Institut de Statistique, Biostatistique et Actuariat

Rough stochastic volatility modeling

Membres du jury:

Prof. Jérôme Barbarin, *Promoteur*
Prof. Donatien Hainaut

Mémoire présenté en vue de
l'obtention du mastère
en sciences actuarielles
(orientation sciences actuarielles)
par:

Jean-Loup Dupret

Louvain-La-Neuve
Août 2020

Preface.

This dissertation has been prepared in partial fulfillment of the requirements for the Master Degree in actuarial sciences delivered by the Université Catholique de Louvain.

Acknowledgment.

First and foremost, I wish to express my deepest gratitude to Jérôme Barbarin, for accompanying me throughout this master's thesis. Thanks to his experience, he gave me lots of high quality advice and ideas to carry out this work.

I thank the reader of this thesis, Donatien Hainaut, for the time he will take to review my work.

I would also like to thank my family and my friends for their unfailing support and continuous encouragement throughout my years of study.

Contents

Chapter 1

Introduction.

This master's thesis presents an overview of volatility modeling with a focus on equity models based on fractional Brownian motions with Hurst exponent $H < 1/2$. These stochastic volatility models are called *rough* due to the sample paths properties of the volatility process driven by such fractional Brownian motions.

We first introduce the general concept of volatility and how the celebrated model of Black and Scholes can be improved by assuming that the volatility of the underlying price is a stochastic process rather than a constant. However, such models are still limited and cannot reproduce some important empirical facts of the observed volatility time series and of the observed volatility surface. In order to address these shortcomings, Gatheral et al. (2014) propose a model called *"rough fractional stochastic volatility"* (RFSV) where the process of log-volatility is modeled in terms of fractional Brownian motions. More precisely, a fractional Ornstein-Uhlenbeck process is used with $H < 1/2$, in contrast with the *"fractional stochastic volatility"* model (FSV) previously introduced by Comte and Renault (1998), where the Hurst coefficient is assumed to satisfy $H > 1/2$. Gatheral et al. (2014) find a highly consistent model with empirical estimates of the volatility time series as well as with risk-neutral data. We then present more tractable pricing models such as the rBergomi or the rough Heston which are derived from the RFSV model and which enable us to price European options in a more effective way. We also briefly introduce a more general class of models based on Brownian semistationary (BSS) processes that accommodate both for roughness ($H < 1/2$) at short time scales and long-memory ($H > 1/2$) at longer time horizons.

Finally, in the practical part of this thesis, we first confirm the results presented by Gatheral et al. (2014) in their paper but with other indices and other time windows. In a second step, we calibrate and compare four models (Heston, Bates, rBergomi and rough Heston) in terms of fit to the observed European option prices and implied volatility surface for the CAC 40. We finally use these calibrated models to price a life insurance contract embedding path-dependent options in it.

Chapter 2

What is volatility ?

In this chapter, we first introduce the notion of volatility. We give a hint towards the possible ways to model volatility, the most important parameter in derivatives pricing. We especially focus on classical stochastic volatility models which assume that the volatility of the underlying price is a stochastic process rather than a constant.

2.1 Definition

Usually, the volatility denoted σ is used to address the changing fluctuation intensity of asset prices and hence is a measure of the uncertainty about the returns provided by the stock. The precise definition of volatility can only be given within an appropriate stochastic model. On this account, this notion depends on the framework and on the purpose of the modeling. In most situations, however, the volatility is related to the standard deviation of the logarithmic price increments, conditioned on recent price observations, which are usually described by a stochastic process. Indeed, let's assume in a first step that the stock price process follows a Black and Scholes (B&S) model :

$$dS_t = \mu\, S_t\, dt + \sigma\, S_t\, dW_t \tag{2.1.1}$$

Where μ is the stock's expected rate of return and W_t a standard Brownian motion defined on $(\Omega, \mathbb{P}, \mathcal{F})$. The discretized version of this process becomes with $\varepsilon \sim N(0,1)$:

$$(S_{t+\Delta t} - S_t)/S_t = \mu\, \Delta t + \sigma\, \varepsilon\, \sqrt{\Delta t} \tag{2.1.2}$$

When Δt is small, equation (2.1.2) shows that $\sigma^2 \Delta t$ is approximately equal to the variance of the percentage change in the stock price in time Δt and $\sigma\sqrt{\Delta t}$ approximately equal to its standard deviation in time Δt. However, equation (2.1.2) only holds approximately when Δt is small. Thanks to Itô's lemma (see Concept 7.2.1 in Appendix), we now have more precisely :

$$d \ln S_t = (\mu - \frac{1}{2}\sigma^2)\, dt + \sigma\, dW_t \tag{2.1.3}$$

Hence, we have $\ln(\frac{S_t}{S_0}) \sim N\left((\mu - \frac{1}{2}\sigma^2)\, t\, , \sigma^2 t\right)$ so that the logarithmic returns are normally distributed. Indeed, recall that logarithmic returns between time 0 and time t are defined by $\ln(S_t/S_0)$ and are a good approximation of the percentage change (or rate of return) : $(S_t - S_0)/S_0$. Therefore, one could define uncertainty with regard to future stock values by concentrating on the standard deviation of logarithmic returns; that is, the relative (or lognormal) volatility σ. In the B&S framework, stocks typically have a relative volatility between 15% and 60%.

All potential practical applications of the B&S model hinge on knowledge of the volatility parameter of the stock price returns. Indeed, if we consider the risk-neutral version of equation (2.1.1) given by :

$$dS_t = r\, S_t\, dt + \sigma\, S_t\, dW_t^{\mathbb{Q}} \tag{2.1.4}$$

where $W_t^{\mathbb{Q}}$ is a Brownian motion under the unique risk-neutral measure $\mathbb{Q}$. We can see that all the parameters are observable except for the relative volatility σ, making the volatility the most crucial parameter of the B&S equation.
However, empirical observations show that the volatility parameter σ is not the simple constant assumed in the original B&S model and that it varies through time. This can indeed easily be seen by introducing the concept of historical volatility and implied volatility.

2.1.1 Historical volatility

The historical volatility is the volatility as observed in the past over a certain time interval at a certain granularity (e.g. daily observations). Indeed, the most natural approach to the volatility specification uses an estimate of the standard deviation based upon a past series of returns from the underlying stock. It is therefore a statistical measure based on real-world observations. As explained above, we can concentrate on the annualized standard deviation of logarithmic returns. Hence, a traditional annualized estimator of σ is simply given by the square root of :

$$\hat{\sigma}^2 = \frac{1}{n\,\Delta t}\sum_{i=1}^{n}\left(\ln\left(\frac{S_i}{S_{i-1}}\right)\right)^2$$

where n is the number of observations of the stock price process and Δt the time step between two observations. For daily observations, we usually have $\Delta t = 1/252$ if we consider 252 trading days per year. However, it can clearly be shown that this estimator of volatility is not constant trough time. For example, if we consider daily observations of the S&P500 with $n = 21$ (i.e. a monthly window), we obtain the following graph of the annualized estimator $\hat{\sigma}$ in function of time :

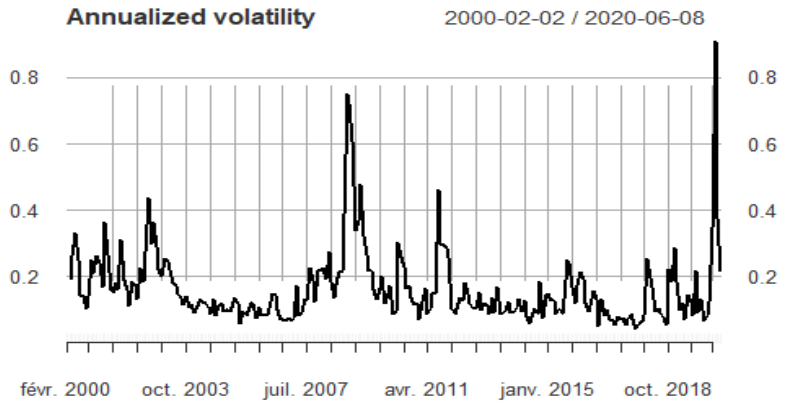

Figure 2.1.1: Annualized volatility $\hat{\sigma}$, S&P500.

Since the volatility is unstable through time, historical precedent is a poor guide for estimating future volatility. Moreover, estimates of option prices based on historical volatilities are systematically biased. Therefore, we will consider in thesis other methods to estimate and measure volatility.

2.1.2 Implied volatility

We now introduce the concept of implied volatility which is based on risk-neutral data and we show that it is not constant either. Let us consider at time t a European call option on the underlying S_t with exercise price K and a duration $\tau = T - t$ till the expiration date T. Recall that the B&S price of a call option is given under the unique risk-neutral measure $\mathbb{Q}$ by :

$$C(S_0, K, r, T, \sigma) = S_0\,N(d_1) - Ke^{-rT}\,N(d_2) \tag{2.1.5}$$

with $N(\cdot)$ the standard normal CDF, $d_1 = \frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T\right]$ and $d_2 = d_1 - \sigma\sqrt{T}$.

Now suppose that the current market price C_t^m of the option is observed, and thus it can be taken as an input. Then the implied volatility at time t, denoted as $\hat{\sigma}_t^{imp}$, is the **unique** parameter σ of the B&S formula such that the option price C_t given by equation (2.1.5) equals the market price C_t^m, i.e. :

$$C_t = C(S_t,\, K,\, r,\, T-t,\, \hat{\sigma}_t^{imp}) = C_t^m$$

This means that the implied volatility $\hat{\sigma}_t^{imp}$ is the only value, when put in the Black-Scholes formula, that results in a model price equal to the current market price of a call option. There is then a one-to-one relationship between the market price of a European option and its implied volatility.

In a study by Szakmary et al. (2003), the authors focus on futures prices and options. They conclude that implied volatility outperforms historical volatility as a predictor of realized volatility in the underlying futures prices. Furthermore, the actual value of the implied volatility $\hat{\sigma}_t^{imp}$ of a call (or a put) option determined in this way depends, in general, on the strike K and the time to maturity $\tau = T - t$. Given a family $C_0^m(T, K)$ of current market prices at time $t = 0$ of call options with different strikes and maturities, we therefore find a parametrized family of implied volatilities $\hat{\sigma}_0^{imp}(T, K)$. We call volatility surface the plot depicting the dependence of the implied volatility against these two parameters (as in Figure 2.1.2). In some sense, the volatility surface describes the deviation of market option prices

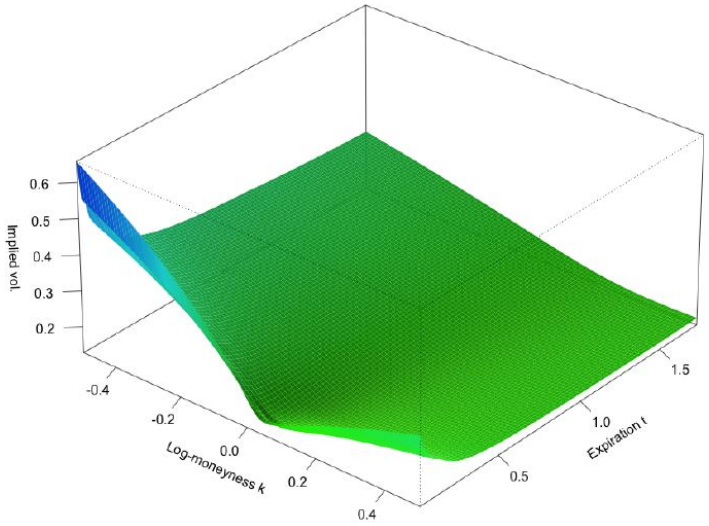

Figure 2.1.2: Example of implied volatility surface $\hat{\sigma}_0^{imp}(T, K)$, S&P500 as of June 20, 2013.

from their theoretical values, obtained within the B&S model. Namely, if the B&S model was able to exactly explain the observed option prices, the volatility surface would be representable as a constant function. The implied volatility can hence be summarized as the wrong parameter to put in the wrong formula to get the right price (Leoni (2014)). Moreover, if we replace the family of observed prices C_t^m by prices of a stochastic model (see Section 2.3), the implied volatility is said to be model-based.

Moreover, if we assume that the maturity date T is fixed, we obtain the mapping $K \mapsto \hat{\sigma}_0^{imp}(T, K)$, called the implied volatility curve. We have a *skew* if implied volatilities are higher for low strikes than for high strikes and a *smile* if the curve has a minimum. This minimum occurs usually for K lying not far away from the forward value $F_0 = S_0\, e^{rT}$ of the underlying. An inverse smile is called a *frown* but it is not common in equity markets. These behaviors are described in the following figure :

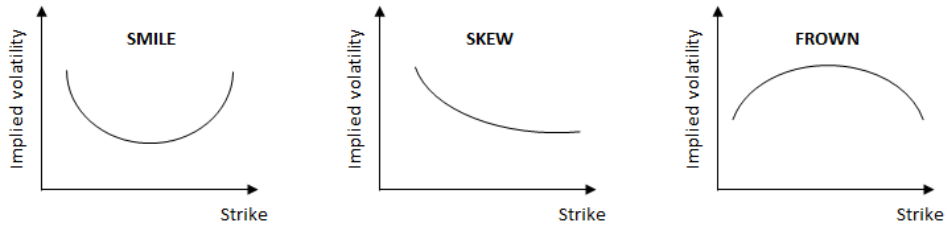

Figure 2.1.3: Implied volatility smile, skew and frown for T fixed.

We can also see on Figure 2.1.2 a decreasing amplitude of the smile being a function of time to maturity. For short maturities the smile (or skew) effect is very pronounced but it almost completely disappears for longer maturities, which is typical for equity data. Another important characteristic of the equity volatility surface is the term structure of at-the-money (ATM) volatility skews which is defined at time t as $\psi(\tau) := |\frac{\partial}{\partial K}\hat{\sigma}_t^{imp}(\tau, K)|_{K=S_t}$. We will come back on it later in Section 4.2 but empirically, it can be well-approximated by a power-law function $\psi(\tau) = A\tau^{-\alpha}$, $0 < \alpha < 1/2$.

We can again conclude that the concept of implied volatility is not constant and depends on K and T. Moreover, this volatility surface is not constant through time either and its level and orientation tend to move according to changes in market prices and market conditions. In order to reflect these characteristics, it is necessary to introduce models of volatility. We first start with local volatility models that are deterministic and then review stochastic volatility models in Section 2.3.

2.2 Local volatility

We know that the volatility surface is not flat with respect to K nor T. Therefore, we may want to construct a complete diffusion-type model that yields an exact fit to the observed term structure of volatility smiles/skews, provided that no static arbitrage opportunities are present in the market. We therefore want to find a **deterministic function** $\sigma(S_t, t)$ of the stock price and the date such that the model exactly provides the observed prices $C_0^m(T, K)$ of plain vanilla options (European calls and puts). Specifically, we may postulate that the martingale dynamic of S_t under the unique risk-neutral measure $\mathbb{Q}$ is :

$$dS_t = S_t \left(r\, dt + \sigma(S_t, t)\, dW_t^{\mathbb{Q}}\right) \tag{2.2.1}$$

The diffusion coefficient $\sigma(S_t, t)$ is referred to as the *local volatility function* and $W_t^{\mathbb{Q}}$ is a standard Brownian motion under $\mathbb{Q}$.
We first need the Breeden-Litzenberger formula in order to construct such diffusion-type models. Breeden and Litzenberger (1978) observed that, for any given future date T, the one-dimensional risk-neutral probability law (since T is fixed) of the underlying asset is uniquely determined by prices of European call options with all possible strikes K and maturity date T. In other words, if we fix T and we assume that we observe prices of all call options with different strikes, we can prove that, under mild technical conditions, we can recover the probability density function of S_T uniquely under the risk-neutral probability measure $\mathbb{Q}$. Mathematically, we postulate that for a fixed maturity T and every strike K, the call price $C_0(T, K)$ at time 0 given by the model (2.2.1) is :

$$C_0(T, K) = B_0\, \mathbb{E}^{\mathbb{Q}}(B_T^{-1}(S_T - K)^+) = e^{-rT} \int_0^\infty (s - K)^+ f(s, T)\, ds$$

with $B_T = e^{rT}$ and r a constant risk-free rate. It can be easily shown thanks to Leibniz integral rule that the density function $f(., T)$ satisfies $\forall K$:

$$f(K, T) = e^{rT}\, \frac{\partial^2 C_0(K, T)}{\partial K^2} \tag{2.2.2}$$

which proves the proposition of Breeden and Litzenberg. Obviously, $C_0(., T)$ must be a twice continuously differentiable function with respect to K. If we dispose of a continuum of market option prices in both strike and time dimension, we can replace $C_0(K, T)$ by $C_0^m(K, T)$ to find the density of the underlying stock process S_t. However, the knowledge of all one-dimensional marginal laws is not sufficient for the **unique** specification of the stochastic process S_t.

Nevertheless, Dupire (1993) managed to overcome this issue and to find a unique fully specified stochastic process S_t that is capable of reproducing the family of observed market prices of call options $C_0^m(T, K)$. To ensure the uniqueness of such an implied process, Dupire proposed to restrict attention to the class of diffusion processes that follows the dynamic given by equation (2.2.1). We finally assume that the observed market prices $C_0^m(T, K)$ of vanilla options correspond to an arbitrage-free model.

Formally, the following definition follows. A function $\sigma(S_t\,, t) : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}^+$ such that the prices of all plain-vanilla call options $C_0(K, T)$ given by the model (2.2.1) coincide with the market prices $C_0^m(T, K)$, that is

$$C_0^m(T, K) := C_0(K, T) = B_0\, \mathbb{E}^{\mathbb{Q}}(B_T^{-1}(S_T - K)^+), \tag{2.2.3}$$

is called the *local volatility*.
Dupire proved the following result. If we assume that $C_0(T, K)$ is of class $C^{1,2} : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+$,

then the implied volatility function is given by the **deterministic** expression :

$$\sigma^2(K,T) = \frac{2\left(\frac{\partial C_0(K,T)}{\partial T} + r\,K\frac{\partial C_0(K,T)}{\partial K}\right)}{K^2\,\frac{\partial^2 C_0(K,T)}{\partial K^2}} = \frac{2\left(\frac{\partial C_0^m(K,T)}{\partial T} + r\,K\frac{\partial C_0^m(K,T)}{\partial K}\right)}{K^2\,\frac{\partial^2 C_0^m(K,T)}{\partial K^2}}$$

and

$$\sigma^2(S_t\,,t) = \sigma^2(K,T)|_{K=S_t\,,T=t}$$

which proves that there exists a unique implied local volatility $\sigma(S_t,t)$ compatible with a full collection $C_0^m(T,K)$ of market prices of European calls under some regularity conditions for $\sigma(S_t,t)$.

The advantage of local volatility (LV) models is that under reasonable assumptions, there exists a unique LV model reproducing market prices of standard option and leading to a complete model (and hence a unique risk-neutral measure $\mathbb{Q}$). It means that any contingent claim can be hedged by a unique replicating strategy and that all derivatives assets are redundant. However, these models hinge on the unrealistic assumption that we can observe call option prices for all maturities and all strikes, which is not the case in practice, leading to ill-posed problems. Furthermore, it is based on the belief that all contingent claims can be uniquely determined by market prices of European calls, which is obviously too simplistic. This explains that these models have the inherent flaw of incorrectly predicting the future movements of volatilities, leading to unrealistic future dynamics of the volatility surface through time.
To improve pricing and hedging properties of LV method, Dupire (1996) and Derman and Kani (1998) proposed to introduce dynamical behavior of the local volatility. In this version of the LV model, the local volatility surface specified by the function $\sigma(S_t,t)$ is no longer *static*, but it is assumed to obey a stochastic process, driven by some additional Brownian motions. However, we will not deepen these models and will focus instead in this thesis on stochastic volatility models, introduced in the next section.

2.3 Stochastic volatility models

Stochastic volatility (SV) models are useful because they explain in a self-consistent way why volatility is not constant trough time and why options with different strikes K and maturities T have different B&S implied volatilities. Indeed, in SV models, the volatility σ_t is modeled as a continuous Brownian semi-martingale (see Concept 7.2.7 in Appendix). Notable amongst such stochastic volatility models are the Hull and White model, the Heston model and the SABR model. Whilst stochastic volatility dynamics are more realistic than local volatility dynamics, generated option prices are not consistent with observed European option prices and lead to volatility surfaces whose shapes can differ substantially from that of the empirically observed volatility surfaces. Adding jumps to these models (such as the Merton or Bates model) can help to better fit the volatility surface as explained at the end of the section.

Further simple visual arguments can also explain the use of SV models. Figure 2.3.1 shows the distribu-

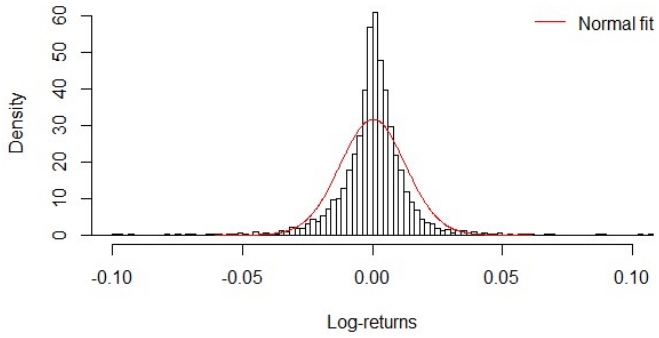

Figure 2.3.1: Histogram of log-returns, S&P500.

tion of the S&P500 log-returns between January 2000 and July 2020, with the normal fit superimposed in red. We see that this distribution is highly peaked and fat-tailed relative to the normal distribution. The Q-Q plot of these log-returns in Figure 7.1.1 (Appendix) shows just how extreme the tails of the empirical distribution of returns are relative to the normal distribution. These are characteristics of mixtures of distributions with different variances. This motivates us to model the volatility as a random variable. Finally, Figure 2.3.2 shows that large moves in returns tend to be followed by large moves (of either sign) and small moves by small moves. This feature is known as volatility clustering

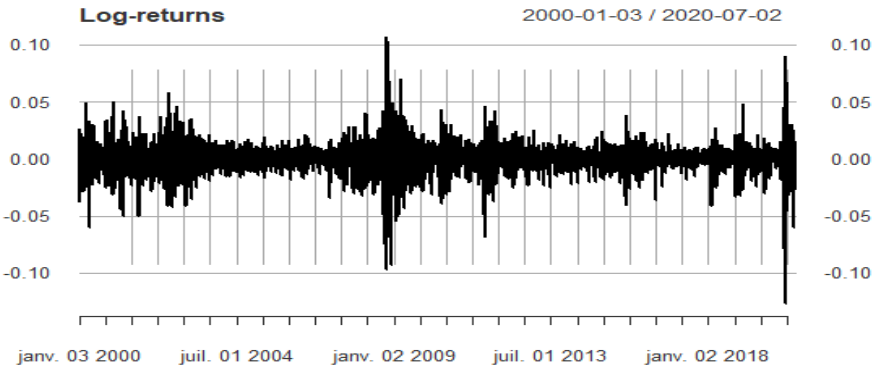

Figure 2.3.2: Time series of log-returns, S&P 500.

and results from the fact that the volatility is auto-correlated. This is confirmed by the autocorrelation function (ACF) of the squared log-returns (which approximate σ_t^2) in Figure 7.1.2 (Appendix), where we observe a very slow decay of the autocorrelation. We will translate this in our stochastic models by a mean reversion of volatility. Clearly, the B&S model with a constant σ does not allow to take into account non-normally distributed log-returns and the slowly varying autocorrelation function of volatility (i.e. the volatility clustering).

In a continuous-time framework, the random volatility σ_t is usually assumed to obey a diffusion-type process. Let the stock price S_t be given under the real-world measure $\mathbb{P}$ by :

$$dS_t = \mu(S_t, t)\, dt + \sigma_t\, S_t\, dW_t \tag{2.3.1}$$

with the stochastic volatility σ_t (called *instantaneous volatility*) satisfying :

$$d\sigma_t = a(\sigma_t, t)\, dt + b(\sigma_t, t)\, d\hat{W}_t \tag{2.3.2}$$

where $\mu(S_t, t)$, $a(\sigma_t, t)$ and $b(\sigma_t, t)$ are adapted processes. W_t and $\hat{W}_t$ are one-dimensional Brownian motions defined on $(\Omega, \mathbb{P}, \mathcal{F})$ with $d\langle W, \hat{W}\rangle_t = \rho\, dt$, that is, a constant instantaneous correlation. We can write $\hat{W}_t = \rho\, W_t + \sqrt{1-\rho^2}\, W_t'$ where W_t' is an independent Brownian motion from W_t under $\mathbb{P}$.

Under non-unique equivalent risk-neutral measures denoted $\mathbb{P}^*$ and defined by Girsanov theorem (see Appendix - Demo. 1), we have that S_t/B_t is martingale and :

$$dS_t = r\, S_t\, dt + \sigma_t\, S_t\, dW_t^* \tag{2.3.3}$$

$$d\sigma_t = \tilde{a}(\sigma_t, t)\, dt + b(\sigma_t, t)\, d\hat{W}_t^* \tag{2.3.4}$$

where W_t^* and $\hat{W}_t^*$ are one-dimensional correlated standard Brownian motion under $\mathbb{P}^*$ and $\tilde{a}(\sigma_t, t)$ is called the risk-neutral drift rate. If we adopt the following commonly standard convention, we have :

$$\tilde{a}(\sigma_t, t) = a(\sigma_t, t) - \frac{\mu(S_t, t) - r}{\sigma}\, \rho\, b(\sigma, t) - \lambda(\sigma_t, t)\, \sqrt{1-\rho^2}\; b(\sigma_t, t) \tag{2.3.5}$$

for some sufficiently regular function $\lambda(\sigma_t, t)$ called the market price of risk[1]. This additional term comes directly from the Girsanov theorem (see Appendix - Demo. 1 for the complete proof). Since

[1]Note that we can rewrite this equation as $\tilde{a}(\sigma_t, t) = a(\sigma_t, t) - b(\sigma_t, t)\Lambda(\sigma_t, t)$, see Demo. 1

S_t/B_t is martingale under $\mathbb{P}^*$, we can write that under $\mathbb{P}^*$ and with respect to $\mathcal{G}_t$, the augmented filtration generated by the volatility process σ_t and the Brownian motion W_t^* :

$$\mathbb{E}^{\mathbb{P}^*}\left[\frac{S_T}{B_T}\middle|\mathcal{G}_t\right] = \frac{S_t}{B_t} = S_0 \exp\left(\int_0^t \sigma_u\, dW_u^* - \frac{1}{2}\int_0^t \sigma_u^2\, du\right) \tag{2.3.6}$$

It is also possible to derive, under mild additional assumptions, the partial differential equation satisfied by the value function of a European contingent claim. Consider a European contingent claim $\pi(S_t, \sigma_t, t)$ whose payoff is $g(S_T)$ at maturity T. The price of this derivative is given by its discounted expected payoff under a risk-neutral measure $\mathbb{P}^*$:

$$\pi(S_t, \sigma_t, t) = B_t\, \mathbb{E}^{\mathbb{P}^*}(B_T^{-1} g(S_T) \mid \mathcal{F}_t)$$

Then the price $\pi(S_t, \sigma_t, t)$ satisfies the Feymann-Kac PDE :

$$\frac{\partial \pi}{\partial t} + r\, S \frac{\partial \pi}{\partial S} + \tilde{a}\frac{\partial \pi}{\partial \sigma} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \pi}{\partial S^2} + \frac{1}{2} b^2 \frac{\partial^2 \pi}{\partial \sigma^2} + \rho\, \sigma S\, b \frac{\partial^2 \pi}{\partial S \partial \sigma} - r\pi = 0$$

with terminal condition $\pi(s, \sigma, T) = g(s)$ and for $\rho_t = \rho\ \forall t$.

Generally speaking, stochastic volatility models are not complete since there are two sources of risk (two Brownian motions) and only one risky asset traded (the stock). Hence, $\mathbb{P}^*$ is not unique and there exist other risk-neutral measures under which S_t/B_t is martingale. Therefore, one needs to specify exogeneously the market price of volatility risk $\lambda(\sigma, t)$ (or[1] $\Lambda(\sigma_t, t)$) since it cannot be determined by arbitrage considerations.

Finally, it is a stylized fact that, at least in equity markets, although the level and orientation of the volatility surface do change 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, i.e. as a process whose parameters $a(\sigma_t, t)$ and $b(\sigma_t, t)$ are independent of price and time.

2.3.1 Hull and White model

Consider the diffusion-type process described above and particularly the equation (2.3.3) and (2.3.4) under the (non-unique) martingale measure $\mathbb{P}^*$. As a first step, Hull and White (1987) postulate that W_t and $\hat{W}_t$ are independent Brownian motions under the real-world probability $\mathbb{P}$. That is, the stochastic volatility σ_t is independent of the stock price. By construction, W_t^* and $\hat{W}_t^*$ are also independent under $\mathbb{P}^*$. Moreover, it is rather clear from equation (2.3.6) that dynamics of the stock price S_t, when conditioned on the augmented filtration $\mathcal{G}_t$, is lognormal under $\mathbb{P}^*$. They also postulate that the processes $a(\sigma_t, t)$ and $b(\sigma_t, t)$ of equation (2.3.2) may depend of time and of σ_t but they do not depend on S_t. Mathematically, if we denote by A_t the stochastic process representing the integrated squared volatility, we have :

$$A_t = \int_0^t \sigma_u^2\, du$$

Let then F_S stand for the cumulative distribution function of S_T under $\mathbb{P}^*$, so that :

$$F_S(s) = \int_0^\infty F_{S\,|\,A}(s\,|\,w)\, dF_A(w)$$

where $F_{S|A}(s\,|\,w)$ denotes the conditional cumulative distribution function of S_T given $A_T = w$ and is lognormally distributed as said above. Specifically, it comes $dF_{S\,|\,A}(s\,|\,w) = f_{S\,|\,A}(s\,|\,w)\, ds$ where :

$$f_{S|A}(s\,|\,w) = \frac{1}{\sqrt{2\pi w} s}\, \exp\left\{ - \frac{(\ln(s/S_0) - rT + \frac{1}{2}w)^2}{2w} \right\}$$

since the conditional distribution of $\log(S_T/S_0)$ knowing $A_t = w$ is $N(rT - \frac{w}{2}, w)$ under $\mathbb{P}^*$. Therefore, it is clear that the constant parameter σ of the B&S model is replaced by $\sqrt{w/T}$ since the standard deviation of $\log(S_T/S_0)$ under the classical B&S model is $\sigma\sqrt{T}$. Then, Hull and White obtain the price at time 0 of any contingent claim $X = g(S_T)$ settling at time T, with :

$$\pi_0(X) = \int_0^\infty g(s)\, dF_S(s) = \int_0^\infty g(s) f_{S|A}(s \mid w)\, ds\, dF_A(w) \tag{2.3.7}$$

From that, the valuation of any contingent claim is rather straightforward if we know how to price this claim analytically in the B&S framework and if the cumulative distribution function F_A of A_t is given explicitly. Indeed, it comes easily :

$$\pi_0(X) = \int_0^\infty \pi_0^{BS}(X|w)\, dF_A(w) \tag{2.3.8}$$

where $\pi_0^{BS}(X|w)$ denotes the price of X in the B&S model with the volatility parameter σ replaced by $\sqrt{w/T}$ (for instance, in equation (2.1.5) of a European call option with $g(s) = (s-K)^+$).

Therefore, the price of a path-independent contingent claim in a stochastic volatility framework with independent driving Brownian motions is the expected Black-Scholes price, where the expected value is taken over the probability distribution of the integrated squared volatility. See Romano and Touzi (1997) for the case of correlated Brownian motions between the volatility process and the stock process. One of the stochastic volatility models considered by Hull and White (1987) for the instantaneous variance $v_t = \sigma_t^2$ is :

$$dv_t = a(v_t, t)\, v_t\, dt + b(v_t, t)\, v_t\, d\hat{W}_t$$

However, there is no closed form solutions for such models and the value of an option must be found numerically (even for constant a and b). One solution proposed by Hull and White is to compute the moments of A_t and use them to derive an approximate formula for the option price. This solution is not perfect which explains why the following model with a closed-form solution at disposal is more popular.

2.3.2 Heston model

We consider here the following mean-reversion process under the real-world measure $\mathbb{P}$ introduced by Heston (1993) :

$$dS_t = S_t\, \mu_t\, dt + S_t\, \sqrt{v_t}\, dW_t \tag{2.3.9}$$

$$dv_t = \kappa(\eta - v_t)\, dt + \theta\, \sqrt{v_t}\, d\hat{W}_t \tag{2.3.10}$$

with W_t and $\hat{W}_t$ two correlated Brownian motions defined on $(\Omega, \mathbb{P}, \mathcal{F})$ such that $\langle W, \hat{W} \rangle_t = \rho\, t$ with $-1 \leq \rho \leq 1$, the correlation between the stock and the variance processes. We also have that $S_0 \geq 0$ is the initial stock price, $v_0 = \sigma_0^2 \geq 0$ the initial variance, η is the long-run variance of v_t, κ denotes the speed of mean-reversion and θ the volatility of variance parameter. The instantaneous variance v_t here is a CIR process (square root process).

