t e^{\lambda u} dB_u^H\right), \qquad t \geq 0,$$

is the unique $\mathbb{P}$-almost surely continuous process that solves Langevin equation (2.15) with fractional Brownian noise $(\sigma B_t^H)_{t\geq0}$. Notice that, since $e^{\lambda u}$ is Lipschitz-continuous (1-Hölder continuous), thanks to the results in Section 2.1.5, we can conclude that $\int_0^t e^{\lambda u} dB_u^H$ is well-defined for every $H \in (0,1)$.

Analogously to the classical Brownian motion, the two-sided fractional Brownian motion $(B_t^H)_{t\in\mathbb{R}} \in L^0(\Omega)$ is defined as

$$B_t^H = \begin{cases} B_t^{H,1} & \text{if } t \geq 0 \\ B_{-t}^{H,2} & \text{if } t < 0 \end{cases},$$

where $B^{H,1}$ and $B^{H,2}$ are independent (one-sided) fractional Brownian motions.

---

[4]$L^0(\Omega)$ is the space of measurable functions on $(\Omega, F, \mathbb{P})$.

With initial condition $\xi = \sigma \int_{-\infty}^{0} e^{\lambda u} dB_u^H$, we find the following solution of Langevin equation:

$$X_t^H := \sigma \int_{-\infty}^{t} e^{-\lambda(t-u)} dB_u^H.$$

It is easy to show that $(X_t^H)_{t\in\mathbb{R}}$ is a Gaussian process, and it follows immediately from the stationarity of the increments of fractional Brownian motion that it is stationary. Furthermore, for every $\xi \in L^0(\Omega)$, almost surely

$$X_t^H - X_t^{H,\xi} = e^{-\lambda t}(X_0^H - \xi) \xrightarrow{t\to\infty} 0,$$

which implies that every stationary and almost surely continuous process that solves (2.15) has the same distribution as $(X_t^H)_{t\geq 0}$.

We call $(X_t^{H,\xi})_{t\geq 0}$ a *fractional Ornstein-Uhlenbeck process with initial condition* ξ and $(X_t^H)_{t\in\mathbb{R}}$ a *stationary fractional Ornstein-Uhlenbeck process*.

Autocovariance function and long-range dependence

Cheridito et al. show in [CKM03] that for $H \in (0,1), H \neq 1/2$ and $t > 0$, the decay of the autocovariance function of X^H is very similar to that of the autocovariance function of the increment process of B^H with time lag t.

In fact, it can be derived from (2.1) that, for fixed $N \in \mathbb{N}$, $h > 0$, as $s \to \infty$

$$\text{Cov}(B_{h+t}^H - B_h^H, B_{h+s+t}^H - B_{h+s}^H) = \sum_{n=1}^{N} \frac{t^{2n}}{(2n)!} \Big(\prod_{k=0}^{2n-1} (2H-k) \Big) s^{2H-2n} + O(s^{2H-2N-2}).$$

On the other hand, for fixed $N \in \mathbb{N}$, $h > 0$, as $s \to \infty$

$$\text{Cov}(Y_h^H, Y_{h+s}^H) = \frac{1}{2}\sigma^2 \sum_{n=1}^{N} \lambda^{-2n} \Big(\prod_{k=0}^{2n-1} (2H-k) \Big) s^{2H-2n} + O(s^{2H-2N-2}).$$

In particular, using the results in section 2.1.2 on long-range dependence, for $H \in (1/2, 1)$, Y^H tends to exhibit long-range dependence and persistence (positively correlated consecutive increments) as $s \to \infty$, while for $H \in (0, 1/2)$, X^H looses long-range dependence and the process becomes antipersistent (negatively correlated consecutive increments) as $s \to \infty$. The enhanced negative correlation with smaller H gives a relatively rougher paths with a more irregular behavior.

Therefore, at least asympotically, even for a fOU process, trajectories tend to be more regular for higher Hurst parameters.

Chapter 3

Volatility models with fractional Brownian motion

In this chapter, we introduce some more advanced stochastic volatility models based on fractional Brownian motion. We present at first the FSV model, based on the fBm with Hurst parameter $H \in (1/2, 1)$, the first fractional volatility model to appear in the literature. Then, we turn to more recent models that assume the Hurst parameter to be less than $1/2$. Because of the irregularity of their sample paths, these models are called 'rough'. For each of the models we will investigate, it is important to have in mind some basic properties that the model needs to satisfy in order to be considered acceptable. Some of these properties are derived from empirical observations, and therefore it seems logical to begin our study from these.

We work in a filtered probability space $(\Omega, F, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ satisfying the usual conditions[1]. We assume $(\mathcal{F}_t)_{t \geq 0}$ to be a Brownian filtration, that is $\mathcal{F}_t = \sigma\{(B_s)_{s \leq t}\}$, $t \geq 0$, is the sigma-algebra generated by $(B_s)_{s \leq t}$, where B is a Brownian motion on the probability space $(\Omega, F, \mathbb{P})$.

3.1   Stylized facts about volatility

Stylized facts are observations that have been made in so many contexts that they are widely understood to be empirical truths. Due to their generality, they are often qualitative and, although essentially true, may have inaccuracies in the details. In developed countries, we observe from the data analysis of financial time series (such as daily stock returns) a number of stylized facts about the volatility of financial asset prices. The purpose of this section is

---

[1] $(\Omega, F, \mathbb{P})$ is complete and $(\mathcal{F}_t)_{t \geq 0}$ is complete and right-continuous.

to analyze qualitatively the various stylized facts that have been shown to affect the volatility (here in the sense of *Historical Volatility*) of stock prices. During the presentation, we will revise some concepts already encountered in Chapter 1 and Chapter 2, to make the review more systematic.

First, we establish the notation. Let S_t be the stock price at time t and $r_t = \log(S_t) - \log(S_{t-1})$ the logarithmic return over the period $t-1$ to t. The mean and variance are defined as usual as

$$\mu_t = \mathbb{E}[r_t], \qquad \sigma_t^2 = \mathbb{E}[(r_t - m_t)^2]. \tag{3.1}$$

We can also define the conditional mean and conditional variance as:

$$m_t = \mathbb{E}_{t-1}\big[r_t\big] \tag{3.2}$$

$$h_t = \mathbb{E}_{t-1}\big[(r_t - m_t)^2\big] \tag{3.3}$$

where $\mathbb{E}_{t-1}[U] := \mathbb{E}[U|\mathcal{F}_{t-1}]$ is the conditional expectation of the random variable U given the information set at time $t-1$. Intuitively, the conditional variance tells us 'how much variance is left' if, given the information at time $t-1$, we use m_t to 'predict' r_t.

Higher moments of the process $(r_t)_{t\geq 0}$ are often used in volatility models. The skewness and kurtosis are defined by

$$\xi_t = \frac{\mathbb{E}[(r_t - \mu_t)^3]}{\sigma^3}, \qquad \zeta_t = \frac{\mathbb{E}[(r_t - \mu_t)^4]}{\sigma^4}. \tag{3.4}$$

Skewness is a measure of asymmetry in a probability distribution. It can be negative, zero, positive or undefined. Since the normal distribution is symmetric, the skewness is zero. An example of positive skewness is the log-normal model, which has a right skewed distribution.

Kurtosis measures the 'thickness' of the tail of a distribution; a distribution is fat-tailed if it has a higher kurtosis than that of the normal distribution, which is 3. Therefore, a distribution with a higher kurtosis than 3 is considered fat-tailed. Consequently, rare events, measured in the tail, have a higher probability to occur in a fat-tailed distribution than in a normal distribution.

(a) Clustering

There is substantial evidence that volatility is not constant (see, e.g., [Fam65], [Off73], [TW92]); the idea is that as trading activity fluctuates, so does volatility. In particular, any casual observation of financial time series reveals bunching of *high* and *low* volatility episodes. This results from the

fact that large changes in the price of an asset are often followed by other large changes, and small changes are often followed by small changes (see [Man63] and [Con01]). This feature is known as *clustering* or *intermittency*, and results from the fact that volatility is autocorrelated, in the sense that the autocorrelation function decays slowly. This establishes a strong relationship between volatility clustering and long memory of volatility.

(b) Persistence

Intuitively, volatility persistence measures the strength of the volatility feedback effect. High persistence means that volatility shocks will be felt further in the future, albeit to a lesser extent. Generally speaking, volatility is highly persistent. This behaviour has been reported by numerous studies, such as [Bai96] and [Sch89]. To make a precise definition of volatility persistence, let the conditional variance of returns k periods in the future, also called *forecast variance*, be defined as

$$h_{t+k|t} = \mathbb{E}_t\big[\big(r_{t+k} - m_{t+k}\big)^2\big]. \tag{3.5}$$

Volatility is said to be persistent if information about today's returns (which, just to simplify conceptually, we can consider to be the whole information in today's information set, that is, $\mathcal{F}_t$ coincides with the sigma-algebra generated by r_t) has a large effect on the forecast variance many periods in the future. More precisely, taking partial derivative, we define the *forward variance*[2]:

$$\theta_{t+k|t} = \frac{\partial h_{t+k|t}}{\partial r_t^2}. \tag{3.6}$$

The higher $\theta_{t+k|t}$, the more persistent the volatility is.
Another traditional method to quantify persistence is to consider the correlations bewteen two increments of the volatility process. If they are positively correlated, then we have persistence, otherwise, in the case of negative correlation, we say that the process is antipersistent. From this last 'definition', we see that persistence in intimately connected with *clustering* and *long memory*, in that all these three properties can be measured by analyzing the decay of the autocorrelation function.

---

[2]By taking the derivative with respect to r_t^2 instead of r_t, we obtain a dimensionless number, as squared returns and conditional variance are in the same unit

(c) Long vs short memory

Changes in volatility typically have a long-lasting impact on its subsequent evolution. We say that volatility has a *long memory* (or long-range dependence, LRD for brevity). This was one of the conclusions of the influential analyses of Ding et al. [DGE93] and Andersen et al. [And+03]. Mathematically, as we have seen in Chapter 2, this means that the volatility process $(\sigma_t)_{t \geq 0}$ has a slowly decaying autocorrelation function (ACF). An important issue, still debated, is the relationship between the persistence and the LRD property of the volatility process. The authors above argue that the phenomenon behind their empirical findings of a high persistence is the long memory property, suggesting that the latter is a necessary condition for the former. We will develop this idea in the following section thanks to the FSV model of Comte and Renault, which enjoys this LRD property. However, we will also show in the RFSV model of Gatheral et al., that we can also explain the empirical evidence of voletility persistence with models that do not exhibit long memory, in the sense that their autocorrelation function does not decay as a power-law (see Definition 2.3, Chapter 2). Indeed, as the famous mathematician Rama Cont [Con07] writes : "The econometric debate on the short range or long range nature of dependence in volatility still goes on (and may probably never be resolved)".

(d) Mean reversion

Mean reversion in volatility is generally interpreted as meaning that there is a normal level of volatility to which the volatility process will eventually return. In other words, the volatility process $(\sigma_t)_{t \geq 0}$ is 'pushed' towards a value $\bar{\sigma} \geq 0$ that can be interpreted as a long-term mean or equilibrium. Stock prices are generally viewed as consistent with mean reversion of volatility (see, e.g. [FPS99]), which implies that current information has no effect on the long-run forecast. More precisely, this implies that[3]

$$\theta_{t+k|t} \xrightarrow[k \to \infty]{p} 0, \qquad \text{for all t.} \tag{3.7}$$

(e) Fat tails

It is well established that the (unconditional) distribution of asset returns has fat tails (see, e.g., [Man63], [Fam65]). Fat tails mean a higher probability of large losses (and gains) than the normal distribution would suggest. Typical

---

[3]The convergence is in probability.

kurtosis estimates range from 4 to 50 indicating very extreme non-normality (also known as 'leptokurtic distribution'). Therefore, probability of observing an extreme event (either a downturn or a takeoff) is larger than what is hypothesized by normally distributed data generating processes.

(f) Leverage effect

A phenomenon coined by Black in 1976 ([Bla76]) as the *leverage effect* (also known as the 'asymmetric volatility phenomenon') suggests that stock price returns are negatively correlated with volatility. Moreover, this relation is asymmetric: when returns are negative, volatility increases rapidly; but, when returns are positive, volatility decreases to a much lesser extent. There are two competing theories that aim at explaining the occurrence of such a phenomenon (see [BW00] for a detailed discussion of these two theories). The first one, known as the *financial leverage* hypothesis, says that when a stock declines, its debt-to-equity ratio increases, making the company riskier and thus leading to a higher volatility of returns. The second theory, the *volatility feedback* hypothesis ([CH92]), avers that volatility may trigger a risk premium effect. News of increasing volatility reduces the demand for a stock because of risk aversion, leading to a decline in stock value.

3.2   Fractional stochastic volatility model

We now present the fractional stochastic volatility (FSV) model introduced by Comte and Renault in [CR98], based on the Hull-White SV model. These will form the prerequisites to investigate more advanced models introduced in the next section, where we will exploit the modelling power of the fractional Brownian motion in order to fix some shortcomings related to the volatility surface fitting.

3.2.1   Hull-White SV model

In the stochastic volatility model proposed by Hull and White in [HW87] in 1987, the underlying asset S_t and the instantaneous variance $\nu_t = \sigma_t^2$ are assumed to obey the following SDEs under the physical measure $\mathbb{P}$:

$$dS_t = \mu(S_t, t)S_t dt + \sqrt{\nu_t}S_t dB_t^1 \tag{3.8}$$

$$d\nu_t = \alpha(\nu_t, t)\nu_t dt + \beta(\nu_t, t)\nu_t dB_t^2 \tag{3.9}$$

where μ, α, β are real deterministic functions defined on $\mathbb{R}^2_{>0}$ and $B_t = (B^1_t, B^2_t)$ is, as usual, a 2-dimensional correlated Brownian motion.

Following the notation of Section 1.4.2, under the (non-unique) risk-neutral measure $\mathbb{Q}$ the dynamics becomes:

$$dS_t = rS_t dt + \sqrt{\nu_t} S_t dB^{1,\mathbb{Q}}_t \tag{3.10}$$

$$d\nu_t = \hat{\alpha}(\nu_t, t)\nu_t dt + \beta(\nu_t, t)\nu_t dB^{2,\mathbb{Q}}_t. \tag{3.11}$$

By the Itô formula, the solution of (3.10) is

$$S_t = S_0 \exp\left(\int_0^t \sqrt{\nu_s} dB^1_s + \int_0^t \left(r - \frac{\nu_s}{2} ds\right) \right) \tag{3.12}$$

Hull and White prove the following lemma (see [HW87]):

Lemma 3.1. *Assume the stock price $(S_t)_{t\geq 0}$ and its instantaneous variance $(\nu_t)_{t\geq 0}$ follow the dynamics $(3.10) - (3.11)$, where $\hat{\alpha}$ and β are independent of S_t, and $(B^{1,\mathbb{Q}}_t)_{t\geq 0}$ and $(B^{2,\mathbb{Q}}_t)_{t\geq 0}$ are independent. Let $\bar{\sigma}^2_{t,T}$ be the mean variance over the time interval $[t, T]$ defined by*

$$\bar{\sigma}^2_{t,T} = \frac{1}{T-t} \int_t^T \nu(s) ds. \tag{3.13}$$

Then, under $\mathbb{Q}$, the conditional distribution of $\log(S_T/S_t)$ given $\bar{\sigma}^2_{t,T}$ is

$$\mathcal{N}\left(\left(r - \frac{\bar{\sigma}^2_{t,T}}{2}\right)(T-t), \ \bar{\sigma}^2_{t,T}(T-t) \right). \tag{3.14}$$

The European Call option at time t of underlying stock $(S_t)_{t\geq 0}$ with strike price K and expiration T takes the form

$$C\left(t, S_t, K, T\right) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}(S_T - K)^+ \big| \mathcal{F}_t\right]. \tag{3.15}$$

Using the tower property of conditional expectation, i.e. $\mathbb{E}\left[X | \mathcal{G}\right] = \mathbb{E}\left[\mathbb{E}[X|\tilde{\mathcal{G}}] | \mathcal{G}\right]$ a.s., for $\mathcal{G} \subset \tilde{\mathcal{G}}$, we obtain

$$C\left(t, S_t, K, T\right) = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}(S_T - K)^+ \big| \mathcal{F}_t \vee \sigma\{\sigma_s, s \in [t,T]\}\right] \Big| \mathcal{F}_t\right] \tag{3.16}$$

The inner expectation is the value of the Call option with expiration T given the knowledge of all the information before time t and of how volatility will move during the interval $[t, T]$ (but not the knowledge of how the stock will move during that interval).