It is important to note that the Heston model allows to take into account the leverage effect. The leverage effect comes from the observation that positive and negative returns do not have the same impact on volatility. Indeed, positive returns increase volatility more than negative ones. This leverage effect influences the implied volatility and we shall include it in our model to ensure a good fit of the observed implied volatilities. More precisely, the observed anticorrelation between price moves and volatility moves may then be modeled naturally by taking $\rho < 0$. Furthermore, instantaneous correlation between returns and volatility allows to take into account asymmetric smiles (i.e. skews) whereas the independence assumption leads tot a U-shaped symmetric volatility curve (i.e. a smile).

Moreover, in order to ensure that the volatility process never hits zero (theoretically), we impose the so-called Feller condition :

$$2\kappa\eta \geq \theta^2 \tag{2.3.11}$$

Using Girsanov theorem with a constant risk-free rate r, we have under a risk-neutral measure $\mathbb{P}^*$:

$$dS_t = S_t\, r\, dt + S_t\, \sqrt{v_t}\, dW_t^* \tag{2.3.12}$$

The risk-neutral process for the variance is then obtained by introducing the function $\Lambda(v_t, t)$ (see footnote 1) into the drift of dv_t as follows (using equation (2.3.4) and Girsanov as in Demo. 1) :

$$dv_t = \big(\kappa(\eta - v_t) - \Lambda(v_t, t)\big)\, dt + \theta\, \sqrt{v_t}\, d\hat{W}_t^*$$

where $\hat{W}_t^* = \hat{W}_t + \frac{\Lambda(v_t,t)}{\theta\sqrt{v_t}}\, t$.
In the Heston model, we consider the following market price of risk $\Lambda(v_t, t) = \lambda\, v_t$, which leads to the risk-neutral version of the variance process :

$$dv_t = \kappa^*(\eta^* - v_t)\, dt + \theta\sqrt{v_t}\, d\hat{W}_t^* \tag{2.3.13}$$

where $\kappa^* = \kappa + \lambda$ and $\eta^* = \kappa\,\eta/(\kappa + \lambda)$. We clearly see that equations (2.3.10) and (2.3.13) have the exact same form (up to different constants) which explains that the Heston model is usually directly specified in the risk-neutral world. Since the Heston model is a SV model, recall that it is not complete and therefore there exists no unique equivalent martingale measure.

There also exists a closed-form solution for the characteristic function of the Heston model. We refer to the Concept 7.2.2 of characteristic function in Appendix for a theoretical remainder of the characteristic function $\Phi_X(u)$ of a distribution function. The Fast Fourier approach by Carr and Madan (more details on this method in Concept 7.2.3 in Appendix) which uses the characteristic function of the log-stock price at maturity speeded up and extended the practical use of the Heston model by its ability to facilitate the calibration of this model to the observed European option prices. The Heston characteristic function of the log-stock price $\log(S_t)$ at time t and point u is given by :

$$\begin{aligned}\Phi(u,t) &= \mathbb{E}\left[\exp\left(iu\log(S_t)\right)|S_0, v_0\right] \\ &= \exp\left(iu(\log S_0 + r\,t)\right) \times \exp\left(\eta\,\kappa\,\theta^{-2}\left((\kappa - \rho\,\theta u i - d)\,t - 2\log\left(\left(1 - ge^{-dt}\right)/(1-g)\right)\right)\right) \\ &\times \exp\left(v_0\,\theta^{-2}\,(\kappa - \rho\,\theta\,iu - d)(1 - e^{-dt})/(1 - ge^{-dt})\right)\end{aligned} \tag{2.3.14}$$

where :

$$\begin{aligned} d &= \left((\rho\,\theta\,u\,i - \kappa)^2 - \theta^2\left(-iu - u^2\right)\right)^{1/2} \\ g &= (\kappa - \rho\,\theta\,u\,i - d)\,/\,(\kappa - \rho\,\theta\,u\,i + d)\end{aligned}$$

Finally, recall that the Heston model does not fit the observed implied volatility surface since it is a SV model and hence, generated option prices are not consistent with currently observed European option prices. More precisely, the Heston model is usually unable to capture the high curvature/slope of the implied volatility curve (smile or skew) usually observed for short-term options, as in Figure 2.1.2 of the S&P500 volatility surface. Hence, the Heston model cannot fit the term structure of ATM skews $\psi(\tau)$ either. This is a direct consequence of the fact that the conditional distribution of the short-term log-returns (under the risk-neutral measure) does not depart sufficiently from a normal distribution. More advanced SV models based on jumps or on fractional Brownian motions will be built in the following sections of this thesis to fill this gap and generate consistent volatility surfaces.

2.3.3 Jump-diffusion type model

Moreover, the Heston model and Hull and White models presented so far are not able to reproduce the sudden large changes in equity prices observed empirically (=crash-like events). In order to model

these sudden large changes, the Merton and Bates models introduce *jumps* in the equity price process. The jump times are then modeled by a Poisson process. Once a jump occurs, the jump height is drawn from its jump distribution, which is supposed to be lognormal. Therefore, we first define the compound Poisson process :

$$J_t = \sum_{i=1}^{N_t} Y_i$$

where $N_t \sim Poi(\lambda t)$, a Poisson process with jump intensity $\lambda \in \mathbb{R}^+$ and Y_i i.i.d. random variables with distribution F_Y. J_t is in general not a martingale and we hence introduce the compensated process $\tilde{J}_t = J_t - \mathbb{E}[J_t] = J_t - \lambda\, t\, \mathbb{E}[Y]$, which is martingale. We can now define jump-diffusion type processes which are processes consisting of a diffusion process with drift and a compensated Poisson process $\tilde{J}_t$:

$$dX_t = \mu_t\, dt + \sigma_t\, dW_t + d\tilde{J}_t$$

where μ_t and σ_t are adapted processes. We now first introduce the Merton model under a risk-neutral measure $\mathbb{P}^*$, which is given by the following SDE :

$$\frac{dS_t}{S_t} = (r - \lambda\,\xi)\, dt + \sigma\, dW_t^* + (Y_t - 1)\, dN_t \tag{2.3.15}$$

where $\log(Y_t) \overset{\text{i.i.d.}}{\sim} N(\mu_j, \sigma_j^2)$ and where $\xi = \mathbb{E}[Y_t - 1] = e^{\mu_j + \frac{1}{2}\sigma_j^2} - 1$ and $\mathbb{V}[Y_t - 1] = \left(e^{\sigma_j^2} - 1\right) e^{2\mu_j + \sigma_j^2}$. We also have that $\langle W^*, Y\rangle_t = 0$, $\langle W^*, N\rangle_t = 0$ and $\langle N, Y\rangle_t = 0$. Using the Itô formula for jump-diffusion processes (see Tankov (2003)), we find :

$$S_t = S_0\, \exp\left((r - \sigma^2/2 - \lambda\,\xi)t + \sigma W_t^*\right) \prod_{i=1}^{N_t} Y_i$$

and the characteristic function $\Phi(u,t)$ of the log-stock price at t and point u is given by :

$$\begin{aligned}\Phi(u,t) &= \mathbb{E}[\exp(iu\, \log(S_t))|S_0, v_0] \\ &= \exp(iu(\log S_0 + r\, t)) \times \exp\left\{ i\, u\, \omega\, t - \frac{1}{2}u^2\sigma^2 t + \lambda\, t\left(e^{iu\mu_j - u^2\sigma_j^2/2} - 1\right)\right\}\end{aligned}$$

where $\omega = -\frac{1}{2}\sigma^2 - \lambda\left(e^{\mu_j + \frac{1}{2}\sigma_j^2} - 1\right) = -\frac{1}{2}\sigma^2 - \lambda\xi$ in order to have that S_t/B_t is a martingale under a (non-unique) risk-neutral measure $\mathbb{P}^*$. We can rewrite $\Phi(u,t)$ as :

$$\Phi(u,t) = \exp\left(iu(\log S_0 + r\, t) - \frac{1}{2}u(u+i)\sigma^2 t\right) \times \exp\left\{-\lambda\, t\left[iu\left(e^{\mu_j + \frac{1}{2}\sigma_j^2} - 1\right) - \left(e^{iu\mu_j - u^2\sigma_j^2/2} - 1\right)\right]\right\}$$

Written in this from, it comes that $\Phi(u,t)$ in the Merton model is the product of the characteristic function of $\log(S_t)$ in the B&S model and of the jump characteristic function.

We can then extend the Merton model by introducing stochastic volatility in it. In fact, we can combine the Heston model of stochastic volatility (Section 2.3.2) with the Merton model of independent normally distributed jumps in the log asset price. We obtain the Bates model (1996), also called SVJ, given by the following SDE under a risk-neutral measure $\mathbb{P}^*$:

$$\begin{aligned}\frac{dS_t}{S_t} &= (r - \lambda\,\xi)\, dt + \sqrt{v_t}\, dW_t^* + (Y_t - 1)\, dN_t \\ dv_t &= \kappa^*(\eta^* - v_t)\, dt + \theta\sqrt{v_t}\, d\hat{W}_t^*\end{aligned}$$

with $\langle W^*, \hat{W}^*\rangle_t = \rho\, t$ and W_t^*, $\hat{W}_t^*$ are independent from N_t and from Y. We also have that N_t and Y are independent like in the Merton model. The characteristic function $\Phi(u,t)$ is obtained in the exact same way as for the Merton model, by multiplying the characteristic function of $\log(S_t)$ in the Heston model (equation (2.3.14)) with the jump characteristic function above. Having the characteristic function of the log-price $\log(S_t)$ under the Bates model in closed-form, we can again use the

Carr-Madan formula described in Concept 7.2.3 to compute prices of European options and derive the implied volatility surface.

Under the Bates model, the distribution of the conditional short-term log-returns now departs sufficiently from the Gaussian distribution (presence of excess kurtosis and additional skewness). As a result of this non-normal distribution, the Bates model usually improves the fit of the implied volatilities for short-term options compared to the Heston model. Indeed, jumps generate a steep short-dated skew that dies quickly with time to expiration, which compensates the fact that SV models do not generate enough skew for very short expirations. The presence of jumps then allows to better fit the term structure of ATM skew $\psi(\tau)$. We can conclude that the SVJ succeeds in generating a more realistic volatility surface than the Heston model although the fit is still not perfect, especially for short expirations.

Finally, we can note that it is also possible to add jumps in the instantaneous volatility process v_t of the Bates model (called the SVJJ model) and also find the characteristic function of this model in quasi closed-form. Indeed, it seems unrealistic to suppose that the instantaneous volatility would not jump if the stock price were to jump. Furthermore, adding a simultaneous upward jump in volatility to jumps in the stock price allows us to maintain the clustering property of SV models as shown in Figure 2.3.2. However, Gatheral (2011) concludes that the SVJJ model does not provide a better fit of the volatility surface compared to the Bates model and in addition, has more parameters to calibrate. We therefore decide to only consider the Bates model and will compare it to fractional Brownian motion models in the following chapter which have far less parameters to calibrate and allow an even better fit of the implied volatility surface, especially of the ATM skew term structure $\psi(\tau)$ for short maturities.

2.3.4 Other SV models

- SABR model :
Another widely used SV model is the SABR model. However, this model is especially used by practitioners in interest rate derivative markets or in foreign exchange markets. Since we mostly consider equity derivatives in this thesis, we will not deepen this model.

- ARCH and GACH type models :
Neither will we consider for this thesis such statistical models for time series. Indeed, the non-degenerate feature of $(W_t, \hat{W}_t)$ is characteristic of the SV paradigm, in contrast to ARCH-type (and GARCH) models where the volatility process is a deterministic function of past values of the underlying asset price. Furthermore, we only consider here continuous-time processes whereas ARCH and GARCH-type models are discrete processes. For an econometric approach of volatility modeling, we refer to Ghysels, Harvey and Renault (1996).

Note that all the SV models we have reviewed can be extended using stochastic interest rates r_t (e.g. Hull and White type models for the short-term rate r_t). However, since we essentially focus on the volatility risk and not the interest rate risk, we do not cover these models in this master's thesis and only consider a deterministic interest rates curve. Furthermore, we did not explicitly introduce the dividend yield q in all the formula and models above but it can simply be done by replacing the terms in rt by $(r-q)t$.

Chapter 3

Self-similarity and fractional Brownian motions

In this chapter, we introduce some mathematical and statistical concepts that will help us to build more advanced volatility models later in this thesis. We mainly here focus on fractional Brownian motions and their properties. Based on these findings, we will finally study the so-called fractional Ornstein-Uhlenbeck process.

3.1 Self-similarity

A stochastic process $(X_t)_{t\geq 0}$ is self-similar (SS) if $\exists\, H > 0$ such that we have the following equality in distribution $\forall c > 0$:

$$(X_{ct})_{t\geq 0} \overset{d}{=} (c^H X_t)_{t\geq 0} \tag{3.1.1}$$

Intuitively, self-similarity means that a stochastic process scaled in time (that is, plotted with a different time scale) looks statistically the same as the original process when properly rescaled in space. H is called the *Hurst* exponent of the stochastic process.
Even though self-similar processes cannot be (strictly) stationary, it can have (strictly) stationary increments. In fact, we will focus on self-similar processes with stationary increments. One reason is that stationary increments of self-similar processes provide models of long-range dependent series (see Section 3.3). A process $(X_t)_{t\geq 0}$ has (strictly) stationary increments, if for any $h \in \mathbb{R}^+$:

$$(X_{t+h} - X_t)_{t\geq 0} \overset{d}{=} (X_t - X_0)_{t\geq 0} \tag{3.1.2}$$

If $(X_t)_{t\geq 0}$ is Gaussian[1] with mean 0 and $X_0 = 0$, then the stationarity of increments is equivalent to :

$$\mathbb{E}\left((X_t - X_s)^2\right) = \mathbb{E}\left((X_{t-s})^2\right), \quad \text{for all } s, t \in \mathbb{R}^+$$

3.2 Fractional Brownian motion

3.2.1 Definition

The previous basic results for self-similar processes with stationary increments lead to a natural definition of fractional Brownian motions.
First, let $(X_t)_{t\geq 0}$ be a real value H-selfsimilar process with stationary increments (H-SSSI) and suppose that $\mathbb{E}\left((X_1)^2\right) < \infty$. Then, it is easy to show (see Demo. 2 - Appendix) that :

$$\mathbb{E}(X_t X_s) = \frac{1}{2}\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\}\mathbb{E}\left((X_1)^2\right)$$

An example of such H-SSSI processes are fractional Brownian motions. Let $0 < H < 1$, a real-valued Gaussian process[1] $(B_t^H)_{t\geq 0}$ is called *fractional Brownian motion* (fBm) if $\mathbb{E}(B_t^H) = 0$ and :

$$\mathrm{Cov}(B_t^H, B_s^H) = \mathbb{E}(B_t^H B_s^H) = \frac{1}{2}\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\}\mathbb{E}\left((B_1^H)^2\right) \tag{3.2.1}$$

[1] $(X_t)_{t\geq 0}$ is a Gaussian process if $a_1 X_{t_1} + \ldots + a_n X_{t_n}$ is a Gaussian random variable with $a_i \in \mathbb{R}$, $t_i \in \mathbb{R}^+$, $i = 1, \ldots, n$.

Zero-mean H-SSSI Gaussian processes therefore provide a natural way to generate fBm. Moreover, since the distribution of a zero-mean Gaussian process is entirely determined by its covariance function, equation (3.2.1) implies that, for a given H, all H–SSSI zero-mean Gaussian processes differ by a multiplicative constant only, the term $\mathbb{E}\left((B_1^H)^2\right)$.
Following MandelBrot and Van Ness (1968), when $0 < H < 1$, fBm have a stochastic integral representation in terms of standard Brownian motions given by :

$$B_t^H = C_H\Big\{\int_{-\infty}^{0}\Big((t-s)^{H-1/2} - (-s)^{H-1/2}\Big)\,dW_s + \int_0^t (t-s)^{H-1/2}\,dW_s\Big\} \tag{3.2.2}$$

where W_t a standard Brownian motion and $C_H = \Big(\int_{-\infty}^{0}\left((1-s)^{H-1/2} - (-s)^{H-1/2}\right)^2 ds + \frac{1}{2H}\Big)^{-1/2}$
We can also write as in Bayer et al. (2016) :

$$B_t^H = C_H\Big\{\int_{-\infty}^{t}\frac{dW_s}{(t-s)^{\gamma}} - \int_{-\infty}^{0}\frac{dW_s}{(-s)^{\gamma}}\Big\} \tag{3.2.3}$$

with $\gamma = 1/2 - H$ and $C_H = \sqrt{\frac{2H\,\Gamma(3/2-H)}{\Gamma(H+1/2)\,\Gamma(2-2H)}}$. Strictly speaking, this expression is only formal.

Note that the Mandelbrot-Van Ness representation uses the entire history of the Brownian motion $(W_s)_{s\leq t}$, which has an impact on the memory of the process (see Section 3.3). Furthermore, a fBm $(B_t^H)_{t\geq 0}$ has stationary increments as shown above but has independent increments if and only if $H = 1/2$ since in this case $(B_t^{1/2})_{t\geq 0} = (W_t)_{t\geq 0}$ is Brownian motion. If $H > 1/2$, increments are positively correlated and the sample paths are then smoother and if $H < 1/2$, increments are negatively correlated and the sample paths are then more irregular, rougher.

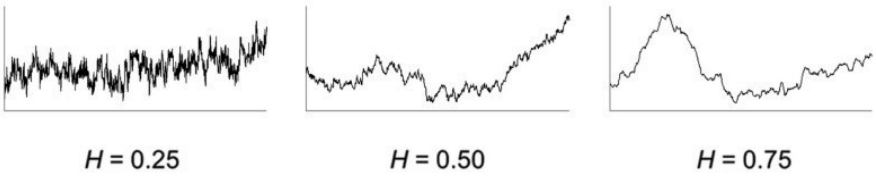

Figure 3.2.1: Fractional Brownian motions for different values of H.

3.2.2 Sample paths properties

First, we say that a stochastic process $(X_t)_{t\geq 0}$ is Hölder continuous (or Hölder regular) of order $\gamma \in (0,1)$ if for $s,t \in \mathbb{R}^+$:

$$P\Big\{\omega \in \Omega, \sup_{0<t-s<h(\omega)} \frac{|X_t(\omega) - X_s(\omega)|}{|t-s|^{\gamma}} \leq \delta\Big\} = 1$$

The Hölder continuity condition is stronger that the classical continuity condition and therefore implies the latter. Secondly, by self-similarity and stationary increments of $(B_t^H)_{t\geq 0}$, we have :

$$\mathbb{E}(\,|B_t^H - B_s^H|^{1/\gamma}) = \mathbb{E}(\,|B_{t-s}^H|^{1/\gamma}) = |t-s|^{H/\gamma}\,\mathbb{E}(\,|B_1^H|^{1/\gamma}) \tag{3.2.4}$$

Hence, we can prove from (3.2.4) and from the Kolmogorov's continuity criterion (see Karatzas and Shreve (1998)) that sample paths of $(B_t^H)_{t\geq 0}$ are Hölder continuous of order β, for any $\beta < H$. Similarly, we can also prove that the sample paths of $(B_t^H)_{t\geq 0}$ are almost surely nowhere Hölder continuous of order β for $\beta > H$. Therefore, the larger the Hurst exponent, the smoother the sample path, which confirms what we can see on Figure 3.2.1.

Secondly, sample paths of fractional Brownian motions have nowhere bounded variation and hence, are nowhere differentiable (see Kawada and Kôno (1971) for the proof), as it is the case for standard

Brownian motions.

Finally, we show in Demo. 3 that fBm are not semi-martingale for $H \neq 1/2$ (see again Concept 7.2.7 in Appendix). Of course, Brownian motions $(B_t^H)_{t\geq 0}$ with $H = 1/2$ are semi-martingale. This allows us to construct the so-called Itô-calculus with respect to Brownian motion. On the other hand, stochastic integrals with respect to fractional Brownian motion (with $H \neq 1/2$) are needed. The non-semimartingale property implies that the classical Itô construction and its properties do not hold. The Itô theory for stochastic integral has thus to be adapted in order to define $\int_0^t \phi(s)\, dB_s^H$.

Because of the sample path properties of fBm, for $\mathbf{1/2 < H < 1}$, the fBm has smoother sample paths than Brownian motion and hence it will be easier to construct a stochastic integral. Indeed, one can show that a pathwise Riemann–Stieltjes construction[2] can be built if $(B_t^H)_{t\geq 0}$ has a finite p-variation for $p \in (1/H, 2)$ (see Concept 7.2.4 of p-variation in Appendix). This comes from the fact that if X_t is β–Hölder continuous then its $1/\beta$-variation is finite. Then, since we have from (3.2.4) that B_t^H is Hölder continuous for $1/2 < \beta < H < 1$, it comes that $1/H < 1/\beta = p < 2$ and therefore that the p-variation is finite with $p \in (1/H, 2)$, which implies that we can use pathwise integration for $1/2 < H < 1$.
For $\mathbf{0 < H < 1/2}$, sample paths of fractional Brownian motion are more irregular than those of Brownian motion and therefore other constructions have to be followed. The most important constructions (for general H) to be found in the literature either restrict $\phi(t)$ to specific classes of functions, use pathwise integration or base a definition on Malliavin calculus. We will not deepen this problematic here but in the following, further arguments will be developed (mainly based on pathwise integration).

3.3 Long-range dependence

Long-range dependence is commonly defined for weakly stationary processes. A stochastic process $(X_t)_{t\geq 0}$ is said to be weakly stationary if it has a constant mean $\mathbb{E}(X_t) = \mu_X \ \forall t$, a finite second moment $\mathbb{E}(X_t^2) \ \forall t$ and an autocovariance function given by :

$$\gamma_X(t-s) = \mathbb{E}(X_t X_s) - \mathbb{E}(X_t)\,\mathbb{E}(X_s) = \mathbb{E}(X_t X_s) - \mu_X^2 = \mathbb{E}(X_{t-s} X_0) - \mu_X^2$$

Note that γ_X depends only on the lag $|t-s|$. For Gaussian processes (which are the processes studied in this thesis), the notion of weak and strict stationarity coincide since Gaussian processes are entirely determined by their covariance function.

We also need to introduce the definition of slowly varying function. A function L is slowly varying at infinity if it is positive on $[c, \infty)$ with $c \geq 0$ and for any $a > 0$:

$$\lim_{u\to\infty} \frac{L(au)}{L(u)} = 1 \tag{3.3.1}$$

For example, the functions $L(u) = K$, $K > 0$ (constant function) and $L(u) = \log u$, $u > 0$, are slowly varying at infinity.

We now turn to two of the main definitions of *long-range dependence* (also called *long memory*) that can be found in the literature. A stochastic process $(X_t)_{t\geq 0}$ is called long-range dependent (LRD) if :

- Condition 1 : The autocovariances of the process are not absolutely summable; that is,

$$\sum_{h=0}^{\infty} |\gamma_X(h)| = \infty \tag{3.3.2}$$

Indeed, the process presents in this case an aggregation behavior and may then describe systems with memory and persistence to past events.

[2] Classical integration theory with almost surely convergence.

or **Condition 2** : The autocovariance function of the process satisfies,

$$\gamma_X(h) = L(h)\, h^{2d-1}, \quad h \geq 0 \tag{3.3.3}$$

where L is a slowly varying function at infinity and $d \in (0, 1/2)$. This condition refers to the slow decay of the autocorrelation function, anything slower than exponential and more precisely that it decays as a power-law with exponent less than 1.

Suppose $(Y_t)_{t\geq 0}$ is an H–SSSI process with $0 < H < 1$, and consider the increments :

$$X_t = Y_t - Y_{t-1}$$

Since Y_t has stationary increments, it is clear that X_t is (strictly) stationary which implies that $X_{t+h} \overset{d}{=} X_t \; \forall h$. Furthermore, the process of increments X_t has zero mean, $\mathbb{E}(X_t^2) = \mathbb{E}(Y_1^2)$ and its autocovariance function is obtained from equation (3.2.1) and given by :

$$\gamma_X(h) = \frac{\mathbb{E}(Y_1^2)}{2}\big\{(h+1)^{2H} - 2h^{2H} + (h-1)^{2H}\big\} \sim \mathbb{E}(Y_1^2)\, H(2H-1)h^{2H-2}, \text{ as } h \to \infty \tag{3.3.4}$$

The approximation for $h \to \infty$ is obtained using a Taylor development of second order for x^{2H}. Equation (3.3.4) shows that if $1/2 < H < 1$, then the process X_t is long-range dependent in the sense of condition 1 and in the sense of condition 2 with $d = H - 1/2 \in (0, 1/2)$.

Finally, since a fBm $(B_t^H)_{t\geq 0}$ is a H-SSSI process (from equation (3.2.1)), we have that the process of increments $X_t = B_t^H - B_{t-1}^H$ is long-range dependent for $1/2 < H < 1$. In this case, the process X_t is called *fractional Gaussian noise*. If $0 < H < 1/2$, X_t loses the long-range property and is then short-range dependent. We therefore have a one to one correspondence in the case of fBm between sample path regularity and long-memory through the Hurst exponent H. Note that the fact that fBm can be rewritten as an integral with respect to Brownian motion (equation (3.2.2)) highlights the memory of the process.

3.4 Fractional Ornstein-Uhlenbeck processes

We now expand in this part the well-known Ornstein-Uhlenbeck process using the previous sections about fractional Brownian motions and their properties. This so-called *fractional Ornstein-Uhlenbeck* (fOU) process will be used in the next chapter to build volatility models based on fBm. Cheridito et al. (2003) show that the Langevin equation with fractional Gaussian noise has a stationary solution (precisely the fOU process) from which we now analyze the properties.
Let $\lambda, \nu \in \mathbb{R}^+$ and $\xi \in L^0(\Omega)$ a random variable. The Langevin equation with initial condition $X_0 = \xi$,

$$X_t = \xi - \lambda \int_0^t X_s\, ds + N_t \quad \text{or} \quad dX_t = -\lambda X_t\, dt + dN_t, \quad t \geq 0$$

can be simply solved pathwise for several general noise processes $(N_t)_{t\geq 0}$ and particularly for fBm $(B_t^H)_{t\geq 0}$, with dB_t^H being a fractional Gaussian noise as said above. Specifically, it can be shown that for each $H \in (0, 1)$ and $a \in (-\infty, \infty)$, the integral $\int_a^t e^{\lambda u} dB_u^H$ exists as a pathwise Riemann-Stieltjes integral, which is almost surely continuous in t, and that,

$$X_t^{H,\xi} := e^{-\lambda t}\Big(\xi + \nu \int_0^t e^{\lambda u} dB_u^H\Big), \quad t \geq 0$$

is the unique almost surely continuous process that solves the Langevin equation given by :

$$X_t = \xi - \lambda \int_0^t X_s ds + \nu B_t^H$$

In particular, with initial condition $\xi = X_0^H = \nu \int_{-\infty}^0 e^{\lambda u} dB_u^H$, we find the following following solution:

$$X_t^H := \nu \int_{-\infty}^t e^{-\lambda(t-u)} dB_u^H, \quad t \in \mathbb{R} \tag{3.4.1}$$

It is clear that the process $(X_t^H)_{t\geq 0}$ is a Gaussian process, and it follows immediately from the stationarity of the increments of fractional Brownian motion that this process is stationary.

Furthermore, Cheridito et al. (2003) show that for $H \in (0, 1/2) \cup (1/2, 1)$, the decay of $\mathrm{Cov}(X_t^H, X_{t+s}^H)$ when $s \to \infty$ is very similar to the decay of $\mathrm{Cov}(B_{h+t}^H - B_h^H, B_{h+t+s}^H - B_{h+s}^H)$ when $s \to \infty$. Then using the property of long-range dependence for fractional Gaussian noise with $H \in (1/2, 1)$ (see Section 3.3), we have that the process $(X_t^H)_{t\geq 0}$ exhibits long-range memory for $H \in (1/2, 1)$ in the sense of condition 1 and 2. This regime corresponds to a persistent process where consecutive increments of the fBm are positively correlated. The relatively stronger positive correlation for the consecutive increments of the associated fBm process with higher H values gives a relatively smoother process $(X_t^H)_{t\geq 0}$. For the case $0 < H < 1/2$, the fOU process loses the long-range property. In this regime consecutive increments of the fBm process are negatively correlated. The enhanced negative correlation with smaller H gives a relatively rougher process with a more irregular behavior.

Chapter 4

Fractional Brownian motion volatility models

In this chapter, we will introduce more advanced volatility models based on fractional Brownian motions to address the shortcomings of the classical SV models reviewed in Chapter 2. We will mainly focus on models where the logarithm of the volatility is assumed to follow a fractional Ornstein–Uhlenbeck process, which ensures that the instantaneous volatility process is stationary and reverses to the mean.

Seminal analyses of Ding et al. (1993), Andersen and Bollerslev (1997) and Andersen et al. (2001) have claimed that a stylized fact of the volatility process $(\sigma_t)_{t\geq 0}$ is to enjoy the long memory/LRD property. The presence of this LRD property is used in these analyses to justify the volatility persistence (i.e. the strong dependence at long time scales) that we can observe on Figure 2.3.2 and on the autocorrelogram 7.1.2 (Appendix). More precisely, the authors above argue that the phenomenon behind the empirical findings of slowly decaying sample ACF and high persistence of volatility is the stationary LRD property of the volatility process. We will develop this idea in the following section thanks to the FSV model of Comte and Renault (1998), which enjoys this LRD property. However, we will also show in the RFSV model of Gatheral et al. (2014) that we can also explain these empirical findings of volatility persistence with models that are not long-range dependent in the sense that their autocorrelation function does not decay as a power-law. For these models, we will provide explicit expressions enabling us to analyze thoroughly the dependence structure of the volatility process. We will therefore consider in this chapter models that enjoy the LRD property as well as models that do not. Indeed, the famous mathematician Rama Cont (2007) even quotes : *"The econometric debate on the short range or long range nature of dependence in volatility still goes on (and may probably never be resolved)"*.

4.1 Fractional stochastic volatility model (FSV)

Comte and Renault (1998) built this long-memory model (called FSV model) with the following SDE under the real-world measure $\mathbb{P}$:

$$dS_t = \mu(S_t, t)\, S_t\, dt + \sigma_t\, S_t\, dW_t \tag{4.1.1}$$

$$d\log(\sigma_t) = \lambda(\eta - \log(\sigma_t))dt + \nu\, dB_t^H \tag{4.1.2}$$

We clearly see that equation (4.1.2) is a fOU process as described above. In order for $\log(\sigma_t)$ to have the LRD property, the fBm B_t^H is built with $1/2 < H < 1$. Furthermore, Comte and Renault consider a Hull & White framework where B_t^H is independent from W_t. For notation purposes, we will denote from now on the process $(\log \sigma_t)_{t\geq 0}$ by $(X_t)_{t\geq 0}$. Therefore, equation (4.1.2) becomes :

$$dX_t = \lambda(\eta - X_t)dt + \nu\, dB_t^H \tag{4.1.3}$$

We can then use path-wise integration to find the solution of this fOU process. According to equation (3.4.1), we have that X_t^H is the stationary solution of this equation and the volatility process satisfies:

$$\begin{aligned} X_t^H = \log \sigma_t &= \eta + \nu \int_{-\infty}^{t} e^{-\lambda(t-u)} dB_u^H \\ \Leftrightarrow \exp(X_t^H) = \sigma_t &= \exp\left(\eta + \nu \int_{-\infty}^{t} e^{-\lambda(t-u)} dB_u^H\right) \end{aligned} \tag{4.1.4}$$

Comte and Renault (1998) also show that the volatility process itself (and not only its logarithm) is asymptotically stationary and that it entails long-memory properties for $H \in (1/2, 1)$. More precisely,

they show that $\text{Cov}[\sigma_t, \sigma_{t+h}]$ is of order $O(|h|^{2d-1})$ when $h \to \infty$ with $d = H - 1/2$, in accordance with condition 2 (3.3.3).

We now present the justification developed by Comte and Renault (1998) for using a fBm with $H > 1/2$ in the model (4.1.3). First, recall that under the assumption of independence between W_t and $\hat{W}_t$, we showed in Section 2.3.1 (Hull and White) that the conditional distribution of $\log(S_T/S_0)$ knowing $A_t = \int_0^t \sigma_u^2 \, du$ is $N(rT - \frac{A_T}{2}, A_T)$ under $\mathbb{P}^*$. If we define $B_{t,T} = \frac{1}{T-t} \int_t^T \sigma_u^2 \, du$, the conditional distribution becomes $N\left(T(r - \frac{B_{0,T}}{2}), TB_{0,T}\right)$. With this form similar to the one of equation (2.1.3), we see that the constant σ from the B&S model is replaced with $B_{t,T}$.
Such standard SV models are then able to reproduce only partially the following empirical stylized fact regarding implied volatilities. Studies have shown that the amplitude of the smile (or slope of the skew) is a decreasing function of the time to maturity. Indeed, on the volatility surface of the S&P500 in Figure 2.1.2, we can clearly see that the smile effect is very pronounced for short maturities but it almost completely disappears for longer maturities. When using the general Hull and White framework above, a simple application of the law of large numbers to the volatility process (assumed to be stationary) leads to the vanishing of the volatility's randomness and a decrease of the smile/skew phenomenon when time to maturity T increases. Indeed, we have $B_{t,T} \to \mathbb{E}[\sigma^2]$ for T large.
However, the decrease of the smile amplitude (or skew slope) when time to maturity increases appears to be much slower than implied by standard SV models. The occurrence of fairly pronounced smiles and skew slopes even for rather long maturity options is due to volatility persistence. Long-memory mean reverting volatility processes with fractional Brownian motions such as in (4.1.3) allow to better capture this effect of volatility persistence than standard SV models.

We now turn to the RFSV model introduced by Gatheral et al. (2014) and explain the advantages of such models compared to standard SV and FSV models.

4.2 Rough fractional stochastic volatility model (RFSV)

Gatheral et al. (2014) build the so-called RFSV model for the log-volatility process with the same SDE as in the FSV of Renault and Comte (1998), based on the fOU process :

$$dX_t = \lambda(\eta - X_t)dt + \nu \, dB_t^H \tag{4.2.1}$$

However, the Hurst exponent H of the fBm B_t^H is now in $(0, 1/2)$ in order to introduce a short-range dependence for $\log \sigma_t$. This explains why this model is called *Rough fractional stochastic volatility*. We now present the two main reasons cited by Gatheral et al. that justify the modeling of $\log \sigma_t$ using a fBm with $H < 1/2$, one based on historical data and the other on risk-neutral data. The graphs of this section are taken from the paper of Gatheral et al. (2014), for the S&P500 from January 2000 to March 2013. We will reconstruct these plots ourselves for other indices in the practical application of Chapter 5.

- **Historical Data :**

We will now show how Gatheral et al. derive the RFSV model and show that it is consistent with the empirical statistical properties of the volatility time series under the physical measure, especially when estimating the smoothness of the volatility process.
Let's first pretend we observe the discrete values on $[0, T]$ of the volatility process : σ_0, $\sigma_\Delta, ..., \sigma_{k\Delta}, ...$ with $\Delta = T/N$ the time step and $k \in \{0, T/\Delta\}$. Then for $q > 0$, Gatheral et al. define :

$$m(q, \Delta) = \frac{1}{N} \sum_{k=1}^{N} |\log(\sigma_{k\Delta}) - \log(\sigma_{(k-1)\Delta})|^q \tag{4.2.2}$$

Assuming the increments of the log-volatility process are stationary and that the law of large numbers can be applied with $N \to \infty$ (and $\Delta \to 0$), $m(q, \Delta)$ is the estimator of :

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

You can find on Figure 4.2.1 a plot of $\log(m(q,\Delta))$ against $\log(\Delta)$ for the S&P500. Gatheral et al. find empirically for each given q a linear relation between $\log(m(q,\Delta))$ and $\log(\Delta)$. We have then the following scaling property with K_q a constant :

$$\mathbb{E}[\,|\log(\sigma_\Delta) - \log(\sigma_0)\,|^q\,] = K_q\,\Delta^{\zeta_q} \tag{4.2.3}$$

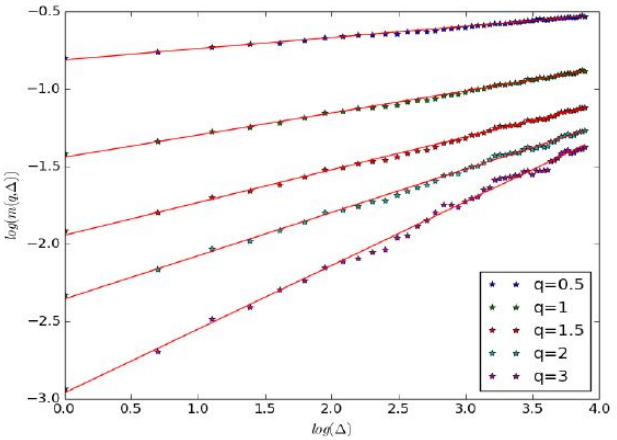

Figure 4.2.1: $\log m(q,\Delta)$ as a function of $\log\Delta$.

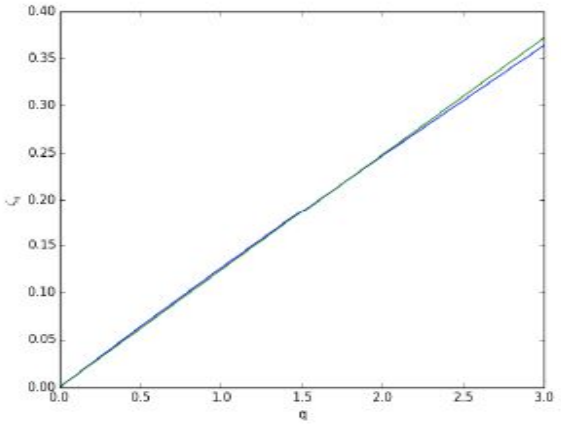

Figure 4.2.2: ζ_q as a function of q.

Moreover, they again find a linear relation on Figure (4.2.2) when plotting ζ_q against q which leads to the following relation : $\zeta_q = c\,q$, with c a constant. Then, if we recall the equation (3.2.4) for fractional Brownian motions, we see that both the processes $(\log\sigma_t)_{t\geq 0}$ and $(B_t^H)_{t\geq 0}$ enjoy this scaling property with $c = H$, $q = 1/\gamma$ and $K_q = \mathbb{E}(\,|B_1^H|^{1/\gamma})$.

From this scaling property, combined with the stylized fact highlighted by Andersen et al. (2001) that the distribution of increments of log-volatility $(\log\sigma_{t+\Delta} - \log\sigma_t)_{t\geq 0}$ is close to Gaussian, it seems logical to let the process $(\log\sigma_t)_{t\geq 0}$ be a fBm with Hurst exponent H. The Gaussianity of log-volatilty increments will be confirmed in Chapter 5, Figure 5.1.3. H can then really be seen as the smoothness parameter of the log-volatility process. When estimating this parameter, Gatheral et al. (2014) often find low values of H (≈ 0.1) which contradicts the FSV of Comte and Renault. These findings are confirmed by Bennedsen et al. (2016) with a study on two thousands US equities. In their article, they also propose alternative methods to estimate the roughness parameter H based on OLS and NLLS (ordinary and non-linear least squares) regressions.

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

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

However, it appears that $\sigma_t = \beta\exp(\nu B_t^H)$ is not stationary (where β is another constant). This leads us to impose stationarity by modeling the log-volatility as a fOU process with $0 < H < 1/2$:

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

We hence retrieve equation (4.2.1) with $(X_t)_{t\geq 0} = (\log\sigma_t)_{t\geq 0}$ and thus confirm the empirical validity of the RFSV model with historical data. Therefore, as in the FSV model, we have that X_t^H is the stationary solution of this equation and that the volatility process satisfies :

$$X_t^H = \log\sigma_t = \eta + \nu\int_{-\infty}^t e^{-\lambda(t-u)}dB_u^H \;\Leftrightarrow\; \exp(X_t^H) = \sigma_t = \exp\left(\eta + \nu\int_{-\infty}^t e^{-\lambda(t-u)}dB_u^H\right)$$

The volatility process $(\sigma_t)_{t\geq 0}$ itself in the RFSV is now stationary, like in the FSV. However, this time, when $H < 1/2$, the volatility process does not decay like a power function and the property of long

memory is lost.

We can now verify whether this stationary log-volatility process $(X_t^H)_{t\geq 0}$ enjoys the properties of a fBm since we said above that both the empirical discrete log-volatility process and fractional Brownian motions exhibit the same scaling property (4.2.3). Gatheral et al. show, for $\lambda \ll 1/T$ and $0 < H < 1$, that the log-volatility process $(X_t^H)_{t\geq 0}$ behaves locally (at time scales smaller than T) as a fBm. Indeed, they prove as $\lambda \to 0$:

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

and

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

Hence, when λ is small, we have that our RFSV model (4.2.1) approximately reproduces the scaling property of the fBm (and of the discrete spot log-volatility process) and is still locally Hölder continuous of exponent β for all $\beta < H$. Note that a small λ implies a slow speed of mean reversion and thus a very long reversion time scale.

Finally, if we analyze more closely the autocorrelation function of the volatility process itself, we show in Demo. 4 that when λ is small :

$$\mathbb{E}[\sigma_{t+\Delta}\sigma_t] = \mathbb{E}[e^{X_t^H + X_{t+\Delta}^H}] \approx e^{2\mathbb{E}[X_t^H] + 2\mathrm{Var}[X_t^H]} e^{-\nu^2 \frac{\Delta^{2H}}{2}} \tag{4.2.4}$$

It follows that in the RFSV model, $\log(\mathbb{E}[\sigma_{t+\Delta}\sigma_t])$ is linear in Δ^{2H}. This property is very well satisfied on data, as shown by Figure 4.2.3, where the logarithm of the empirical counterpart of $\mathbb{E}[\sigma_{t+\Delta}\sigma_t]$ against Δ^{2H} is plotted for the S&P500 with $H = 0.14$.

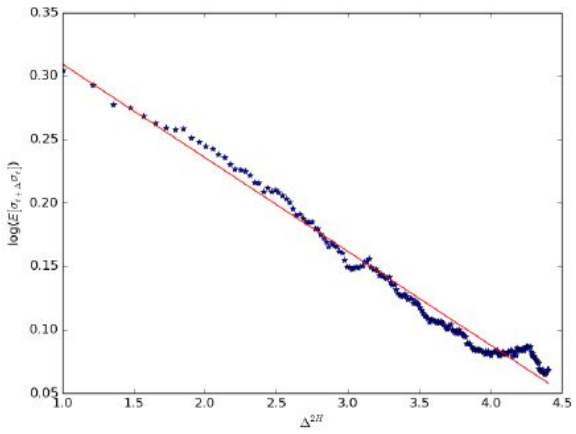

Figure 4.2.3: Empirical counterpart of $\log(\mathbb{E}[\sigma_{t+\Delta}\sigma_t])$ against $\boldsymbol{\Delta^{2H}}$.

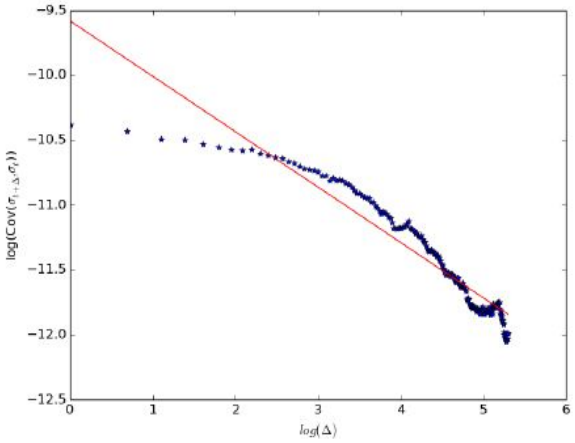

Figure 4.2.4: Empirical counterpart of $\log(\mathrm{cov}[\sigma_{t+\Delta}\sigma_t])$ against $\boldsymbol{\log(\Delta)}$.

In particular, a corollary of the equation (4.2.4) is that the autocovariance function of the volatility does not decay as a power-law as widely believed (see introduction of the chapter). Indeed, on Figure 4.2.4, Gatheral et al. show that a log-log plot of the autocovariance function does not yield a straight line. Therefore, the historical data are in accordance with the RFSV and the absence of long-range dependence of the volatility process (see Condition 2, Section 3.3).

- **Risk-neutral data :**

Moreover, the RFSV model is not only consistent with empirical statistical properties of the volatility time series but also with the observed volatility surface under $\mathbb{P}^*$ (and is hence consistent with risk-neutral data).

First, as said previously in Chapter 2, standard time-homogenous models of volatility such as the Hull and White and Heston models do not provide a good fit of the volatility surface (especially for short expirations). Particularly, these models generate a term structure of at-the-money (ATM) skew[1] under a risk-neutral measure $\mathbb{P}^*$ that is constant for small time to maturity τ and behaves as a sum of decaying exponentials for larger τ, which is not confirmed empirically. Indeed, empirical studies show that this observed term structure of ATM skew is well approximated by power-law functions of τ, as pictured in Figure 4.2.5, which can be generated by SV models where the (log-)volatility is driven by fractional Brownian motions with values of the Hurst exponent in $(0, 1/2)$ (unlike Comte and Renault). This has been studied more closely by Fukasawa (2011) who has built a model where the volatility process is driven by a fBm and which is indeed consistent to the power-law with a term structure of ATM volatility skews of the form $\psi(\tau) \sim \tau^{H-1/2}$ with $0 < H < 1/2$. We will not focus on this model but this is interesting in and of itself in that it provides a counterexample to the widespread belief (mentioned in Section 2.3.3 on Jump-diffusions) that the explosion of the volatility skew as $\tau \to 0$ necessarily implies the presence of jumps. We can indeed model this behavior using fBm instead.

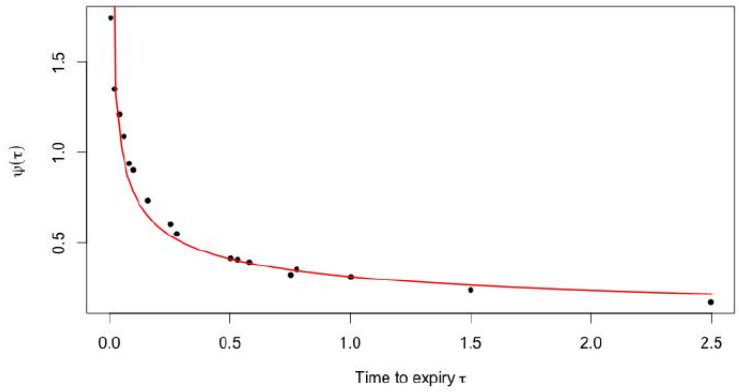

Figure 4.2.5: The black dots are estimates of the S&P ATM volatility skews as of June 20, 2013; the red curve is the power-law fit $\psi(\tau) = A\,\tau^{-0.4}$

4.2.1 FSV versus RFSV

We now compare the FSV and RFSV described above. The FSV model built by Comte and Renault is a fractional Ornstein-Uhlenbeck process with $H \in (1/2, 1)$ in order to ensure the long-memory property of the log-volatility process and this way to capture the so-called volatility persistence. However, this model generates a term structure of volatility skews $\psi(\tau)$ that is increasing with τ (at least for small values of τ), which is inconsistent with the observed term structure. Comte and Renault are then forced to take $\lambda \gg 1/T$ to generate a decreasing term structure of the skew for lager values of τ and hence, the log-volatility process does not behave as a fBm. Furthermore, this huge level of volatility persistence λ would not be consistent with empirical evidence for volatility smiles/skews of very short-term options. Finally, if we consider the quantity $m(2, \Delta)$ for the FSV model, we see on Figure 4.2.6 below that this model is not consistent with S&P volatility data for which we observe the scaling property $\log m(2, \Delta) \approx \zeta_2 \log \Delta + k$ for a constant k (that is, a particular case of (4.2.3) with $p = 2$).

The RFSV is a particular case of the FSV model but with $H < 1/2$ and $\lambda \ll 1/T$. This fOU process has this time very long reversion time scale since λ is small (λ being the parameter that governs the speed of mean-reversion). This choice of H allows to reproduce both the observed smoothness of the volatility process and generate a term structure of volatility skew in agreement with the observed one (exploding term structure of ATM skews when τ goes to zero). Moreover, taking λ very small implies that the dynamics of our process is close to that of a fBm as shown by Gatheral et al. (2014). For time scales of practical interest (T large), the authors consider that we may even proceed as if λ were exactly zero. This simplification will be used in the rBergomi and in the rough Heston models below.

[1] Recall that the term structure of ATM volatility skew is defined as $\psi(\tau) := |\frac{\partial}{\partial K}\hat{\sigma}_t^{imp}(K, \tau)|_{K=S_t}$.

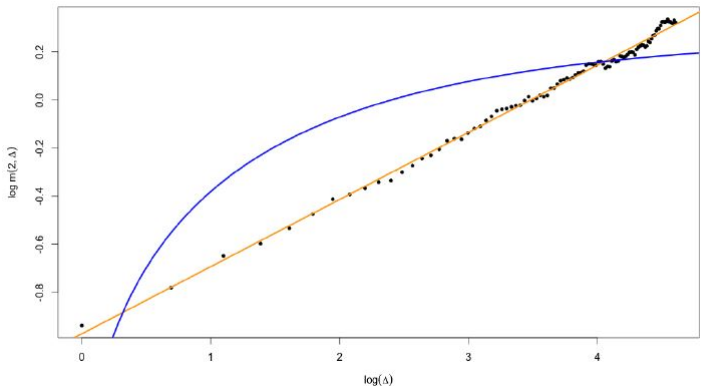

Figure 4.2.6: Incompatibility of long memory models such as the FSV model. Black points are empirical estimates of $m(2, \Delta)$ for S&P500; the blue line is the FSV model with $\lambda = 0.5$ and $H = 0.53$; the orange line is the RFSV model with $\lambda = 0$ and $H = 0.14$.

We also said that the data tend to not exhibit long-range dependence which is in accordance with the RFSV model but not the FSV model. We now want to know if the RFSV can still capture the effect of persistence of the observed volatility time series (and hence the slower decrease of the smile amplitude as highlighted by Comte and Renault - Section 4.1) despite of the short memory property of the RFSV. If we use standard procedures (see Gatheral et al. (2014)) to detect long memory in the data or in the volatility generated by the RFSV model, one can be wrongly convinced that the volatility time series exhibits long memory. Indeed, Gatheral et al. show that these estimation procedures identify spurious long memory of volatility in the RFSV model. This is due to the fact that the (log-)volatility process in the RFSV exhibits some persistence even if does not present long memory in the sense that its autocorrelation does not decay as a power-law. Therefore, if we come back to the introduction of the chapter, we can now explain why the long memory of volatility is widely accepted as a stylized fact by several authors even if the volatility in the data (and generated by the RFSV) is not long-range dependent.

Although empirically and theoretically grounded, the RFSV does not allow to price rapidly option prices and therefore to calibrate effectively a pricing model. Indeed, we need Monte-Carlo simulations which are fairly slow, especially for calibration. Therefore, we now show how to adapt the RFSV to obtain more tractable models that can be used to price claims on both the underlying and the integrated volatility. We first analyze in more details a simple case of the RFSV model when $\lambda = 0$, the rBergomi model, based on the forward variance curve. Secondly, we derive an extension of the classical Heston model based on a rough fractional stochastic volatility model. This rough Heston model has the nice property of having a quasi-closed form characteristic function of the log-stock price.
It is also important to mention that we will not implement nor calibrate the FSV model in the practical part (Chapter 5) of this thesis. Indeed, no efficient calibration method exists for this model. The paper of Chronopoulou and Viens (2012) presents a method for estimating its parameters and for pricing under the FSV but this method is far from being tractable and is quite complex to implement. Furthermore, we just showed that this model is not capable of reproducing the statistical and risk-neutral properties of the observed spot volatility.

4.2.2 rBergomi model

First, let denote $v_u = \sigma_u^2$ the instantaneous variance at time u and the forward variance curve (also called variance swap curve) $\xi_t(u) = \mathbb{E}[v_u|\mathcal{F}_t],\ u \geq t$ with $(\mathcal{F}_t)_{t\geq 0}$ the filtration generated by a Brownian motion W_t. We can easily recover v_t from $\xi_t(u)$ as $v_t = \lim_{u\to t} \xi_t(u)$. Following Bergomi and Guyon (2012), forward variance models are models that can be written (under the real measure $\mathbb{P}$) as :

$$dS_t = \mu(S_t, t) + S_t \sqrt{v_t} \left(\rho \, dW_t + \sqrt{1-\rho^2} \, dW'_t \right)$$
$$d\xi_t(u) = \eta_t(u\,; \omega) dW_t$$

where W_t, W_t' are independent one-dimensional Brownian motions defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and $\eta_t(u\,;\omega)$ is a measurable stochastic process for all $u > 0$. These authors proved that all conventional Markovian (see Concept 7.2.5) stochastic volatility models, such as all the SV models of Section 2.3, may be written as forward variance models. Apart from them, one well-known example of a stochastic volatility model expressed in forward variance curve form is the $n-$factor Bergomi model :

$$\xi_t(u) = \xi_0(u)\,\mathcal{E}\left(\sum_{i=1}^{n} \nu_i \int_0^t e^{-\lambda_i(u-s)} dW_s^{(i)}\right) \tag{4.2.5}$$

where $\mathcal{E}(\Phi) = \exp\big(\Phi - 1/2\,\mathbb{E}[|\Phi|^2]\big)$ (called Wick stochastic exponential) and Φ is a zero-mean Gaussian process. We see that the entire forward variance curve is determined by n-factors, each of the OU-type with correlated Brownian motions $W_t^{(i)}$. However, the term structure of the volatility skew $\psi(\tau)$ is not consistent with the observations. Hence, it would be tempting to replace the exponential kernels in (4.2.5) with a power-law kernel. This would give a model of the form :

$$\xi_t(u) = \xi_0(u)\,\mathcal{E}\left(\nu \int_0^t \frac{dW_s}{(u-s)^\gamma}\right) \tag{4.2.6}$$

Note that $\int_0^t \frac{dW_s}{(u-s)^\gamma}$ is known as the *Volterra* fBm with parameter $\gamma = 1/2 - H$ and has, similar to classical fBm, $(H-\varepsilon)$-Hölder sample paths, for any $\varepsilon > 0$. Fukasawa (2011) shows that such models indeed generate a term structure of the volatility skew of the form $\psi(\tau) = 1/\tau^\gamma$ in accordance with the observed one. Finally, using a discrete Laplace transform, it can easily be shown that $(t-s)^{-\gamma} \approx \sum_{i=1}^n \alpha_i\, e^{-\lambda_i(t-s)}$ for some coefficients α_i. Therefore, we retrieve the general n-factor Bergomi model but this time with all Brownian motions perfectly correlated. Hence, this model (4.2.6) is also a special case of a OU process.

Following Bayer et al. (2016), we now generalize this Bergomi model (4.2.6) by introducing a fBm. This model is called the rBergomi model and it is based on the simplifying assumption $\lambda = 0$ of the RFSV model. Indeed, we start from the RFSV model with $\lambda = 0$ which gives us the simple model already introduced before :

$$\log \sigma_{t+\Delta t} - \log \sigma_t = \nu\,(B^H_{t+\Delta t} - B^H_t)$$

Using the Mandelbrot-Van Ness representation (3.2.3) for B_t^H, we have under the physical measure $\mathbb{P}$:

$$\begin{aligned}
\log v_u - \log v_t &= 2\,\nu\,C_H\Big\{\int_{-\infty}^t \frac{dW_s}{(t-s)^\gamma} - \int_{-\infty}^0 \frac{dW_s}{(-s)^\gamma}\Big\} \\
&= 2\,\nu\,C_H\Big\{\int_t^u \frac{1}{(u-s)^\gamma} dW_s + \int_{-\infty}^t \Big[\frac{1}{(u-s)^\gamma} - \frac{1}{(t-s)^\gamma}\Big] dW_s\Big\} \\
&=: 2\,\nu C_H\{M_t(u) + Z_t(u)\}.
\end{aligned}$$

with $\gamma = 1/2 - H$. It is shown in Demo. 5 that if we let $\tilde{W}_t(u) := \sqrt{2H}\int_t^u \frac{dW_s}{(u-s)^\gamma}$ and $\eta := 2\,\nu C_H/\sqrt{2H}$, we have :

$$\mathbb{E}^{\mathbb{P}}[v_u|\mathcal{F}_t] = v_t \exp\Big\{2\,\nu\,C_H\,Z_t(u) + \frac{1}{2}\eta^2\,\mathbb{E}^{\mathbb{P}}[|\tilde{W}_t(u)|^2]\Big\}$$

Under an equivalent measure $\mathbb{P}^*$ where the change of measure is deterministic and given by $W_t^* = W_t + \lambda_t\,t$, we have that W_t^* is martingale under $\mathbb{P}^*$. From that, we finally have for the rBergomi model under $\mathbb{P}^*$ (again see Demo. 5) :

$$v_u = \mathbb{E}^{\mathbb{P}}[v_u|\mathcal{F}_t]\,\mathcal{E}\left(\eta\,\tilde{W}_t^*(u)\right) \exp\Big\{\eta\,\sqrt{2H}\int_t^u \frac{\lambda_s}{(u-s)^\gamma} ds\Big\} = \mathbb{E}^{\mathbb{P}^*}[v_u|\mathcal{F}_t]\,\mathcal{E}\left(\eta\,\tilde{W}_t^*(u)\right) \tag{4.2.7}$$

$$= \mathbb{E}^{\mathbb{P}^*}[v_u|\mathcal{F}_t]\,\exp\Big\{\eta\sqrt{2H}\int_t^u \frac{1}{(u-s)^\gamma} dW_s^* - \frac{\eta^2}{2}(u-t)^{2H}\Big\} \tag{4.2.8}$$

where $\mathcal{E}(\Phi) = \exp\big(\Phi - 1/2\,\mathbb{E}^{\mathbb{P}^*}[|\Phi|^2]\big)$ and Φ is a zero-mean Gaussian process under $\mathbb{P}^*$. $\mathbb{E}^{\mathbb{P}^*}[v_u|\mathcal{F}_t]$ is the forward variance curve $\xi_t(u)$ observed on the market and is the product of two terms : $\mathbb{E}^{\mathbb{P}}[v_u|\mathcal{F}_t]$

which depends on the history of the driving Brownian motion, and a term which depends on the deterministic price of risk λ_t. It is important to note the presence of the Volterra fBm in equation (4.2.8) with $\gamma = 1/2 - H > 0$ that induces the rough behavior of the variance process v_u.

This model is non-Markovian, contrarily to the SV models of Section 2.3. Indeed, $\mathbb{E}^{\mathbb{P}^*}[v_u|\mathcal{F}_t] \neq \mathbb{E}^{\mathbb{P}^*}[v_u|v_t]$. However, given the state vector $\xi_t(u)$, which can in principle be computed from observed option prices, the dynamics of the model are well-determined. Finally, the pricing model to be simulated under $\mathbb{P}^*$ is :

$$\begin{aligned} S_T &= S_0 \exp\left(rT - \frac{1}{2}\int_0^T v_u\, du + \int_0^T \sqrt{v_u}\, dW_u^{*,S}\right) = S_0\, e^{rT}\, \mathcal{E}\left(\int_0^T \sqrt{v_u}\, dW_u^{*,S}\right) \\ v_t &= \xi_0(t)\, \mathcal{E}\left(\eta\, \tilde{W}_t^*\right) = \mathbb{E}^{\mathbb{P}^*}[v_t]\, \mathcal{E}\left(\eta\, \tilde{W}_t^*\right) \end{aligned} \tag{4.2.9}$$

where $W_t^{*,S} = \rho\, W_t^* + \sqrt{1-\rho^2}\, W_t^{'*}$ and $W_t^{'*}$ a Brownian motion under $\mathbb{P}^*$, independent from W_t^*. We can hence model the leverage effect in the rBergomi model by anticorrelating the Brownian motion that drives the volatility process with the Brownian motion driving the price process such that $\mathrm{Corr}(W_t^*, W_t^{*,S}) = \rho t$ with $\rho < 0$.

For the simulation and the calibration of this model, we refer to Section 5.4.1. As the RFSV, the rBergomi model generates a realistic term structure of ATM volatility skews due to the power-law kernel in $\tilde{W}_t(u)$, highly similar to the form of (4.2.6). Moreover, Bayer et al. (2016) show that the rBergomi model is surprisingly consistent with the observed volatility surface. However, as shown in the Section 4.2, the RFSV model with $\lambda = 0$ is not stationary and hence neither is the rBergomi model. It would therefore not exhibit realistic long-term behavior of volatility. Finally, in the case studied in the article, the authors argue from the shape of VIX implied volatility smiles that the change of measure cannot be deterministic. A more thorough investigation of the market implied change of measure $d\mathbb{P}^*/d\mathbb{P}$ is needed in future work.

We could not doubt consider other models where the **log-volatility** behaves as a fBm at reasonable time scales. One of them was proposed by Mendes and Oliveira (2008). The log-volatility is this time stationary (as the RFSV) and is modeled as a fractional Gaussian noise, with fractional Gaussian noise being the increment process of a fBm for a certain increment length (see section 3.3), δ is the observation time scale ($\delta = 1$ for daily data), ν a volatility of volatility parameter and $0 < H < 1/2$:

$$\log \sigma_t = \beta + \frac{\nu}{\delta}\{B_t^H - B_{t-\delta}^H\}$$

or equivalently,

$$\sigma_t = \exp\left(\beta + \frac{\nu}{\delta}\int_0^t \mathbb{1}_{s>t-\delta}\, dB_s^H\right)$$

We note that this model looks very similar to the RFSV (4.2.1). Indeed, the indicator function kernel $\mathbb{1}$ acts like the exponential kernel in the RFSV to force stationarity for long times. However, this model is mathematically more complex and requires for example Malliavin Calculus for the derivation of option prices. Hence, the choice of the fOU process with $0 < H < 1/2$ is probably the simplest way to accommodate this rough behavior together with the stationarity property.

Finally, we derive a second pricing model from the RFSV based on an extension of the classical Heston model. This rough Heston model has the nice property of having a closed form characteristic function of the log-stock price which allows the fast-pricing of vanilla call options, when using the Carr-Madan formula described in Appendix in Concept 7.2.3. However, this time, the log-volatility is not explicitly modeled as a fBm (or a fOU process with $\lambda \to 0$).

4.2.3 Rough Heston model

Recall from Section 2.3.2 that the Heston model under a risk-neutral measure $\mathbb{P}^*$ is given by[2] :

$$dS_t = S_t\, r\, dt + S_t\, \sqrt{v_t}\, dW_t^*$$
$$dv_t = \lambda(\eta - v_t)\, dt + \nu\, \sqrt{v_t}\, d\hat{W}_t^*$$

Again, we have that $\langle W^*, \hat{W}^* \rangle_t = \rho\, t$. We can rewrite the Heston model model in terms of forward variance curve (or variance swap curve) as said above (see Demo. 6) :

$$d\xi_t(u) = \nu e^{-\lambda(u-t)} \sqrt{v_t}\, d\hat{W}_t^* \tag{4.2.10}$$

Furthermore, from the Mandelbrot-van Ness representation of the fBm (equation (3.2.2)), we clearly see that the kernel $(u-s)^{H-1/2}$ plays a central role in the rough dynamic of the fBm for $H < 1/2$. Indeed, as said above, one can show that the Volterra fBm $\int_0^t (u-s)^{H-1/2} d\hat{W}_u^*$ has Hölder regularity $H - \varepsilon$ for any $\epsilon > 0$. Therefore, in order to allow for a rough behavior of the volatility in a Heston-type model, El Euch and Rosenbaum (2019) naturally introduce the Volterra kernel $(u-s)^{H-1/2}$ in the Heston SV model as follows :

$$v_u = \theta^t(u) - \frac{1}{\Gamma(H+1/2)} \int_t^u (u-s)^{H-1/2} \lambda\, v_s\, ds + \frac{1}{\Gamma(H+1/2)} \int_t^u (u-s)^{H-1/2} \nu \sqrt{v_s}\, d\hat{W}_s^* \tag{4.2.11}$$

where $\theta^t(.)$ is assumed to be continuous, $\mathcal{F}_t$-measurable and represents a time-dependent mean reversion level. When $H = 1/2$, we can verify that we indeed recover the classical Heston model above. It can also be shown that the trajectories of the volatility itself are almost surely Hölder regular/continuous of order $H - \varepsilon$, for any $\varepsilon > 0$.