Recalling from Section 1.1.1 the distribution of the log-returns in the Black-Scholes model, it is clear from (3.14) that, under the assumptions of Lemma 3.1, the conditional distribution of $\log(S_T/S_t)$ given $\bar{\sigma}_{t,T}^2$ in the Hull-White model equals the distribution of $\log(S_T/S_t)$ in a Black-Scholes model with constant volatility $\sigma = \bar{\sigma}_{t,T}$. Hull and White show that the inner conditional expectation in (3.16) is the price of the Black-Scholes Call option

$$C_{BS}(t, \bar{\sigma}_{t,T}, S_t, K, T) = S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2), \qquad (3.17)$$

where Φ is the standard normal CDF and

$$\frac{1}{\bar{\sigma}_{t,T}\sqrt{T-t}} \left[\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\bar{\sigma}_{t,T}^2}{2}\right)(T-t) \right], \qquad d_2 = d_1 - \bar{\sigma}_{t,T}\sqrt{T-t}. \ (3.18)$$

Therefore, the option price formula in the Hull-White model is deduced as follows:

$$C(t, S_t, K, T) = \mathbb{E}^{\mathbb{Q}}\left[C_{BS}\big(t, \bar{\sigma}_{t,T}, S_t, K, T\big) \big| \mathcal{F}_t \right]. \qquad (3.19)$$

To conclude, the price at time t of a European Call option of expiration T in the Hull-White model is the conditional expectation of the B-S option pricing formula where the constant volatility is replaced by $\bar{\sigma}_{t,T}$. We observe that, at time t, $\bar{\sigma}_{t,T}$ is the root-mean-square time average of the instantaneous volatility over the remaining period $[t, T]$, and the Call option price is the average of the B-S Call option prices over all possible volatility paths. In other words, the square of the implied B-S volatility $\sigma_{t,T}^{imp}$ appears to be a forecast of the average of temporal aggregation of $\bar{\sigma}_{t,T}^2$, where the instantaneous volatility is viewed as a flow variable.

It is now well known that, when the assumptions of Lemma 3.1 are satisfied, this model is able to reproduce some stylized facts regarding implied volatilities. For example, a symmetric smile is well explained by this stochastic volatility model. According to Renault and Touzi [RN96], assuming that the two driving Brownian motions are uncorrelated, provided $\bar{\sigma}_{t,T}$ is an L^2 random variable, the implied volatility curve[4] $I(K)$ for fixed t, S_t, T, is a smile, that is, it is locally convex around the minimum $K_{min} = S_t e^{r(T-t)}$, which is the forward price of the stock.

Moreover, a striking empirical evidence that emerges from numerous studies (see, e.g. [HW87] and [RN96]) is the decreasing amplitude of the smile as a function of time to maturity. For short maturities the smile is very

---

[4]From now on, when we will talk about implied volatility referring to a model other than the Black-Scholes one, we will understand that the Call option price for that model is an invertible function of the stock price (or in some cases of the strike price), with all the other parameters fixed. In this setting, the implied volatility is defined, as in the Black-Scholes case, by inverting the Call option function.

pronounced, but it almost completely disappears for longer maturities.

However, the decrease of the smile amplitude typically observed when time to maturity increases, appears to be much slower than implied by the Hull-White model. This evidence is clearly related to the volatility persistence, and shows that the volatility as modeled in the Hull-White model is not persistent enough to conform to empirical evidence, a direct consequence of the fact that the autocorrelation function of volatility decays too fast. Thus, the model exhibits short memory. These shortcomings motivate the need to extend such model.

3.2.2   FSV model

In a pioneering paper, Comte and Renault [CR98] proposed to model the logarithm of the volatility using a stationary fractional Ornstein-Uhlenbeck process. This was the first fractional volatility model in the literature, and it was then named *fractional stochastic volatility* (FSV) model. The main goal was to extend the Hull and White SV model in order to capture the well-documented evidence of persistence, long memory and mean-reversion of volatility in the time series of logarithmic returns. For this reason, the fBm B_t^H is chosen with $1/2 < H < 1$. Indeed, recall from Section 1.2 of Chapter 2 that the stationary fOU exhibits LRD and persistence for $H \in (1/2, 1)$, while loses LRD and becomes antipersistent for $H \in (0, 1/2)$. The dynamics under the physical measure $\mathbb{P}$ of the risky asset $(S_t)_{t \geq 0}$ and of the log-volatility $(\log(\sigma_t))_{t \geq 0}$ is described by the following SDE:

$$dS_t = \mu(t, S_t)S_t dt + \sigma_t S_t dB_t \tag{3.20}$$

$$d\log(\sigma_t) = \lambda(\eta - \log(\sigma_t))dt + \nu dB_t^H \tag{3.21}$$

where $\eta \in \mathbb{R}$, λ and ν are positive parameters and $\mu : [0, T] \times \mathbb{R} \to \mathbb{R}$ is a deterministic function. Furthermore, Comte and Renault consider a Hull and White framework where B_t^H and B_t are independent. For notation purpose, we will denote from now on the process $(\log(\sigma_t))_{t \geq 0}$ by $(X_t)_{t \geq 0}$. Therefore, Equation (3.21) becomes:

$$dX_t = \lambda(\eta - X_t)dt + \nu dB_t^H \tag{3.22}$$

A solution to (3.22) is mean-reverting. In fact, for $\lambda \geq 0$, the drift is positive if $X_t < \eta$, while it is negative if $X_t > \eta$. Thus, η can be interpreted as the long-term mean value of log-volatility and λ as the speed of mean reversion. Equation (3.22) is in a slightly more general form than the one we introduced

in Chapter 2, in that the constant $\lambda\eta$ has been added to the drift term. It can be shown that the theory developed in Chapter 2 is immediately adapted to this new setting, without changing the probabilistic properties of the process. The stationary solution $(X_t^H)_{t\geq 0}$ of (3.22), called *fractional Ornstein-Uhlenbeck process*, is given by:

$$X_t^H = \eta + \nu \int_{-\infty}^t e^{-\lambda(t-u)} dB_u^H \qquad (3.23)$$

or equivalently

$$\sigma_t = \exp\Big(\eta + \nu \int_{-\infty}^t e^{-\lambda(t-u)} dB_u^H\Big) \qquad (3.24)$$

Note that, as $(X_t^H)_{t\geq 0}$ is stationary, $(\sigma_t)_{t\geq 0}$ is stationary as well. Comte and Renault [CR98] show that the volatility process itself (and not only its logarithm) entails the long memory property. More precisely, they show that

$$\rho(t) = \text{Cov}[\sigma_h, \sigma_{h+t}] \sim O(t^{2\alpha-1}), \qquad \text{for } t \to \infty \qquad (3.25)$$

where $\alpha = H - 1/2$, in accordance with Definition 2.4.

The FSV model of Comte and Renault is able to fix some of the limitations of the Hull-White model by incorporating mean-reversion, high persistence and long memory. Also, thanks to a higher persistence, the decrease in the smile amplitude of the implied volatility surface appears to be slower than implied from the Hull-White model, which is another good point in favor of the FSV. However, this is not sufficient to provide a good fit of the volatility surface (especially for short expirations). Particularly, this model generates a term structure of at-the-money (ATM) volatility skew[5] that is increasing with τ (at least for small values of τ), which is inconsistent with the observed stylized fact that the ATM skew is well approximated by power-law functions of τ proportional to $\tau^{-\alpha}$, for $0 < \alpha < 1/2$.
This suggests that we adopt a different model.

---

[5]$\psi(\tau) = \Big|\dfrac{\partial}{\partial k}\hat{\sigma}^{impl}(k,\tau)\Big|_{k=0}.$

3.3 RFSV model

In this section we turn to the Rough Fractional Stochastic Volatility (RFSV) model introduced by Gatheral, Jaisson and Rosenbaum [GJR14] in 2014 for the log-volatility process. The RFSV model is a variant of the FSV model of Comte and Renault, where the Hurst parameter H of the fractional Brownian motion is now $H \in (0, 1/2)$. The consequent increased irregularity of the trajectories of the sample paths explains the name 'Rough FSV' given to the model. The SDE of the model is the same as in the FSV model, with the log-volatility under the physical measure $\mathbb{P}$ satisfying

$$d\log(\sigma_t) = \lambda(\eta - \log(\sigma_t))dt + \nu dB_t^H \qquad (3.26)$$

We now present the main reasons, based on both historical and risk-neutral data, that justify the use of a fBm with $H \in (0, 1/2)$. The graphs of this section are taken from [GJR14], and refer to the S&P500 from January 2000 to April 2013. We end the section by comparing the FSV and RFSV models with respect to risk-neutral data.

3.3.1 Historical data

Gatheral et al. derive the RFSV model starting with the observation that both the increments of log-volatility and the fBm satisfy two important properties:

- Scaling property;

- Gaussian distribution,

for which they provide empirical evidence. Moreover, when estimating the Hurst parameter they find low values of $H(\approx 0.1)$.

To begin, let us suppose we observe the volatility process $(\sigma_t)_{t \geq 0}$ at equidistant discrete points $0 = t_0, t_1, ..., t_N = T$, and let $\Delta = t_i - t_{i-1}$ be the time step. We thus observe the values $\sigma_0, \sigma_\Delta, ..., \sigma_{N\Delta}$. For $q > 0$, define

$$m(q, \Delta) = \frac{1}{N} \sum_{k=1}^{N} \big| \log(\sigma_{k\Delta}) - \log(\sigma_{(k-1)\Delta}) \big|^q. \qquad (3.27)$$

Assuming that the increments of the log-volatility process are stationary and that the law of large numbers can be applied with $N \to \infty$ (e.g. supposing

that the increments of the log-volatility are i.i.d.), $m(q, \Delta)$ is an unbiased and consistent estimator of

$$\mathbb{E}[|\log(\sigma_\Delta) - \log(\sigma_0)|^q] \qquad (3.28)$$

and thus $m(q, \Delta)$ can be seen as the empirical counterpart of (3.28). Gatheral et al. find empirically (see Figure 3.1), for each given $q > 0$, a linear relation between $\log(m(q, \Delta))$ and $\log(\Delta)$:

$$\log(m(q, \Delta)) = \log(K_q) + \zeta_q \log(\Delta) \qquad (3.29)$$

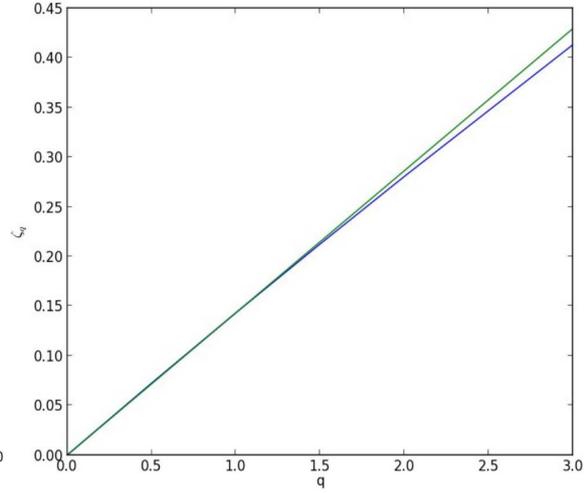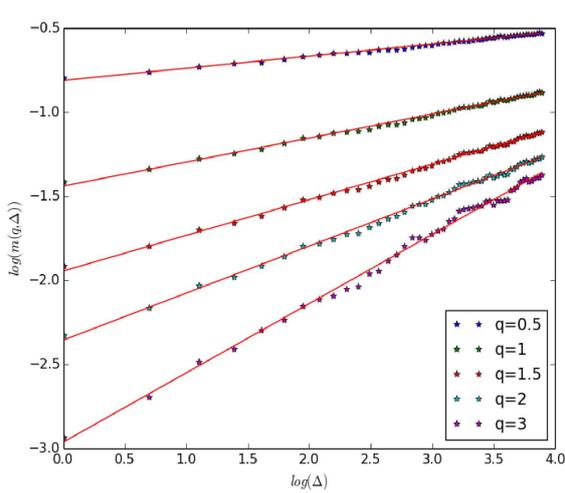where K_q is a constant and ζ_q is the slope of the line associated to q.

Figure 3.1: $\log(m(q, \Delta))$ as a function of q Figure 3.2: ζ_q as a function of q

Under the assumptions above (stationarity and LLN), $m(q, \Delta)$ is (approximately) equal to $\mathbb{E}[|\log(\sigma_\Delta) - \log(\sigma_0)|^q]$. We have then the following property for the expectation, called *scaling property*:

$$\mathbb{E}[|\log(\sigma_\Delta) - \log(\sigma_0)|^q] = K_q \Delta^{\zeta_q}. \qquad (3.30)$$

Moreover, when plotting ζ_q against q, they find again a linear relation: $\zeta_q = cq$, with c a constant (see Figure 3.2). Now, if we recall Proposition 1.1 of Chapter 2, thanks to the self-similarity and stationarity of increments of fBm, we have that

$$\mathbb{E}\left[|B_t^H - B_s^H|^q\right] = \mathbb{E}\left[|B_{t-s}^H|^q\right] = |t - s|^{qH} \mathbb{E}\left[|B_1^H|^q\right]. \qquad (3.31)$$

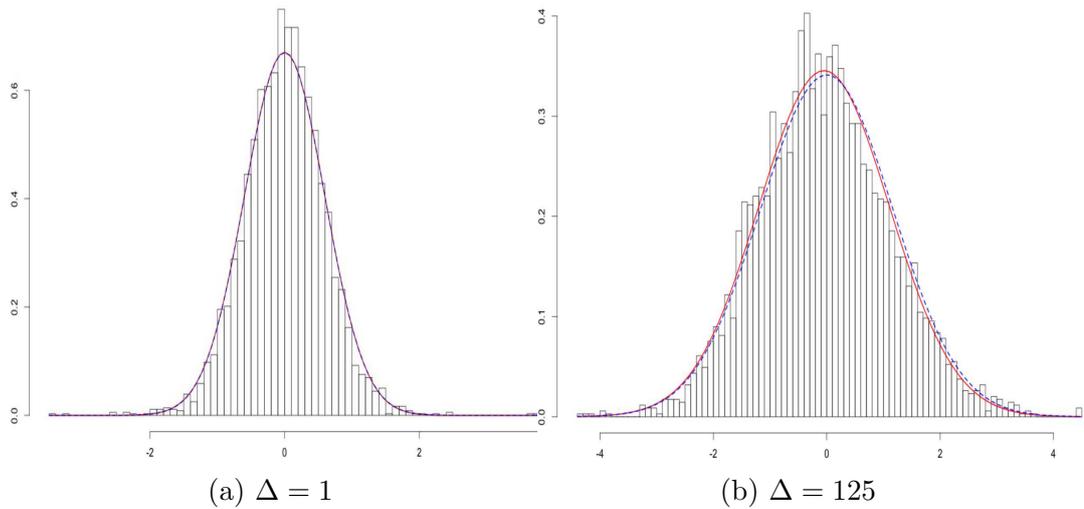(a) $\Delta = 1$ (b) $\Delta = 125$

Figure 3.3: Histograms for different time lags Δ of the increments of log-volatility $\log(\sigma_{t+\Delta}) - \log(\sigma_t)$ of the S&P500; normal fit in red; normal fit rescaled by Δ^H in blue.

For $t - s = \Delta$, we see that the processes $(\log(\sigma_t))_{t \geq 0}$ and $(B_t^H)_{t \geq 0}$ enjoy the same scaling property with $c = H$ and $K_q = \mathbb{E}\big[|B_1^H|^q\big]$.

In a study by Andersen et al. [AB97], the authors report the well-established stylized fact that the distribution of increments of log-volatility process $(\log(\sigma_{t+\Delta}) - \log(\sigma_t))_{t \geq 0}$ is very close to a Gaussian distribution. The Gaussianity of log-increments is confirmed with great accuracy by Gatheral et al. for different values of lag Δ (1, 5, 25 and 125 days) (see Figure 3.3 for two values of Δ). However, the fit is not perfect, as the distributions of the increments of log-volatility exhibit slightly thicker tails than the normal distribution, another well-known stylized fact.

From the Gaussianity of log-volatility increments, combined with the scaling property, it seems logical to let the process $(\log(\sigma_t))_{t \geq 0}$ be a fBm. When estimating the Hurst parameter H (by evaluating the slope of the linear regression of ζ_q against q), Gatheral et al. find consistently low values, between 0.06 and 0.20. These findings have been confirmed by Bennedsen et al. [BLP21] in 2016 with a study on 2,000 US equities.