As for the rBergomi, we consider the limiting case $\lambda \to 0$ for the rough Heston. Indeed, it is shown in El Euch et al. (2019) that $\lambda\theta^t(.)$ can be directly inferred from the forward variance curve observed at time t. By doing so, it allows us to have $\xi_t(u) = \mathbb{E}^{\mathbb{P}^*}[v_u|\mathcal{F}_t] = \theta^t(u)$ with $\lambda \to 0$ and then to write (see again Demo. 6) the rough Heston in the forward variance form as :

$$d\xi_t(u) = \frac{\nu}{\Gamma(H+1/2)} (u-t)^{H-1/2} \sqrt{v_t}\, d\hat{W}_t^*$$

The link with the Heston model in forward variance form (4.2.10) can be easily made by setting $\lambda = 0$ and by adding the Volterra kernel. Since the forward variance curve at time $t = 0$ is observed (or at least can be retrieved) from the market, we only have three parameters left to estimate : H, ν and ρ. Moreover, we note that the limit $u \to t$ of the rough Heston makes no sense. This reflect the fact that this model (as the rBergomi) is not Markovian with respect to the current variance state v_t. However, the rough Heston is Markovian in the forward variance curve $\xi_t(u)$, which means that future states of the process v_t depends only upon the present state of $\xi_t(u)$ (see Concept 7.2.5 of Markov property). Indeed, equation (4.2.11) can be rewritten with $\lambda = 0$:

$$v_u = \xi_t(u) + \frac{1}{\Gamma(H+1/2)} \int_t^u (u-s)^{H-1/2} \nu \sqrt{v_s} d\hat{W}_s^* \tag{4.2.12}$$

From Gatheral and Keller-Ressel (2019), we have that the forward variance model of Heston and rough Heston are said to be affine. In such cases, they show that the characteristic function of the log-asset price $X_T = \log S_T$ at time t can be written as :

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

where $g(t,u)$ is called the Riccati equation. In the classical Heston model, we have with $a = iu$ that :

$$g(t, iu) = \partial_t\, h(t,a) + \lambda\, h(t,a)$$

[2]The notations of the Heston model in risk-neutral form are here adapted to match the ones of the RFSV.

where $\partial_t h(t,a) = -\frac{1}{2}a(a+i) - (\lambda - i\rho\,\nu\,a)h(t,a) + \frac{1}{2}\nu^2\,h(t,a)^2$. It is possible to solve this equation and to prove that the characteristic function is the same as the one presented in Section 2.3.2 on the Heston model.

In the rough Heston model with $\lambda = 0$ (and again $a = iu$), we have that

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

with D^α the fractional derivative as defined in Concept 7.2.8. This equation is exactly the same as in the classical Heston model (with zero-mean reversion) but with the time derivative replaced by a fractional derivative that makes the model rough.

To solve the rough Heston Riccati equation and find the solution $h(t,a)$, we use the method described in Gatheral & Radoičić (2019) based on the combination of a 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, simple and accurate to compute approximate solutions of fractional ordinary differential equation. One other available method is the Adams scheme, see El Euch et al. (2019) for further details. If we apply the approximate scheme introduced by Gatheral & Radoičić, we can therefore obtain a characteristic function of the log-asset price under the rough Heston in quasi-closed form and then price efficiently European options using the Carr-Madan formula (7.2.1) combined with a Fast Fourier Transform algorithm. We refer to Section 5.4.1 for the simulation of the rough Heston model.

It is important to note that in the rough Heston model, we no longer model the increments of log-volatility $(\log\sigma_{t+\Delta} - \log\sigma_t)$ as normally distributed, which is a stylized fact observed on the financial markets (see Section 5.1, Figure 5.1.3). Instead, in the rough Heston, the variance in difference $v_{t+\Delta} - v_t$ exhibits a mixture of distributions. However, after having simulated the variance process under the rough Heston model, we have that $(\log\sigma_{t+\Delta} - \log\sigma_t)$ is empirically close to a Gaussian distribution for various sets of parameters H, ρ and ν (for example, see Q-Q plot 7.1.3 in Appendix with $H = 0.12$, $\nu = 0.3$ and $\rho = -0.8$).

However, all these models based on fractional Brownian motions are still considered by several authors as limited models, as in Bennedsen et al. (2016). Indeed, when designing a realistic model of volatility that allows for both rough sample paths and flexible long-term properties, it is important to be aware of the principle, pointed out by Gneiting and Schlather (2004), that self-similar processes, such as the fBm, have either long memory or rough sample paths. Even if we showed that models which are not long range dependent such as the RFSV can take into account some volatility persistence, long-memory models are an easy way to capture these strongly persistent dynamics and strong autocorrelations of the volatility at long time scales. We hence introduce in the next section more general models that encompass some models based on fBm such as the rBergomi and that conveniently decouple the behavior of volatility at short and long time scales; in particular, they accommodate both roughness and long-memory.

4.3 Brownian semistationary processes (BSS)

The authors in Bennedsen et al. (2016) seek stationary stochastic Gaussian processes with roughness index H such that $0 < H < 1$ and a long-term memory structure that is independent of the value of H. Given a process $(X_t)_{t\geq 0}$ with these properties, we define our model for volatility as :

$$\sigma_t = \beta\exp(X_t) \tag{4.3.1}$$

The Brownian semistationary process (BSS) is able to capture all of the properties above and is introduced in Barndorff-Nielsen and Schmiegel (2009) under an equivalent martingale measure $\mathbb{P}^*$ as :

$$X_t = \int_{-\infty}^{t} g(t-s)\,v_s\,d\hat{W}_s^* \tag{4.3.2}$$

where $\hat{W}_t^*$ is Brownian motion under $\mathbb{P}^*$, $g : (0, +\infty) \to \mathbb{R}$ is a square integrable kernel function and v_t is an adapted, stationary volatility of volatility process. Here, we fix $v_t = v \ \forall t$ and v a constant in such a way to make X_t Gaussian. This way, the increments $(\log \sigma_{t+\Delta} - \log \sigma_t)$ are again Gaussian as in the RFSV, which is a stylized fact in equity data (Andersen et al. (2001)).

Bennedsen et al. (2016) introduce additional assumptions concerning the properties of the kernel function g, which enables us to derive some theoretical results for X_t. We now let $\alpha = H - 1/2$ in this section for the roughness parameter and the three following assumptions :

***A*(1)** For some $\alpha \in (-1/2, 1/2)\backslash\{0\}$, $x \in (0, 1]$ and L_0 a slowly varying function at zero[3] :

$$g(x) = x^{\alpha} L_0(x)$$

***A*(2)** The function g is continuously differentiable with derivative g' that is ultimately monotonic and satisfies $\int_1^{\infty} g'(x)^2 \, dx < \infty$

***A*(3)** For L_1 a slowly varying function at infinity, $\lambda \geq 0$ and $\gamma \in \mathbb{R}$ such that $\gamma > 1/2$ when $\lambda = 0$:

$$g(x) = e^{-\lambda x} x^{-\gamma} L_1(x)$$

- **Short-memory properties :**

Under these assumptions, Bennedsen et al. (2016) show that the BSS process[4] X_t has α as roughness index controlling the memory of the process for short time horizons h in the sense that :

$$1 - \rho_X(h) := 1 - \text{cov}(\log \sigma_t, \log \sigma_{t+h}) = 1 - \text{cov}(X_t, X_{t+h}) \sim c \, L_0(|h|)^2 \, |h|^{2\alpha+1}, \quad |h| \to 0$$

for c a constant and $\alpha \in (-1/2, 1/2)$. Gatheral et al. (2014) find the same structure for the ACF of $(\log \sigma_t)$ in the RFSV model when the level of volatility persistence λ tends to zero (as shown in equation (7.3.1) in Demo. 4). Therefore, by choosing $\alpha \in (-1/2, 0)$ (or $H \in (0, 1/2)$), we obtain processes of log-volatility with short-memory properties (and rough behavior) for small horizons h. Note that the condition 2 (3.3.3) for LRD is indeed not verified.

- **Long-memory properties :**

Secondly, the tail behavior of the kernel function g at infinity, as specified by (A3), controls the long-term memory properties of X_t. There are two cases to consider :

1. The first case $\lambda = 0$:

(i) If $\gamma \in (1, \infty)$: $\rho_X(h) \sim L_1(|h|) \, |h|^{-\gamma}, \quad h \to \infty$

(ii) If $\gamma \in (1/2, 1)$: $\rho_X(h) \sim L_1(|h|) \, |h|^{1-2\gamma}, \quad h \to \infty$

Then, we let the rate of polynomial decay be denoted by β where $\beta = \gamma$ if $\gamma \in (1, \infty)$ and $\beta = 2\gamma - 1$ if $\gamma \in (1/2, 1)$. From condition 2 of LRD (3.3.3) with $\gamma = 1 - d$, we have that X_t is long-range dependent if $\gamma \in (1/2, 1)$ for long horizons h.

2. The second case $\lambda > 0$:
The assumption (A3) with the second case $\lambda > 0$ allows for models where autocorrelations decay to zero exponentially fast, leading to processes with short-range dependence (i.e. short memory) for long horizons h. Indeed, when $|h| \to \infty$: $\rho_X(h) \sim L_1(|h|) e^{-\lambda |h|} |h|^{-\gamma}$.

[3] See definition 3.3.1 and replace $u \to \infty$ by $u \to 0$.
[4] All the following results are also true for the volatility process itself σ_t if we let $v_t = v$.

We now introduce two kernel functions g that satisfy (A1), (A2), and (A3). The first one is **the power-law kernel** :

$$g(x) = x^{\alpha}(1+x)^{-\gamma-\alpha}, \quad x > 0, \quad \alpha \in (-1/2, 1/2), \quad \gamma \in (1/2, \infty), \quad \lambda = 0$$

With this kernel function, the BSS process X has a roughness index α and asymptotic memory properties controlled by γ as in the first case $\lambda = 0$. By choosing $\alpha \in (-1/2, 0)$ and $\gamma \in (1/2, 1)$, this power-law kernel hence generates processes with short-memory property for small h and with long-range memory for long horizons h. The second kernel is **the Gamma kernel** :

$$g(x) = x^{\alpha}\, e^{-\lambda x}, \quad x > 0, \quad \alpha \in (-1/2, 1/2), \quad \lambda \in (0, \infty)$$

With this kernel function, the process X_t has also a roughness index α but asymptotic memory properties controlled by λ, as in the second case $\lambda > 0$. If $\alpha \in (-1/2, 0)$, the Gamma kernel thus generates processes with short-range memory both for short and long horizons h.
We refer to Bennedsen et al. (2016) for the estimation and simulation of the model:

$$\begin{aligned} S_t &= S_0 \exp\left(rt - \frac{1}{2}\int_0^t \sigma_s^2\, ds + \int_0^t \sigma_s\, dW_s^* \right) \\ \sigma_t &= \beta \exp(X_t) \end{aligned}$$

In Figure 7.1.4 (Appendix), you can find the ACF of the estimated BSS models (Power and Gamma kernels) versus the empirical ACF of log-volatility data given in Bennedsen et al. (2016). We observe that the models fit the empirical ACF's very well and that there is no difference in goodness of fit between the two kernels. Indeed, even if the Gamma-BSS process has an asymptotic short memory when $h \to \infty$, by selecting very small values of λ (which is the case when estimating this parameter), it is possible to specify processes with a very high degree of persistence and a slow decay of the ACF, mimicking long memory on finite time intervals such as the Power-BSS with $\gamma \in (1/2, 1)$. This indicates that we are not able to confirm that log-volatility has the long memory property by inspecting the fit of the ACF, as it has been sometimes suggested in the literature (and as mentioned in the introduction of this chapter). It echoes with the explanation of Gatheral et al. (2014) of "spurious" long memory property (see Section 4.2.1). However, in terms of forecasting performance, the authors show that there is a gain from using a model with strict long memory, as opposed to a highly persistent model with (technically) short memory (as the RFSV). Indeed, the BSS models outperform the RFSV model in terms of volatility forecasting when the step size $\Delta = T/n$ is small (15 to 130 minutes) but it is the opposite when $\Delta \geq 1$ day as in the case studied in Gatheral et al (2014).

To conclude, this model is able to disentangle the short-term and long-term properties of the volatility process (for example by using a power-law kernel). Indeed, with just two parameters controlling the short- and long-term behavior, Bennedsen et al. find a model that outperforms standard volatility models (and the RFSV) at intraday time scales ($\Delta < 1$). However, we will not implement this model in the practical part of this thesis. Indeed, as the RFSV, such general BSS models are tractable for volatility forecasting but not for option pricing. As a matter of fact, the rBergomi model introduced above is also a special case of BSS processes with a power-law kernel and is more suitable and tractable for option pricing. We then decide to focus on this simpler and more tractable rBergomi model for calibration purposes. Note that we will use the paper *Hybrid scheme for Brownian semistationary processes* of Bennedsen et al. (2017) for simulating and calibrating the rBergomi. Nevertheless, it would be interesting in future work to implement these more general BSS models and compare them with models strictly based on fBm in terms of option pricing and implied volatility surface fit.

Chapter 5

Practical applications

In this practical application, we first confirm that rough fractional volatility models are highly consistent with the statistical properties of the observed spot volatility time series for other indices than the S&P500. In a second step, we show that these models are also highly consistent with risk-neutral data. Indeed, we calibrate and compare four models (Heston, Bates, rBergomi and rough Heston) in terms of fit to the observed European option prices and implied volatility surface for the CAC 40. We finally use these calibrated models to price a life insurance contract embedding exotic options in it.

5.1 Estimation of the realized spot volatility process

In order to perform statistical investigations of volatility, one requires data of volatility on a time grid. In Gatheral et al. (2014) and for this thesis, we will consider that the mesh of this grid is one day. We seek to extract the unobserved realized spot volatility process $(\sigma_t)_{t\geq 0}$ from high-frequency observations of the asset price S_t. Indeed, σ_t is not directly observable and we need to construct a proxy for it. In particular, we are interested in the integrated variance which is defined for some large $n \in \mathbb{N}$:

$$IV_t^{\Delta} := \int_{t-\Delta}^{t} \sigma_s^2 \, ds \,, \quad t = \Delta, 2\Delta, ..., n\Delta$$

As juste mentioned, we will consider a Δ equal to 1 day. The realized variance RV_t appears to be a natural estimator of the integrated variance. Let Δ^* be the interval in minutes over which the returns of an index (or a stock) are computed. The Δ^*-minute log-return is the log-difference of two consecutive index values over Δ^*-minutes :

$$r_{t+i\Delta^*, t+(i-1)\Delta^*} := p_{t+i\Delta^*} - p_{t+(i-1)\Delta^*}$$

where $p_t = \log(S_t)$ denotes the logarithm of an index price at t and where Δ^* will be equal to 5 minutes throughout this thesis. The 5-minute RV is calculated by the sum of squared 5-minute log-returns within a day. We then define the daily realized volatility at day t as :

$$RV_t^{5-min} := \sum_{i=1}^{m} r^2_{t+i\Delta^*, t+(i-1)\Delta^*}$$

where $m = 390/5 = 78$ is the number of 5-minute returns per day (between 9:30 a.m. and 4:00 p.m.). It is shown in Andersen et Bollerslev (2001) and in Barndoff-Nielsen and Shephard (2002) that if prices are observed without noise, then as $\min_{\Delta^*} |t + i\Delta^*, t + (i-1)\Delta^*| \to 0$ (or equivalently as $m \to \infty$), the 5-min RV estimator consistently estimates the integrated variance IV_t^{Δ} of the price process on the t-th day. Since Δ^* is rather small (5 minutes), we can consider that the 5-min RV is a good approximation of the integrated variance. Finally, since Δ is sufficiently small, the spot volatility process in daily terms (i.e. not annualized) $(\sigma_t)_{t\geq 0}$ is estimated using :

$$\hat{\sigma}_t^2 = \Delta^{-1} \, \widehat{IV}_t^{\Delta} = \Delta^{-1} \, RV_t^{5-min}$$

We will focus for this thesis on the CAC 40 index and therefore use the 5-minute realized variance of the Oxford-Man Institute of Quantitative Finance Realized Library for this index. This proxy is close to the RV_t^{5-min} described above. However, since in practice market microstructure noise is not negligible, our estimator RV_t^{5-min} is biased and further statistical and econometric methods have to be applied to our return data in order to find a robust estimator to this market noise. More details on these methods

can be found on `https://realized.oxford-man.ox.ac.uk/documentation/econometric-methods`. Finally, the data can be downloaded on `http://realized.oxford-man.ox.ac.uk/data/download` and will be used throughout the thesis. Note that one has to take the square root of these estimators to obtain an estimate of the daily spot volatility.

5.1.1 Statistical properties of the spot volatility

We decide to focus first our analysis on the spot volatility of the CAC 40 index from January 2004 to July 2020 in order to see whether our models are consistent with periods of high volatility (e.g. 2008 crisis, Covid-19 outbreak) as well as with periods of low volatility. In this section, we will confirm that the properties described in Gatheral et al. (2014) also hold for our time series of interest. The following plot depicts the behavior of our daily spot volatility proxy for this period.

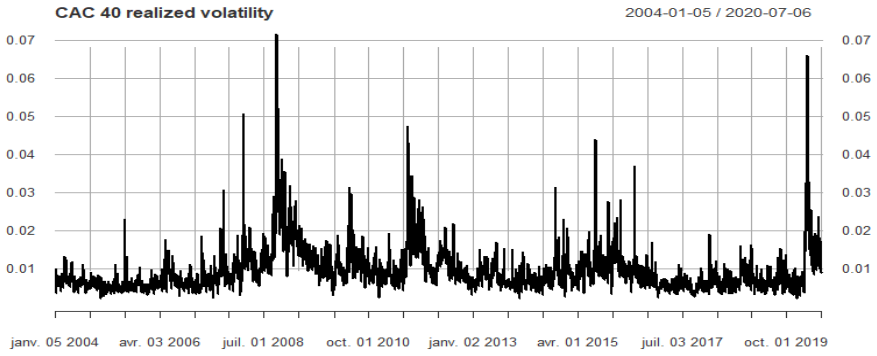

Figure 5.1.1: Daily realized volatility for the CAC 40 based on the RV^{5-min} estimate.

We can clearly see the phenomenon of volatility clustering, with persistent periods of high volatility that alternate with low volatility periods as mentioned in Section 2.3 on SV models. We have thus that large price changes are followed by large price changes (of either sign) and small changes by small changes. This implies a serial dependence and that the volatility process is predictable, although noisy. In particular, the autocorrelation function of the realized volatility is positive and slowly decays. This shows the strong persistence of the volatility process. On the contrary, if we plot the autocorrelogram of the CAC 40 log-returns, we will not find significant autocorrelations between them. Therefore, what matters is the amplitude of the log-return and not the direction/sign of it.

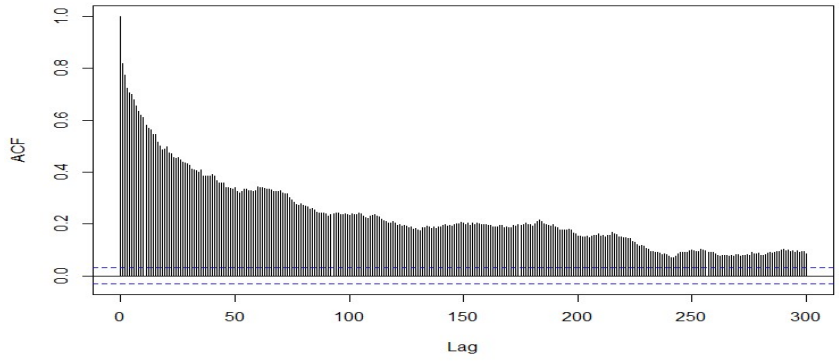

Figure 5.1.2: Autocorrelogram of the daily realized volatility time series.

Applying the augmented Dickey-Fuller (ADF) test for stationary on realized volatility, we find a p-value of $1.1e^{-10}$ which rejects the presence of a unit root and we then go for the alternative hypothesis of stationarity of the volatility process. The ADF test on the log-volatility process gives the same result. Moreover, on closer inspection of the empirical volatility series, we also observe that the sample path of the volatility on a restricted time window (e.g. one month) seems to exhibit the same kind of qualitative properties as those of the global sample path on the whole time horizon (same shape,

same volatility clustering, same mean-reversion, ...). This so-called fractal-type behavior is reproduced by rough models of volatility (small H and λ). This feature of volatility has been investigated both empirically and theoretically in Bacry and Muzy (2003), for example.

We now verify that the log-volatility increments are approximately normally distributed for different values of lag Δ (1, 5, 25 and 50 days) in order to be able to model the log-volatility by a fOU process. We see in Figure 5.1.3 that the distributions of differences in the log-realized volatility are

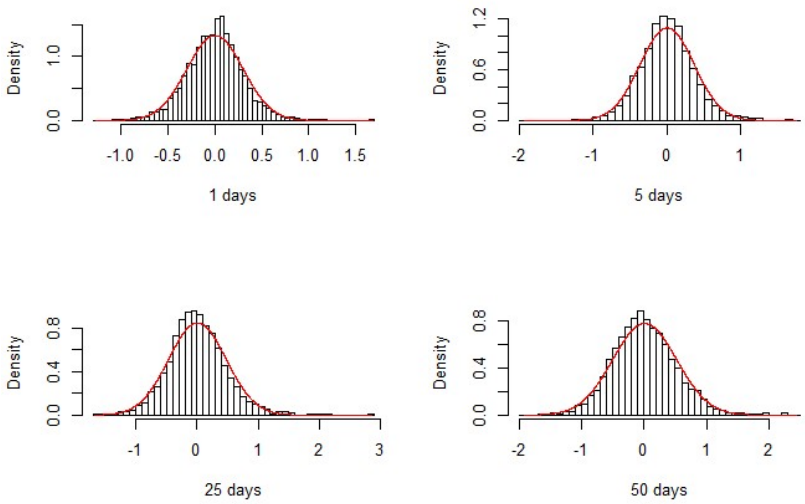

Figure 5.1.3: Histograms of log-volatility increments with normal distribution superimposed in red.

indeed close to Gaussian. However, the fit is not perfect and you can see on the Q-Q plots 7.1.5 (Appendix) that these distributions exhibit fat-tails and are therefore not perfectly Gaussian. Even if we consider other time windows without crises and other indices than the CAC 40, the presence of fat-tails in the histograms and in the Q-Q plots of log-volatility increments for different lags Δ is systematic. Moreover, it is shown in Bennedsen et al. (2017) and in Barndorff-Nielsen (1997) that the normal-inverse Gaussian (a distribution with semi-heavy tails) provides a better fit of the log-volatility data. This can thus be considered as a slight flaw of models where log-volatility is modeled as a fBm.

We now try to estimate the smoothness of the volatility process for the CAC 40 and see if it enjoys the same scaling property as shown for the S&P in the paper of Gatheral et al (2014) (i.e. the same behavior as in Figure 4.2.1). On the plot 5.1.4, we clearly see the same linear relations between $\log m(q, \Delta)$ and $\log \Delta$ for each given q. We also note that the higher q, the more noise there is around the straight line. Moreover, plotting ζ_q against q for the CAC 40, we again find a linear relation $\zeta_q = c\, q$ with $c = H$, the smoothness parameter :

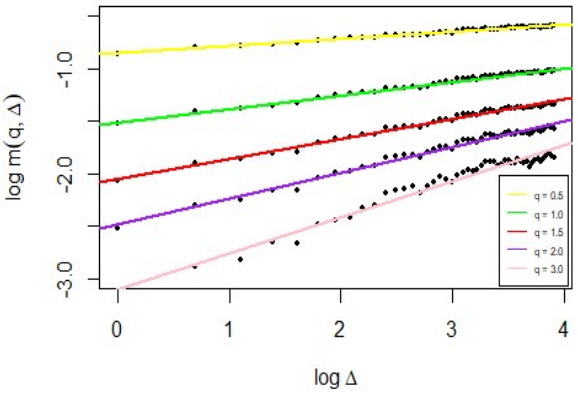

Figure 5.1.4: $\log m(q, \Delta)$ as a function of $\log \Delta$.

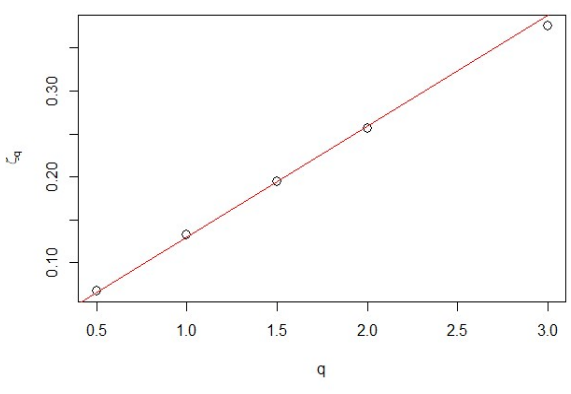

Figure 5.1.5: Linear regression of ζ_q against q

We can then simply estimate H as the slope of the linear regression of ζ_q against q with intercept set to 0. We obtain $\hat{\mathbf{H}} = \mathbf{0.12923}$. We can also easily estimate ν^2 as the exponential of the intercept in the linear regression of $\log(m(2,\Delta))$ on $\log(\Delta)$. Indeed, from (4.2.1) we have :

$$m(2,\Delta) = \Delta^{2H}\,\nu^2\,\mathbb{E}[\,|B_1^H|^2]$$

Then, from the basic properties of fBm, we have that $\mathbb{E}[\,|B_t^H|^2] = t^{2H}$ and hence $\mathbb{E}[\,|B_1^H|^2] = 1$. Finally, it comes :

$$\log(m(2,\Delta)) = 2H\log(\Delta) + \log(\nu^2)$$

The estimated $\hat{\nu}$ for the CAC 40 spot volatility process is $\mathbf{0.30977}$.

Finally, we confirm in Figure 5.1.6 that $\log\left(\mathbb{E}[\sigma_{t+\Delta}\sigma_t]\right)$ is a linear function of $\mathbf{\Delta^{2H}}$ (in accordance with the RFSV model) and in Figure 5.1.7 that the autocovariance function of the volatility does not decay as a power-law as widely believed. Indeed, the log-log plot of the autocovariance function does not yield a straight line, which confirms the absence of long-range dependence of the realized spot volatility process for the CAC 40.

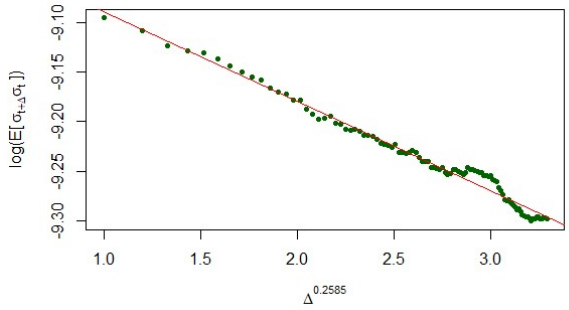

Figure 5.1.6: Empirical counterpart of $\log\left(\mathbb{E}[\sigma_{t+\Delta}\sigma_t]\right)$ against $\mathbf{\Delta^{2H}}$.

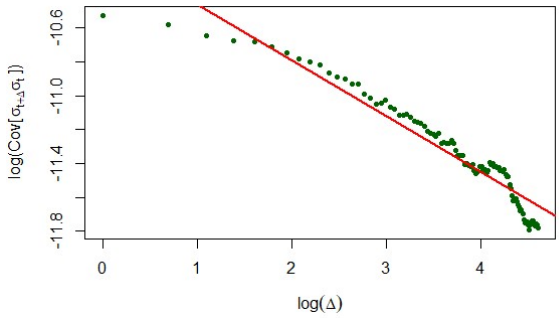

Figure 5.1.7: Empirical counterpart of $\log\left(\mathrm{cov}[\sigma_{t+\Delta}\sigma_t]\right)$ against $\mathbf{\log(\Delta)}$.

It is clear that only the RFSV (including the rBergomi and rough Heston) as well as some of the BSS models can reproduce all the statistical properties we have just shown for the CAC 40 spot volatility process. Indeed, the SV models from Section 2.3 are Markovian processes (with respect to v_t) and do not allow to take the memory property of the log-volatility into account and cannot reproduce the smoothness and scaling properties observed for the estimated spot volatility. The FSV with $H > 1/2$ cannot reproduce the smoothness of the CAC 40 volatility process either and do not yield a straight line for the regression of $\log\left(\mathbb{E}[\sigma_{t+\Delta}\sigma_t]\right)$ against Δ^{2H}.

5.1.2 A universal property ?

We now compare the scaling property and estimates of H and ν for different indices and different time windows. We obtain the following Table 5.1.1 :

$\hat{H} \mid \hat{\nu}$	**01/2004 - 07/2020**		**01/2010 - 01/2020**		**01/2018 - 07/2020**	
BEL20	0.1295	0.2866	0.1143	0.2964	0.1696	0.3070
CAC40	0.1292	0.3098	0.1132	0.3190	0.1657	0.3374
FTSE100	0.1106	0.3514	0.0984	0.3604	0.1391	0.3738
NI225	0.1271	0.3346	0.1068	0.3641	0.1566	0.3493
S&P500	0.1477	0.3589	0.1339	0.3896	0.2298	0.3984

Table 5.1.1: Estimates of H and ν for various indices and time windows.

We clearly find evidence of rough volatility for all the observed indices and time windows with a Hurst exponent H systematically lower than $1/2$. For all the indices, we observe higher estimate of H during

periods that include crises (subprime crisis (2008) and Covid-19 crisis). At first glance, it may be somehow counter-intuitive to have smoother sample paths of the spot volatility during crises. However, recall that there is a one-to-one correspondence between regularity and long-memory through this parameter H. Since crises have a strong influence on the volatility process even at long time scales, the volatility process tends to be more persistent after such crises. Indeed, after crashes, volatility tends to more slowly decrease to go back to previous levels. This can be seen on Figure 5.1.1 and on the VIX (Volatility Index) in Figure 7.1.6 (Appendix), where we see jumps of volatility and then a more gradual decline to pre-crash levels. Secondly, the level of the volatility of volatility parameter ν tend to be higher during the period 2018-2020 due to the Covid-19 outbreak. Indeed, we expect this parameter to be higher during periods of financial uncertainty. However, note that this is not necessarily confirmed for the period 2004-2020, including both crises.

Moreover, from further graphical analyses similar to Figure 5.1.4 (not plotted in this thesis), we observe slightly stronger linear relationships between $\log(m, q)$ and $\log(\Delta)$ and less noise for time windows without crises. Graphically, the scaling property is verified the best between 2010-2020. Finally, we can also add that the length of the window does not seem to influence the level of the estimate of H or ν, nor the quality of the linear fit for the scaling property.

We can conclude that roughness, strong persistence and scaling are universal features of financial market volatility, even during crises. Rough fractional stochastic volatility models are therefore extremely robust to different types of market conditions.

5.2 Calibration of pricing models

However, the statistical properties we have just reviewed tell us nothing about the ability of the different models to fit the implied volatility surface and to price vanilla and exotic options. This comparison will be developed in the coming sections, where we will first calibrate several models on the observed vanilla option prices and then compare these models in terms of volatility surface fit and exotic option pricing.

5.2.1 General method

We first explain the difference between parameter estimation (as in previous sections) and calibration:

- **Parameter estimation** is based on historical (real-word) data. We compute model parameters from past values of observed time series, such as the past asset prices or the historical volatility for example. However, as said above, past in no representation for the future and there is a huge uncertainty of *real-world* projections. Who can say where the asset value of the CAC40 will be over an horizon of 5 years ? What is its probability distribution up to this 5 years ?

- **Calibration** is based on risk-neutral data. We compute model parameters from market prices of financial derivatives and from the observed implied volatility matrix. Indeed, implied volatilities are reflective of the following characteristics of each participant : risk appetite, market sentiment, estimate of the historical volatility, known and unknown events, open interest of buyers versus sellers, etc (Leoni, 2014). Therefore, there is a lower uncertainty in risk-neutral projections. Moreover, the information is more available, accurate and shared in the risk-neutral world (at least for market makers).

We now review and explain more precisely the calibration method for the following models : Heston, Bates (SVJ), rBergomi and rough Heston models. The general idea for calibrating all these models is the same. From these pricing models, we can compute derivatives prices. We then want to find the model parameters that match derivative model prices as best as possible with the observed market prices of these derivatives. In financial application, the derivatives on which we calibrate the model are the available (vanilla) derivatives in the market, for which we have fast model pricers available (e.g.

Carr-Madan fromula). In the equity world, these are European calls and puts observed in the market.

It then boils down to an optimization problem where we have to find the minimum distance between model prices/model-based implied volatility and market prices/market implied volatility. For this thesis, we choose to minimize the distance between model-based implied volatility $\widehat{\sigma}_j^{imp}$ and market implied volatility σ_j^{imp} since it leads to more stable results. Finally, we introduce the weighted Root Mean Square Error as the loss function, which is defined as :

$$\text{RMSE} = \sqrt{\sum_{j=1}^{N} w_j \left(\sigma_j^{imp} - \widehat{\sigma}_j^{imp}\right)^2} \tag{5.2.1}$$

where $w_j = \frac{1}{\text{Ask}_j - \text{Bid}_j}$, i.e. the weight is the inverse of the bid-ask spread, expressed in terms of implied volatility. This way we give more importance to observations with small bid-ask spread. Indeed, if an option is liquid, then its bid-ask spread is supposed to be small and its price/implied volatility is more accurate. However, we can note that taking the proportional weight $w_j = \frac{1}{N}$ almost leads to the same results in our analysis. Mathematically, it boils down to find the optimal set $\Theta^* \in \mathbb{R}^p$ of the p model parameters such that :

$$\Theta^* = \arg\min_{\Theta} \mathcal{L}\left(\sigma^{imp}, \widehat{\sigma}^{imp}(\Theta)\right)$$

where $\mathcal{L}$ is the RMSE loss function. The minimization problem is solved using the *nlminb* function on *R*. Finally, we mention that the result of the optimization procedure can be highly dependent on the initial parameters. Good intuition for the starting values is hence of utmost importance. We then use a grid search method for picking the initials parameters that minimizes the loss function. The calibration process is summarized in Figure 5.2.1 below.

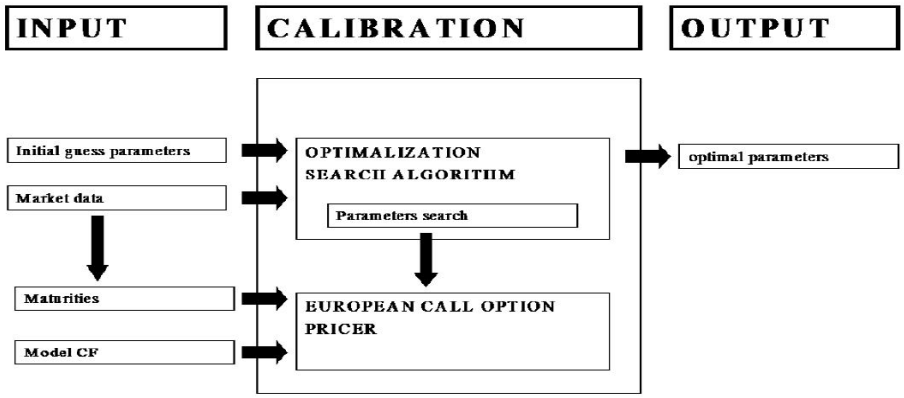

Figure 5.2.1: Calibration methodology.

5.2.2 Market data

We first choose as starting date $t = 0$, the 6^{th} of July 2020. The maturity of vanilla options as well as the risk-free term structure will be defined with respect to this date. The unit of time is the year and we consider that there are 252 trading days in a year (convention : trading days/total trading days).

5.2.2.1 Interest rates and dividend yield

The LIBOR curve is the most widely used proxy (with government bond yield) for the risk-free interest rates by derivatives dealers (see Hull and White (2013)). The LIBOR is the short-term borrowing rate of AA-rated financial institutions, i.e. the rate of short-term loans between banks. However, the use of LIBOR to value derivatives was called into question by the credit crisis in 2008. Banks became increasingly reluctant to lend to each other because of credit concerns. Therefore, LIBOR cannot be considered as perfectly risk-free. Indeed, most derivatives dealers now use interest rates based on

overnight indexed swap (OIS) rates rather than LIBOR (especially for collaterized transactions) but market data for such rates cannot be found easily and for free on the internet. Hence, we finally decide to use the EUR LIBOR (= EURIBOR) rates as an easy proxy for the risk-free rate.

Note that LIBOR rates are quoted in annual simple rates since the maturities are lower than or equal to one year. Since we price our derivatives using continuously compounded interest rates $r^c(T)$, we have to transform the quoted rates in the following way : denote the quoted rate $r(T)$, with T the corresponding maturity expressed in year (0.25, 0.5, etc.). We have :

$$r^c(T) = \frac{1}{T}\ln(1 + r\,T)$$

We finally obtain the following continuously compounded risk-free term structure :

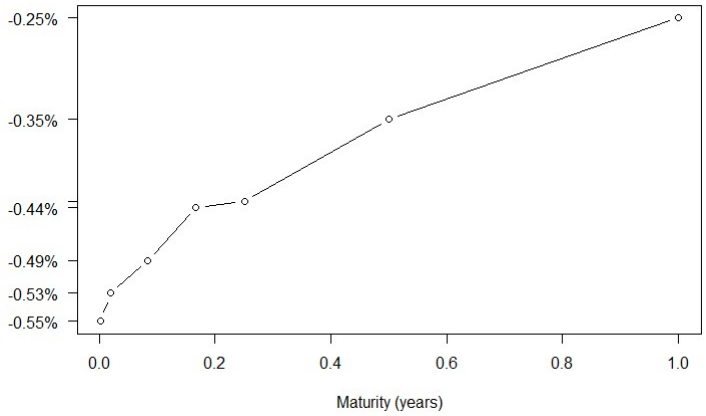

Figure 5.2.2: EURIBOR term structure on the 6^{th} of July, 2020.

Note that the risk-free rates obtained with this proxy are negative and small in absolute value (due to the current low-rate economic environment). Therefore, they will not impact significantly the price of the considered derivatives (at least for sufficiently short maturities).
However, the EURIBOR rates are only available up to a maturity of one year and we will need risk-free interest rates for maturities above one year when pricing options or insurance contracts with $T > 1$. Therefore, we calibrate a Neslon-Siegel model on the EURIBOR term structure in $t = 0$ to obtain risk-free rates for higher maturities. We are well aware that extrapolating rates in such a way for long maturities is not correct but the lack of data for such risk-free rates forces us to use this method instead. Hence, the rate at maturity T in the Nelson-Siegel model is given by :

$$r_0^{NS}(T) = \beta_0 + \beta_1\, w_1(T) + \beta_2\, w_2(T)\,,\ \text{with } w_1(T) = \left(\frac{1 - e^{-\frac{T}{\tau}}}{(\frac{T}{\tau})}\right) \text{ and } w_2(T) = \left(\frac{1 - e^{-\frac{T}{\tau}}}{\frac{T}{\tau}} - e^{\frac{T}{\tau}}\right)$$

If we calibrate the parameters of the NS model on the EURIBOR curve of Figure 5.2.2, we find $(\beta_0, \beta_1, \beta_2, \tau) = (-0.002270, -0.003136, -0.003127, 0.032787)$. With this model, we obtain the final plot of the continuously compounded interest rate term structure given in Figure 7.1.7 (Appendix).

Concerning the dividend yield, we suppose that it is equal to 1% (from Bloomberg forecasts). However, this assumption does not change anything for the conclusions of our calibration.

5.2.2.2 European option data

First, we have that the price S_0 at time 0 of the CAC40 index is equal to 5028.56 €. Secondly, as mentioned above, we need market prices of European options that are traded on exchange markets in order to calibrate later our model. We use Bloomberg terminals in order to gather a database that includes 366 option prices. We obtain from Bloomberg the option bid price and the option ask

price for different strikes and different maturities. The true market price is here considered as the mid price between the bid price and the ask price, which is a common assumption. Using the one-to-one relationship between market prices and implied volatility from B&S formula (2.1.5), we can build a database with the bid, ask and mid implied volatility for each strike and maturity. We then apply the following common filters (as in Moyaert and Petitjean (2011)) and select the following options :

- **Out-of-the-money options** : Since out-of-the-money options are more actively traded than in-the-money options (as a protection), the quotes on out-of-the-money options are usually more reliable. Moreover, from put-call parity, we have that the put and the call at the same maturity and strike have the same time value and thus the same implied volatility. For model estimation, it suffices (theoretically) to pick one of them (the out-of-the money one) as the two options contain identical information about the underlying stock price dynamics. Note that, in practice, the put-call parity does not always hold (transaction cost, slippage, ...).
- **The bid-ask spread** (on the option prices) is less than 5%. Otherwise, we consider that there is no sufficient liquidity to take the option into account.
- **Maturity** is less than one year. Indeed, options with an expiration date longer than 1 year are rejected since they are less sensitive to volatility (Moyaert and Petitjean (2011)).

We have a final database of 140 OTM options for which we can observe the strike, the maturity, the bid, ask and mid implied volatility. We want to insist on the fact that 140 observed option data is no a lot compared to calibrations done in most of scientific papers. These calibrations are usually done on the S&P500, for which there are more than 1000 different options available (see the calibration done in El Euch et al. (2019) for example). The lack of option data for the CAC 40 is especially true for larger maturities ($T > 0.7$).

We now turn to the specific calibration method for each of the model above. Even if the general method is the same, the way we price our European options will vary in function of the model. Therefore, the calibration process will not be exactly the same if we use a Fast Fourier Transform algorithm or an Hybrid scheme for example.

5.2.3 Calibration of Heston and Bates model

As explained in chapter 2 (Sections 2.3.2 and 2.3.3), we have at disposal for these two models a closed-form for their characteristic function. Then, using the Carr-Madan formula (7.2.1), we can compute prices of European options for a whole range of strikes in one run very fast. We then just have to iterate on the different maturities to find the prices of all our options. Next, it only remains to apply the general calibration method described above to find the optimal parameters of these models.

- **Heston :**

We find a minimal RMSE of 0.1192938 with the initial set of parameters (0.22, 0.5, 0.05, 0.2, -0.75) which leads to the following final parameter values Θ^*:

σ_0^*	κ^*	η^*	θ^*	ρ^*
0.29335	2.76902	0.09757	1.44711	-0.70011

We see that we have a high speed of mean reversion κ and a high level of the vol-of-vol parameter θ which implies large swings in the variance process. We also obtain a large negative correlation between the stock process and the variance process, which is consistent with the leverage effect. Finally, the mean reversion level towards which the variance will converge is 9.76% (which represents a volatility of 31,23%). It is no wonder to obtain this high level of volatility given the current state of the economy due to the COVID-19 outbreak (note that the level of the VIX index was at 29,76% on the 6^{th} of July). These effects can be seen in Figures 7.1.8 and 7.1.9 (Appendix) depicting one sample path of the index

price and one sample path of the variance process obtained via the simulation methodology which we will describe in Section 5.4.1. Moreover, the Heston model with the parameters obtained above leads to the Figure 5.2.3 below when compared to the market prices of European options on the CAC 40 index. We see that the quality of the fit is quite good even if for higher maturities, the model can

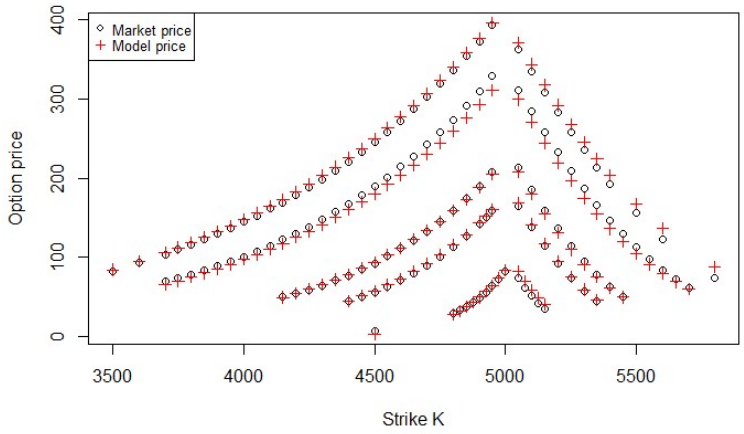

Figure 5.2.3: Heston option prices vs. market option prices.

provide prices that are substantially off market (especially for call options).

- **Bates :**

We find a minimal RMSE of 0.08626908 with the initial set of parameters (0.2, 0.5, 0.05, 0.2, -0.75, 0, 0.13, 0.13) which leads to the following final values Θ^*:

σ_0^*	κ^*	η^*	θ^*	ρ^*	μ_j^*	σ_j^*	λ^*
0.26237	0.4135164	0.1385266	1.1847156	-0.6498814	-0.11896	0.11063	0.6928121

Firstly, the RMSE is substantially lower than the one obtained with the Heston model, which means that the Bates model provides a better fit to the observed European option prices. The mean-reversion speed κ and the volatility of variance parameters θ are also lower which implies that the swings in the variance process are less pronounced than in the Heston model. Moreover, we now observe jumps in the price process that are on average negative since $\mathbb{E}[Y-1] = e^{\mu_j+\sigma_j^2/2} - 1 < 0$. The annual frequency of these jumps is equal to 69.28%. Finally, we observe that the long-run variance η is higher than in the Heston model. The following plot depicts the fit of the Bates options prices compared to the observed

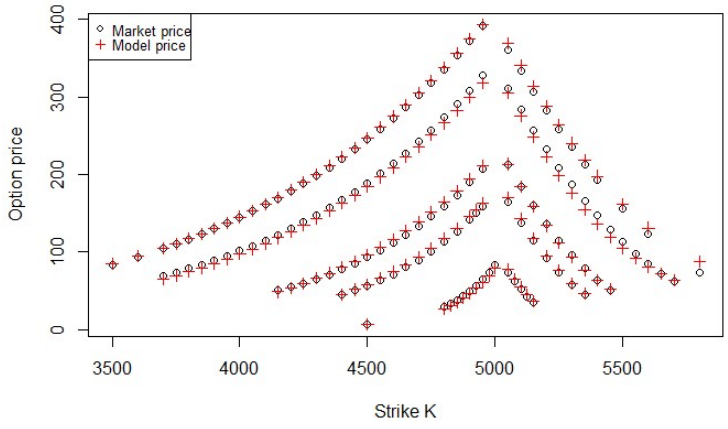

Figure 5.2.4: Bates option prices vs. market option prices.

prices. The fit is clearly better than in the Heston model as said above, even if for larger maturities it

is not perfect yet. One sample path of the price process and one sample path of the variance process under our calibrated Bates model are plotted in Figure 7.1.10 and 7.1.11 (Appendix). We can clearly see the presence of two down jumps.

Finally, if we test the sensitivity of the calibration of the Heston and Bates models to the Carr-Madan parameters $(\tilde{\eta}, N, \alpha)$, we see that the final parameters of the Heston and the Bates models are robust for different values of $\tilde{\eta}$, N and α. We hence retain the standard values given in the Carr-Madan's article (1999) since these parameters lead to the lowest RMSE :

$$\tilde{\eta} = 0.25 \quad N = 4096 \quad \alpha = 1.5$$

5.2.4 Calibration of the rough Heston model

As explained in Section 4.2.3 on the rough Heston model, we have a quasi-closed form for the characteristic function of the log-price X_t. We can then apply exactly the same calibration technique as for the Heston and Bates models described in the previous section. However, recall that the rough Heston is not Markovian in the current variance state v_t but Markovian in the (forward) variance swap curve $\xi_t(u)$. Hence, the rough Heston's characteristic function depends on this variance swap curve. In practice however, the variance swap curve $\xi_0(u)$ at time $t = 0$ is not easily retrieved from the financial markets. Indeed, it is hard to obtain high quality variance swap quote data (since variance swaps are OTC contracts) and in any case, the bid/ask spread is wide as explained in Bayer et al. (2016). We thus choose to proxy the value of a variance swap of maturity T by the value of a log-contract (also of maturity T) as explained for example in Chapter 11 of Gatheral (2011, pp. 135-137). We do not deepen this problematic in this thesis. The variance swap curve at time $t = 0$ is therefore calibrated on our database of 140 European options using the code provided by this author. We hence obtain the following variance swap curve $\xi_0(t)$ in Figure 5.2.5 (when interpolated between observed maturities).

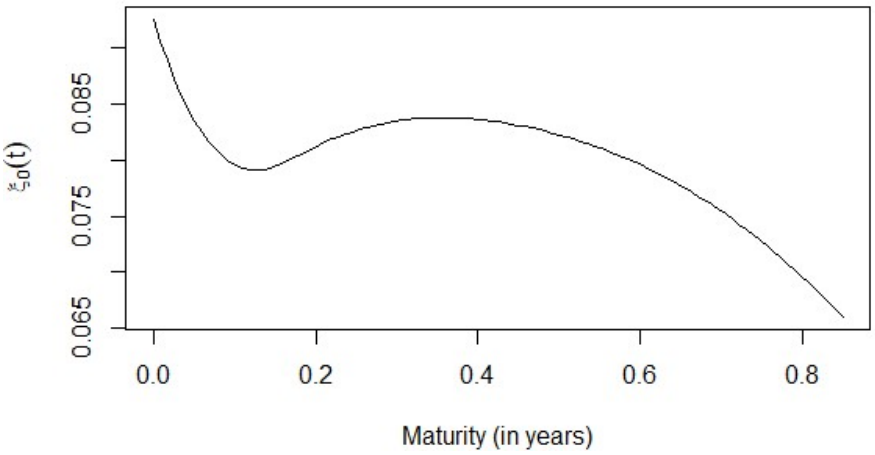

Figure 5.2.5: Variance swap curve $\xi_0(t)$.

Using again the Carr-Madan formula (7.2.1) augmented with a FFT algorithm, we find a RMSE of 0.092938 with the initial set of parameters (0.1, 0.1, -0.7) which leads to the following final values Θ^*:

H^*	ν^*	ρ^*
0.122261	0.4143393	-0.8024252

Based on the RMSE, we then have a slightly less good fit than with the Bates model. However, recall that the rough Heston has only three parameters to estimate compared to the eight parameters of the Bates model. We will also analyze in the next section which of the models best fits the behavior of the implied volatility surface. Moreover, it is important to note that the Hurst index H obtained via the risk-neutral calibration method is consistent with the Hurst index obtained via statistical estimation (under the real-world measure) of Section 5.1, which was equal to 0.1292. Yet, we have a higher ν than the one estimated based on the historical time series (= 0.3098) due to the market price of risk

included in the risk-neutral valuation. Finally, we observe a stronger leverage effect with this model than the previous ones. The following Figure 5.2.6 depicts the fit of the rough Heston options prices compared to the observed market prices of European options. The fit is clearly better than in the

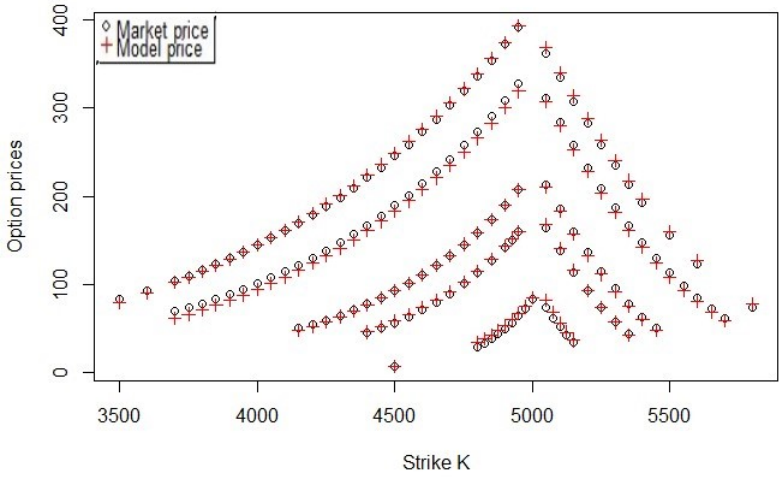

Figure 5.2.6: Rough Heston option prices vs. market option prices.

Heston model even if some OTM put options are still not well fitted by the model.

As a remark, we can add that we tried to model the variance swap curve $\xi_0(t)$ instead of deriving it from the market to see whether the variance curve of Figure 5.2.5 was optimal and consistent. We used the quadratic form $\hat{\xi}_0(t) = \hat{a} + \hat{b} \times T + \hat{c} \times T^2$ (where T is the maturity of the variance swap) as a simple approximation of this curve. We observed that the shape and level of the two curves were very similar (see Figure 7.1.12 in Appendix). Moreover, the RMSE of the rough Heston with the market variance swap curve was very close to the model with the estimated quadratic curve. We can conclude that our results are therefore coherent and that our method is consistent.

One sample path of the price process and one sample path of the variance process under our calibrated rough Heston model are plotted in Figure 7.1.13 and 7.1.14 (Appendix). Again the simulation method used for the rough Heston is described in Section 5.4.1. We clearly see the rough-type behavior of the variance process compared to the sample paths of the two previous models. We also have that the volatility of volatility parameter ν is quite high, leading to large movements in the variance process.

5.2.5 Calibration of the rBergomi model

The calibration of this model departs significantly from the three models described above. Indeed, no characteristic function is available in (quasi-)closed form. We will then rely on two different methods to calibrate the parameters ρ, η and H of the rBergomi model under $\mathbb{P}^*$. Note that the variance swap curve is exactly the same as the one derived in the calibration of the rough Heston model.

- Guessing :
The calibration method presented in Bayer et al. (2016) is based on guessed parameters. These guessed parameters are chosen in such a way to best fit the term structure of ATM skews $\psi(\tau)$. First, following these authors, both the roughness parameter H and the volatility of volatility η should be the same under $\mathbb{P}$ and $\mathbb{P}^*$. Recall that in Section 5.1, we found the historical parameter estimates $H = 0.12923$ and $\nu = 0.30977$. From Section 4.2.2 on the rBergomi model, we have that $\eta = 2\nu\, C_H/\sqrt{2H} = 0.5109$. However, as explained in Gatheral (2019), we define in a first step :

$$\tilde{\eta} := \eta\sqrt{2H} = 2\nu\, C_H = 0.2597$$

due to the fact that we have $\eta\sqrt{2H}\int_t^u \frac{1}{(u-s)^\gamma} dW_s^*$ in the right-hand side of equation (4.2.8). $\tilde{\eta}$ seems to be inconsistent with the implied estimate derived in Bayer et al. (2016), which was close to 0.7.

However, the historical estimate is in daily terms and the implied estimate in annualized terms. To convert the historical estimate, Gatheral (2019) multiplies it by the annualization factor $(252)^H$, to get :

$$\tilde{\eta}_{ann} = \tilde{\eta} \times (252)^H = 0.5306 \;\Rightarrow\; \eta_{ann} = \frac{0.5306}{\sqrt{2H}} = 1.0439$$

This way, the historical and implied estimates seem to be more consistent. The remaining difference can be explained by the fact that is not unexpected for the implied volatility of volatility η to be higher than the historical one to reflect the volatility of the volatility risk premium.

Finally, following again Bayer et al. (2016), the product $\rho\eta$ sets the level of ATM skews $\psi(\tau)$ for longer times to expiration τ. In the next section, we will see (in Figure 5.3.2) that this level is around -0.51. It then gives a value of ρ close to $-0.51/1.0439 = -0.4886$. We will confirm in Section 5.3.4 that these parameters provide satisfactory results in terms of fit of the observed term structure of ATM volatility skews $\psi(\tau)$. We can now analyze on Figure 5.2.7 whether these guessed parameters ($H = 0.129$, $\eta = 1.0439$, $\rho = -0.4886$) are able to fit the observed option prices. The fit is clearly poorer compared to the previous models even if for guessed parameters, it is not that bad, especially for put options. The RMSE is equal to 0.427614, which is obviously well above the RMSE of previous models. Longer maturity calls are also substantially off market, as in the Heston model[1].

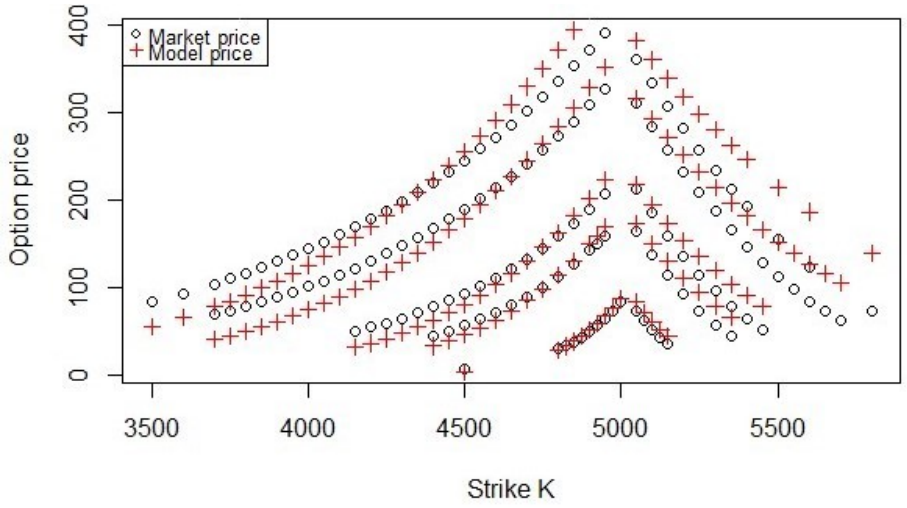

Figure 5.2.7: rBergomi option prices vs. market option prices (guessed parameters).

- Hybrid scheme : This method uses a Monte-Carlo type approach in order to calibrate the rBergomi model. We no longer try to fit the ATM skew term structure but try to fit as best as possible the observed European option prices (as for the previous models). For a parameter set $\Theta = (\rho, \eta, H)$, we simulate m realizations of the variance process and of the price process given in equation (4.2.9). We can then compute the price of European calls (and puts in the same way) using :

$$C_0(T,K) = \mathbb{E}^{\mathbb{P}^*}[(S_T - K)^+] = \frac{1}{m}\sum_{i=1}^{m}(S_{T,i} - K)^+$$

We then apply the general calibration method above and the one-to-one correspondence between European option prices and implied volatility to find the model parameters that minimizes the distance between model-based implied volatility and market implied volatility. Since the rBergomi variance process is a special case of Brownian Semistationary (BSS) processes, we use the papers of Bennedsen et al. (2017) and of McCrickerd and Pakkanen (2018) to improve the efficiency of the simulations.

[1]It is not impossible that this systematic poorer fit for higher maturities may be due to data quality issues and to the small number of European options available for larger maturities T of the CAC 40.

Their idea is roughly to build the following hybrid scheme for the simulation of $\tilde{W}_t(u)$:

$$\begin{aligned}\int_t^u \frac{dW_s}{(u-s)^\gamma} &= \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \frac{dW_s}{(u-s)^\gamma} \\ &\approx \sum_{k=1}^{\kappa} \int_{t_{k-1}}^{t_k} \frac{dW_s}{(u-s)^\gamma} + \sum_{k=\kappa+1}^{n} \frac{1}{(u-s_k)^\gamma} \int_{t_{k-1}}^{t_k} dW_s \\ &= \sum_{k=1}^{\kappa} \int_{t_{k-1}}^{t_k} \frac{dW_s}{(u-s)^\gamma} + \sum_{k=\kappa+1}^{n} \frac{1}{(u-s_k)^\gamma} Z_k \sqrt{\frac{u-t}{n}} \end{aligned} \tag{5.2.2}$$

where $s_k = u - \frac{k}{n}(u-s)$, the Z_k are i.i.d. $N(0,1)$ random variables and the s_k are such that $\int_{t_{k-1}}^{t_k} \frac{ds}{(u-s)^\gamma} = \frac{1}{(u-s_k)^\gamma}$. The choice $\kappa = 1$ works well in practice. Note that $\kappa = 0$ corresponds to the classical Euler scheme (which performs poorly).

We find a RMSE of 0.11315 with the initial set of parameters equal to the guessed parameters. This leads to the following final values Θ^*:

H^*	η^*	ρ^*
0.150329	1.489926	-0.93283

We see that the calibrated Hurst exponent H is close to the guessed H (the historical estimate) even if H^* is slightly higher. However, the volatility of volatility η^* clearly needs to be higher than the guessed η to better fit the European option data (due to the volatility of volatility risk premium as said above). The anti-correlation between the stock process and the variance process is also stronger than the guessed ρ. Moreover, it is important to add that these results are highly unstable with respect to the initial parameters as well as with the number of paths m and the time step $\Delta t = t_k - t_{k-1}$ used in the simulations. The hybrid scheme is not yet fast enough to provide a reliable calibration method, at least in our implementation. We used $m = 100,000$ and $\Delta t = T/n$ with $n = 200$ and the optimization was done in more than 20 minutes with these parameters. We need to increase m and decrease Δt to obtain more precise and stable results but it then requires far more computing time and memory, which is not suitable and tractable in practice. We obtain the following fit:

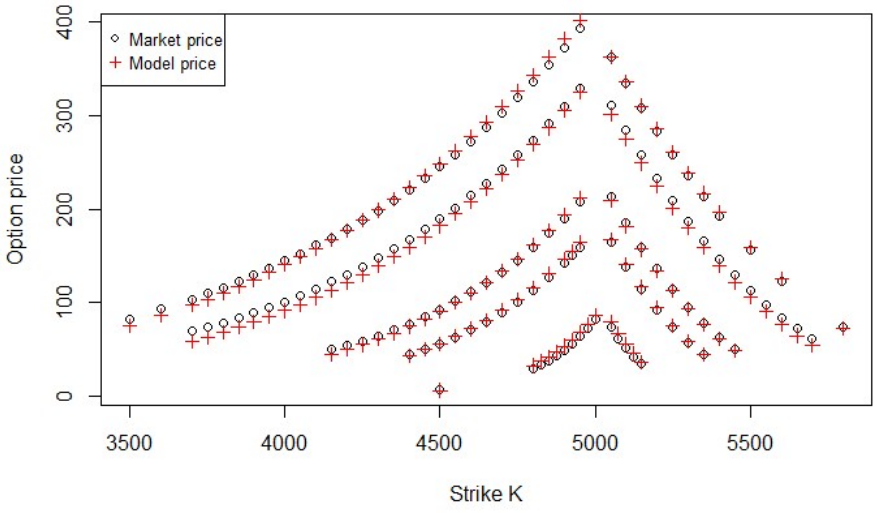

Figure 5.2.8: rBergomi option prices vs. market option prices with hybrid scheme.

You can also find in Figure 7.1.15 and 7.1.16 (Appendix), one simulated sample path of the rBergomi price process and of the variance process using this Hybrid scheme. The roughness of the variance process is again clearly visible. We conclude that the rBergomi model, even with calibrated parameters, provides a higher RMSE than in the Bates and in the rough Heston model and hence a less good fit to the observed option prices (at least in our implementation). Furthermore, this model is fairly

slow in terms of running time. To quote El Euch et al. (2019) : *"Even with the introduction of the efficient hybrid scheme [of Bennedsen et al. (2017)], practical implementation [of the rBergomi model] has proved to be difficult."* One solution to improve the efficiency of this pricing model is to use asymptotics such as in Fukasawa (2017), in Forde and Zhang (2017) or in Bayer et al. (2019, a). Finally, Bayer et al. (2019, b) propose a neural network approach to calibrate more accurately the rBergomi and to approximate the implied volatility surface. We will however not deepen these more advanced methods in this thesis.

5.3 Volatility surface fit

We now compare the Heston, Bates, rough Heston and rBergomi models in terms of implied volatility surface fit. Particularly, we will look at the fit of the volatility skews for different maturities (and their decreasing slope) as well as the fit of the term structure of ATM volatility skews $\psi(\tau)$, on which we have insisted throughout this thesis.

We first plot the volatility surface on the 6^{th} of July obtained from Bloomberg on Figure 5.3.1. We clearly see the presence of volatility skews with higher implied volatility for higher strikes K (or lower moneyness S_0/K) and lower implied volatility for lower strikes. This observation holds for most of the maturities. Moreover, as explained in Section 4.1 and in Comte and Renault (1998), we see that the slope of the volatility skew slowly decreases as a function of the time to maturity. This phenomenon is due to the effect of volatility persistence. We can also see that the term structure of the implied volatility (for all the different strikes) is decreasing, implying lower expected uncertainty for options with longer times to maturity.

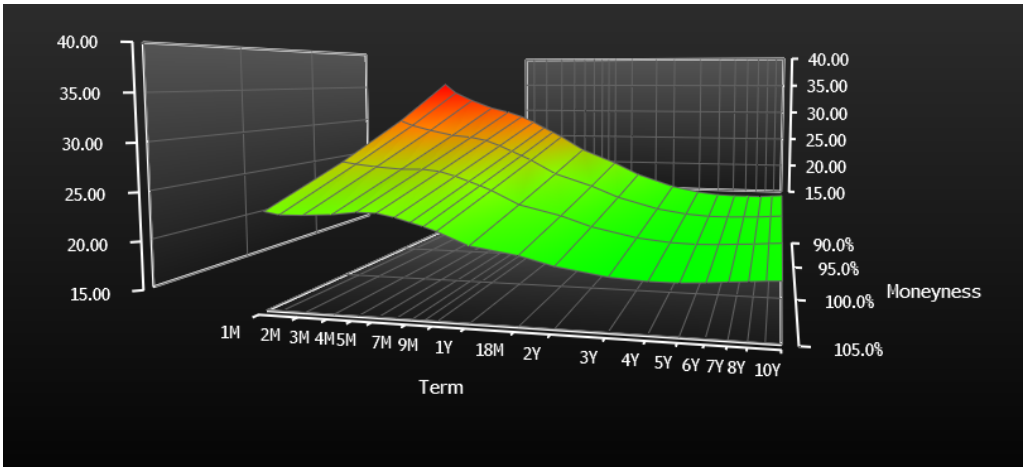

Figure 5.3.1: Implied volatility surface, CAC 40.