3.3.2   Model specification

If we let $(\log(\sigma_t))_{t \geq 0}$ be a fBm, we find in a first step the simple model:

$$\log(\sigma_{t+\Delta}) - \log(\sigma_t) = \nu(B_{t+\Delta}^H - B_t^H), \tag{3.32}$$

where $\nu \in \mathbb{R}_{>0}$ is a positive constant. However, in this model the volatility process is not stationary, and this may render mathematical tractability and modeling rather difficult. Therefore, we are lead to impose stationarity by modeling the log-volatility as a stationary fractional Ornstein-Uhlenbeck process, with $H \in (0, 1/2)$:

$$d\log(\sigma_t) = \lambda(\eta - \log(\sigma_t))dt + \nu dB_t^H, \tag{3.33}$$

where $\eta \in \mathbb{R}$ is the long-term mean and $\lambda \in \mathbb{R}_{\geq 0}$ is the velocity of mean reversion. Hence, as in the FSV model, the stationary solution $(X_t^H)_{t \geq 0} = (\log(\sigma_t))_{t \geq 0}$ of (3.33) is

$$X_t^H = \eta + \nu \int_{-\infty}^{t} e^{-\lambda(t-u)} dB_u^H, \tag{3.34}$$

or equivalently

$$\sigma_t = \exp\Big(\eta + \nu \int_{-\infty}^{t} e^{-\lambda(t-u)} dB_u^H\Big). \tag{3.35}$$

The volatility process is now stationary. However, remember that the two motivations that lead us to model $(\log(\sigma_t))_{t \geq 0}$ as a fBm were that both processes have Gaussian increments and exhibit the same scaling property. We need to verify that these two properties are still valid (at least to a first approximation) when we model $(\log(\sigma_t))_{t \geq 0}$ as a fOU. In other words, we want $(X_t^H)_{t \geq 0}$ to 'behave' as a fBm. Actually, this is exactly what happens (at time scales smaller than T), if we assume $\lambda << 1/T$. More precisely, Gatheral et al. show that as $\lambda \to 0$:

$$\mathbb{E}\big[\sup_{t \in [0,T]} |X_t^H - X_0^H - \nu B_t^H|\big] \to 0 \tag{3.36}$$

and

$$\mathbb{E}\big[|X_t^H - X_0^H|^q\big] \to \nu^q K_q \Delta^{qH}. \tag{3.37}$$

When λ is small, the first limit suggests that we can proceed as if the log-volatility were a fBm, while the second limit implies that $(X_t^H)_{t \geq 0}$ approximately reproduces the scaling property of the fBm. Notice that a small λ implies a slow speed of mean reversion. As we will see, this is in line with a slower decay than traditional SV models of both the ACF and the ATM skew.

3.3.3 Autocorrelation function

We now analyze the autocorrelation function (ACF) of the volatility process, and we will reach a somewhat surprising conclusion about the RFSV model. We aim at proving that, when λ is small:

$$\mathbb{E}[\sigma_{t+\Delta}\sigma_t] = \mathbb{E}[e^{X_t^H + X_{t+\Delta}^H}] \approx e^{2\mathbb{E}[X_t^H] + 2\mathrm{Var}[X_t^H]}e^{-\nu^2\frac{\Delta^{2H}}{2}}. \qquad (3.38)$$

We start by showing that the ACF of the log-volatility process $(X_t^H)_{t\geq 0}$ has the following form, when $\lambda \to 0$:

$$\mathrm{Cov}[X_t^H, X_{t+\Delta}^H] = \mathrm{Var}[X_t^H] - \frac{1}{2}\nu^2\Delta^{2H} + o(1). \qquad (3.39)$$

By the stationarity of $(X_t^H)_{t\geq 0}$ follows that

$$\begin{aligned}
\mathbb{E}[(X_t^H - X_{t+\Delta}^H)^2] &= \mathbb{E}[(X_t^H)^2] + \mathbb{E}[(X_{t+\Delta}^H)^2] - 2\mathbb{E}[X_t^H X_{t+\Delta}^H] \\
&= \mathrm{Var}[X_t^H] + \mathrm{Var}[X_{t+\Delta}^H] - 2\mathrm{Cov}[X_t^H, X_{t+\Delta}^H] \qquad (3.40) \\
&= 2\mathrm{Var}[X_t^H] - 2\mathrm{Cov}[X_t^H, X_{t+\Delta}^H]
\end{aligned}$$

and that, for $\lambda \to 0$

$$\mathbb{E}[(X_t^H - X_{t+\Delta}^H)^2] \to \nu^2 K_2 \Delta^{2H} \qquad (3.41)$$

with K_2 a constant, which follows directly from (3.37).

Now, since $(X_t^H)_{t\geq 0}$ is a Gaussian process, we have that:

$$\mathbb{E}[\sigma_t \sigma_{t+\Delta}] = \mathbb{E}[e^{X_t^H + X_{t+\Delta}^H}] = e^{\mathbb{E}[X_t^H] + \mathbb{E}[X_{t+\Delta}^H] + \mathrm{Var}[X_t^H]/2 + \mathrm{Var}[X_{t+\Delta}^H]/2 + \mathrm{Cov}[X_t^H, X_{t+\Delta}^H]}. \qquad (3.42)$$

Replacing (3.39) in the last expression, we find (3.38).

It follows that, in the RFSV model, $\log(\mathbb{E}[\sigma_t \sigma_{t+\Delta}])$ is (approximately) linear in Δ^{2H}. This property is very well satisfied on data, as shown by Figure 3.4, where the logarithm of the empirical counterpart of $\mathbb{E}[\sigma_{t+\Delta}\sigma_t]$ (i.e. the sample mean estimator) is plotted against Δ^{2H}, with $H = 0.14$.

A consequence of Equation (3.38) is that the autocovariance function (and thus the ACF too) of the volatility process decays exponentially, opposing the widely believed stylized fact of a power-law decay. In particular, an exponential decay implies that the volatility process does not exhibit long memory. This is confirmed empirically by Gatheral et al., who show that the log-log plot of the autocovariance function $\rho(\Delta) = \mathrm{Cov}(X_t, X_{t+\Delta})$ does not yield a straight line (see Figure 3.5).

Now, if volatility has short memory, how can we explain the effect of persistence of the observed volatility time series? Gatheral et al. argue that
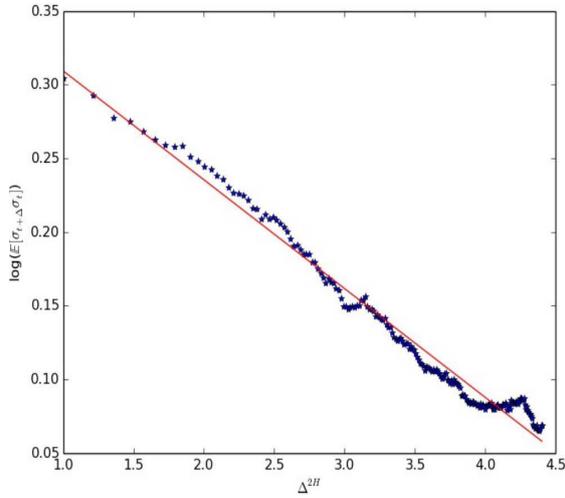
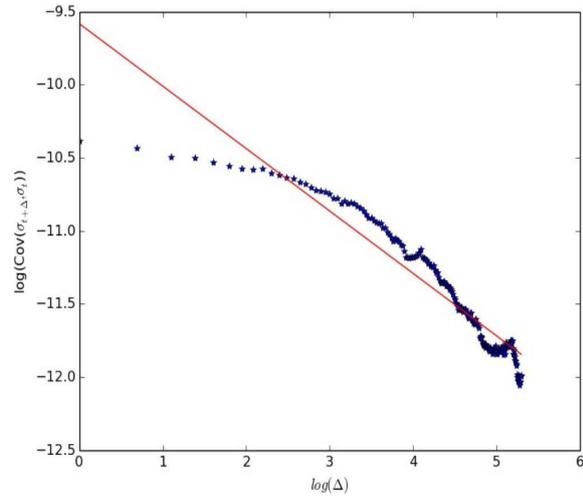Figure 3.4: Empirical counterpart of $\log(\mathbb{E}[\sigma_t \sigma_{t+\Delta}])$ against Δ^{2H}

Figure 3.5: Empirical counterpart of $\log(\text{Cov}[\sigma_t, \sigma_{t+\Delta}])$ against $\log(\Delta)$

the general consensus around the long memory of volatility stems from the fact that classical estimation procedures often either misinterpret spurious long memory or make rather strict modeling assumptions. To show this, by applying standard financial econometric procedures (e.g. those employed by Andersen et al. [And+03]), for estimating long memory to the RFSV model, they indeed identify long memory, which is clearly incompatible with the structure of the ACF. Moreover, many authors consider persistence and long memory to be the same, and thus they often draw conclusions about long memory based on evidence about persistence. The RFSV model is a clear example that the two concepts are distinct, and the phenomenon of persistence can occur even in models with short memory.

3.3.4 Risk-neutral data

The modeling power of the RFSV reveals not only in its consistency with empirical observations, but also in its capacity to reproduce faithfully and with a remarkable accuracy the shape of the volatility surface. We already stressed in Chapter 2 that traditional stochastic volatility models generally fail to generate an ATM volatility skew consistent with the observed one. This is true for some fractional volatility models as well, like the FSV of Comte and Renault, as said in the previous section. On the other hand, SV models with rough volatility (i.e. models driven by fBm with Hurst parameter less than 1/2) provide a good fit of the volatility surface. This was one of the conclusions of the analysis of Fukasawa [Fuk11], who shows

that a stochastic volatility model where the volatility process is driven by a fractional Brownian motion with Hurst parameter H generates a term structure of ATM volatility skew of the form $\psi(\tau) \sim \tau^{H-1/2}$, at least for τ small. The analysis of Fukasawa implies that, since the ATM skew tipically observed is proportional to $\tau^{-\alpha}$ with $\alpha \in (0, 1/2)$, SV models with rough volatility (like the RFSV) are consistent with the observed term structure (see Figure 3.6), while for models with $H \in (1/2, 1)$ (like the FSV), the ATM skew is increasing in time to expiration τ, which is completely inconsistent with the observed skew. As a consequence, in order for a fractional SV model to generate a volatility surface with a reasonable shape, we need to have $H \in (0, 1/2)$.
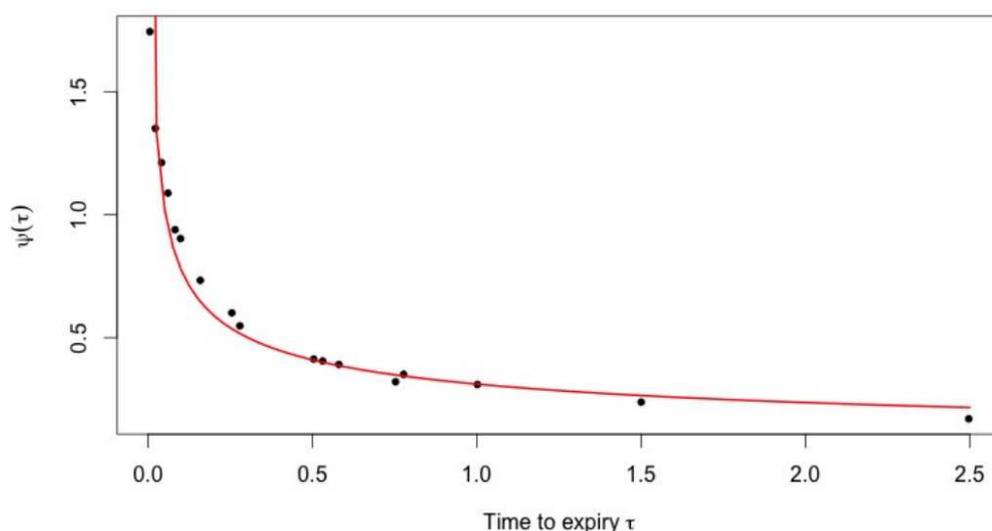
Figure 3.6: The black dots are non-parametric estimates of the S&P500 ATM volatility skew as of June 20, 2013; the red curve is the power-law fit $\psi(\tau) \sim 0.3\tau^{-0.4}$.

3.4   The Rough Heston model

The analyses of Fukasawa [Fuk11] discussed above show that rough volatility models are generally successfull in reproducing the volatility surface. On the other hand, the Heston stochastic volatility model (see Section 1.4.3 of Chapter 1) is able to reproduce several important stylized facts of historical data, including mean reversion, the leverage effect (see [BLT06]) and fat tails ([DBJ03]). Moreover, the existence of a fast and easily implemented closed form characteristic function of the log-stock price (as we will see in the next

chapter) allows for fast-pricing of European options.

In this section, we present an extension of the classical Heston model proposed by Euch and Rosenbaum [ER16] in 2016 that incorporates rough volatility, offering the best of the characteristics from both rough volatility and the classical Heston model.

Forward variance curve

First, recall from Section 1.4.3 that, under the risk-neutral measure $\mathbb{Q}$, the Heston model is given by[6]:

$$dS_t = rS_t dt + \sqrt{\nu_t} S_t dB_t^{1,\mathbb{Q}} \tag{3.43}$$

$$d\nu_t = \lambda(\eta - \nu_t)dt + \nu\sqrt{\nu_t}dB_t^{2,\mathbb{Q}} \tag{3.44}$$

where $\nu_t = \sigma_t^2$ is the instantaneous variance at time t. We define the forward variance curve (also called variance swap curve) as $\xi_t(u) = \mathbb{E}^{\mathbb{Q}}[\nu_u|\mathcal{F}_t]$, with $u \geq t$. Integrating (3.44) on $[t, u]$, we obtain:

$$\nu_u = \nu_t + \int_t^u \lambda(\eta - \nu_s)ds + \int_t^u \nu\sqrt{\nu_s}dB_s^{2,\mathbb{Q}}. \tag{3.45}$$

For future purpose, we express the Heston model in forward variance curve form. This is obtained by taking the conditional expectation from both sides of equation (3.45) (see [DF06] pg. 66):

$$d\xi_t(u) = \lambda(\eta - \xi_t(u))du \tag{3.46}$$

whose explicit solution is

$$\xi_t(u) = (\nu_t - \eta)e^{-\lambda(u-t)} + \eta. \tag{3.47}$$

Taking the differential form of (3.47) and using (3.44), we find:

$$d\xi_t(u) = \left(\lambda(\eta - \nu_t)dt + \nu\sqrt{\nu_t}dB_t^{2,\mathbb{Q}}\right)e^{-\lambda(u-t)} + \nu_t\lambda e^{-\lambda(u-t)}dt - \eta\lambda e^{-\lambda(u-t)}dt \tag{3.48}$$

which gives

$$d\xi_t(u) = \nu e^{-\lambda(u-t)}\sqrt{\nu_t}dB_t^{2,\mathbb{Q}}. \tag{3.49}$$

---

[6]We adapt the notation to match those used in the RFSV model.

Roughening Heston

El Euch, Gatheral and Rosenbaum [EGR19] introduce roughness in the volatility proces using the Mandelbrot-Van Ness representation of the fBm. We clearly see from Proposition 1.3 and Proposition 1.4 in Chapter 2 that the kernel $(u-s)^{H-1/2}$ plays a central role in the representation of the fBm, for $H \in (0, 1/2)$. Indeed, such kernel determines the roughness of the sample paths of the fBm. More precisely, given a Brownian motion $(B_t)_{t \geq 0}$, the process $\left(\int_0^t (u-s)^{H-1/2} dB_s \right)_{t \geq 0}$ known as *Volterra* fBm, has almost surely γ-Hölder continuous sample paths, for any $\gamma \in (0, H)$. Hence, in order to get a rough behaviour of the volatility in the Heston model, El Euch, Gatheral and Rosenbaum extend the Heston model by modeling the instantaneous variance using the kernel $(u-s)^{H-1/2}$:

$$\nu_u = \theta_t(u) - \frac{\lambda}{\Gamma(H+1/2)} \int_t^u \nu_s(u-s)^{H-1/2} ds + \frac{\nu}{\Gamma(H+1/2)} \int_t^u \sqrt{\nu_s}(u-s)^{H-1/2} dB_t^{2,\mathbb{Q}}$$
(3.50)

where $\theta_t(u)$ is a $\mathcal{F}_t$-measurable random variable that makes the model time consistent and plays the role of a time-varying mean-reversion value, and Γ denotes the Gamma function. It can be shown that, in the limit $H \to 1/2$, we retrieve the classical Heston model, and that, for any $\gamma \in (0, H)$, the sample paths of $(\nu_u)_{u \geq 0}$ are almost surely γ-Hölder continuous.

El Euch and Rosenbaum [ER17] observe that we can simplify (3.50) if we consider the forward variance curve introduced above and set $\lambda = 0$. In fact, they show that, if we assume that the function $u \to \xi_t(u)$ admits a fractionl derivative[7] of order $\alpha \in (0, 1)$, then the mean-reversion function $\theta_t(.)$ can be chosen by taking

$$\lambda \theta_t(u) = D^\alpha(\mathbb{E}[\xi_t(u)] - \nu_t)(u) + \lambda \mathbb{E}[\xi_t(u)]. \quad (3.51)$$

This means that, for any $\lambda > 0$, provided the existence of a fractional derivative, we can choose the function $\theta_t(.)$ so that the model is consistent with the forward variance curve.