From our market option prices database, we can also build the term structure of ATM volatility skews $\psi(\tau)$. First, recall that $\psi(\tau) := \left|\frac{\partial}{\partial K}\hat{\sigma}_t^{imp}(\tau, K)\right|_{K=S_t}$. We can approximate it for $t = 0$ by :

$$\widehat{\psi}(\tau) = \left|\frac{\hat{\sigma}_0^{imp}(\tau, K+\delta) - \hat{\sigma}_0^{imp}(\tau, K-\delta)}{2\delta}\right|_{K=S_0}$$

for small enough δ and for each available time to maturity $\tau = T - 0$. However, this estimate is not always very precise and we prefer using a numerical approximation of the derivative with respect to K at point $K = S_0\,(= 5028.56)$ with the R function *numderiv*, which provides a better convergence. We obtain the following term structure of ATM volatility skews for our data :

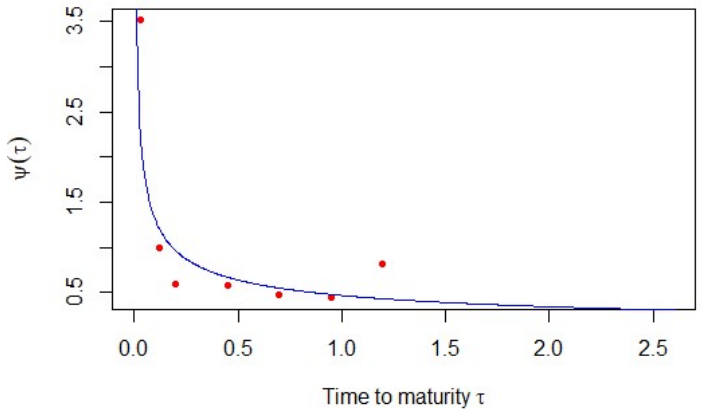

Figure 5.3.2: ATM volatility skew term structure, CAC 40.

The red dots are the estimated ATM skews for each of the available maturities. We clearly see that we obtain an ATM volatility skew which is consistent to the power-law. Indeed, the blue line is the power-law fit to the data, obtained by the following linear regression :

$$\log \psi(\tau) = -\alpha \log \tau$$

We find $\hat{\alpha} = 0.4492$. Since we said in Section 4.2 that for RSFV models $\psi(\tau) \sim \tau^{H-1/2}$ (see Fukasawa (2011)), we have that $\hat{H}_{ATM} = 0.0508$, which is lower than the Hurst exponent estimated with historical data in Section 5.1. However, we do not have a lot of option data in our database and only a few different maturities, which prevents us from having a consistent estimator $\hat{H}$ using this power-law fit. Indeed, there are almost no options traded on the CAC 40 for high maturities (especially above one year). For example, the point $\widehat{\psi}(1.20)$ is estimated from only 5 option prices and the precision of this estimator is thus poor, which can explain why it does not lie on the fitted blue line. Therefore, the analyses on the ATM volatility skew term structure will be based mostly on the approximate shape of the fit (and whether a power-law fit is suitable) and less on finding a precise estimate of the coefficient α (or H) from the risk-neutral data. We will now compare our different models and see which of them approximates the best this blue line. Note that the fit is also verified for $T = 1.2$ which was removed from the calibration.

5.3.1 Heston fit

We first plot in Figure 5.3.3 below the CAC 40 volatility skews for each of the available maturity (in terms of log-strike for a better readability). The red triangles are the market implied volatility for the different strikes and different maturities and the green line is the fit with the Heston model calibrated in Section 5.2.3. We see that the observed empirical skews are almost perfectly fitted by the Heston model for all maturities (except for $T = 0.027$ and $T = 1.20$ where we can observe slight differences). This can somehow contradict the fact that the decrease of the skew slope when time to maturity increases is much slower than implied by standard SV models. However, recall that we only have option data up to a maturity of $T = 1.20$, which is not enough to clearly appreciate this effect.

We now analyze in Figure 5.3.4 the fit of the Heston model to the empirical ATM skew term structure derived in the previous section. The blue line is therefore again the power-law fit to the observed market implied volatilities (with $\hat{\alpha} = 0.4492$). The green dots are the $\widehat{\psi}(\tau)$ derived from the Heston model for the available maturities. We clearly observe that the Heston model does not allow to capture the high ATM skews for short maturities and hence the power-law shape of $\psi(\tau)$. This was indeed one of the issue of this model that we described at the end of Section 2.3.2. We can however note that for $T > 0.20$, we obtain the same shape with the Heston fit as the blue line. We can finally conclude that the Heston model is consistent with the observed skews for the available maturities but not with the estimated ATM skew term structure $\widehat{\psi}(\tau)$, especially for short expirations.

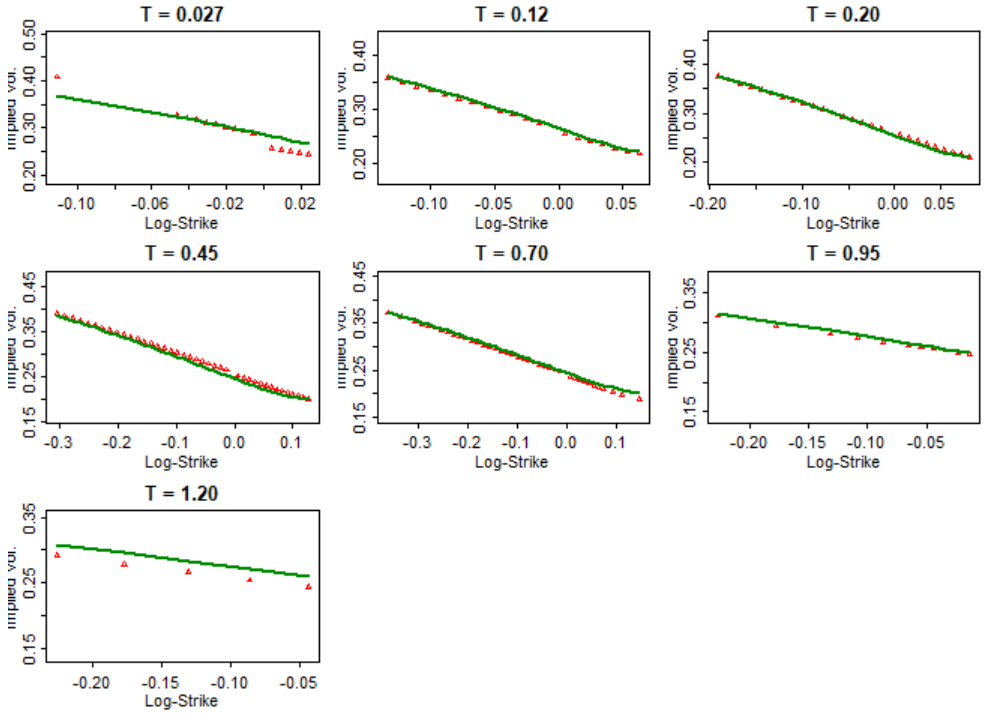

Figure 5.3.3: Empirical (red) vs. Heston (green) implied volatility skews, CAC 40.

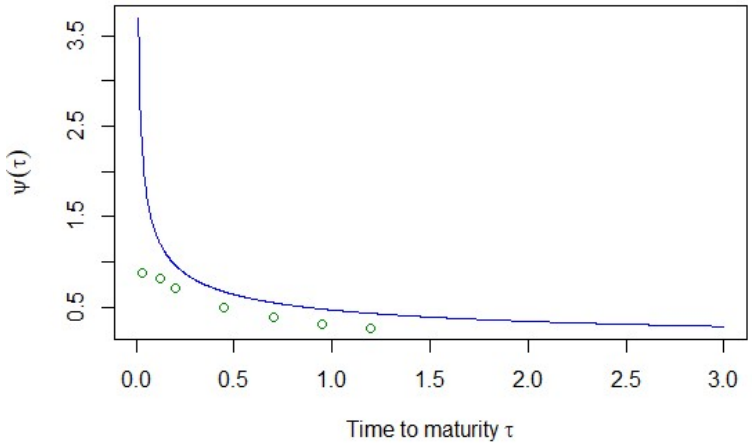

Figure 5.3.4: Empirical fit (blue) vs. Heston fit (green) of $\psi(\tau)$, CAC 40.

5.3.2 Bates model

We reproduce the same analysis for the Bates model. We plot in Figure 7.1.17 (Appendix) the skews for all the available maturities. We see that the fit is even better than in the Heston model, especially for $T = 0.027$ and $T = 1.20$, which is consistent with the lower RMSE of the Bates model.

We then have the following plot in Figure 5.3.5 for the term structure of ATM skews $\widehat{\psi}(\tau)$. The green points are slightly closer to the power-law fit compared to the Heston model, especially for short times to maturity. The presence of jumps hence provides a better fit of the term structure of ATM volatility skew $\psi(\tau)$ as explained in Section 2.3.3. However, the Bates model is still not able to capture the explosion of the skew for very short expirations. This may be due to the fact that the parameter λ controlling the frequency of jumps is only of 0.6928 and that the average size of the jumps is also rather low. A higher value of λ or of $|\mu_j|$ would allow to better capture this explosion phenomenon for $\tau \to 0$.

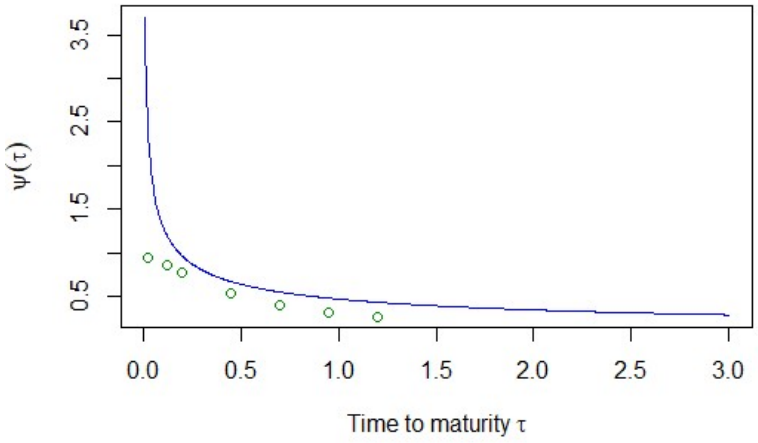

Figure 5.3.5: Empirical fit (blue) vs. Bates fit (green) of $\psi(\tau)$, CAC 40.

5.3.3 Rough Heston fit

In Figure 7.1.18 (Appendix), we plot the skews derived from the rough Heston model for all available maturities. The fit is seemingly perfect for maturities up to $T = 0.7$. For larger T, we have a slightly steeper slope than observed with the market data (but recall that there are only a few observations for these maturities). What is more interesting to analyze is the fit to the ATM skew term structure since the rough Heston model should be able to capture the high curvature of $\widehat{\psi}(\tau)$ for small time to maturities thanks to the rough behavior of its volatility process. We obtain the following plot :

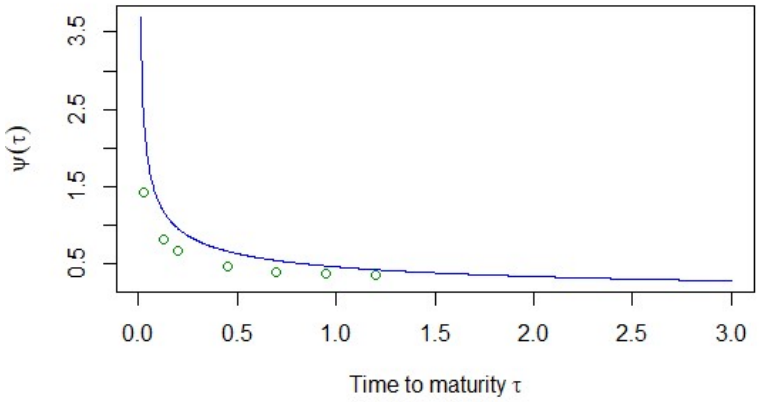

Figure 5.3.6: Empirical fit (blue) vs. rough Heston fit (green) of $\psi(\tau)$, CAC 40.

We clearly see that the green dots (i.e. the rough Heston fit for available maturities) exhibit the exact same shape as the blue line and can therefore be fitted by a power-law function $\tau^{H-1/2}$. If we had at disposal even shorter maturities, the rough Heston model would be able to better capture the explosion of $\widehat{\psi}(\tau)$ for $\tau \to 0$. We can finally conclude that the rough Heston model is highly consistent with the empirical implied volatility surface even if the explosion phenomenon is still not perfectly captured.

5.3.4 rBergomi fit

- Guessed : The fit of the observed skews with the guessed parameters is obviously not as good as with the previous models since the obtained prices with these guessed parameters are substantially off market as shown in the calibration. However, we confirm with the next Figure 5.3.7 that these guessed parameters allow to almost perfectly match the observed ATM skew term structure $\widehat{\psi}(\tau)$. Indeed, it better captures the explosion of the ATM skews for short-maturities than the rough Heston model.

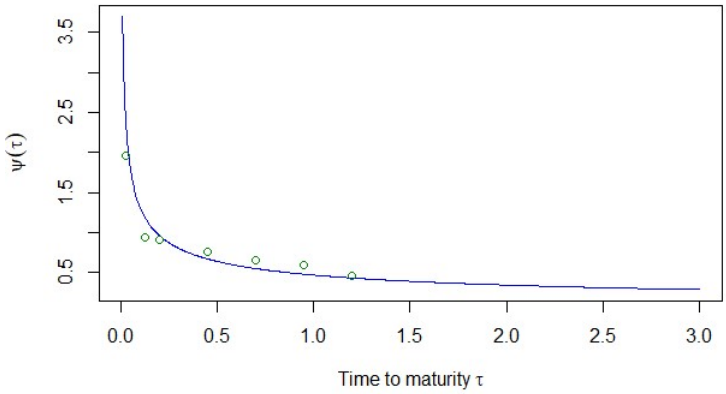

Figure 5.3.7: Empirical fit vs. rBergomi fit (guessed) of $\psi(\tau)$.

- Hybrid scheme : Since the RMSE was equal to 0.113, the rBergomi with parameters calibrated on European option prices provides a good fit of the observed skews for the available maturities (except for $T = 0.45$, see Figure 7.1.19 in Appendix). In terms of the fit of the ATM skew term structure, we obtain the following plot :

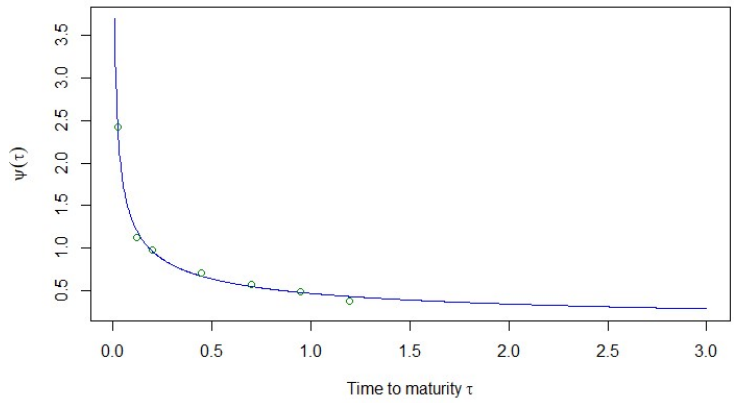

Figure 5.3.8: Empirical fit vs. rBergomi fit (hybrid scheme) of $\psi(\tau)$.

We obtain a perfect fit of the observed ATM skew term structure $\widehat{\psi}(\tau)$ with green dots that almost exactly lie on the blue curve. This rBergomi model based on the Hybrid scheme is even better than the guessed rBergomi to capture the explosion phenomenon for $\tau \to 0$.
We conclude that the rBergomi model with guessed parameters is highly consistent with the observed ATM skew term structure but not with the market prices of European calls and puts. The rBergomi model based on the Hybrid scheme provides a perfect fit of the observed ATM skew term structure while having a low RMSE and hence a good fit to the market option data. This model is however not stable and is quite time-intensive as already explained.

5.3.5 Local volatility fit

Finally, Bloomberg terminals provide a fit of the current volatility surface based on the local volatility model introduced by Dupire (see Section 2.2). Recall that the local volatility $\sigma(S_t, t)$ is a deterministic function of S_t and t such that the model exactly provides the observed prices $C_0^m(T, K)$ of European options (provided that there is no static arbitrage opportunities). The fit calibrated automatically by Bloomberg is given in Figure 5.3.9 below.
The shape of the volatility surface does not appear as realistic compared to the observed one (no presence of skews and no decreasing term structure for example) due to the interpolations between the observed market prices of European options. Furthermore, this surface is highly irregular and not

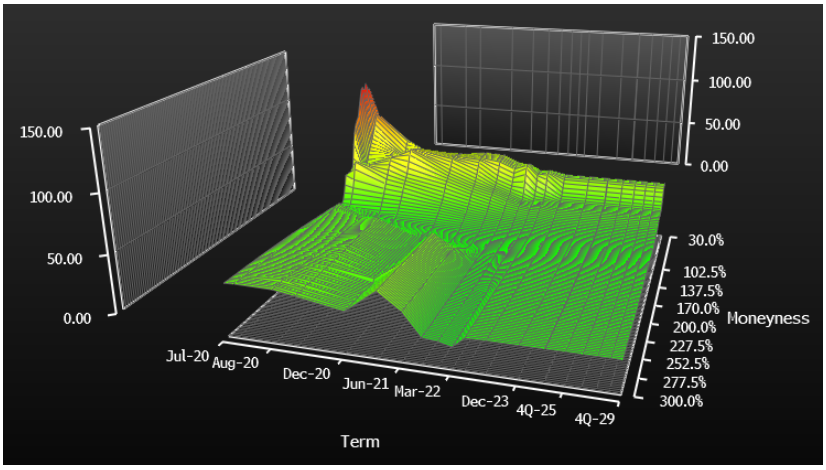

Figure 5.3.9: Local volatility surface, CAC 40.

smooth, which leads to unstable results, especially when taking the derivative of the implied volatility.

5.3.6 Conclusion of the fit

From the analyses of the implied volatility surface, we can conclude that rough volatility models with the rough Heston and the rBergomi clearly outperform standard volatility models in terms of goodness of fit.
With a RMSE of 0.0929, an almost perfect fit of volatility skews for available maturities and a good fit of the ATM volatility skew term structure, we conclude that the rough Heston is the best of the considered models. Combined with the fact that this model exhibits a rough variance process (with a H very close to the historical estimate) and can hence reproduce numerous statistical properties of the spot volatility process, the rough Heston model provides highly accurate and consistent results. Moreover, the rough Heston is also particularly tractable (only three parameters to calibrate and a characteristic function available in quasi-closed form) which is very important in practice for option pricing. To quote Gatheral (2019) : *"For perhaps the first time, we have a simple consistent model of historical and implied volatilities"*. The main drawback of this model is that the log-volatility is not modeled explicitly as a fBm even if we showed empirically that the distribution of $(\log \sigma_{t+\Delta} - \log \sigma_t)$ is close to Gaussian under the rough Heston for various lags Δ.
Even if the rBergomi model provides the best fit of the term structure of ATM volatility skews, its RMSE is above the rough Heston's one and it is much less tractable and stable than the latter. As mentioned above, further work is needed to provide a better implementation of this rBergomi model.

5.4 Life insurance contract valuation

Although all models are well calibrated and hence vanilla options have approximately the same prices under all models, exotic option prices can differ dramatically. It is important to point out that vanilla options determine the marginal distribution (at maturity T of the option), not the process. Indeed, the underlying fine-grain properties of the process have an important impact on path-dependent option prices. We now highlight the impact of exotic price ranges by valuating a life insurance contract, embedding a cliquet option in it. This section is mainly based on Bacinello et al. (2010).

5.4.1 Simulation and discretization methodology

In order to price this life insurance contract, we first need to explain how we can discretize and simulate sample paths of the price process and the variance process under our different models.

- Heston :
Once the Heston model has been calibrated and the optimal parameters $\sigma_0, \kappa, \eta, \theta$ and ρ have been

found (where we drop the * from the notations), we can use a Monte Carlo approach to simulate the sample paths of the Heston model. We first simulate the stock price process and the volatility process by generating correlated standard normal random numbers ε_S and ε_v (with correlation ρ). The relationship between ε_S and ε_v can be written as :

$$\varepsilon_{v,t_i} = \rho\, \varepsilon_{S,t_i} + \sqrt{1-\rho^2}\, \varepsilon'_{t_i}$$

where ε_{S,t_i} and ε'_{t_i} are independent $N(0,1)$ random variables. Using an Euler discretization scheme, we have :

$$S_{t_{i+1}} = S_{t_i}\left(1 + r\Delta t + \sqrt{v_{t_i}}\sqrt{\Delta t}\,\varepsilon_{S,t_i}\right) \tag{5.4.1}$$

$$v_{t_{i+1}} = v_{t_i} + \kappa(\eta - v_{t_i})\Delta t + \theta\sqrt{v_{t_i}}\sqrt{\Delta t}\,\varepsilon_{v,t_i} \tag{5.4.2}$$

for $i = 0, ..., n-1$, with $t_0 = 0$ and $t_n = T$, where n is the number of time steps and $\Delta t = T/n$, for a given maturity T. However, we prefer using a Milstein scheme for the variance process instead since it provides a better convergence, which gives :

$$v_{t_{i+1}} = v_{t_i} + \left(\kappa(\eta - v_{t_i}) - \theta^2/4\right)\Delta t + \theta\sqrt{v_{t_i}}\sqrt{\Delta t}\,\varepsilon_{v,t_i} + \theta^2\Delta t\,(\varepsilon_{v,t_i})^2/4 \tag{5.4.3}$$

We then simulate m sample paths of this stock price process and of this volatility process by generating m vectors ε_S and ε_v (of dimension n) of standard random numbers $N(0,1)$. We now have the full path of $(S_{t_i})_{i=0,...,n}$ for each m, which allows us to price exotic options, such as path-dependent options (see Section 5.4.2).
Even if theoretically the process never hits zero thanks to Feller condition (2.3.11), in practice, the variance can become negative under a simulation because of the discretization. We then use absorption (if negative variance, set it equal to zero) or reflection techniques (if negative variance, take absolute value) to overcome this problem even if these solutions are wrong from a theoretical point of view.

- Bates :
We use the same Milstein scheme for the variance process as in the Heston model. However, the discretization scheme (5.4.1) of the price process is replaced by :

$$S_{t_{i+1}} = S_{t_i}\left(1 + r\Delta t + \sqrt{v_{t_i}}\sqrt{\Delta t}\,\varepsilon_{S,t_i} + (Y_{t_i} - 1)\,\Delta N_{t_i}\right) \tag{5.4.4}$$

where $Y_{t_i} \sim \text{Log}N(\mu_j, \sigma_j^2)$ and $\Delta N_{t_i} \sim Poi(\lambda\,\Delta t)$.

- rBergomi :
The price process is exactly the same as in the Heston model with (5.4.1). However, the variance process is now simply given by :

$$v_{t_i} = \xi_0(t_i)\,\mathcal{E}\left(\eta\,\tilde{W}^*_{t_i}\right)$$

where we use the calibrated forward variance curve for $\xi_0(t_i)$ $(= \mathbb{E}^{\mathbb{P}^*}[v_{t_i}])$ and the hybrid scheme (5.2.2) for the simulation of $\tilde{W}^*_{t_i}$, with $i = 0, ..., n$.

- Rough Heston :
The discretized price process is again given by (5.4.1). We then simply use an Euler discretization scheme for the variance process of the rough Heston given by equation (4.2.12). At current time $t = 0$, we hence write for $i = 0, ..., n$ with $t_0 = 0$ and $t_n = T$:

$$v_{t_i} = \xi_0(t_i) + \frac{1}{\Gamma(H+1/2)}\sum_{k=0}^{i}(t_i - t_k)^{H-1/2}\nu\sqrt{v_{t_k}}\,\sqrt{\Delta t}\,\varepsilon_{v,t_k} \tag{5.4.5}$$

5.4.2 Valuation

We now describe and price the following endowment life insurance contract. This endowment insurance provides a lump sum payment at maturity T in case of survival, or a payment at the time of death t if

it occurs before T. We consider an equity-linked case, where payment amounts depend on the market value of a reference fund $F(t)$ and where an amount $F(0) = D$ is invested in it at time $t = 0$. We then introduce a **yearly** minimal guarantee κ_g for the policyholder and a maximal **yearly** return κ_m that this policyholder can earn on the fund. The survival benefit is then given by $B_T^s = F_T^s \mathbb{1}_{t \geq T}$ and the death benefit by $B_t^d = F_t^d \mathbb{1}_{t<T}$, where F_t^e for $e = s, d$ is given by a series of guarantees (= cliquet) :

$$\begin{aligned} F_t^e = F_0 \prod_{u=1}^{\lfloor t \rfloor} & \min \left\{ e^{\kappa_m} \, ; \, \max \left\{ 1 + \eta \left(\frac{S_u}{S_{u-1}} - 1 \right) \, ; \, e^{\kappa_g} \right\} \right\} \\ & \times \min \left\{ e^{\kappa_m (t - \lfloor t \rfloor)} \, ; \, \max \left\{ 1 + \eta \left(\frac{S_t}{S_{\lfloor t \rfloor}} - 1 \right) \, ; \, e^{\kappa_g (t - \lfloor t \rfloor)} \right\} \right\} \end{aligned} \tag{5.4.6}$$

In the equation above, we have that $F_0 = D$, the initial value of the reference fund and S_t is the price process of each fund unit defined on a filtration $(\mathcal{F}_t)_{t \geq 0}$. The cliquet above implicitly assumes yearly resettlements. Finally, the rate η identifies the portion of yearly return recognized to the policyholder with $\eta \in (0, 1]$.

We then denote by τ the random residual lifetime of the policyholder aged x at time $t = 0$. We also introduce the non-homogeneous Poisson process $N_t := \mathbb{1}_{\tau < t}$ defined on a filtration $(\mathcal{G}_t)_{t \geq 0}$, which is equal to zero as long as the individual is alive and jumps to one at death. The intensity of N_t is the deterministic Makeham force of mortality $\mu_x(t) = c + ab^{x+t}$ and hence her survival probability at time T, being alive at $t = 0$, is given by ${}_Tp_x = P(N_T = 0 \,|\, \mathcal{G}_0) = \exp\left(- \int_0^T \mu_x(u) \, du \right)$.

The fair value at time $t = 0$ of the cash-flow F_t^e is given by its discounted expected value under a risk-neutral measure $\mathbb{P}^*$. However, this cash flow also depends on the path of the Poisson process $(N_t)_{t \geq 0}$ which is not known at time $t = 0$. Therefore, we take the following expectation conditionally to $\mathcal{G}_0$:

$$FV_0 = \mathbb{E} \left[(1 - N_T) \, e^{-rT} \, \mathbb{E}^{\mathbb{P}^*}[F_T^s \,|\, \mathcal{F}_0] + \int_0^T e^{-rt} \, \mathbb{E}^{\mathbb{P}^*}[F_t^d \,|\, \mathcal{F}_0] \, dN_t \,\middle|\, \mathcal{G}_0 \right]$$

The processes $(N_t)_{t \geq 0}$ and $(S_t)_{t \geq 0}$ being independent, we find that (see Demo. 7) :

$$FV_0 = {}_Tp_x \, e^{-rT} \, \mathbb{E}^{\mathbb{P}^*}[F_T^s] + \int_0^T e^{-rt} \, \mathbb{E}^{\mathbb{P}^*}[F_t^d] \, {}_tp_x \, \mu_x(t) \, dt$$

We now want to price this insurance contract with the four models we calibrated above (Heston, Bates, rBergomi and rough Heston). We can then use a Monte-Carlo approach to simulate their price process $(S_t)_{t \geq 0}$ (as well as their respective variance process) and take the average of the paid benefit under each simulation j to derive the expectation under $\mathbb{P}^*$, i.e. :

$$FV_0 = \frac{1}{m} \sum_{j=1}^{m} \left(e^{-rt_n} F_{t_n,j}^s \, \mathbb{1}\Big\{ \sum_{i=0}^{n} \Delta N_{t_i,j} = 0 \Big\} + \sum_{i=0}^{n} e^{-rt_i} \Delta N_{t_i,j} F_{t_i,j}^d \right)$$

where the price process $(S_t)_{t \geq 0}$ in F_t^e is defined by the risk-neutral pricing equation of each model. The discretization of their price process and variance process is given in the previous Section 5.4.1. Moreover, the use of a Poisson process above makes it easy to add the mortality effect in the simulations to compute FV_0. Indeed, for each simulation $j \in m$, we generate $n = T/\Delta t = 5/0.001 = 5000$ Poisson random variables $\Delta N_{t_i,j}$ with parameter $\lambda = \mu_x(t_i) \, \Delta t$ for $i = 0, ..., n$. If $\Delta N_{t_i,j} = 1$ for some i, we compute a death benefit $F_{t_i,j}^d$ and if $\Delta N_{t_i,j} = 0$ for all i, we compute a survival benefit $F_{t_n,j}^s$.

It is important to point out that the volatility of volatility parameter for each of our models is quite high due to the current uncertainty in the economical environment and on the financial markets. The variance can then become negative under some simulations because of the discretization. We therefore use the reflection technique which consists in taking the absolute value of v_{t_i} if negative, even if this solution is wrong theoretically (since we change the properties of the process). It also happens due to this high volatility of volatility parameter to have an explosion in the variance process (above 500% for some paths). This is the case in each model but especially for the rBergomi and rough Heston

models where the variance process tends to explode after 6 years. This can be considered as a flaw of our rough volatility models, especially when pricing long-term contracts with maturities higher than 10 years. Note that this can also be explained by the fact that our models are only calibrated on option data with maturities lower than one year. The comparison we did in Chapter 5 should hence be carried out in other market conditions and with other stocks/indices (with higher maturities available).

We consider the following parameters :

- The policyholder is a women aged $x = 50$ at time $t = 0$ and the maturity T of the endowment insurance is 5 years.
- We use the Belgian regulatory life table FR for the Makeham force of mortality $\mu_x(t)$.
- The initial amount D invested in the fund equals 10 000 €
- The policyholder pays a unique premium P at time $t = 0$ (different from D)
- The minimal yearly guaranteed rate is $\kappa_g = 0.01$ and the maximal yearly rate $\kappa_m = 0.2$ with a participation rate $\eta = 80\%$

We obtain the following fair value of our life insurance contract for the different models using Monte-Carlo simulations with $m = 30\,000$:

	Heston	Bates	rBergomi	rough Heston
FV_0	14 194.93 €	14 280.03 €	13 840.93 €	13 210.61 €

The pure premium P then must be equal to the fair value at time $t = 0$ in order for the contract to be fair. We see that the rough volatility models tend to have a lower price than standard SV models. It is especially true for the rough Heston model which provides a price substantially lower. We therefore see how the sample path properties impact the price of exotic options. Since we said that the rough Heston model is the most consistent with risk-neutral data, we retain 13 210.61 € as being the contract's price which is the most market-consistent. We confirm these effects with the two following Tables 5.4.1 and 5.4.2, displaying the fair value FV_0 in function of the minimum yearly guaranteed rate k_g and the maturity of the contract T.

$T = 5$	Heston	Bates	rBergomi	rough Heston
$k_g = 0.5\%$	14 037.21 €	14 115.89 €	13 642.21 €	12 993.38 €
$k_g = 1\%$	14 194.93 €	14 280.03 €	13 840.93 €	13 210.61 €
$k_g = 2\%$	14 528.76 €	14 625.53 €	14 254.81 €	13 661.89 €
$k_g = 3\%$	14 888.85 €	14 995.56 €	14 692.21 €	14 136.66 €
$k_g = 5\%$	15 696.26 €	15 817.39 €	15 644.92 €	15 162.59 €

Table 5.4.1: FV_0 for various k_g with maturity $T = 5$.

We observe the highest differences between the rough and the standard SV models when k_g tends to be small. Indeed, in these cases, the minimum yearly guarantee is exercised less often and therefore, the yearly returns which are highly influenced by the sample paths play a more important role, leading to stronger differences in fair value between models with divergent sample path properties.

$k_g = 1\%$	Heston	Bates	rBergomi	rough Heston
$T = 1$	10 736.96 €	10 757.02 €	10 126.02 €	10 701.57 €
$T = 5$	14 194.93 €	14 280.03 €	13 840.93 €	13 210.61 €
$T = 10$	20 349.01 €	20 606.18 €	20 179.69 €	20 043.04 €
$T = 15$	30 132.37 €	30 766.19 €	27 783.27 €	27 663.06 €

Table 5.4.2: FV_0 for various maturities T with $k_g = 1\%$.