Considering the model in the asymptotic setting $\lambda \to 0$ is particularly convenient. The first reason is that, as explained in the RFSV model, a small value of λ implies a slow speed of mean reversion, and this helps to better fit the volatility surface. In fact, in the classical Heston model, as in other classical SV models (see, e. g., Hull-White), the decrease of the smile amplitude is too fast. Therefore, taking λ small (or even 0) allows for a slower decay

---

[7]The definitions from Fractional Calculus are provided in the next section; for the moment we can accept the fact that in most 'practical' situations, the condition (3.51) is satisfied.

and thus makes the model more adapt in reproducing the volatility surface. Moreover, taking $\lambda = 0$ gives us a mathematically simpler a more tractable model. In fact, assuming $\lambda = 0$ and taking the conditional expectation from both sides of (3.50) we have that the stochastic integral vanishes and thus $\xi_t(u) = \mathbb{E}^{\mathbb{Q}}[\nu_u | \mathcal{F}_t] = \theta_t(u)$. Therefore we obtain:

$$\nu_u = \xi_t(u) + \frac{\nu}{\Gamma(H + 1/2)} \int_t^u \sqrt{\nu_s}(u - s)^{H-1/2} dB_s^{2,\mathbb{Q}} \qquad (3.52)$$

and, for $h > 0$,

$$\nu_u = \xi_{t+h}(u) + \frac{\nu}{\Gamma(H + 1/2)} \int_{t+h}^u \sqrt{\nu_s}(u - s)^{H-1/2} dB_s^{2,\mathbb{Q}}. \qquad (3.53)$$

Subtracting the two equations we obtain

$$\xi_{t+h}(u) - \xi_t(u) = \frac{\nu}{\Gamma(H + 1/2)} \int_t^{t+h} \sqrt{\nu_s}(u - s)^{H-1/2} dB_s^{2,\mathbb{Q}}. \qquad (3.54)$$

Taking the limit $h \to 0$ gives

$$d\xi_t(u) = \frac{\nu}{\Gamma(H + 1/2)} \sqrt{\nu_t}(u - t)^{H-1/2} dB_t^{2,\mathbb{Q}}. \qquad (3.55)$$

The link with the classical Heston model in forward variance form (3.49) can be made by setting $\lambda = 0$ and multiplying by the kernel $\frac{(u-t)^{H-1/2}}{\Gamma(H+1/2)}$.

Chapter 4

Option Pricing methods under Stochastic Volatility

The final part of this thesis concerns the numerical methods needed for implementing some of the models described in the preceding chapters. We will confine ourwelves to option pricing, for which many efficient techniques exist. In this context, there are methods that give exact solutions or semi-exact solutions, but unfortunately this is only possible in a limited amount of 'fortunate' cases. In the other cases, numerical techniques are necessary in order to approximate the solution. Organizing these techniques, roughly speaking, leads to a first distinction between a 'stochastic' approach, based on *Monte Carlo techniques*, and a 'deterministic' approach, based on solving *Partial (Integro) Differential Equations* or on *Numerical Integration*. For the first approach, we will present methods for the numerical solution of stochastic differential equations; for the second approach we will only analyze Numerical Integration techniques based on Fourier methods. We will also implement on the computer some algorithms for the simulation of the Heston stochastic volatility model, and show the pros and cons of each approach.

4.1 Monte Carlo simulation

The Monte Carlo method is a simple technique of numerical approximation of the expected value of a random variable. It is based on the strong law of large numbers: if $(X_n)_{n\in\mathbb{N}}$ is a sequence of integrable i.i.d random variables, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mathbb{E}[X_1], \qquad \text{a.s.} \tag{4.1}$$

Consequently, the first step to approximate the mean value of a random variable whose distribution is known, consists of generating M independent realizations of that random variable. This poses the problem of how to 'artificially' create randomness. Fortunately, at least for the most well-known distributions (and in particular for the Normal distribution), 'pseudo-random' number generators exist, i.e. algorithms that output a sequence of numbers that can be used as a replacement for an independent and identically distributed sequence of 'true random numbers'. We will not discuss these interesting algorithms. Once M i.i.d. realizations of the random variable are generated, using the *law of large numbers* and the *central limit theorem* we can calculate confidence intervals for the mean value of the random variable. In the context of financial mathematics it is used in many circumstances, particularly for the computation of the Greeks and the pricing of derivatives. Regarding this last issue, Monte Carlo method is particularly useful because, when pricing options, we want to calculate an expression like $V_T = e^{-rT}\mathbb{E}[F(S_T)]$ under an appropriate probability measure (the risk-neutral measure). More precisely, Monte Carlo method can be used to approximate the value of an option as follows:

1. Partition the time interval $[0, T]$ with a stepsize $\delta = T/N$ for some integer N, and define the time points $t_i = i\delta$, $i = 0, ..., N$.

2. For each $i = 0, ..., N$ (the time steps) and $j = 0, ..., M$ (the j-th realization of the value of the underlying asset), generate asset values $S_{i,j}$. Here the realization of the value of the asset is made according to its distribution under the risk-neutral measure.

3. For each $j = 0, ..., M$, compute the payoff value H_j of the option. Notice that, in the case of a path dependent derivative, the value of the option that we can calculate is $H_j = H(T; S_{0,j}, ..., S_{N,j})$, which could be different from the true value of the option, and thus may be considered as an approximation (the approximation becomes more accurate for grater values of N).

4. Compute the *Monte Carlo estimate*

$$\mathbb{E}^{\mathbb{Q}}[H(T, S)|\mathcal{F}_{t_0}] \approx \frac{1}{M} \sum_{j=1}^{M} H_j. \tag{4.2}$$

5. Approximate the value of the option at time t_0 by

$$V(t_0, S) \approx e^{-r(T-t_0)} \frac{1}{M} \sum_{j=1}^{M} H_j. \tag{4.3}$$

6. Use probability theory, (typically the *law of large numbers* and the *central limit theorem* are sufficient), to find estimates and confidence region for the convergence of the approximation (4.2).

Point 2. is often problematic. In fact, usually, the distribution of the underlying asset (i.e. the distribution of the solution of the SDE of the model) is not known explicitly. In this case, in order to obtain some realizations of the underlying asset we will use discretized versions of the SDE, and we will therefore build a numerical scheme. In this way, the discretization error of the SDE must be added to the errors given by the approximations above.
In the next section we present two of these numerical schemes, the Euler-Maruyama method and the Milstein method.

4.2 Simulating Stochastic Differential Equations

In this section we discuss the approximate numerical solution of stochastic differential equations. A solution X of the SDE will satisfy an equation of the form $LX = 0$, with L an operator. In this context, a numerical scheme can be thought as a collection of indexed operators L^δ acting on stochastic processes, where δ indicates the time step discretization. As δ approaches 0 we want that the operator L^δ 'generates' approximating solutions X^δ 'converging' to X (we will give precise definitions of 'convergence'). There are a number of properties the scheme should satisfy, in order to generate a good approximation to the original SDE:

- *regularity* of the solution X, derived from hypotheses on the coefficients;

- *consistency* of the scheme, i.e. the fact that, in an appropriate norm, $L^\delta \xrightarrow{\delta \to 0} L$.

- *stability* of the scheme, i.e. the fact that, for two solutions X^δ, Y^δ of $L^\delta Z = 0$, we require that $\max_{t \in [0,T]} ||X_t^\delta - Y_t^\delta|| \leq F(X_0, Y_0, \delta, T)$, with F an appropriate function and $||.||$ an appropriate norm. In other words, we want that 'if two solutions start close to each other, they remain close to each other' and that 'small changes in the initial conditions have little effects on the evolution of the solution';

We will present the Euler-Maruyama scheme, some improvements to it, the Milstein method, and we will discuss the notions of weak and strong convergence. In option pricing it is tipically only weak convergence that is required.

4.2.1   The Euler-Maruyama scheme

We begin with the simplest, most widely used method for SDEs, and illustrate its basic convergence properties.

In this Section $T > 0$ and $K > 0$ are fixed constants.

Suppose we have an SDE of the form

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t \tag{4.4}$$

where $\mu : [0, T] \times \mathbb{R} \to \mathbb{R}$ and $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ verify the assumption

$$|\mu(t, x) - \mu(s, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leq K(|t - s| + |x - y|^2) \tag{4.5}$$

for $x, y \in \mathbb{R}$ and $t, s \in [0, T]$. By definition, the exact solution of (4.4) satisfies

$$X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s \qquad t \in [0, T]. \tag{4.6}$$

If we define the operator

$$L_t := X_t - X_0 - \int_0^t \mu(s, X_s)ds - \int_0^t \sigma(s, X_s)dB_s \qquad t \in [0, T], \tag{4.7}$$

then Assumption (4.5) ensures (see [Pas11], Section 9.1) the existence of a strong solution $X = (X_t)_{t \in [0,T]}$ of $L_t X = 0$, $t \in [0, T]$, such that

$$||X||_T := \sqrt{\mathbb{E}\left[\sup_{t \in [0,T]} X_t^2\right]} < \infty. \tag{4.8}$$

It can be shown (see [Pas11], Section 9.1) that, if we define $\mathbb{M}$ as the vector space of continuous $\mathcal{F}_t$-adapted processes $(X_t)_{t \in [0,T]}$ such that $||X||_T < \infty$, then $(\mathbb{M}, ||.||_T)$ is a semi-normed complete space.

Following the standard discretization approach used for deterministic ODEs, we define a stepsize $\delta = T/N$ for some integer N, and compute approximation solutions at times $t_i = i\delta$, $i = 0, ..., N$. For $Z \in \mathbb{M}$, we define the discretized operator L_t^δ as

$$
\begin{aligned}
L_t^\delta Z := Z_t - Z_0 &- \int_0^t \sum_{n=0}^{N-1} \mu(t_n, Z_{t_n}) \mathbb{1}_{[t_n, t_{n+1}]}(s)ds \\
&- \int_0^t \sum_{n=0}^{N-1} \sigma(t_n, Z_{t_n}) \mathbb{1}_{[t_n, t_{n+1}]}(s)dB_s, \qquad t \in [0, T].
\end{aligned}
\tag{4.9}
$$

Then, given an initial datum $X_0 \in \mathbb{R}$, the approximating process X^δ, defined as the stochastic process in $\mathbb{M}$ which solves $L_t^\delta X^\delta = 0$, for $t \in [0, T]$, with initial condition $X_0^\delta = X_0$, satisfies

$$X_{t_{n+1}}^\delta = X_{t_n}^\delta + \mu(t_n, X_{t_n}^\delta)\delta + \sigma(t_n, X_{t_n}^\delta)(B_{t_{n+1}} - B_{t_n}), \qquad (4.10)$$

called Euler-Maruyama (E-M) scheme. Comparing (4.6) and (4.10), we see that, given a realization ω, over $[t_n, t_{n+1}[$ the EM method is using the a left-end point Riemann-Stieltjes sum to approximate the two integrals. Thanks to the fact that X^δ is an adapted process, this ensures that the stochastic integral in (4.9) is consistent with the definition of Itô integral.

Simulation of the Euler-Maruyama scheme is easily performed by noting that the term $B_{t_{n+1}} - B_{t_n} \sim \mathcal{N}(0, \delta)$. The discretized SDEs (4.10) can therefore be simulated by producing N independent normally distributed random variables $Z_i \sim \mathcal{N}(0, 1)$, $i = 0, ..., N - 1$, and calculating the approximating solution X^δ by means of

$$X_{t_{n+1}}^\delta = X_{t_n}^\delta + \mu(t_n, X_{t_n}^\delta)\delta + \sigma(t_n, X_{t_n}^\delta)\sqrt{\delta}Z_n. \qquad (4.11)$$

We now quickly show that the Euler-Maruyama method satisfies the properties, highlighted at the beginning of the chapter, of regularity, consistency and stability. The proofs can be found in [Pas11].

Proposition 4.1 (Regularity)**.** *The solution X of $L_t X = 0$ is such that, for each t, t' with $0 \leq t < t' \leq T$,*

$$\mathbb{E}\left[\sup_{s \in [t,t']} |X_s - X_t|^2 \right] \leq K_1(t' - t), \qquad (4.12)$$

where K_1 is a constant that depends only on $T, \mathbb{E}[X_0^2]$ and K. [1]

This implies in particular (just take $s = t'$ in (4.12)) that the function $t \to X_t$ from $([0, T], |.|)$ onto the space $L^2(\Omega, F, \mathbb{P})$ with the L^2 norm is $\frac{1}{2}$-Hölder continuous.

The consistency of the discretized operator L^δ derives from the following proposition:

Proposition 4.2 (Consistency)**.** *Let $Y \in \mathbb{M}$ and for each t, t' with $0 \leq t < t' \leq T$*

$$\mathbb{E}\left[\sup_{s \in [t,t']} |Y_s - Y_t|^2 \right] \leq K_2(t' - t). \qquad (4.13)$$

---

[1]Recall that T and K are fixed from the beginning of the Section.

Then

$$||LY - L^\delta Y||_T \le C\delta^{1/2}, \tag{4.14}$$

where the constant C depends only on K, K_2 and T.

In the language of funtional analysis, the proposition above implies in particular that, as $\delta \to 0$, the operator L^δ converges to L in the strong operator topology (here the operators take values in the space of processes Y satisfying the assumptions of Proposition 4.2).

Finally, the property of stability derives from the following proposition, also known as Maximum Principle:

Proposition 4.3 (Stability)**.** *There exists a constant C_0, depending only on K and T such that, for every pair of processes $Y, Z \in \mathbb{M}$, we have*

$$||Y - Z||_T^2 \le C_0\big(\mathbb{E}[|Y_0 - Z_0|^2] + ||L^\delta Y - L^\delta Z||_T^2\big). \tag{4.15}$$

This kind of result guarantees the 'stability' of the numerical scheme, in the sense that, for two solutions Y^δ, Z^δ of $L^\delta Y = 0$, (4.15) becomes

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y_t - Z_t|^2\right] \le C_0\mathbb{E}[|Y_0 - Z_0|^2], \tag{4.16}$$

and this gives an estimate of the sensitivity of the solution with respect to some perturbation of the initial datum.

4.2.2 Weak convergence

There are two approaches for measuring the error in a discretized scheme, namely the *weak error* and the *strong error*.
Given a measurable function f, the *weak error*

$$e_\delta^{weak} := \sup_{n\in\{0,...,N\}} \big|\mathbb{E}[f(X_{t_n}^\delta)] - \mathbb{E}[(f(X_{t_n})]\big| \tag{4.17}$$

measures how well the method can approximate the mean of $f(X_t)$ (in the discretized points). It is typical to restric the test function f to be a member of a class of functions $\mathcal{C}$ such as polynomials of degree at most l.
We say that the method *converges weakly* if, for any function f in this class,

$$e_\delta^{weak} \to 0, \qquad \text{as } \delta \to 0. \tag{4.18}$$

Furthermore, we say that the method has a *weak order of convergence p* if, for any f in the class, there exists a constant k and a stepsize δ^* (both depending on f), such that[2]

$$e_\delta^{weak} \leq k\delta^p, \qquad \text{for all } 0 < \delta < \delta^*. \tag{4.19}$$

Other definitions of *weak order of convergence p* can be found in the literature, most of them being less restrictive. One of them requires that the error is calculated only at the final time, that is:

$$\left| \mathbb{E}[f(X_{t_N}^\delta)] - \mathbb{E}[f(X_{t_N})] \right| = O(\delta^p) \tag{4.20}$$

There exist several different conditions for the Euler-Maruyama scheme one can require in order to guarantee the weak convergence of order 1, in the sense of (4.20). For example, Theorem 14.5.2 of Kloeden and Platen [KP92] assumes that the functions μ and σ and the test functions be four times continuously differentiable with various polynomially bounded derivatives[3]. More generally, they prove that the Euler-Maruyama scheme has weak order of convergence β in the sense of (4.20), provided we assume μ and σ and the test functions, to be $2\beta + 2$ times differentiable with polynomially bounded derivatives.

4.2.3 Strong convergence

Whereas weak convergence measures the 'error of the means', strong convergence measures the 'mean of the errors'. The *strong error*

$$e_\delta^{strong} = \sup_{n=0,...,N} \mathbb{E}[|X_{t_n}^\delta - X_{t_n}|] \tag{4.21}$$

is found by taking the expectation of this error at each discretization time point. We say that the method *converges strongly* if

$$e_\delta^{strong} \to 0, \qquad as \ \delta \to 0, \tag{4.22}$$

and that the method has *strong order of convergence p* if there exists a constant k and δ^* such that[4]

$$e_\delta^{strong} \leq k\delta^p, \qquad for \ all \ 0 < \delta \leq \delta^* \tag{4.23}$$

---

[2]More precisely, we take p to be the largest value for which this holds.

[3]A function $f : \mathbb{R}^d \to \mathbb{R}$ is polynomially bounded if $|f(x)| \leq k(1 + ||x||^q)$, for some constants k and q and for all $x \in \mathbb{R}^d$.

[4]More precisely, we take p to be the largest value for which this holds.

Strong convergence typically requires less restrisctive hypotheses than the weak convergence in order to be guaranteed, even if for some methods the weak order is higher than the strong order. For example, the Euler-Maruyama scheme converges strongly under much simpler assumptions than the weak order, but with a strong order of $1/2$. To see this, we use the three propositions in the previous section. In fact, by the maximum principle, Proposition 4.3, we have that[5]

$$||X - X^\delta||_T^2 \leq C_0 ||L^\delta X - L^\delta X^\delta||_T^2 = C_0 \leq C\delta \qquad (4.24)$$

where the last inequality follows from the properties of consistency and regularity results above. This proves the following result:

Theorem 4.1 (Strong convergence of Euler-Maruyama). *Assuming* $X, X^\delta \in \mathbb{M}$, *there exists a constant* C *depending only on* K, T *and* $\mathbb{E}[X_0^2]$, *such that*

$$||X - X^\delta||_T \leq C\delta^{\frac{1}{2}}. \qquad (4.25)$$

For applications in option pricing, weak error criteria are most relevant because we want to ensure that prices (which are expectation under the risk-neutral measure) computed from X^δ are close to prices computed from X, and so weak error is more appropriate. It is nevertheless useful to be aware of strong error criteria to appreciate the merits of some discretization methods over others.

4.2.4 The Milstein scheme

The Milstein scheme improves the Euler-Maruyama discretization by adding a second diffusion term. The idea is that the approximation

$$\int_0^t \sum_{n=0}^{N-1} \sigma(t_n, X_{t_n}) \mathbb{1}_{[t_n, t_{n+1}[}(s) dB_s \approx \int_0^t \sigma(s, X_s) dB_s \qquad (4.26)$$

is the main source of error for the Euler-Maruyama scheme. To improve it, we estimate the error by using Itô's formula for $\sigma(s, X_s)$.
We start by the so-called stochastic Taylor expansion.

---

[5]In order for the first equality in (4.24) to be satisfied we need to extend the process X^δ to the whole interval $[0, T]$ in a way such that the hypotheses of Proposition (4.3) are satisfied. Such an extension exists (see Chapter 10.2 of Kloeden Platen [KP92]).

Stochastic Taylor expansion

Starting from the Itô's formula for a scalar valued function $U(t, X_t)$ of the solution $X = (X_t)_{t \in [0,T]}$ of the SDE (4.4), we have the integral representation:

$$U(t, X_t) = U(t_0, X_{t_0}) + \int_{t_0}^t M^0 U(s, X_s) ds + \int_{t_0}^t M^1 U(s, X_s) dB_s, \quad (4.27)$$

where the differential operators M^0 and M^1 are defined by

$$M^0 = \frac{\partial}{\partial t} + \mu \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2}, \qquad M^1 = \sigma \frac{\partial}{\partial x}. \quad (4.28)$$

Let us now apply the above formula to the integrand functions $U(t, x) = \mu(t, x)$ and $U(t, x) = \sigma(t, x)$. We obtain

$$
\begin{aligned}
X_t = & X_{t_0} \\
& + \int_{t_0}^t \left[\mu(t_0, X_{t_0}) + \int_{t_0}^s M^0 \mu(u, X_u) du + \int_{t_0}^s M^1 \mu(u, X_u) dB_u \right] ds \\
& + \int_{t_0}^t \left[\sigma(t_0, X_{t_0}) + \int_{t_0}^s M^0 \sigma(u, X_u) du + \int_{t_0}^s M^1 \sigma(u, X_u) dB_u \right] dB_s \\
= & X_{t_0} + \mu(t_0, X_{t_0}) \int_{t_0}^t ds + \sigma(t_0, X_{t_0}) \int_{t_0}^t dB_s + R_1(t, t_0),
\end{aligned}
$$
$$(4.29)$$

with the remainder

$$
\begin{aligned}
R_1(t, t_0) = & \int_{t_0}^t \int_{t_0}^s M^0 \mu(u, X_u) du ds + \int_{t_0}^t \int_{t_0}^s M^1 \mu(u, X_u) dB_u ds \\
& + \int_{t_0}^t \int_{t_0}^s M^0 \sigma(u, X_u) du dB_s \\
& + \int_{t_0}^t \int_{t_0}^s M^1 \sigma(u, X_u) dB_u dB_s.
\end{aligned}
$$
$$(4.30)$$

If we replace t_0 by t_n and t by t_{n+1}, and discard the remainder, then we retrieve the Euler-Maruyama method. Higher-order stochastic Taylor expansions are obtained by successively applying the Itô formula to the integrand functions in the remainder R_1. Unlike the deterministic case (where $\sigma \equiv 0$), there are now a number of different alternatives depending on which term we choose to expand. Now, recalling the euristic 'equality' $''dB^2 = dt''$, it may seem logical to expand the integrand $M^1 \sigma$ in the fourth double integral

of the remainder R_1 in (4.30), which is the one whose infinitesimal is of the lowest order. By doing this, we obtain the stochastic Taylor expansion

$$
\begin{aligned}
X_t =& X_{t_0} + \mu(t_0, X_{t_0}) \int_{t_0}^t ds + \sigma(t_0, X_{t_0}) \int_{t_0}^t dB_s \\
&+ M^1 \sigma(t_0, X_{t_0}) \int_{t_0}^t \int_{t_0}^s dB_u dB_s + R_2(t, t_0)
\end{aligned}
\tag{4.31}
$$

with the remainder

$$
\begin{aligned}
R_2(t, t_0) =& \int_{t_0}^t \int_{t_0}^s M^0 \mu(u, X_u) du\, ds + \int_{t_0}^t \int_{t_0}^s M^1 \mu(u, X_u) dB_u ds \\
&+ \int_{t_0}^t \int_{t_0}^s M^0 \sigma(u, X_u) du\, dB_s \\
&+ \int_{t_0}^t \int_{t_0}^s \int_{t_0}^u M^0 M^1 \sigma(v, X_v) dv\, dB_u dB_s \\
&+ \int_{t_0}^t \int_{t_0}^s \int_{t_0}^u M^1 M^1 \sigma(v, X_v) dB_v dB_u dB_s.
\end{aligned}
\tag{4.32}
$$

Then, replacing t_0 by t_n and t by t_{n+1}, and discarding the remainder R_2, we have the approximating solution

$$
\begin{aligned}
X_{n+1}^\delta =& X_n^\delta + \delta\mu(t_n, X_{t_n}^\delta) + \sigma(t_n, X_{t_n}^\delta)(B_{t_{n+1}} - B_{t_n}) \\
&+ M^1 \sigma(t_n, X_{t_n}^\delta) \int_{t_n}^{t_{n+1}} \int_{t_n}^s dB_u dB_s.
\end{aligned}
\tag{4.33}
$$

Since, by Itô's formula, the double integral can be evaluated as

$$
\int_{t_n}^{t_{n+1}} \int_{t_n}^s dB_u dB_s = \frac{1}{2}\left((B_{t_{n+1}} - B_t)^2 - \delta\right),
\tag{4.34}
$$

we finally obtain the Milstein scheme

$$
\begin{aligned}
X_{n+1}^\delta =& X_n^\delta + \delta\mu(t_n, X_{t_n}^\delta) + \sigma(t_n, X_{t_n}^\delta)(B_{t_{n+1}} - B_{t_n}) \\
&+ \sigma(t_n, X_{t_n}^\delta)\frac{1}{2}\left((B_{t_{n+1}} - B_t)^2 - \delta\right)\frac{\partial}{\partial_x}\sigma(t_n, X_{t_n}^\delta).
\end{aligned}
\tag{4.35}
$$

Since $(B_{t_{n+1}} - B_{t_n})^2 \sim \delta Z^2$, with $Z \sim \mathcal{N}(0, 1)$, the discretized SDEs (4.35) can be simulated by producing N independent normally distributed random

variables $Z_i \sim \mathcal{N}(0,1)$, $i = 0, ..., N-1$, and calculating the approximating solution X^δ by means of

$$X^\delta_{t_{n+1}} = X^\delta_{t_n} + \mu(t_n, X^\delta_{t_n})\delta + \sigma(t_n, X^\delta_{t_n})\sqrt{\delta}Z_n + \sigma(t_n, X_{t_n})\frac{\delta(Z_n^2 - 1)}{2}\frac{\partial}{\partial_x}\sigma(t_n, X_{t_n}).$$
(4.36)

Under appropriate assumptions on μ and σ, (see, e.g. Theorem 10.3.5 and Theorem 10.6.3 of Kloeden and Platen ([KP92]), the Milstein scheme has a strong order of convergence 1 and weak order 1. Therefore, in this case, the Milstein method is an improvement of the Euler-Maruyama scheme in the sense that it gives a higher order of convergence than the Euler-Maruyama method in the strong sense. However, on the other hand, it gives no improvement in the weak sense and requires the knowledge of the first derivative of the volatility function, and this increases considerably the computational complexity. For this reason, although the Milstein scheme is definetely manageable in the one-dimensional case, its extension to multi-dimensional SDEs problems (e.g. more than one risk factors with correlated brownian motions) is far from being trivial.

4.2.5 Applications to the Heston model

We adapt now the discretization schemes above to the Heston model. We will see that, in some circumstances, this may lead to undesired and unrealistic path realizations because while the Heston model, under the Feller property, has been proved to be non-negative (see Section 1.4.3 os Chapter 1), an Euler-Maruyama discretization does not guarantee the non-negativity. This will lead us to construct a more advanced scheme, presented at the end of this section, known as the Quadratic Exponential scheme.

Difficulties with standard discretization schemes

Recall the Heston model

$$dS_t = \mu S_t dt + \sqrt{\nu_t}S_t dB_t^1 \tag{4.37}$$

$$d\nu_t = k(\hat{\nu} - \nu_t)dt + \eta\sqrt{\nu_t}dB_t^2 \tag{4.38}$$

with B_t^1, B_t^2 correlated Brownian motions. Although here we have a system of SDEs, since the equation for ν_t is uncoupled, we can at first compute an approximation for the process ν_t and then using it to approximate S_t. Therefore, we can think in terms of two scalar SDEs.

The Euler-Maruyama scheme for the instantaneous variance ν_t becomes

$$\nu^\delta_{t_{n+1}} = \nu^\delta_{t_n} + \delta k(\hat{\nu} - \nu^\delta_{t_n}) + \sqrt{\delta}Z_n\eta\sqrt{\nu^\delta_{t_n}}. \tag{4.39}$$

Recall from Proposition 1.1 of Section 1.4.3 that, almost surely, the instantaneous variance process $(\nu_t)_{t \geq 0}$ cannot reach zero, provided the Feller condition $\eta^2 \leq 2k\hat{\nu}$ is satisfied. On the other hand, if this condition does not hold, then, almost surely, the process reaches the origin at a finite time. If this happens the process is immediately reflected, so that the process cannot become negative. However, the non-negativity conditions becomes problematic when the Euler-Maruyama discretization is employed.

In Figure 4.1 we simulated sample paths using the Euler-Maruyama scheme in order to compare a situation where the Feller condition $\eta^2 \leq 2k\hat{\nu}$ is satisfied ($\eta = 0.1$), with one where it is not satisfied ($\eta = 0.3$). We can see that the PDF of the discretized instantaneous variance process is very high in a neighborhood of zero. This causes the variance to become negative with high probability, and is clearly undesirable.

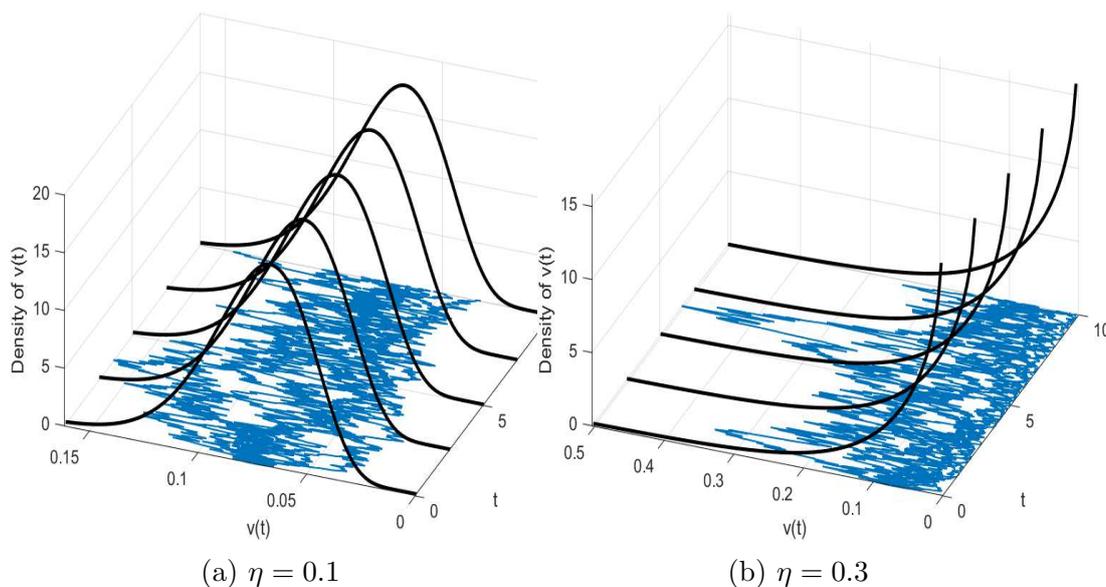(a) $\eta = 0.1$ (b) $\eta = 0.3$

Figure 4.1: Simulation of 10 paths of the variance process ν_t, using the Euler-Maruyama scheme with 500 time steps with the following parameters: $\nu_0 = 0.075, \quad k = 0.5, \quad \hat{\nu} = 0.075, \quad T = 10$. The sample paths are in blue. The black curves represent the probability density function of ν_{t_i}, at five different times t_i.

Probability of negative realization

Assuming $\nu_{t_n}^\delta > 0$, we can calculate the probability that the next realization $\nu_{t_{n+1}}^\delta$ becomes negative:

$$
\begin{aligned}
\mathbb{P}(\nu_{t_{n+1}}^\delta | \nu_{t_n}^\delta) &= \mathbb{P}[\nu_{t_n}^\delta + \delta k(\hat{\nu} - \nu_{t_n}^\delta) + \sqrt{\delta} Z_n \eta \sqrt{\nu_{t_n}^\delta} < 0 | \nu_{t_n} \delta > 0] \\
&= \mathbb{P}[\sqrt{\delta} Z_n \eta \sqrt{\nu_{t_n}^\delta} < -\nu_{t_n}^\delta - \delta k(\hat{\nu} - \nu_{t_n}^\delta) | \nu_{t_n} \delta > 0] \\
&= \mathbb{P}\left[Z_n < -\frac{\nu_{t_n}^\delta + \delta k(\hat{\nu} - \nu_{t_n}^\delta)}{\sqrt{\delta} Z_n \eta \sqrt{\nu_{t_n}^\delta}} \Big| \nu_{t_n} \delta > 0 \right] > 0,
\end{aligned}
\tag{4.40}
$$

where the last inequality comes from the fact that, since $Z_n \sim \mathcal{N}(0,1)$, its PDF does not have compact support.

Following the computations in (4.40), Figure 4.2 shows that, even when the Feller condition is satisfied, the Euler-Maruyama discretization can give rise to negative values, especially for low numbers of steps.
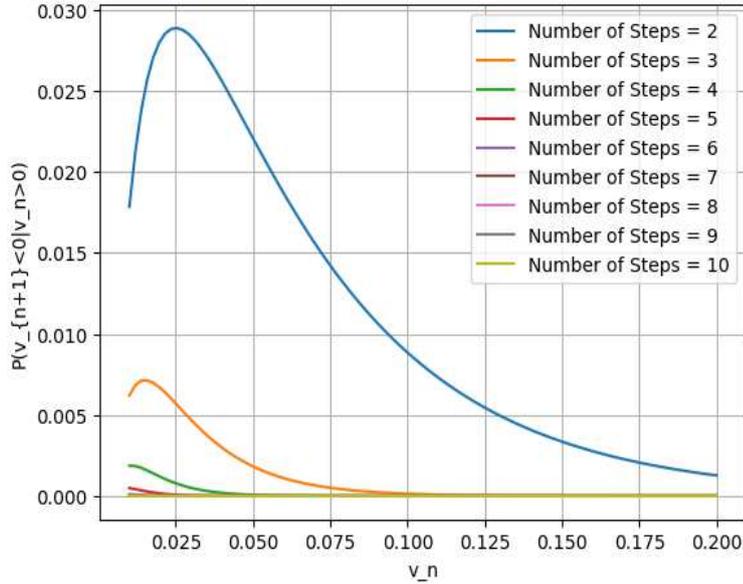

Figure 4.2: $\mathbb{P}(\nu_{t_{n+1}}^\delta | \nu_{t_n}^\delta)$ as a function of $\nu_{t_n}^\delta$, for 9 different number of time steps. The parameters are $T = 1, k = 0.5, \hat{\nu} = 0.075, \eta = 0.25$ and the Feller condition is satisfied.