The differences in fair value between models are also the most exacerbated when the maturity of the contract is larger. The impact of the sample paths properties of each model is indeed stronger at longer time scales. The fact that the volatility process generated by the rough models tends to explode after 6 years also explains these higher differences between standard and rough volatility models.

We finally want to highlight that the numbers given in the Tables above can slightly change for each run of simulations but are quite stable with our chosen m, equal to 30 000. However, changing the discretization step Δt can lead to significant variations between results. $\Delta t = 0.001$ was the smallest value with $m = 30\,000$ allowed by our computer.
We again want to draw attention on the fact that the risk-free rate used for the valuation of the contract is extrapolated based on data we have between $t = 0$ and $t = 1$ thanks to a Nelson-Siegel model. Data on risk-free rates for higher maturities are thus necessary to build a more realistic term structure.

Chapter 6

Conclusion

This master's thesis reviewed different pricing models that go beyond the Black & Scholes model by introducing non-constant volatility σ_t. We especially focused on standard stochastic volatility models (SV) and on more advanced volatility models based on fractional Brownian motions (and to a lesser extent on Brownian semistationary processes). Indeed, standard SV models such as the Heston or the Bates model are not able to reproduce numerous statistical properties of the spot volatility process and cannot fit the observed implied volatility surface, such as the term structure of ATM volatility skews $\psi(\tau)$. Therefore, rough fractional volatility models were introduced by Gatheral et al. (2014) from fractional Ornstein-Uhlenbeck processes in order to provide a consistent model both with real-world and risk-neutral data. More precisely, these rough volatility models are based on fractional Brownian motions with a Hurst index $H < 1/2$ and therefore, the volatility process generated by such models exhibits rough sample paths and short-range dependence, in contrast to the FSV model introduced by Comte and Renault (1998) with $H > 1/2$.

The main contribution of this thesis is to show that these rough stochastic volatility models with $H \approx 0.1$ tend to outperform standard SV models in terms of European option calibration and in terms of fit of the volatility surface while reproducing the observed statistical properties of the spot volatility process. We especially retain the rough Heston model as a tractable implementation of rough fractional stochastic volatility models. This rough Heston model provides highly accurate and consistent results while having only three parameters to estimate and a characteristic function available in quasi-closed form. Combined with the Carr-Madan formula (7.2.1), the pricing of European options with the rough Heston model is hence extremely fast and accurate. Moreover, since the sample path properties of the observed spot volatility time series are better reflected with rough-type models than with SV or FSV models, we obtain a more accurate fair value of life insurance contracts embedding path-dependent options.

Future work is however needed to have a more thorough comparison of all the models introduced in this thesis :

- The implementation of the rBergomi model should be improved by considering more advanced calibration methods based for example on neural networks or on asymptotics.

- It would be interesting to add in the practical part a comparison with the general version of BSS models based on Gamma and power-law kernels. Bennedsen et al. (2016) made this comparison for volatility forecasting but what about the fit to the observed European option prices and to the ATM volatility skew term structure ? What about the tractability of the implementation and its stability ? Can these models beat the Bates and RFSV models in terms of RMSE ?

- All the models reviewed in this master's thesis should also be re-calibrated on other market data of European options. Choosing an other index/stock with a larger set of available maturities and at another point in time should help us to confirm and validate the results derived in this dissertation.

Chapter 7

Appendix

7.1 Figures

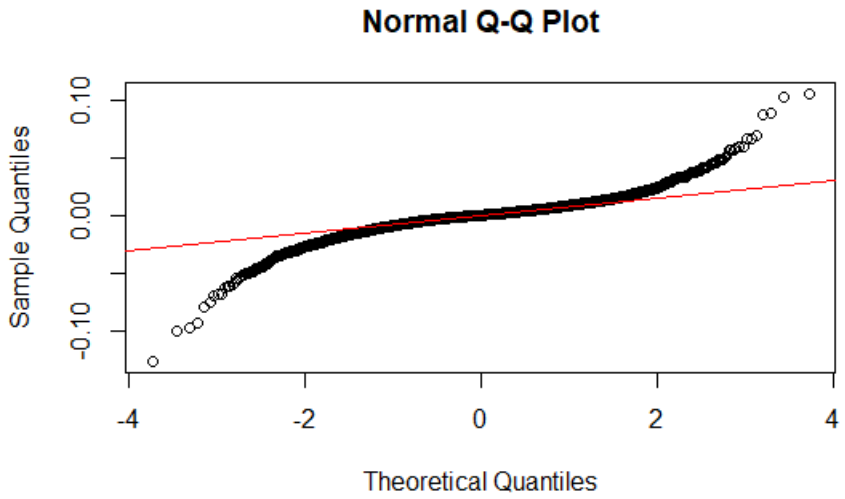

Figure 7.1.1: Q-Q plot log-returns, S&P 500.

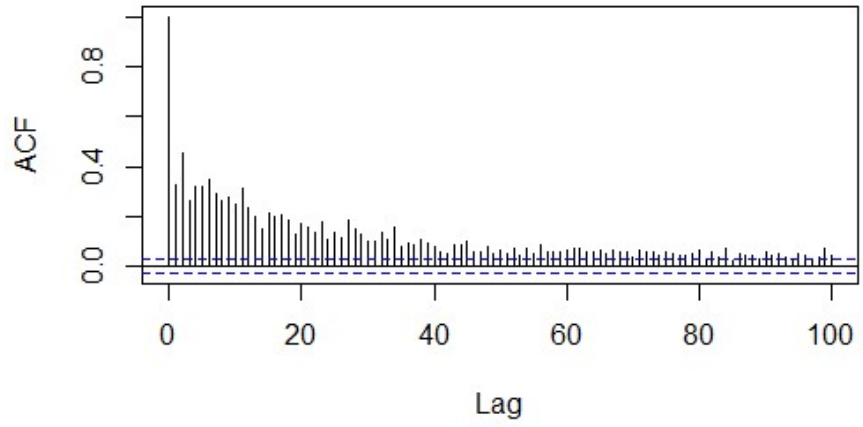

Figure 7.1.2: Autocorrelogram of the squared log-returns, S&P 500.

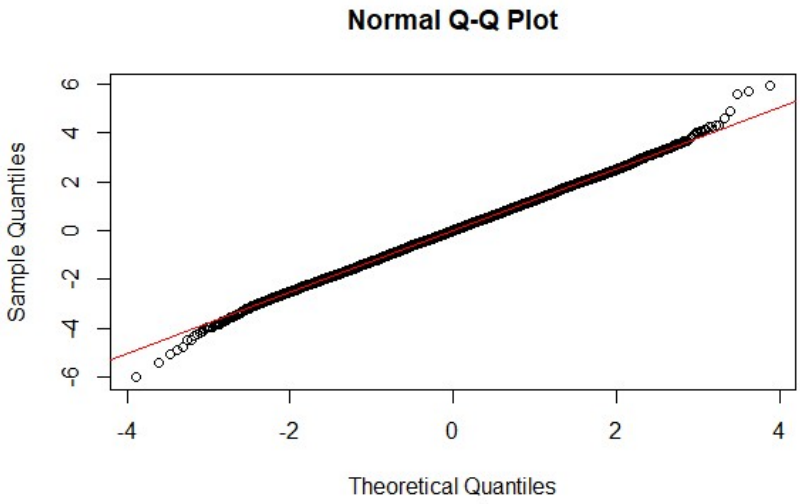

Figure 7.1.3: Q-Q plot of the distribution of $\log(\sigma_{t+\Delta}) - \log(\sigma_t)$ for Δ =1 under the rough Heston model.

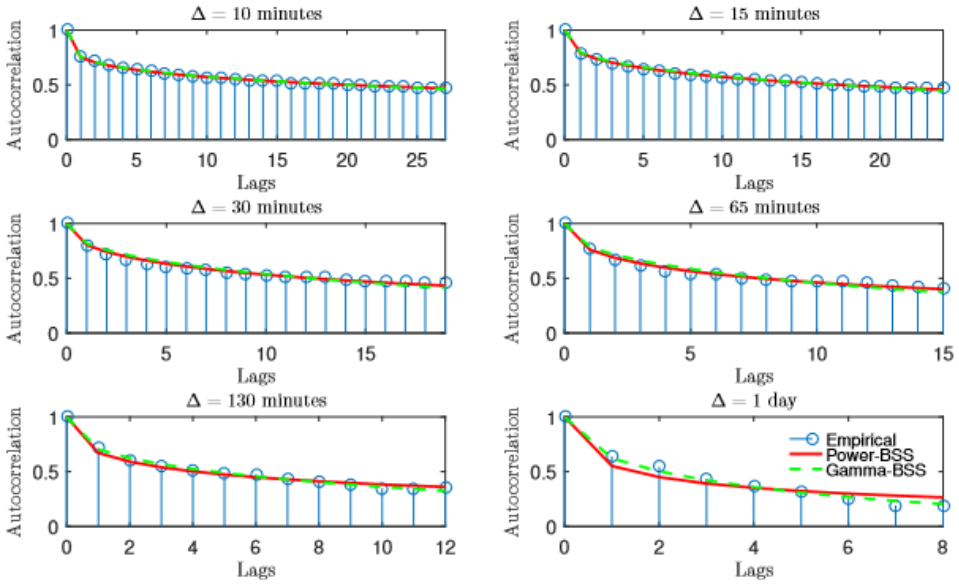

Figure 7.1.4: Empirical autocorrelation functions of log-volatility (light blue bars and circles) and fitted ACF's of the Power-BSS (solid red line) and Gamma-BSS (dashed green line) models.

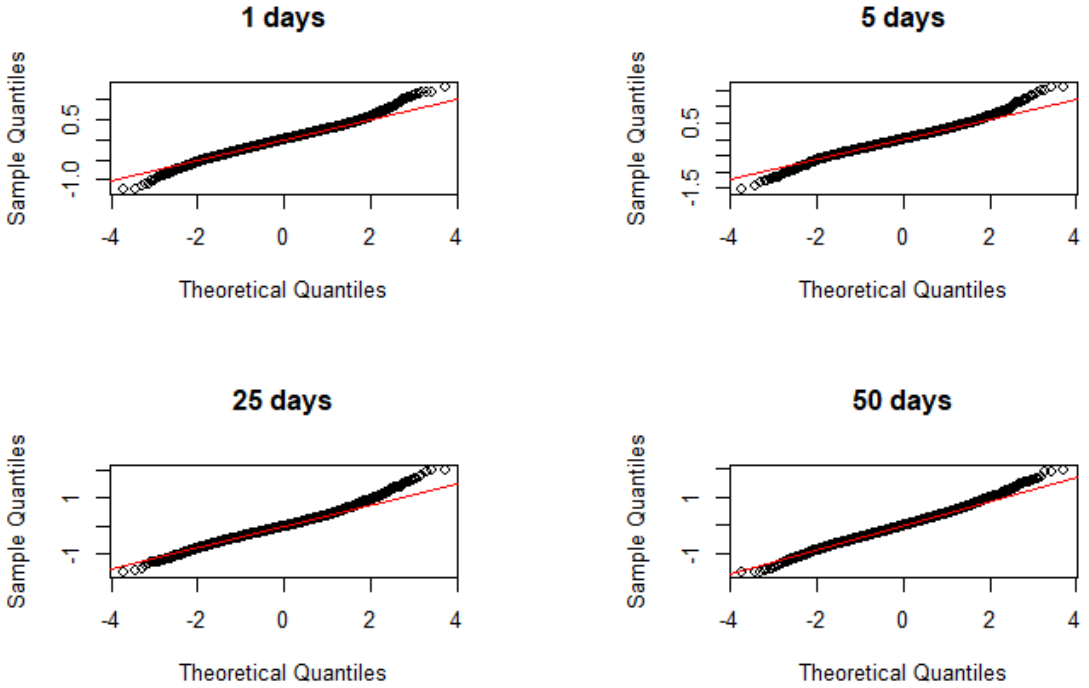

Figure 7.1.5: Q-Q plot of the distribution of $\log(\sigma_{t+\Delta}) - \log(\sigma_t)$ for Δ =1, 5, 25 and 50 days.

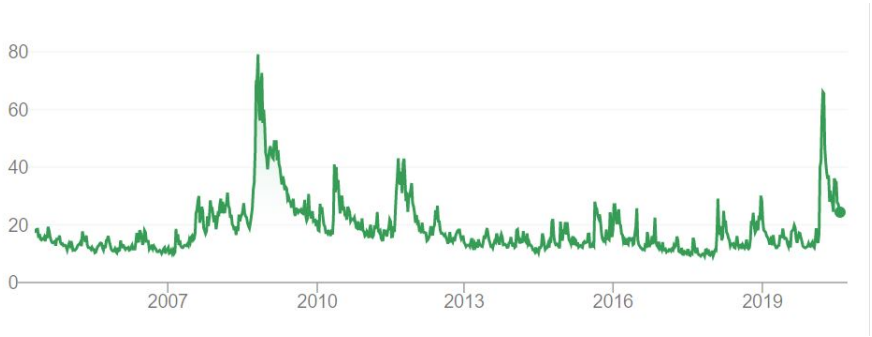

Figure 7.1.6: VIX (Volatility Index), average of the annualized volatilities of European calls and puts on the S&P500.

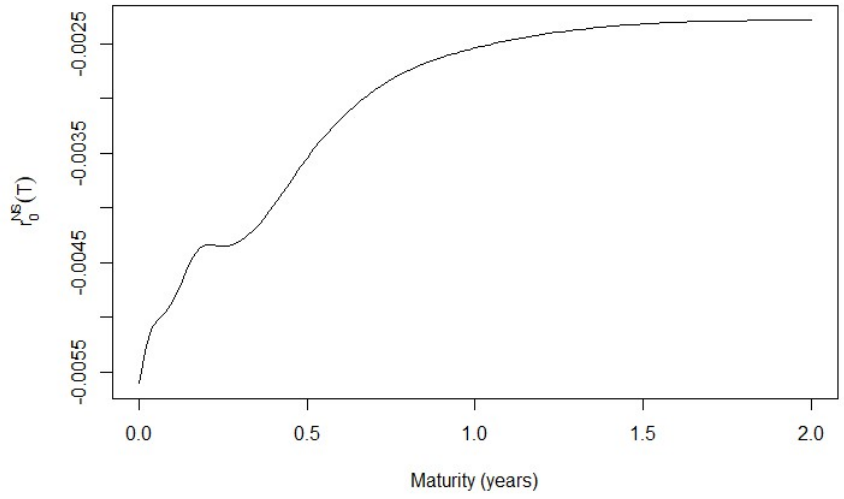

Figure 7.1.7: Nelson-Siegel EURIBOR term structure at time $t = 0$.

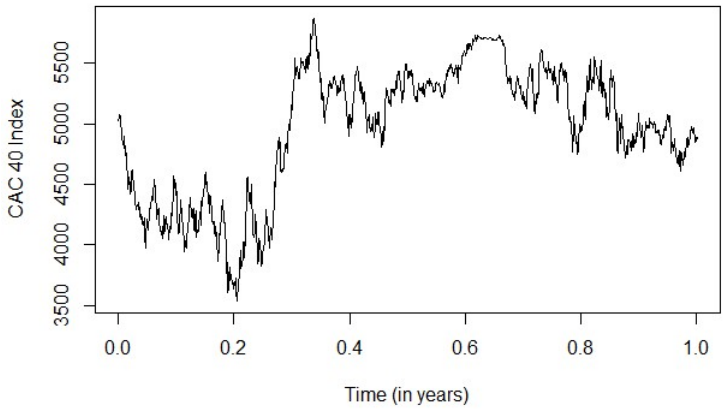

Figure 7.1.8: One sample path of the CAC 40 price under our Heston model.

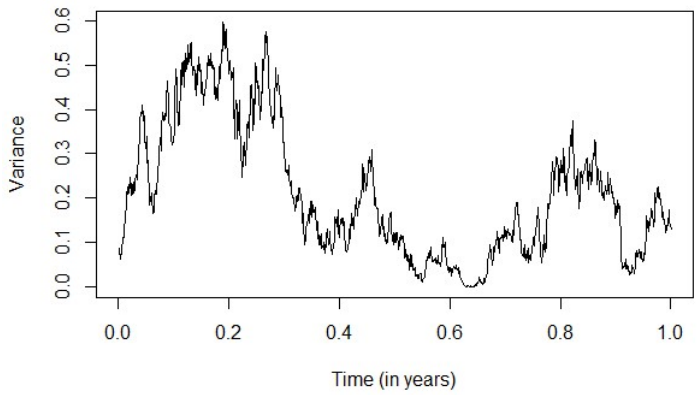

Figure 7.1.9: One sample path of the CAC 40 variance process under our Heston model.

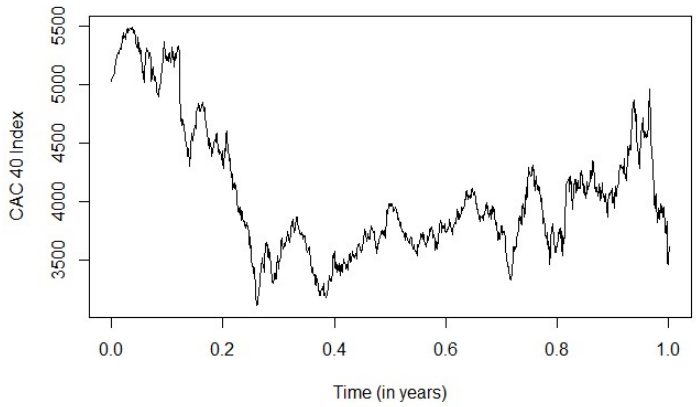

Figure 7.1.10: One sample path of the CAC 40 price under our Bates model.

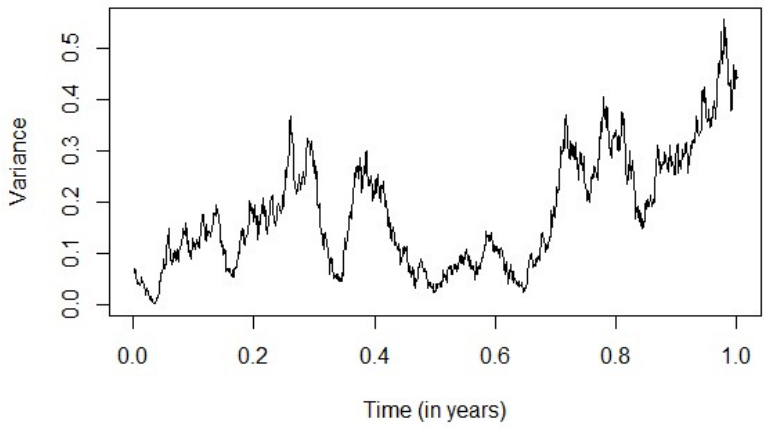

Figure 7.1.11: One sample path of the CAC 40 variance process under our Bates model.

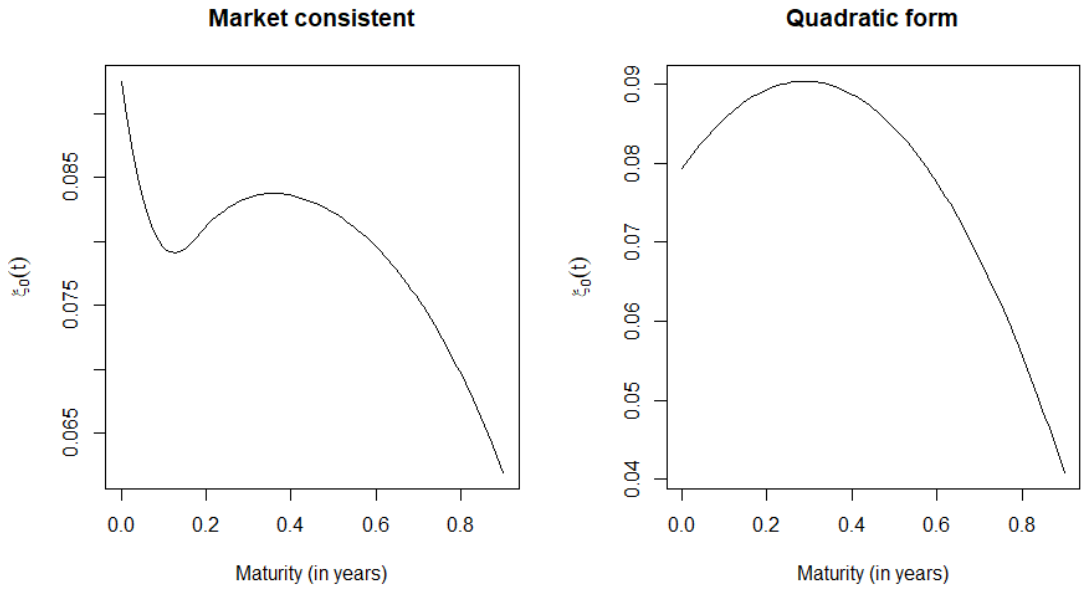

Figure 7.1.12: $\xi_0(t)$ derived from market data (left) and estimated with a quadratic function (right).

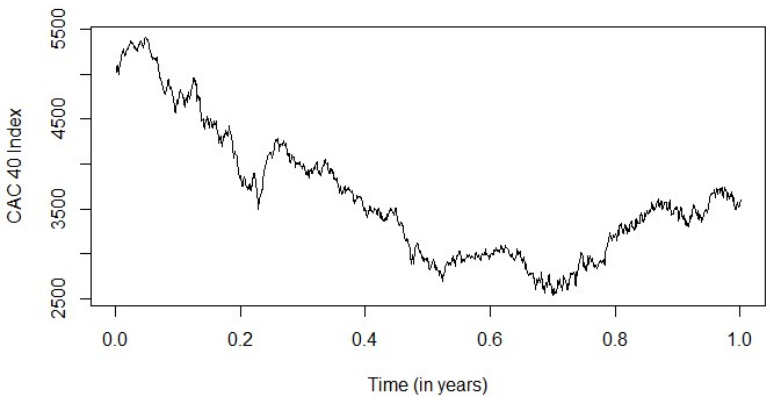

Figure 7.1.13: One sample path of the CAC 40 index price under our rough Heston model.

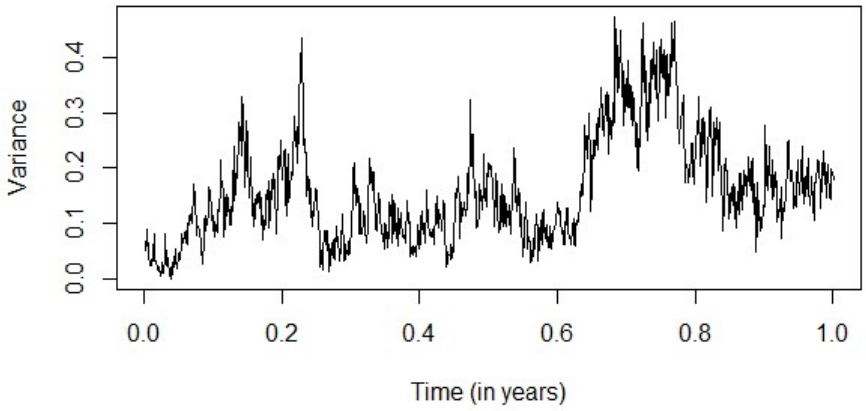

Figure 7.1.14: One sample path of the CAC 40 variance process under our rough Heston model.

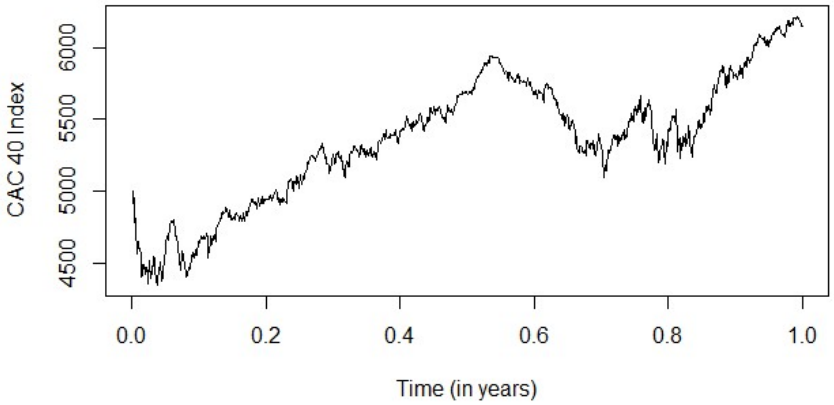

Figure 7.1.15: One sample path of the CAC 40 index price under the rBergomi model (Hybrid scheme).

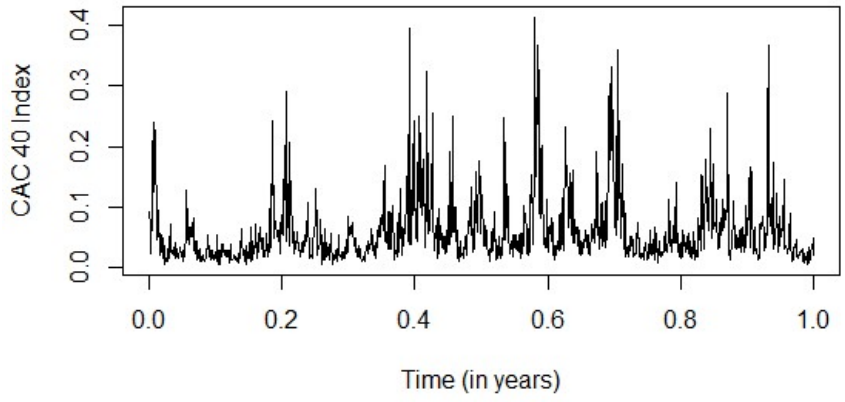

Figure 7.1.16: One sample path of the CAC 40 variance process under the rBergomi model (Hybrid scheme).

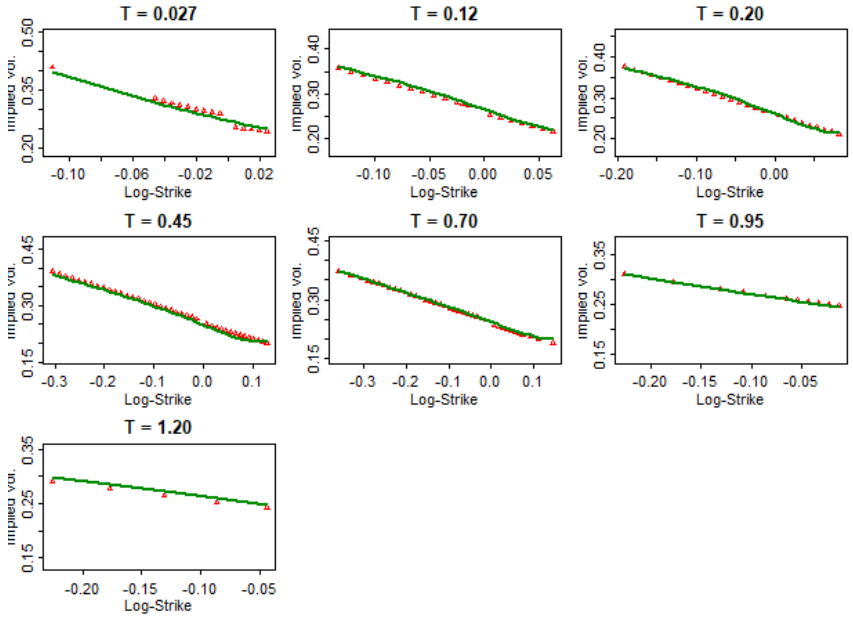

Figure 7.1.17: Empirical (red) vs. Bates (green) implied volatility skews, CAC 40.

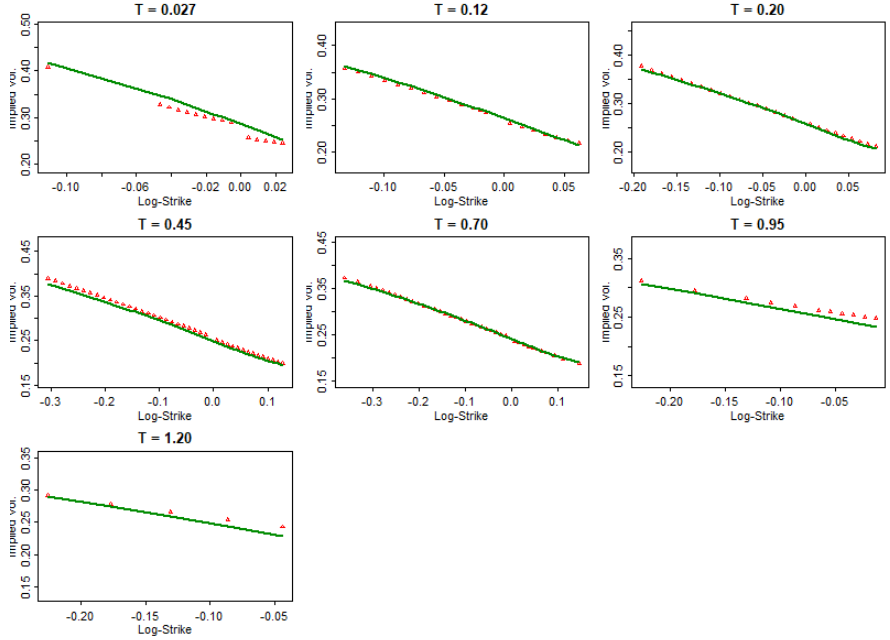

Figure 7.1.18: Empirical (red) vs. rough Heston (green) implied volatility skews, CAC 40.

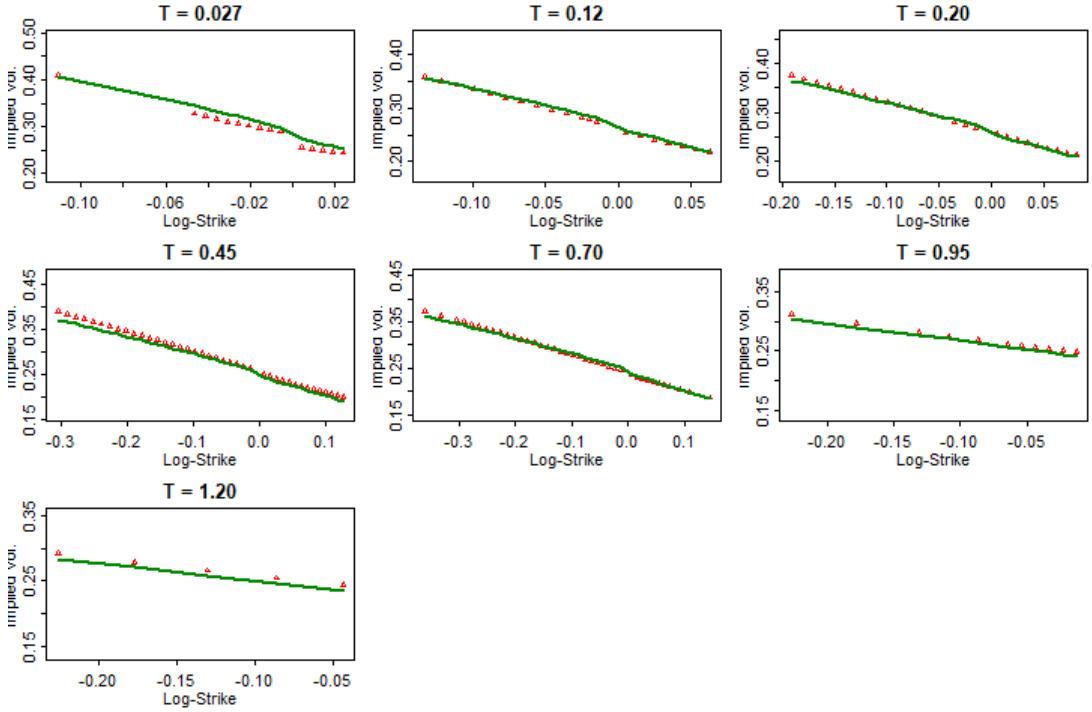

Figure 7.1.19: Empirical (red) vs. rBergomi with Hybrid scheme (green) implied volatility skews, CAC 40.