A similar problem occurs when we use the Milstein scheme. To overcome this difficulty new approaches have been developed in the literature.

Quadratic Exponential Discretization Scheme

An alternative discretization scheme has been proposed by Andersen in 2006 [And07], the so-called Quadratic Exponential (QE) scheme. The QE scheme has the advantage that it determines ν_t accurately while using more information of the probability density function than the above-mentioned schemes. Basically, the approximation for ν_t is derived in the following way.
First

$$\nu_{t_{n+1}}^{\delta} = a_n(b_n + Z_n)^2, \tag{4.41}$$

with Z_i, $i = 0, ..., N-1$ i.i.d. standard Gaussian random variables, and a_n, b_n defined by:

$$a_n = \frac{m_n}{1 + b_n^2}, \qquad b_n^2 = 2\psi_n^{-1} - 1 + \sqrt{2\psi_n^{-1}}\sqrt{2\psi_n^{-1} - 1}, \quad \psi_n = \frac{s_n^2}{m_n^2} \tag{4.42}$$

where s_n and m_n are functions of the time step δ and of ν_{t_n} (we avoid their expression, see [And07]). This scheme works well for large values of ν_t, but not for small values. Then the author finds an optimal threshold value ν_c and defines $\nu_{t_{n+1}}$ as in (4.41) if $\nu_{t_{n+1}} > \nu_c$; otherwise he takes $\nu_{t_{n+1}}$ such that $\mathbb{P}(\nu_{t_{n+1}} \in [x, x + dx]) \approx (p\delta(0) + \beta(1 - p)e^{-\beta x})$, for $x \geq 0$, where $\delta(0)$ is the Dirac delta function and p and β are constant.

4.3 Fourier methods for Option Pricing

At a financial institution, one can distinguish a number of tasks that must be performed in order to price financial derivative products. First of all, depending on the specific need, a model has to be chosen. This means that a system of SDEs that model the dynamic of the underlying asset (and variables related to it, like volatility and interest rates) has to be determined. Once a model is selected, its parameters have to be determined. There are mainly two techniques used for determining the parameters:

- *Parameter estimation*, which is based on historical (real-world) data. Here the parameters are established from the observation of the time series of past values, such as asset prices or historical volatility.

- *Calibration*, which is based on the observation of the current-world market risk-neutral data. Here the model parameters are computed from market prices of financial derivatives and from the observed implied volatility surface. This is often more accurate, as implied volatilities are generally more reflexive of the characteristics of each participant

(e.g. risk appetite, market sentiment, known and unknown events...) than time series, and therefore there is typically a lower uncertainty in risk-neutral projections. Moreover, the information is more available and accurate in the risk-neutral world.

What is done in the calibration procedure is to 'guess' the parameters by chosing those ones that, when used together with the current European Call option prices, generate implied volatilities that best approximate the shape of the quoted implied volatility surface. For this reason, it is important to value European options quickly and accurately. The class of Fourier-based methods that we will present offers highly efficient pricing techniques for pricing European options.

The two main classes of Fourier Methods are based either on an implementation of the Fast Fourir Transform (FFT) (see [CM01]) or on a Fourier series expansion. Regarding the latter approach, we will present a method based on an expansion in cosine series, known as COS method, and will implement it in the framework of the Heston model.

4.3.1 COS method

Let $f \in L^1(\mathbb{R})$ be a 2π-periodic function, with its Fourier series given by

$$S_f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos(kx) + b_k \sin(kx)), \qquad x \in \mathbb{R}, \qquad (4.43)$$

where the coefficients are defined by

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt)dt, \qquad k \geq 0, \qquad (4.44)$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt)dt, \qquad k \geq 1. \qquad (4.45)$$

We also recall the Fourier inversion formula for a function f such that $f, \hat{f} \in L^1(\mathbb{R})$, where $\hat{f}(u) = \int_{\mathbb{R}} e^{iux} f(x)dx$ is the Fourier transform of f:

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ixu} \hat{f}(u)du \qquad (4.46)$$

For a random variable X with probability density funciton f_X, since the characteristic function Φ_X is the Fourier transform of f_X, we have that:

$$f_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ixu} \Phi_X(u)du \qquad (4.47)$$

In what follows we will assume that the Fourier series of f converges pointwise and its sum is f. There are several conditions that guarantee this assumption. For example, an important classical result about the pointwise convergence of fourier series, known as Dirichlet-Jordan Theorem, states that for a periodic function f of bounded variation, as $n \to \infty$, at each point x of the domain, the Fourier series converges to

$$\lim_{\epsilon \to 0} \frac{f(x + \epsilon) + f(x - \epsilon)}{2}. \tag{4.48}$$

In particular, if f is continuous at x, then the Fourier series calculated in x converges to $f(x)$.
We also assume that $\hat{f} \in L^1(\mathbb{R})$ is known explicitly.

By setting $b_k = 0$ for each $k \geq 1$ in (4.43) we obtain the so-called Fourier cosine expansion. In particular, for even functions the Fourier expansion coincides with the Fourier cosine expansion. Clearly, the choice of domain $\mathbb{R}$ and 2π-periodicity for f is not restrictive. In fact, if $f : [a, b[\to \mathbb{R}$, then we can consider the simple change of variable $\theta := \pi \frac{x-a}{b-a}$, define the function

$$g(\theta) = f\left(\frac{b-a}{\pi}\theta + a\right), \qquad \theta \in [0, \pi[, \tag{4.49}$$

and extend by parity g to $[-\pi, \pi[$. Now we have a function defined on $[-\pi, \pi[$ and thus we extend by 2π-periodicity to the whole real line. In this way, by expanding g in Fourier series[6] and going back to the original variable x, we obtain

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos\left(k\pi \frac{x-a}{b-a}\right), \tag{4.50}$$

where[7]

$$
\begin{aligned}
a_k &= \frac{2}{b-a} \int_a^b f(t) \cos\left(k\pi \frac{t-a}{b-a}\right) dt \\
&= \frac{2}{b-a} \int_a^b f(t) \Re\left(e^{ik\pi \frac{t-a}{b-a}}\right) dt \\
&= \frac{2}{b-a} \Re\left(e^{-ik\pi \frac{a}{b-a}} \int_a^b f(t) e^{ik\pi \frac{t}{b-a}} dt\right).
\end{aligned}
\tag{4.51}
$$

---

[6]That $g \in L^1(\mathbb{R})$ follows from the fact that $f \in L^1(\mathbb{R})$.
[7]$\Re(\xi)$ is the real part of $\xi \in \mathbb{C}$.

Now, if the interval $[a, b[$ is sufficiently large, we can approximate $\hat{f}(u) = \int_{\mathbb{R}} e^{iux} f(x) dx \approx \int_a^b e^{iux} f(x) dx$. Substituting into (4.51) we obtain the approximation

$$f(x) \approx \frac{A_0}{2} + \sum_{k=1}^{\infty} A_k \cos\left(k\pi \frac{x-a}{b-a}\right), \qquad x \in [a, b], \qquad (4.52)$$

with

$$A_k = \frac{2}{b-a} \Re\left(e^{-ik\pi \frac{a}{b-a}} \hat{f}\left(\frac{k\pi}{b-a}\right)\right) \qquad (4.53)$$

The Cosine Serie Expansion (COS) method, used to determine the price of an option, is based on the cosine Fourier series expansion presented above. It was developed by Fang and Oosterlee [FO08].

The point of departure for deriving the COS formula for pricing European options is the risk-neutral valuation formula. Given the price S_t of the underlying asset, set $X_t := \log(S_t)$[8] and let $y \to f_X(t, x; T, y)$ denote the transition probability density of X_T, i.e. the conditional probability density function of X_T, given the value of X at time t, $X_t = x$. Also, let $u \to \Phi_X(t, x; T, u)$ be the associated characteristic function[9]. If we let F denote the payoff function of a European option, then the price of the payoff at time t_0 of the option with expiration T, under the risk-neutral measure $\mathbb{Q}$, is given by

$$H(t_0, x) = e^{-r(T-t_0)} \mathbb{E}^{\mathbb{Q}}[F(e^{X_T})|\mathcal{F}_{t_0}]_{|X_{t_0}=x} = e^{-r(T-t_0)} \int_{\mathbb{R}} F(e^y) f_X(t_0, x; y, T) dy$$
$$(4.54)$$

where the constant rate r is assumed constant, for semplicity. To simplify notation, we keep t_0, x, T fixed and use the shortcuts $f_X(y) := f_X(t_0, x; T, y)$ and $\Phi_X(u) := f_X(t_0, x; T, u)$.

The formula (4.54) above is often not very useful because $f_X(y)$ is usually not known explicitly. On the other hand, the expression of the characteristic function of X_T is often available, and from now on we suppose to know it. Recalling that the characteristic function is the Fourier transform of the probability density funciton, the computations above suggest that we apply the Fourier cosine series expansion to $f_X(y)$. First of all we need to truncate the integration range to a finite interval. Therefore we choose an integration interval $[a, b[$ such that the truncated integral approximates very well the infinite counterpart. At this point we expand $f_X(y)$, following the procedure

---

[8]Considering the logarithm of the asset will simplify some calculations
[9]$\Phi_X(t, x; T, u) = E[e^{iuX_T}|X_t = x]$.

explained above, and we plug the expression obtained into (4.54) to obtain:

$$
\begin{aligned}
H(t_0, x) &\approx e^{-r(T-t_0)} \int_a^b F(e^y) f_X(y) dy \\
&= e^{-r(T-t_0)} \frac{\bar{A}_0}{2} \int_a^b F(e^y) dy \\
&\quad + e^{-r(T-t_0)} \sum_{k=1}^{\infty} \bar{A}_k \int_a^b F(e^y) \cos\left(k\pi \frac{y-a}{b-a}\right) dy
\end{aligned}
\tag{4.55}
$$

where

$$
\bar{A}_k = \frac{2}{b-a} \Re\left(e^{-ik\pi \frac{a}{b-a}} \Phi_X\left(\frac{k\pi}{b-a}\right)\right).
\tag{4.56}
$$

Finally, if we define

$$
\bar{B}_k := \frac{2}{b-a} \int_a^b F(e^y) \cos\left(k\pi \frac{y-a}{b-a}\right) dy, \qquad k \geq 0,
\tag{4.57}
$$

we obtain

$$
H(t_0, x) \approx \frac{b-a}{2} e^{-r(T-t_0)} \left(\frac{\bar{A}_0 \bar{B}_0}{2} + \sum_{k=1}^{\infty} \bar{A}_k \bar{B}_k\right).
\tag{4.58}
$$

Notice that $\bar{B}_k$ are the coefficients of the Fourier cosine series expansion of the payoff function F. Thus we have transformed an integral of the product of two functions, $F_X(y)$ and $F(e^y)$, into a series of the products of their Fourier cosine coefficients.

Due to the rapid decaying of the coefficients $\bar{A}_k \bar{B}_k$, we can truncate the series summation and get a last approximation, known as COS formula:

$$
H(t_0, x) \approx \frac{b-a}{2} e^{-r(T-t_0)} \left(\frac{\bar{A}_0 \bar{B}_0}{2} + \sum_{k=1}^{N} \bar{A}_k \bar{B}_k\right).
\tag{4.59}
$$

4.3.2 Application to the Heston model

The goal of this section is to implement on a computer the COS method to price European options and the Greeks in the Heston model, which under the risk-neutral measure $\mathbb{Q}$ take the form

$$dS_t = rS_t dt + \sqrt{\nu_t} S_t dB_t^1 \tag{4.60}$$

$$d\nu_t = k(\hat{\nu} - \nu_t) dt + \eta \sqrt{\nu_t} dB_t^2 \tag{4.61}$$

with the two Brownian motions being ρ-correlated (see Section 1.4.2 of Chapter 1).

In order to use the expression (4.59) for European option, we need to calculate the coefficients $\bar{A}_k, \bar{B}_k$.

For A_k, we need an explicit representation for the characteristic function of $X_T = \log(S_T)$ of the Heston model. In the original paper by Heston [Hes93], the characteristic function is given by :

$$\phi_H(t_0, \nu_{t_0}; T, u) = \exp\left(iur\tau + \frac{\nu_{t_0}}{\eta^2}\left(\frac{1 - e^{-D_1\tau}}{1 - ge^{-D_1\tau}}\right)(k - i\rho\eta u - D_1)\right) \cdot$$
$$\cdot \exp\left[\frac{k\hat{\nu}}{\eta^2}\left(\tau(k - i\rho\eta u - D_1) - 2\log\left(\frac{1 - ge^{-D_1\tau}}{1 - g}\right)\right)\right], \tag{4.62}$$

where $\tau = T - t_0$ and

$$D_1 = \sqrt{\left((k - i\rho\eta u)^2 + (u^2 + iu)\eta^2\right)}, \qquad g = \frac{k - i\rho\eta u - D_1}{k - i\rho\eta u + D_1}. \tag{4.63}$$

In the above formula appear the functions log and $\sqrt{}$ that take complex values, and they are multivalued functions. Therefore a branch must be chosen, i.e. we must select, among the infinite possibilities, one interval of length 2π to which the argument of the complex numbers must belong to. Different choice of the branch determine different values for the characteristic function. [LK10] have proven that, using the above representation, the 'correct' (i.e. does not produce discontinuities in the representation) branch to be chosen is the principal branch for the logarithm.

We now turn to the calculation of coefficients $\bar{B}_k$. We focus on the Call option, and thanks to the Put-Call parity formula this is not restrictive.
We will use the method proposed by Le floch [Le 20], who observed that for short maturities the original COS method was not suitable to approximate

far in-the-money call prices because the cosine coefficients $\bar{B}_k$ were computed relatively to the strike K, but the truncation of the integration interval was made according to the forward price. To overcome this problem we will use $X(T) = \log \frac{S_T}{K}$.

Then the payoff becomes $F(K(e^y - 1)) = [K(e^y - 1)]^+$.

- For $a < b < 0$ we obviously obtain $\bar{B}_k^{call} = 0$.

- By basic computations, for $a < 0 < b$ we have

$$
\begin{aligned}
\bar{B}_k^{call} &= \frac{2}{b-a} \int_0^b K(e^y - 1) \cos\left(k\pi \frac{y-a}{b-a}\right) dy \\
&= \frac{2}{b-a} K(\xi_k(0, b) - \zeta_k(0, b)),
\end{aligned}
$$

(4.64)

- Finally, for $0 < a < b$ we get

$$
\bar{B}_k^{call} = \frac{2}{b-a} K(\xi_k(a, b) - \zeta_k(a, b)),
$$

(4.65)

where we used the following functions:

$$
\begin{aligned}
\xi_k(c, d) := \frac{1}{1 + \left(\frac{k\pi}{b-a}\right)^2} &\left[\cos\left(k\pi \frac{d-a}{b-a}\right) e^d - \cos\left(k\pi \frac{c-a}{b-a}\right) e^c \right. \\
&\left. + \frac{k\pi}{b-a} \sin\left(k\pi \frac{d-a}{b-a}\right) e^d - \frac{k\pi}{b-a} \sin\left(k\pi \frac{c-a}{b-a}\right) e^c \right]
\end{aligned}
$$

(4.66)

and

$$
\zeta_k(c, d) := \begin{cases} \frac{b-a}{k\pi} \left[\sin\left(k\pi \frac{d-a}{b-a}\right) - \sin\left(k\pi \frac{c-a}{b-a}\right) \right], & k \neq 0 \\ d - c, & k = 0 \end{cases}
$$

We will use a simple integration range, i.e.

$$
[a, b] = [-L\sqrt{T - t_0}, L\sqrt{T - t_0}].
$$

(4.67)

The parameters for our computations are $t_0 = 0$, $T = 1$, $L = 10$, $k = 1.56$, $\eta = 0.58$, $\hat{\nu} = 0.04$, $\rho = -0.58$, $\nu_0 = 0.018$ $r = 0$, $S_0 = 100$, $K = 100$, and the result is shown in Figure 4.3.
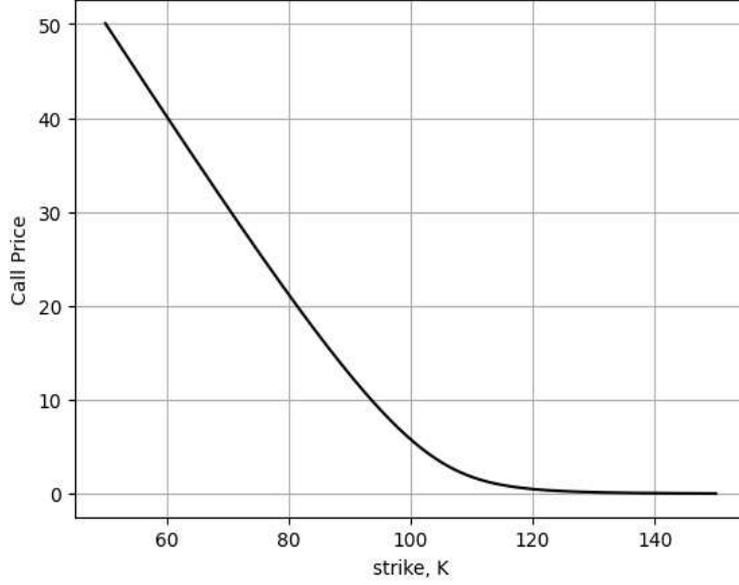
Figure 4.3: Price of European Call option in the Heston model by the COS method.

The series expansion for the Greeks Δ and Γ can be derived similarly. With $S \equiv S_{t_0}$ and $x \equiv X_{t_0}$ we have:

$$\Delta = \frac{\partial V}{\partial S} = \frac{\partial V}{\partial x}\frac{\partial x}{\partial s} = \frac{1}{S}\frac{\partial V}{\partial x}, \qquad \Gamma = \frac{\partial^2 V}{\partial S^2} = \frac{1}{S^2}\left(-\frac{\partial V}{\partial x} + \frac{\partial^2 V}{\partial x^2}\right). \quad (4.68)$$

It then follows that:

$$\Delta \approx \frac{1}{S}e^{-r(T-t_0)}\left[\frac{b-a}{2}\cdot\frac{\bar{A}_0\bar{B}_0}{2} + \sum_{k=1}^{N}\Re\left\{e^{-ik\pi\frac{x-a}{b-a}}\Phi_X\left(\frac{k\pi}{b-a}\right)\cdot\frac{ik\pi}{b-a}\right\}\cdot\bar{B}_k\right],$$

$$(4.69)$$

and

$$\Gamma \approx \frac{1}{S^2}e^{-r(T-t_0)}\left[\frac{b-a}{2}\cdot\frac{\bar{A}_0\bar{B}_0}{2} + \sum_{k=1}^{N}\Re\left\{e^{-ik\pi\frac{x-a}{b-a}}\Phi_X\left(\frac{k\pi}{b-a}\right)\cdot\left(\left(\frac{ik\pi}{b-a}\right)^2 - \frac{ik\pi}{b-a}\right)\right\}\cdot\bar{B}_k\right].$$

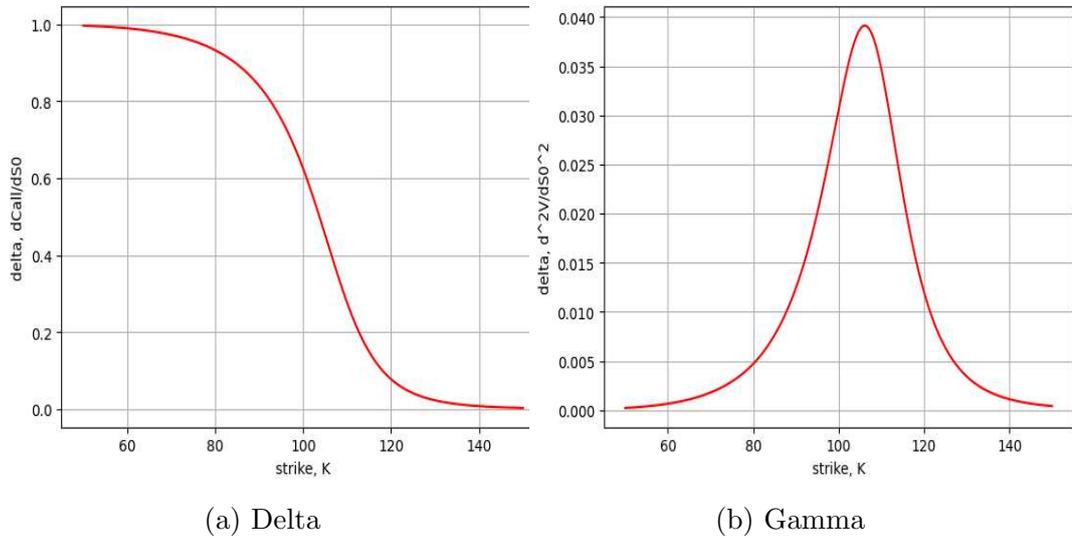$$(4.70)$$

(a) Delta   (b) Gamma

Figure 4.4: Delta and Gamma in the Heston model by the COS method.

4.3.3 Application to the Rough Heston model

The COS method is also useful to price European options in the Rough Heston model. In fact, various expressions for the characteristic function of the Rough Heston model have been recently developed in the literature. We follow the one presented in 2019 from Gatheral and Keller-Ressel [GK18]. Recall from Section 3.4.3 that, supposing $\lambda = 0$ brings the following expression for the instantaneous variance

$$\nu_u = \xi_t(u) + \frac{\nu}{\Gamma(H + 1/2)} \int_t^u \sqrt{\nu_s}(u - s)^{H-1/2} dB_s^{2,\mathbb{Q}} \qquad (4.71)$$

where $\xi_t(u) = \mathbb{E}^{\mathbb{Q}}[\nu_u | \mathcal{F}_t]$, with $u \geq t$ is the forward variance curve. Gatheral and Keller-Ressel study the so-called affine forward variance models, and the classical Heston model in forward variance form and the Rough Heston model are both classified as affine forward variance models. In such case, they show that the characteristic function of the log-price of the asset $X_T = \log S_T$ at time t can be written as:

$$\Phi_X(t; T, u) = E^{\mathbb{Q}}[e^{iuX_T} | \mathcal{F}_t] = \exp\left(iu(X_t + r(T - t)) + \int_t^T \xi_t(s) g(T - s, iu) ds\right),$$
$$(4.72)$$

where $g(t, u)$ is a function expressed in terms of solutions to the Riccati (or fractional Riccati) equation.

In the classical Heston model in forward variance form we have that, with

$a = iu$:

$$g(t, iu) = \frac{\partial}{\partial t}h(t, a) + \lambda h(t, a) \qquad (4.73)$$

where $h(t, u)$ is a solution to the Riccati equation

$$\frac{\partial}{\partial t}h(t, a) = -\frac{1}{2}a(a + i) - (\lambda - i\rho\nu a)h(t, a) + \frac{1}{2}\nu^2 h(t, a)^2. \qquad (4.74)$$

It is possible to solve this equation and to prove that it is indeed equivalent to the formula (4.62) already encountered.

In the rough Heston case (with $\lambda = 0$), we have that

$$g(t, iu) = D^\alpha h(t, a) = -\frac{1}{2}a(a + i) + i\rho\nu a h(t, a) + \frac{1}{2}\nu^2 h(t, a)^2 \qquad (4.75)$$

where D^α is the fractional derivative, with $0 < \alpha < 1$.

We recall, for completeness, the definition of α-fractional derivative. We give a very informal and quick review. For a complete treatment, see [PT17]. The idea is to extend the definition of derivative by allowing expressions like α-th derivative to be defined in such a way that some properties change 'smoothly'.
Let $\Phi : [a, b] \to \mathbb{R}$ be an integrable function and $\alpha \in (0, 1)$. We begin by defining

$$(I^\alpha \Phi)(s) := \frac{1}{\Gamma(\alpha)} \int_a^s \Phi(u)(s - u)^{\alpha - 1} du \qquad (4.76)$$

with $s \in [a, b]$ and Γ the gamma function. Now the idea is to invert the operator I^α. This can be done (see [PT17], pp. 348) and leads to the following definition of fractional derivative:

$$(D^\alpha f)(u) := (I^{-\alpha}f)(u) := \frac{1}{\Gamma(1 - \alpha)}\frac{d}{du}\int_a^u f(s)(u - s)^{-\alpha} ds \qquad (4.77)$$

To find the function $h(t, a)$ for the solution of the fractional Riccati equation for the Rough Heston model, Gatheral and Radoiĉiĉ [GR19] use an approximation scheme based on a combination of short-time expansion of the solution and an asymptotic solution in the long-time limit $\tau = T - t \to \infty$. This method is particularly fast and accurate to compute approximate solutions of fractional ODEs. Another approach was proposed by El Euch et al. [ER16], who used a discretization based on the Adams scheme for ODEs. If we apply one of these approximation scheme, we can therefore obtain a characteristic function of the log-price of the asset under the Rough Heston model using the COS method, and thus price efficiently European options.

Chapter 5

Conclusion

In this thesis we have presented an overview of stochastic volatility modeling with a focus on equity models based on fractional Brownian motion, with a particular attention to the 'rough' case with Hurst exponent $H < 1/2$. We considered various aspects of these models based on historical data, where we stressed the importance of the behavior of the autocorrelation in determining important properties like persistence and long memory. We also reviewed these models from a risk-neutral point of view and we concluded that 'rough' models are generally successful when reproducing the at-the-money (ATM) volatility skew typically observed on the market, in contrast with classical SV models and with models based on fBm with Hurst parameter $H > 1/2$, like the Hull-White model. Therefore, rough models, like the RFSV of Gatheral et al. (2014) and the Rough Heston model, generally provide a more realistic representation of the observed implied volatility surface than traditional SV models. We then applied numerical methods based on Monte Carlo simulation to the famous Heston model and by doing so, we highlighted some of the limitations of classical numerical discretization schemes, analyzing possible alternatives. Finally, a method for option pricing based on Fourier series expansion has been presented and applied to the Heston model, and an overview of a possible implementation in the rough case has been presented.

Appendix A

Main codes

A.1 Simulation of paths and PDF for the Heston model

```matlab
function Heston_Euler


% Parameters setting

Npaths   = 10;
Nsteps   = 500;
T        = 10;
k        = 0.5;
vhat     = 0.075;
v0       = 0.075;
eta      = 0.1;


% Generate Heston paths with Euler-Maruyama scheme

[V,timeGrid] =
    GeneratePathsHestonEM(Npaths,Nsteps,v0,k,vhat,eta,T);


% Figure

figure1 = figure;
axes1 = axes('Parent',figure1);
hold(axes1,'on');
```

```matlab
% Plot for densty of v(t)

plot(timeGrid,V,'linewidth',0.5,'color',[0 0.45 0.74])
PDF = Hestondensity(k,eta,vhat,0.0,T,v0);


% Grid for the PDF of v(t)

GridT=linspace(0.1,T,5);

x_arg=linspace(0,max(max(V(:,:)))*2,250);
for i=1:length(GridT)
    plot3(GridT(i)*ones(length(x_arg),1),x_arg,PDF(x_arg),'k','linewidth',2)
end

axis([0,GridT(end),0,max(max(V))])
grid on;
xlabel('t')
ylabel('v(t)')
zlabel('Density of v(t)')
view(axes1,[-75. 45.]);

function PDF = Hestondensity(k,eta,vhat,s,t,v_s)
a           = eta^2/(4*k)*(1-exp(-k*(t-s)));
b           = 4*k*vhat/(eta^2);
kBar        = 4*k*exp(-k*(t-s))/(eta^2*(1-exp(-k*(t-s))))*v_s;
PDF         = @(x)1/a*ncx2pdf(x./a,b,kBar);


function [V,t] =
    GeneratePathsHestonEM(Npaths,Nsteps,V0,k,vhat,eta,T)


% Initial datum

V=zeros(Npaths,Nsteps);
V(:,1) = V0;


% Gaussian noise
```

```matlab
Z=random('normal',0,1,[Npaths,Nsteps]);
W=zeros([Npaths,Nsteps]);

delta = T/Nsteps;
t = zeros([Nsteps+1,1]);

for i=1:Nsteps
    if Npaths>1
        Z(:,i)  = (Z(:,i) - mean(Z(:,i))) / std(Z(:,i));
    end
    W(:,i+1) = W(:,i) + sqrt(delta).*Z(:,i);
    V(:,i+1) = V(:,i) + k*(vhat-V(:,i))*delta+ eta* sqrt(V(:,i)).*
        (W(:,i+1)-W(:,i));

    V(:,i+1) = max(V(:,i+1),0);
    t(i+1) = t(i) + delta;
end
```

A.2 $\mathbb{P}(\nu_{t_{n+1}}^{\delta} < 0)$ as a function of the previous step

```python
import numpy as np
import matplotlib.pyplot as plt
import scipy.stats as st

    # Probability of negative variance given the preceding step is
        positive
def ProbabilityOfNegative():
    # Parameters specification of Heston model
    k    = 0.5
    eta  = 0.25
    vhat = 0.075
    T    = 1.0
    Nsteps = 10
    delta = T/Nsteps

    l = []
    for Nsteps in range(2,11,1):
        delta = T/Nsteps
```

```python
    F = lambda v_n: st.norm.cdf(
        (-v_n-k*(vhat-v_n)*delta)/(eta*np.sqrt(v_n)*delta))
    v_n = np.linspace(0.01,0.2,100)
    l.append('Number of Steps = {0}'.format(Nsteps))
    plt.plot(v_n,F(v_n))


# Feller condition
Feller = 2.0*k*vhat - eta**2.0
if Feller<0:
    print("Feller condition is not satisfied".format(Feller))
else:
    print("Feller condition is satisfied".format(Feller))

plt.grid()
plt.legend(l)
plt.xlabel('v_n')
plt.ylabel('P(v_{n+1}<0|v_n>0)')

ProbabilityOfNegative()
```

A.3 Price of European Call option in the Heston model by the COS method

```python
import numpy as np
import matplotlib.pyplot as plt
import scipy.optimize as optimize
import scipy.stats as st
import enum

# [a,b]- truncation domain
# S0  - initial asset price
# r   - nterest rate
# L   - simple integration range
# K   - strikes (list)
# tau - time to maturity
# N   - COS number of addends
# phi_H- characteristic function of Heston SV model
```

```python
    # Heston Characteristic function definition
def HestonCHF(r,tau,k,eta,vhat,v0,rho):
    i = np.complex(0.0,1.0)
    d1 = lambda u:
        np.sqrt(np.power(k-eta*rho*i*u,2)+(u*u+i*u)*eta*eta)
    f  = lambda u: (k-eta*rho*i*u-d1(u))/(k-eta*rho*i*u+d1(u))
    C  = lambda u:
        (1.0-np.exp(-d1(u)*tau))/(eta*eta*(1.0-f(u)*np.exp(-d1(u)*tau)))
        \
        *(k-eta*rho*i*u-d1(u))
    A  = lambda u: r * i*u *tau + k*vhat*tau/eta/eta
        *(k-eta*rho*i*u-d1(u))\
        -
            2*k*vhat/eta/eta*np.log((1-f(u)*np.exp(-d1(u)*tau))/(1-f(u)))
    phi_H = lambda u: np.exp(A(u) + C(u)*v0)
    return phi_H


    # xi and zeta definition
def xi_zeta(a,b,c,d,k):
    xi = 1.0 / (1.0 + np.power((k * np.pi / (b - a)) , 2.0))
    expr1 = np.cos(k * np.pi * (d - a)/(b - a)) * np.exp(d) -
        np.cos(k * np.pi
                * (c - a) / (b - a)) * np.exp(c)
    expr2 = k * np.pi / (b - a) * np.sin(k * np.pi *
                    (d - a) / (b - a)) - k * np.pi / (b - a) *
                        np.sin(k
                    * np.pi * (c - a) / (b - a)) * np.exp(c)
    xi = xi * (expr1 + expr2)

    zeta = np.sin(k * np.pi * (d - a) / (b - a)) - np.sin(k * np.pi
        * (c - a)/(b - a))
    zeta[1:] = zeta[1:] * (b - a) / (k[1:] * np.pi)
    zeta[0] = d - c

    value = {"xi":xi,"zeta":zeta }
    return value


     # Definition coefficients B_k
def CallCoefficients(C,a,b,k):
    c = 0.0
    d = b
```

```python
    coef = xi_zeta(a,b,c,d,k)
    xi_k = coef["xi"]
    zeta_k = coef["zeta"]
    if a < b and b < 0.0:
      B_k = np.zeros([len(k),1])
    else:
      B_k      = 2.0 / (b - a) * (xi_k - zeta_k)
    return B_k



    # Value of Heston Call Option definition
def CallCOS(phi_H,C,S0,r,tau,K,N,L):
    if K is not np.array:
        K = np.array(K).reshape([len(K),1])

    i = np.complex(0.0,1.0)
    x0 = np.log(S0 / K)

    a = 0.0 - L * np.sqrt(tau)
    b = 0.0 + L * np.sqrt(tau)

    k = np.linspace(0,N-1,N).reshape([N,1])
    u = k * np.pi / (b - a);

    # Coefficients B_k
    B_k = CallCoefficients(C,a,b,k)
    mat = np.exp(i * np.outer((x0 - a) , u))
    temp = phi_H(u) * B_k
    temp[0] = 0.5 * temp[0]
    value = np.exp(-r * tau) * K * np.real(mat.dot(temp))
    return value



    # Call price definition
def CallPrice():
    S0  = 100
    r   = 0.00
    tau = 1
    K   = np.linspace(50,150,250)
    K   = np.array(K).reshape([len(K),1])


    # COS method settings
```

```
L     = 10
vhat  = 0.04
v0    = 0.018
rho   =-0.58
eta   = 0.58
k     = 1.56
C     =1


# CHF for the Heston model
phi_H = HestonCHF(r,tau,k,eta,vhat,v0,rho)


# COS number of addends
N = 5000
callRef = CallCOS(phi_H, C, S0, r, tau, K, N, L)


# Figure plot
plt.plot(K,callRef,'-k')
plt.xlabel("strike, K")
plt.ylabel("Call Price")
plt.grid()
```

```
CallPrice()
```