7.2 Concept

7.2.1 Itô's Lemma

Let $(\Omega, \mathcal{F}, \mathbb{P})$ and $(W_t)_{t\geq 0}$ a $\mathcal{F}_t-$Brownian motion. Let $(X_t)_{0\leq t\leq T}$ an Itô process of the form :

$$dX_t = \mu_t\, dt + \sigma_t\, dW_t\,, \quad 0 \leq t \leq T$$

where $(\mu_t)_{0\leq t\leq T}$ and $(\sigma_t)_{0\leq t\leq T}$ are $\mathcal{F}_t-$adapted processes, such that $\int_0^T |\mu_s|\, ds < +\infty$ and $\int_0^T |\sigma_s|^2\, ds < +\infty$. Then if $f(t,x)$ is a function $\mathbb{R}^+ \times \mathbb{R} \mapsto \mathbb{R}$, C^1 with respect to time and C^2 with respect to x, then for all $t \in [0,T]$:

$$d(f(t,X_t)) = \left(\frac{\partial f}{\partial t}(t,X_t) + \mu_t \frac{\partial f}{\partial x}(t,X_t) + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2}(t,X_t)\right) dt + \sigma_t \frac{\partial f}{\partial x}(t,X_t)\, dW_t$$

7.2.2 Characteristic function

The characteristic function of a distribution function $F_X(x) = P(X \leq x)$ is given by :

$$\Phi_X(u) = \mathbb{E}[\exp(iuX)] = \int_{-\infty}^{\infty} \exp(iux) dF_X(x)$$

If the density function $f_X(x)$ of X is available, one can write :

$$\Phi_X(u) = \mathbb{E}[\exp(iuX)] = \int_{-\infty}^{\infty} \exp(iux) f_X(x) dx$$

The characteristic function always exists, is continuous and it determines the distribution function uniquely. Out of the characteristic function, one can easily derive the moment of X. If $\mathbb{E}[X^k] < \infty$, then :

$$\mathbb{E}[X^k] = i^{-k} \frac{d}{du^k} \Phi_X(u)\bigg|_{u=0}$$

7.2.3 Carr-Madan formula

We here briefly introduce the Carr-Madan formula in order to :

- Compute the price of European call (and put) options based on the characteristic function of the log-asset price for different SV models.
- Compute the implied volatility of different SV models, again based on their characteristic function.

Let $k = \log(K)$ the log-strike and $C(k,T)$, the price of a European call of log-strike k and maturity T. Using Fourier transform on $c(k,T) = \exp(\alpha k) C(k,T)$ and then its inverse Fourier transform, Carr and Madan (1999) showed that :

$$C(k,T) = \exp(-\alpha k) \frac{1}{\pi} \int_0^{\infty} \exp(-iuk)\, \psi(u;T)\, du \tag{7.2.1}$$

where :

$$\psi(u;T) = \frac{\exp(-rT)\Phi(u-(\alpha+1)i, T)}{\alpha^2 + \alpha - u^2 + i(2\alpha+1)u}$$

with $\Phi(.,T)$ the characteristic function of the log-asset price at maturity T.
The integral is typically calculated using a Fast Fourier Transform. Using the FFT, one can actually calculate the prices for a whole range of strikes in one run. The price in between the computed strikes are obtained via interpolation. For more details, we refer to Schoutens (2009).

In order to derive model-based implied volatilities from the Carr-Madan Formula for different SV models, we just have to compute $\psi(u;T)$ with the B&S characteristic function, which gives :

$$\psi_{BS}(u;T) = \frac{\exp\left(-rT + iu(\log S_0 + rT)\right) \exp\left(-\frac{1}{2}(u-(\alpha+1)i)(u-\alpha\, i)\, \sigma_{BS}^2 T\right)}{\alpha^2 + \alpha - u^2 + i(2\alpha+1)u}$$

where σ_{BS} is the B&S relative volatility. From the definition of model-based implied volatilities, we must have $C_{BS}(k,T) = C(k,T)$ where $C(k,T)$ is the price of a call option obtained via a specified SV model. It implies that :

$$\int_0^\infty \exp(-iuk)(\psi(u;T) - \psi_{BS}(u;T))\, du = 0$$

Solving this equation with respect to σ_{BS} using FFT allows to derive numerically (and efficiently) the volatility surface for different SV models (using their respective characteristic function).
Having a closed (or quasi-closed) form characteristic function of the log-price is therefore a nice property to have for a pricing model.

7.2.4 p-variation

The p-variation of a random process X_t over the interval $[0,T]$ is defined as :

$$V_p(X,[0,T]) = \sup_{\Pi} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i+1}}|^p$$

where Π is a finite partition over $[0,T]$. The index of the p-variation is defined to be :

$$I(X,[0,T]) = \inf\{p > 0\,;\, V_p(X,[0,T]) < \infty\}$$

We can interpret the p-variation as a measure of the regularity or smoothness of a process $(X_t)_{t\geq 0}$ through its link with Hölder continuity.

7.2.5 Markov property

The Markov property refers to the memoryless property of a stochastic process. A stochastic process has the Markov property if the conditional probability distribution of future states of the process (conditionally on both past and present states) depends only upon the present state, not on the sequence of events that preceded it. A process with this property is called a Markov(ian) process.
Mathematically, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and denote $(X_t)_{t\geq 0}$ a stochastic process adapted to the filtration. This process is said to be Markovian if for each event $A \subset \Omega$ and $s < t$:

$$P(X_t \in A\,|\mathcal{F}_s) = P(X_t \in A\,|X_s)$$

7.2.6 Martingale and local martingale

The process $(X_t)_{t\geq 0}$ is a martingale on $(\Omega, \mathbb{P}, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$ if :

- $(X_t)_{t\geq 0}$ is adapted to the filtration :

$$\forall t\, :\, X_t \in \mathcal{F}_t$$

- $(X_t)_{t\geq 0}$ is integrable :

$$\forall t\, :\, \mathbb{E}^{\mathbb{P}}|X_t| < \infty$$

- and :

$$\forall t > s : \mathbb{E}^{\mathbb{P}}[X_t|\mathcal{F}_s] = X_s$$

We then define the stopping time T. T is a stopping time if $\{\omega : T(\omega) \leq t\} = \{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. We have that $(X_t)_{t\geq 0}$ is a martingale if and only if $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ for all stopping times T.

Finally, we can define the notion of local martingale which is a stochastic process that is locally a martingale. Formally, a process $(X_t)_{t\geq 0}$ is a local martingale if there exists a sequence of stopping times T_n with $T_n < T_{n+1}$ a.s (that is, T_n is almost surely increasing), $T_n < T$ a.s. on $\{T > 0\}$, and $\lim_{n\to\infty} T_n = T$ a.s. and moreover, $X_{t\wedge T_n} = X_{\min(t,T_n)}$ is a martingale for each n.

7.2.7 Semi-martingale

A process $(X_t)_{t\geq 0}$ is called a semi-martingale if it admits the Doob-Meyer decomposition :

$$X_t = X_0 + M_t + A_t$$

where M_t is an $\mathcal{F}_t$ local martingale with $M_0 = 0$ and A_t is a càdlàg adapted process of locally bounded variation and X_0 is $\mathcal{F}_0$-measurable. One important properties of semi-martingale is that any semi-martingale has locally bounded quadratic variation.

7.2.8 Fractional integrals and derivatives

We provide here the basis of the Fractional Calculus. We start from the following n-tuple iterated integral with Φ an integrable function on $[a, b]$, with $t_n \in [a, b]$ and $n \geq 1$:

$$\int_a^{t_n} ... \Big\{ \int_a^{t_2} \Big\{ \int_a^{t_1} \Phi(u)\, du \Big\} dt_1 \Big\} ...\, dt_{n-1} = \frac{1}{(n-1)!} \int_a^{t_n} \Phi(u)(t_n - u)^{n-1} du$$

This can be easily shown by induction (Cauchy formula). The idea behind fractional integrals is then to replace the integer n by a real number $\alpha > 0$. It follows the definition

$$(I^\alpha \Phi)(s) := \frac{1}{\Gamma(\alpha)} \int_a^s \Phi(u)(s-u)^{\alpha-1}\, du \tag{7.2.2}$$

with $\Phi \in L^1[a, b]$, $s \in [a, b]$ and Γ the gamma function.

In order to then obtain fractional derivatives, one needs to extend the definition of fractional integrals to $\alpha < 0$. We cannot simply substitute $\alpha < 0$ into (7.2.2) since the resulting integral diverges. We then consider the equation with $0 < \alpha < 1$

$$(I^\alpha \Phi)(s) = \frac{1}{\Gamma(\alpha)} \int_a^s \Phi(u)(s-u)^{\alpha-1}\, du = f(u)$$

We want to solve it for functions Φ assuming f is well-behaved. See Pipiras and Taqqu (2017, p.348) for the proof. The integral is then defined with $0 < \alpha < 1$ by

$$(D^\alpha f)(u) := (I^{-\alpha} f)(u) := \frac{1}{\Gamma(1-\alpha)} \frac{d}{du} \int_a^u f(s)(u-s)^{-\alpha} ds \tag{7.2.3}$$

and is called fractional derivatives of order α. The operator $I^{-\alpha}$ is introduced as the inverse of I^α. The definitions (7.2.2) and (7.2.3) can be extended on the Real line by replacing a with $-\infty$.

7.3 Demonstration

7.3.1 Demo. 1

Remember that we have under the physical measure $\mathbb{P}$:

$$dS_t = \mu(S_t, t)\, S_t\, dt + \sigma_t\, S_t\, dW_t$$

with the stochastic volatility σ_t satisfying :

$$d\sigma_t = a(\sigma_t, t) + b(\sigma_t, t)\, d\hat{W}_t$$

where W_t and $\hat{W}_t$ are one-dimensional Brownian motions defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with $d\langle W, \hat{W}\rangle_t = \rho\, dt$ (where ρ is a constant). We can write :

$$\hat{W}_t = \rho\, W_t + \sqrt{1-\rho^2}\, W_t^{'} = \rho\, W_t + \rho'\, W_t^{'}$$

where $W_t^{'}$ is an independent Brownian motion of W_t.
By the one-dimensional Girsanov's theorem and the independence between W_t and W_t', we have that the following process is a martingale :

$$\left.\frac{d\mathbb{P}^*}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left(-\int_0^t \Phi_s\, dW_s - \frac{1}{2}\int_0^t \Phi_s^2\, ds - \int_0^t \lambda_s\, dW_s' - \frac{1}{2}\int_0^t \lambda_s^2\, ds\right)$$

Where $(\Phi_t)_{t\geq 0}$ and $(\lambda_t)_{t\geq 0}$ are adapted processes such that $\int_0^t \Phi_s^2\, ds < +\infty$ and $\int_0^t \lambda_s^2\, ds < +\infty$ for $0 \leq t \leq +\infty$.
Then the processes $W_t^* = W_t + \int_0^t \Phi_s\, ds$ and $W_t^{'*} = W_t^{'} + \int_0^t \lambda_s\, ds$ are Brownian motions under $\mathbb{P}^*$.
It follows that :

$$dS_t = (\mu(S_t,t) - \Phi_t\,\sigma_t)\, S_t\, dt + \sigma_t\, S_t\, (dW_t + \Phi_t\, dt)$$

$$d\sigma_t = \big(a(\sigma_t,t) - b(\sigma_t,t)\,\rho\,\Phi_t - b(\sigma_t,t)\,\rho'\lambda_t\big)\, dt + b(\sigma_t,t)\left(\rho\, dW_t + \rho'\, dW_t^{'} + \rho\,\Phi_t\, dt + \rho'\lambda_t\, dt\right)$$

Now if we take $\Phi_t = (\mu(S_t,t) - r)/\sigma_t$ and $\tilde{a}(\sigma_t,t) = a(\sigma_t,t) - b(\sigma_t,t)\,\rho\,\Phi_t - b(\sigma_t,t)\,\rho'\lambda_t$, we finally have:

$$dS_t = rS_t\, dt + \sigma_t\, S_t\, dW_t^*$$

$$d\sigma_t = \tilde{a}(\sigma_t,t)dt + b(\sigma_t,t)\,(\rho\, dW_t^* + \rho' dW_t^{'*}) = \tilde{a}(\sigma_t,t)dt + b(\sigma_t,t)\, d\hat{W}_t^*$$

where $\hat{W}_t^* = \rho\, dW_t^* + \rho' dW_t^{'*}$ is again a Brownian motion under $\mathbb{P}^*$.

Note that we could also have written directly $\hat{W}_t^* = \hat{W}_t + \int_0^t \Lambda_s\, ds$ and have :

$$d\sigma_t = (a(\sigma_t,t) - b\,\Lambda_t)\, dt + b(\sigma_t,t) d\hat{W}_t^*$$

Therefore, $a(\sigma_t,t) - b\,\rho\,\Phi_t - b\,\rho'\,\lambda_t = a(\sigma_t,t) - b\,\Lambda_t = \tilde{a}(\sigma_t,t)$. This leads to the footnote 1 p7.

7.3.2 Demo. 2

By self-similarity and stationarity of the increments :

$$\begin{aligned}
\mathbb{E}(X_tX_s) &= \frac{1}{2}\Big\{\mathbb{E}(X_t^2) + \mathbb{E}(X_s^2) - \mathbb{E}\left((X_t - X_s)^2\right)\Big\} \\
&= \frac{1}{2}\Big\{\mathbb{E}(X_t^2) + \mathbb{E}(X_s^2) - \mathbb{E}(X_{|t-s|}^2)\Big\} \\
&= \frac{1}{2}\Big\{t^{2H} + s^{2H} - |t-s|^{2H}\Big\}\mathbb{E}(X_1^2)
\end{aligned}$$

7.3.3 Demo. 3

We relied on Hainaut (2020) for the coming proof. Semi-martingales (see Concept 7.2.7) form the largest class of processes with respect to which the Itô integral can be defined. However, fBm are not semi-martingales and other constructions have to be followed. In order to show this we use the concept of p-variation (see Concept 7.2.4). The following result is given without proof :
The index of the p-variation of the fractional Brownian motion is :

$$I(B^H,[0,T]) = \frac{1}{H}$$

Moreover, $V_p(B^H,[0,T]) = 0$ when $pH > 1$ and $V_p(B^H,[0,T]) = \infty$ when $pH < 1$.

We then use the fact that any semi-martingale has locally bounded quadratic variation. If $H \in (0,1/2)$, then B_t^H cannot even be a martingale since it has infinite quadratic variation, hence, it is not a semi martingale.

If $H \in (1/2,1)$, then the quadratic variation of B_t^H is zero. Let's first suppose that it is a semi-martingale. Then, $M_t = B_t^H - A_t$ has quadratic variation equal to zero. So, $M_t = 0$ for all t almost surely. In this case, $B_t^H = A_t$, but this is not possible since B_t^H has unbounded variation. Hence, B_t^H is not a semi-martingale for any $H \neq 1/2$.

7.3.4 Demo. 4

We start by showing that the autocovariance function of the log-volatility process X_t^H has the following form for $\lambda \to 0$:

$$\mathrm{Cov}[X_t^H, X_{t+\Delta}^H] = \mathrm{Var}[X_t^H] - \frac{1}{2}\nu^2\Delta^{2H} + o(1) \tag{7.3.1}$$

We use the fact that :

$$\begin{aligned}
\mathbb{E}[(X_{t+\Delta}^H - X_t^H)^2] &= \mathbb{E}[(X_{t+\Delta}^H)^2] + \mathbb{E}[(X_t^H)^2] - 2\mathbb{E}[X_{t+\Delta}^H X_t^H] \\
&= \mathrm{Var}[X_{t+\Delta}^H] + \mathrm{Var}[X_t^H] - 2\mathrm{Cov}[X_{t+\Delta}^H, X_t^H] \\
&= 2\mathrm{Var}[X_t^H] - 2\mathrm{Cov}[X_{t+\Delta}^H, X_t^H] \qquad \text{by stationarity of } (X_t^H)_{t\geq 0}
\end{aligned}$$

and that for $\lambda \to 0$:

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

where K_2 is a constant.

Secondly, for the volatility process, we have :

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

Where we use the fact that $(X_t^H)_{t\geq 0}$ is a Gaussian process.
We finally replace equation (7.3.1) in the last expression and we find approximately

$$\mathbb{E}[\sigma_{t+\Delta}\sigma_t] \approx e^{2\mathbb{E}[X_t^H] + 2\,\mathrm{Var}[X_t^H]} e^{-\nu^2 \frac{\Delta^{2H}}{2}}$$

Note that $\mathbb{E}[X_t^H] = \mathbb{E}[X_{t+\Delta}^H]$ and $\mathrm{Var}[X_t^H] = \mathrm{Var}[X_{t+\Delta}^H]$ since $(X_t^H)_{t\geq 0}$ is a stationary process.

7.3.5 Demo. 5

Let $M_t(u) = \int_t^u \frac{1}{(u-s)^\gamma} dW_s$ and $Z_t(u) = \int_{-\infty}^t \left[\frac{1}{(u-s)^\gamma} - \frac{1}{(t-s)^\gamma}\right] dW_s$. We also define $\tilde{W}_t(u) = \sqrt{2H}\int_t^u \frac{dW_s}{(u-s)^\gamma}$ and $\eta = 2\nu\, C_H/\sqrt{2H}$.
If we take the mathematical expectation under $\mathbb{P}$, we have :

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}(v_u|\mathcal{F}_t) &= \mathbb{E}^{\mathbb{P}}\Big[v_t\, \exp\left(2\,\nu\, C_H\{M_t(u) + Z_t(u)\}\right)|\mathcal{F}_t\Big] \\
&= v_t\, \exp(2\,\nu\, C_H Z_t(u))\, \mathbb{E}^{\mathbb{P}}\Big[\exp\left(2\,\nu\, C_H M_t(u)\right)|\mathcal{F}_t\Big] \\
&= v_t\, \exp(2\,\nu\, C_H Z_t(u))\, \exp\left(\mathbb{E}^{\mathbb{P}}\Big[2\,\nu\, C_H M_t(u)|\mathcal{F}_t\Big] + \frac{1}{2}\mathbb{V}^{\mathbb{P}}\Big[2\,\nu\, C_H M_t(u)|\mathcal{F}_t\Big]\right) \\
&= v_t\, \exp(2\,\nu\, C_H Z_t(u))\, \exp\left(\frac{1}{2}\mathbb{E}^{\mathbb{P}}\Big[|\eta\tilde{W}_t(u)|^2\Big]\right) \\
&= v_t\, \exp\left(2\,\nu\, C_H Z_t(u)) + \frac{1}{2}\eta^2\, \mathbb{E}^{\mathbb{P}}\Big[|\tilde{W}_t(u)|^2\Big]\right)
\end{aligned}$$

since $Z_t(u)$ is $\mathcal{F}_t$-measurable and $M_t(u)$ is independent of $\mathcal{F}_t$. We can then rewrite under $\mathbb{P}$:

$$\begin{aligned}
v_u &= v_t \exp\left(2\,\nu\, C_H\{M_t(u) + Z_t(u)\}\right) \\
&= v_t \exp\Big(\eta\,\tilde{W}_t(u) + 2\,\nu\, C_H Z_t(u)\Big) \\
&= \mathbb{E}^{\mathbb{P}}(v_u|\mathcal{F}_t) \exp\left(\eta\,\tilde{W}_t(u) - \frac{1}{2}\mathbb{E}^{\mathbb{P}}\Big[|\eta\tilde{W}_t(u)|^2\Big]\right) \\
&= \mathbb{E}^{\mathbb{P}}(v_u|\mathcal{F}_t)\, \mathcal{E}\left(\eta\tilde{W}_t(u)\right)
\end{aligned}$$

Where $\mathcal{E}(.)$ is the Wick stochastic exponential.
Finally, under the general change of measure $W_t^* = W_t + \lambda_t\, t$ where λ_t is deterministic, we have under $\mathbb{P}^*$:

$$\begin{aligned} v_u &= \mathbb{E}^{\mathbb{P}}[v_u|\mathcal{F}_t]\, \exp\left\{\eta\sqrt{2H}\int_t^u \frac{1}{(u-s)^\gamma} dW_s - \frac{\eta^2}{2}(u-t)^{2H}\right\} \\ &= \mathbb{E}^{\mathbb{P}}[v_u|\mathcal{F}_t]\, \mathcal{E}\left(\eta\tilde{W}_t^*(u)\right)\, \exp\left\{\eta\sqrt{2H}\int_t^u \frac{\lambda_s}{(u-s)^\gamma} ds\right\} \\ &= \mathbb{E}^{\mathbb{P}^*}[v_u|\mathcal{F}_t]\, \mathcal{E}\left(\eta\tilde{W}_t^*(u)\right) \end{aligned}$$

where $\mathbb{E}^{\mathbb{P}^*}[v_u|\mathcal{F}_t] = \xi_t(u) = \mathbb{E}^{\mathbb{P}}[v_u|\mathcal{F}_t]\, \exp\left\{\eta\sqrt{2H}\int_t^u \frac{\lambda_s}{(u-s)^\gamma} ds\right\}$ is the forward variance curve.

7.3.6 Demo. 6

- Heston : The variance process in the Heston model is given by the following CIR process :

$$dv_t = \lambda(\eta - v_t)\, dt + \nu\, \sqrt{v_t}\, d\hat{W}_t^* \tag{7.3.2}$$

With $\xi_t(u) = \mathbb{E}^{\mathbb{P}^*}[v_u|\mathcal{F}_t]$, we take the conditional expectation of the SDE 7.3.2 to get :

$$d\xi_t(u) = \lambda(\eta - \xi_t(u))\, du$$

The solution of this SDE is then :

$$\xi_t(u) = -(\eta - \xi_t(t))e^{-\lambda(u-t)} + \eta = (v_t - \eta)e^{-\lambda(u-t)} + \eta$$

From this last equation, we find :

$$d\xi_t(u) = \left(\lambda(\eta - v_t)\, dt + \nu\, \sqrt{v_t}\, d\hat{W}_t^*\right) e^{-\lambda(u-t)} + v_t\, \lambda\, e^{-\lambda(u-t)} dt - \eta\, \lambda\, e^{-\lambda(u-t)}$$

All the terms in dt cancel each other out and we finally have :

$$d\xi_t(u) = \nu\, e^{-\lambda(u-t)} \sqrt{v_t}\, d\hat{W}_t^*$$

which is well a martingale in t.

- Rough Heston : For the rough Heston model, from equation (4.2.11) with $\lambda = 0$, we have that $\xi_t(u) = \mathbb{E}^{\mathbb{P}^*}[v_u|\mathcal{F}_t] = \theta^t(u)$ and :

$$v_u = \xi_t(u) + \frac{1}{\Gamma(H+1/2)} \int_t^u (u-s)^{H-1/2}\, \nu\, \sqrt{v_s}\, d\hat{W}_s^*$$

and :

$$v_u = \xi_{t+h}(u) + \frac{1}{\Gamma(H+1/2)} \int_{t+h}^u (u-s)^{H-1/2}\, \nu\, \sqrt{v_s}\, d\hat{W}_s^*$$

Substracting these two equations gives :

$$\xi_{t+h}(u) - \xi_t(u) = \frac{1}{\Gamma(H+1/2)} \int_t^{t+h} (u-s)^{H-1/2}\, \nu\, \sqrt{v_s}\, d\hat{W}_s^*$$

Finally, taking the limit $h \to 0$, we obtain :

$$d\xi_t(u) = \frac{\nu}{\Gamma(H+1/2)} (u-t)^{H-1/2} \sqrt{v_t}\, d\hat{W}_t^*$$

7.3.7 Demo. 7

Recall first that the fair value at time $t = 0$ is given by :

$$FV_0 = \mathbb{E}\left[(1 - N_T)\, e^{-rT}\, \mathbb{E}^{\mathbb{P}^*}[F_T^s \,|\, \mathcal{F}_0] + \int_0^T e^{-rt}\, \mathbb{E}^{\mathbb{P}^*}[F_t^d \,|\, \mathcal{F}_0]\, dN_t \;\middle|\; \mathcal{G}_0\right]$$

The processes $(N_t)_{t\geq 0}$ and $(S_t)_{t\geq 0}$ being independent, we can write :

$$FV_0 = \mathbb{E}[1 - N_T \,|\, \mathcal{G}_0]\, e^{-rT}\, \mathbb{E}^{\mathbb{P}^*}[F_T^s \,|\, \mathcal{F}_0] + \int_0^T e^{-rt}\, \mathbb{E}^{\mathbb{P}^*}[F_t^d \,|\, \mathcal{F}_0]\, \mathbb{E}[dN_t \,|\, \mathcal{G}_0]$$

We then have that :

$$\mathbb{E}[1-N_T \,|\, \mathcal{G}_0] = 1-\mathbb{E}[N_T] = 1-(0 \cdot P(N_T = 0) + 1 \cdot P(N_T = 1)) = 1-P(N_T = 1) = P(N_T = 0) = {}_Tp_x$$

and :

$$\begin{aligned}
\mathbb{E}[dN_t \,|\, \mathcal{G}_0] &= 1 \cdot P(dN_t = 1) + 0 \cdot P(dN_t = 0) \\
&= P(dN_t = 1 \,|\, N_{t^-} = 0)P(N_{t^-} = 0) + P(dN_t = 1 \,|\, N_{t^-} = 1)P(N_{t^-} = 1) \\
&= P(dN_t = 1 \,|\, N_{t^-} = 0)P(N_{t^-} = 0) + 0 \\
&= \mu_x(t)\, dt\, {}_tp_x
\end{aligned}$$

which finally gives :

$$FV_0 = {}_Tp_x\, e^{-rT}\, \mathbb{E}^{\mathbb{P}^*}[F_T^s] + \int_0^T e^{-rt}\, \mathbb{E}^{\mathbb{P}^*}[F_t^d]\, {}_tp_x\, \mu_x(t)dt$$

Bibliography

[1] Torben G Andersen et al. "The distribution of realized stock return volatility". In: *Journal of financial economics* 61.1 (2001), pp. 43–76.

[2] Anna Rita Bacinello, Enrico Biffis, and Pietro Millossovich. "Regression-based algorithms for life insurance contracts with surrender guarantees". In: *Quantitative Finance* 10.9 (2010), pp. 1077–1090.

[3] Emmanuel Bacry and Jean-François Muzy. "Log-infinitely divisible multifractal processes". In: *Communications in Mathematical Physics* 236.3 (2003), pp. 449–475.

[4] Ole E Barndorff-Nielsen and Jürgen Schmiegel. "Brownian semistationary processes and volatility/intermittency". In: *Advanced financial modelling* 8 (2009), pp. 1–26.

[5] Ole E Barndorff-Nielsen and Neil Shephard. "Econometric analysis of realized volatility and its use in estimating stochastic volatility models". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64.2 (2002), pp. 253–280.

[6] David S Bates. "Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options". In: *The Review of Financial Studies* 9.1 (1996), pp. 69–107.

[7] Christian Bayer, Peter Friz, and Jim Gatheral. "Pricing under rough volatility". In: *Quantitative Finance* 16.6 (2016), pp. 887–904.

[8] Christian Bayer et al. "On deep calibration of (rough) stochastic volatility models". In: *arXiv preprint arXiv:1908.08806* (2019, b).

[9] Christian Bayer et al. "Short-time near-the-money skew in rough fractional volatility models". In: *Quantitative Finance* 19.5 (2019, a), pp. 779–798.

[10] Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. "Decoupling the short-and long-term behavior of stochastic volatility". In: *arXiv preprint arXiv:1610.00332* (2016).

[11] Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. "Hybrid scheme for Brownian semistationary processes". In: *Finance and Stochastics* 21.4 (2017), pp. 931–965.

[12] Lorenzo Bergomi and Julien Guyon. "Stochastic volatility's orderly smiles". In: *Risk* 25.5 (2012), p. 60.

[13] Peter Carr and Dilip Madan. "Option valuation using the fast Fourier transform". In: *Journal of computational finance* 2.4 (1999), pp. 61–73.

[14] Patrick Cheridito, Hideyuki Kawaguchi, and Makoto Maejima. "Fractional ornstein-uhlenbeck processes". In: *Electronic Journal of probability* 8 (2003).

[15] Alexandra Chronopoulou and Frederi G Viens. "Estimation and pricing under long-memory stochastic volatility". In: *Annals of finance* 8.2-3 (2012), pp. 379–403.

[16] Fabienne Comte and Eric Renault. "Long memory in continuous-time stochastic volatility models". In: *Mathematical finance* 8.4 (1998), pp. 291–323.

[17] Rama Cont. "Volatility clustering in financial markets: empirical facts and agent-based models". In: *Long memory in economics*. Springer, 2007, pp. 289–309.

[18] Emanuel Derman and Iraj Kani. "Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility". In: *International journal of theoretical and applied finance* 1.01 (1998), pp. 61–110.

[19] Zhuanxin Ding, Clive WJ Granger, and Robert F Engle. "A long memory property of stock market returns and a new model". In: *Journal of empirical finance* 1.1 (1993), pp. 83–106.

[20] Bruno Dupire. "A unified theory of volatility". In: *Derivatives pricing: The classic collection* (1996), pp. 185–196.

[21] Bruno Dupire. "Pricing with a smile". In: *Risk* 7.1 (1993), pp. 18–20.

[22] Omar El Euch, Jim Gatheral, and Mathieu Rosenbaum. "Roughening heston". In: *Risk* (2019), pp. 84–89.

[23] Omar El Euch and Mathieu Rosenbaum. "The characteristic function of rough Heston models". In: *Mathematical Finance* 29.1 (2019), pp. 3–38.

[24] Martin Forde and Hongzhong Zhang. "Asymptotics for rough stochastic volatility models". In: *SIAM Journal on Financial Mathematics* 8.1 (2017), pp. 114–145.

[25] Masaaki Fukasawa. "Asymptotic analysis for stochastic volatility: martingale expansion". In: *Finance and Stochastics* 15.4 (2011), pp. 635–654.

[26] Masaaki Fukasawa. "Short-time at-the-money skew and rough fractional volatility". In: *Quantitative Finance* 17.2 (2017), pp. 189–198.

[27] Jim Gatheral. "Rough Volatility". In: *18th Winter school of Mathematical Finance*. Jan. 2019.

[28] Jim Gatheral. *The volatility surface: a practitioner's guide*. Vol. 357. John Wiley & Sons, 2011.

[29] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. "Volatility is rough". In: *Quantitative Finance* 18.6 (2014), pp. 933–949.

[30] Jim Gatheral and Martin Keller-Ressel. "Affine forward variance models". In: *Finance and Stochastics* 23.3 (2019), pp. 501–533.

[31] Jim Gatheral and Radoš Radoičić. "Rational approximation of the rough Heston solution". In: *International Journal of Theoretical and Applied Finance* 22.03 (2019), p. 1950010.

[32] Eric Ghysels, Andrew C Harvey, and Eric Renault. "5 Stochastic volatility". In: *Handbook of statistics* 14 (1996), pp. 119–191.

[33] Tilmann Gneiting and Martin Schlather. "Stochastic models that separate fractal dimension and the Hurst effect". In: *SIAM review* 46.2 (2004), pp. 269–282.

[34] Donatien Hainaut. *Financial modelling for risk management*. Springer, 2020.

[35] Steven L Heston. "A closed-form solution for options with stochastic volatility with applications to bond and currency options". In: *The review of financial studies* 6.2 (1993), pp. 327–343.

[36] John Hull and Alan White. "Libor versus OIS: The Derivatives Discounting Dilemma". In: *Journal of Investment Management (JOIM), Third Quarter* (2013).

[37] John Hull and Alan White. "The pricing of options on assets with stochastic volatilities". In: *The journal of finance* 42.2 (1987), pp. 281–300.

[38] Ioannis Karatzas and Steven E Shreve. "Brownian motion". In: *Brownian Motion and Stochastic Calculus*. Springer, 1998, pp. 47–127.

[39] Takayuki Kawada and Norio Kôno. "A remark on nowhere differentiability of sample functions of Gaussian processes". In: *Proceedings of the Japan Academy* 47.SupplementII (1971), pp. 932–934.

[40] Peter Leoni. *The Greeks and Hedging Explained*. Springer, 2014.

[41] Benoit B Mandelbrot and John W Van Ness. "Fractional Brownian motions, fractional noises and applications". In: *SIAM review* 10.4 (1968), pp. 422–437.

[42] Ryan McCrickerd and Mikko S Pakkanen. "Turbocharging Monte Carlo pricing for the rough Bergomi model". In: *Quantitative Finance* 18.11 (2018), pp. 1877–1886.

[43] Rui Vilela Mendes and Javier Ordóñez Monfort. "A data-reconstructed fractional volatility model". In: *Economics Discussion paper* 2008-22 (2008).

[44] Thibaut Moyaert and Mikael Petitjean. "The performance of popular stochastic volatility option pricing models during the subprime crisis". In: *Applied Financial Economics* 21.14 (2011), pp. 1059–1068.

[45] Marek Musiela and Marek Rutkowski. "Martingale methods in financial modeling". In: *Stochastic Modelling and Applied Probability* 36 (1995).

[46] Vladas Pipiras and Murad S Taqqu. *Long-range dependence and self-similarity*. Vol. 45. Cambridge university press, 2017.

[47] Marc Romano and Nizar Touzi. "Contingent claims and market completeness in a stochastic volatility model". In: *Mathematical Finance* 7.4 (1997), pp. 399–412.

[48] Wim Schoutens. "The World of VG". In: *Actuarial and financial mathematics conference*. 2009, pp. 32–38.

[49] Peter Tankov. *Financial modelling with jump processes*. CRC press, 2003.