A.4 Delta and Gamma in the Heston model by the COS method

```
import numpy as np
import matplotlib.pyplot as plt
import scipy.optimize as optimize
import scipy.stats as st
import enum

    # [a,b]- integration domain
    # S0  - initial asset price
    # r   - nterest rate
    # L   - simple integration range
```

```python
    # K    - strikes (list)
    # tau - time to maturity
    # N    - COS number of addends
    # phi_H- characteristic function of Heston SV model


i = np.complex(0.0,1.0)

    # Heston Characteristic function definition
def HestonCHF(r,tau,k,eta,vhat,v0,rho):
    d1 = lambda u:
        np.sqrt(np.power(k-eta*rho*i*u,2)+(u*u+i*u)*eta*eta)
    f  = lambda u: (k-eta*rho*i*u-d1(u))/(k-eta*rho*i*u+d1(u))
    C  = lambda u:
        (1.0-np.exp(-d1(u)*tau))/(eta*eta*(1.0-f(u)*np.exp(-d1(u)*tau)))
        \
        *(k-eta*rho*i*u-d1(u))
    A  = lambda u: r * i*u *tau + k*vhat*tau/eta/eta
        *(k-eta*rho*i*u-d1(u))\
        -
            2*k*vhat/eta/eta*np.log((1-f(u)*np.exp(-d1(u)*tau))/(1-f(u)))
    phi_H = lambda u: np.exp(A(u) + C(u)*v0)
    return phi_H



    # xi and zeta definition
def xi_zeta(a,b,c,d,k):
    xi = 1.0 / (1.0 + np.power((k * np.pi / (b - a)) , 2.0))
    expr1 = np.cos(k * np.pi * (d - a)/(b - a)) * np.exp(d) -
        np.cos(k * np.pi
                * (c - a) / (b - a)) * np.exp(c)
    expr2 = k * np.pi / (b - a) * np.sin(k * np.pi *
                    (d - a) / (b - a)) - k * np.pi / (b - a) *
                        np.sin(k
                    * np.pi * (c - a) / (b - a)) * np.exp(c)
    xi = xi * (expr1 + expr2)

    zeta = np.sin(k * np.pi * (d - a) / (b - a)) - np.sin(k * np.pi
        * (c - a)/(b - a))
    zeta[1:] = zeta[1:] * (b - a) / (k[1:] * np.pi)
    zeta[0] = d - c

    value = {"xi":xi,"zeta":zeta }
    return value
```

```python
   # Definition coefficients B_k
def CallCoefficients(C,a,b,k):
    c = 0.0
    d = b
    coef = xi_zeta(a,b,c,d,k)
    xi_k = coef["xi"]
    zeta_k = coef["zeta"]
    if a < b and b < 0.0:
      B_k = np.zeros([len(k),1])
    else:
      B_k     = 2.0 / (b - a) * (xi_k - zeta_k)
    return B_k


def COSDelta(phi_H,C,S0,r,tau,K,N,L):
    if K is not np.array:
       K = np.array(K).reshape([len(K),1])

    # integration domain
    a = 0.0 - L * np.sqrt(tau)
    b = 0.0 + L * np.sqrt(tau)

    x0 = np.log(S0 / K)
    k = np.linspace(0,N-1,N).reshape([N,1])
    u = k * np.pi / (b - a);

    B_k = CallCoefficients(C,a,b,k)
    mat = np.exp(i * np.outer((x0 - a) , u))

    temp = phi_H(u) * B_k * u * i
    temp[0] = 0.5 * temp[0]
    value = 1.0/S0 *np.exp(-r * tau) * K * np.real(mat.dot(temp))
    return value


def COSGamma(phi_H,C,S0,r,tau,K,N,L):
    if K is not np.array:
       K = np.array(K).reshape([len(K),1])

    # integration domain
    a = 0.0 - L * np.sqrt(tau)
```

```python
    b = 0.0 + L * np.sqrt(tau)

    x0 = np.log(S0 / K)
    k = np.linspace(0,N-1,N).reshape([N,1])
    u = k * np.pi / (b - a);

    # Determine coefficients
    B_k = CallCoefficients(C,a,b,k)
    mat = np.exp(i * np.outer((x0 - a) , u))

    Gamma_term = (u * i)**2.0 - u * i
    temp = phi_H(u) * B_k * Gamma_term
    temp[0] = 0.5 * temp[0]
    value = 1.0/(S0**2.0) *np.exp(-r * tau) * K *
        np.real(mat.dot(temp))
    return value


    # COS method for Delta and Gamma in Heston model
def DeltaGammaHeston():
    S0  = 100
    r   = 0.00
    tau = 1
    K   = np.linspace(50,150,250)

    # reshape K to a column vector
    K   = np.array(K).reshape([len(K),1])


    # COS method settings
    L    = 10
    vhat = 0.04
    v0   = 0.018
    rho  =-0.58
    eta  = 0.58
    k    = 1.56
    C    =1

    # CHF for the Heston model
    phi_H = HestonCHF(r,tau,k,eta,vhat,v0,rho)


    # COS number of addends
```

```python
N = 5000

DeltaCOS = COSDelta(phi_H,C,S0,r,tau,K,N,L)
GammaCOS = COSGamma(phi_H,C,S0,r,tau,K,N,L)

# Figure plot
plt.figure(2)
plt.plot(K,DeltaCOS,'r')
plt.xlabel("strike, K")
plt.ylabel("delta, dCall/dS0")
plt.grid()

plt.figure(3)
plt.plot(K,GammaCOS,'r')
plt.xlabel("strike, K")
plt.ylabel("delta, d^2V/dS0^2")
plt.grid()


DeltaGammaHeston()
```

Bibliography

[And07]   L. Andersen. "Efficient Simulation of the Heston Stochastic Volatil-
            ity Model". In: *Journal of Computational Finance, vol. 11* (2007).

[AB97]   T. Andersen and T. Bollerslev. "Intraday periodicity and volatil-
            ity persistence in financial markets". In: *Journal of Empirical
            Finance, Volume 4, pp. 115-158* (1997).

[And+03]   T. Andersen et al. "Modeling and Forecasting Realized Volatil-
            ity". In: *Econometrica, vol. 71, 579-625* (2003).

[Bai96]   R. Baillie. "Long memory processes and fractional integration in
            econometrics". In: *Journal of Econometrics, 1996, vol. 73, 5-59*
            (1996).

[BW00]   G. Bekaert and G. Wu. "Asymmetric volatility and risk in equity
            markets". In: *Review of Financial Studies 13, pp. 1–42* (2000).

[BLP21]   M. Bennedsen, A. Lunde, and M. S. Pakkanen. "Decoupling the
            short- and long-term behavior of stochastic volatility". In: *arXiv,
            1610.00332, q-fin.ST* (2021).

[Bia+08]   F. Biagini et al. *Stochastic Calculus for Fractional Brownian Mo-
            tion and Applications*. Springer-Verlag London, 2008.

[Bla76]   F. Black. "Studies of Stock Market Volatility Changes". In: *Pro-
            ceedings of the American Statistical Association, Business and
            Economic Statistics Section, 177-181* (1976).

[BLT06]   T. Bollerslev, J. Litvinova, and G. Tauchen. "Leverage and Volatil-
            ity Feedback Effects in High-Frequency Data". In: *The Journal
            of Financial Econometrics* 4.3 (2006), pp. 353–384.

[BL78]   D.T Breeden and R.H Litzenberger. "Prices of State-Contingent
            Claims Implicit in Option Prices". In: *The Journal of Business,
            University of Chicago Press, vol 51, pp. 621-651* (1978).

[CH92]   J. Y. Campbell and L. Hentschel. "No news is good news: An asymmetric model of changing volatility in stock returns". In: *Journal of Financial Economics, Vol. 31, pp. 281-318* (1992).

[CM01]   P. Carr and D. Madan. "Option Valuation Using the Fast Fourier Transform". In: *Journal of computational Finance* (2001).

[CKM03]   P. Cheridito, H. Kawaguchi, and M. Maejima. "Fractional Ornstein-Uhlenbeck processes". In: *Electronic Journal of Probability, Electron. J. Probab. 8(none), 1-14* (2003).

[CR98]   F. Comte and E. Renault. "Long memory in continuous-time stochastic volatility models". In: *Mathematical Finance, Vol. 8, No. 4, 291–323* (1998).

[Con01]   R. Cont. "Empirical properties of asset returns: Stylized facts and statistical issues". In: *Quantitative Finance 1, pp. 223-236* (2001).

[Con07]   R. Cont. "Volatility Clustering in Financial Markets: Empirical Facts and Agent Based Models". In: *Teyssière, G., Kirman, A.P. (eds) Long Memory in Economics. Springer, Berlin, Heidelberg* (2007).

[DF06]   Brigo D. and Mercurio F. *Interest Rate Models - Theory and Practice*. Springer-Verlag Berlin Heidelberg, 2006.

[DBJ03]   G. Daniel, D. S. Bree, and N. L. Joseph. "Goodness-of-fit of the Heston model". In: (2003). arXiv: `cs/0305055 [cs.CE]`.

[DGE93]   Z. Ding, C. W. J. Granger, and R. F. Engle. "A long memory property of stock market returns and a new model". In: *Journal of Empirical Finance, Volume 1, pp. 83-106* (1993).

[Dup93]   B. Dupire. "Arbitrage Pricing with Stochastic Volatility". In: *Proceeding AFFI Conference, La Baule* (1993).

[EGR19]   O. El Euch, J. Gatheral, and M. Rosenbaum. "Roughening Heston". In: *Risk, pp. 84-89* (2019).

[ER16]   O. El Euch and M. Rosenbaum. "The characteristic function of rough Heston models". In: (2016). arXiv: `1609.02108 [q-fin.MF]`.

[ER17]   Omar El Euch and Mathieu Rosenbaum. "Perfect hedging in rough Heston models". In: (2017). arXiv: `1703.05049 [q-fin.MF]`.

[Fam65]   E. Fama. "The behavior of stock market prices". In: *Journal of Business 38, pp. 34-105* (1965).

[FO08]   F. Fang and C. W. Oosterlee. "A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions". In: *SIAM Journal on Scientific Computing, Vol 31, 826-848* (2008).

[FPS99]   J. Fouque, G. Papanicolaou, and R. Sircar. "Mean-Reverting Stochastic Volatility". In: *International Journal of Theoretical and Applied Finance* (1999).

[Fuk11]   M. Fukasawa. "Asymptotic analysis for stochastic volatility: martingale expansion". In: *Finance and Stochastics, vol. 15, pp. 635-654* (2011).

[GK80]   M. B. Garman and M. J. Klass. "On the Estimation of Security Price Volatilities from Historical Data". In: *The Journal of Business, University of Chicago Press, vol 53, pp. 67-78* (1980).

[GJR14]   J. Gatheral, T. Jaisson, and M. Rosenbaum. "Volatility is rough". In: *Quantitative Finance, Vol 18, pp. 933-949* (2014).

[GK18]   J. Gatheral and M. Keller-Ressel. "Affine forward variance models". In: (2018). arXiv: `1801.06416 [q-fin.MF]`.

[GR19]   J. Gatheral and R. Radoicic. "Rational Approximation of the rough Heston solutionG". In: *International Journal of Theoretical and Applied Finance, vol. 22* (2019).

[Hes93]   S. L. Heston. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options". In: *The Review of Financial Studies, Volume 6, 327–343* (1993).

[HW87]   J. Hull and A. White. "The Pricing of Options on Assets with Stochastic Volatilities". In: *The Journal of Finance, Volume 42, pp. 281-300* (1987).

[Huy03]   D. P. Huy. "A Remark on Non-Markov Property of a Fractional Brownian Motion". In: *Vietnam Journal of Mathematics 31:2, 237-240* (2003).

[IS89]   N. Ikeda and I. S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland Mathcmnticnl Library, 24, Amsterdam, 568pp.Springer, 1989.

[KP92]   P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer Berlin, Heidelberg, 1992.

[Kol40]   A. N. Kolmogorov. "Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum". In: *C.R.(Doklady) Acad. URSS (N.S) 26, 115–118* (1940).

[LL07]   D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance.* Chapman and Hall/CRC, 2007.

[Le 20]   F. Le Floc'h. "More Robust Pricing of European Options Based on Fourier Cosine Series Expansions". In: (2020). arXiv: `2005.13248 [q-fin.CP]`.

[LS01]   W. V. Li and Q. M. Shao. "Gaussian processes: Inequalities, small ball probabilities and applications". In: *Handbook of Statistics Volume 19, Pages 533-597* (2001).

[Lif95]   M. A. Lifshits. *Gaussian Random Functions.* Springer, 1995.

[LK10]   R. Lord and C. Kahl. "Complex logarithm in Heston-like models". In: *Mathematical Finance* 20 (2010).

[Man63]   B. Mandelbrot. "The variation of certain speculative prices". In: *Journal of Business 36, pp. 394-419* (1963).

[MV68]   B. B. Mandelbrot and J.W. Van Ness. "Fractional Brownian motions, fractional noises and applications". In: *SIAM Rev. 10, 422–437* (1968).

[Nou12]   I. Nourdin. *Selected Aspects of Fractional Brownian Motion.* Springer, Milan, 2012.

[Nua02]   D. Nualart. *The Malliavin Calculus and Related Topics.* Springer-Verlag Berlin Heidelberg, 2002.

[Off73]   R. Officer. "The variability of the market factor of the New York Stock Exchange". In: *Journal of Business 46, pp. 434-453.* (1973).

[Par80]   M. Parkinson. "The Extreme Value Method for Estimating the Variance of the Rate of Return". In: *The Journal of Business, University of Chicago Press, pp. 61-65* (1980).

[Pas11]   A. Pascucci. *PDE and Martingale Methods in Option Pricing.* Springer Milano, 2011.

[PT17]   V. Pipiras and M. Taqqu. *Long-Range Dependence and Self-Similarity.* Cambridge University Press. doi:10.1017/CBO9781139600347, 2017.

[RN96]   E. Renault and Touzi N. "Option hedging and implied volatilities in a stochastic volatility model". In: *Mathematical Finance, Vol 6, pp. 279-302* (1996).

[Rog97]   L. C. G. Rogers. "Arbitrage with fractional Brownian motion". In: *Math. Finance 7, 95–105* (1997).

[RSY94]   L. C. G. Rogers, S. E. Satchell, and Y. Yoon. "Estimating the volatility of stock prices: a comparison of methods that use high and low prices". In: *Applied Financial Economics, vol. 3, pp. 241-247* (1994).

[Sch89]   G. W. Schwert. "Why Does Stock Market Volatility Change Over Time?" In: *The Journal of Finance, 44: 1115-1153* (1989).

[Sza+03] A. Szakmary et al. "The predictive power of implied volatility: Evidence from 35 futures markets". In: *Journal of Banking and Finance, vol. 27, pp.2151-2175* (2003).

[TW92]   A. Turner and E. Weigel. "Daily stock market volatility: 1928-1989". In: *Management Science 38, pp. 1586-1609* (1992).

[YZ00]   D. Yang and Q. Zhang. "Drift-Independent Volatility Estimation Based on High, Low, Open, and Close Prices". In: *The Journal of Business, The University of Chicago Press, vol. 73, pp. 477-492* (2000).

[Yos20]   S. Yosr. "Statistical and Numerical Analysis of Rough Volatility Models". In: *PhD Thesis* (2020).

[You36]   L. C. Young. "An inequality of the Hölder type, connected with Stieltjes integration". In: *Acta Math. 67 251 - 282* (1936).