Ecole Polytechnique
Centre de Mathématiques appliquées

VIX calibration using rough volatility models

Under the supervision :
Pr. Stéfano De Marco
Pr. Thibaut Mastrolia

Authors:
Abdelhafid Souilmi
Walid Jebbour

March 16, 2021

Contents

Chapter 1

Bergomi and Rough volatility models

1.1 Introduction and context

It was shown recently [1] that rough volatility models are a very relevant dynamics to model the behavior of the volatility process. In particular, we can note that the classical models do not behave well in regards to the implied volatility surface and the implied volatility-skew at the money, whereas rough volatility models are better at reproducing empirical implied volatility surfaces. Following this finding, multiple adaptations of classical models emerged. These new models incorporated rougher dynamics compared to the usual Brownian motion.

In this work, we are interested in the Bergomi model and its rougher versions. These models have a good behavior in regards to the implied volatility surface of S&P 500 options. Nonetheless, they are unable to reproduce the smile for volatility options. We will use these models to price volatility options, show that neither the Bergomi model nor its rough version can capture the VIX option smile observed in the market, then we would move on to an extented version of the rough Bergomi model (a two factor version) and show that indeed it helps us in successfully capturing the VIX smile.

1.2 Bergomi Forward Variance Model

We denote ξ_t^u the forward variance at time $u > t$ seen from t. The One-factor Bergomi model [2] [3] assumes that the forward variance is subject to the following dynamics :

$$\frac{d\xi_t^u}{\xi_t^u} = \alpha e^{-k(u-t)} dZ_t$$

We note that the instantaneous variance is given by $\sigma_t = \xi_t^t$ and the price process S_t is given by :

$$\frac{dS_t}{S_t} = (r_t - q_t)dt + \sqrt{\xi_t^t} dW_t$$
$$Cov(dZ_t, dW_t) = \rho dt$$

The dynamic of ξ_t^u is special because it allows us to derive a Markov representation, given by:

$$\xi_t^u = \xi_0^u f_u(t, X_t, v_t)$$

With :

$$f_u(t, x, v) = exp\left(\alpha e^{-k(u-t)}x - \frac{\alpha^2}{2}e^{-2k(u-t)}v\right)$$

$$v_t = Var(X_t) = \frac{1 - e^{-2kt}}{2k}$$

and X_t is the Ornstein Uhlenbeck process defined with the following dynamics :

$$\begin{cases} dX_t = -kX_t dt + dZ_t, \\ X_0 = 0 \end{cases}$$

The multiple factor Bergomi model is very similar to the One factor model above. The dynamics of ξ_t^u in the multiple factor Bergomi model are given by :

$$\frac{d\xi_t^u}{\xi_t^u} = \delta N_\alpha \sum_i \alpha_i e^{-k_i(u-t)} dZ_t^i$$

with

$$(Z_t^i)_{i\in\{1,\ldots,N\}} \text{ N brownian motions such that, } Cov(dZ_t^i, dZ_t^j) = \rho_{ij},$$

$$N_\alpha = \frac{1}{\sqrt{\sum_{ij}\alpha_i\alpha_j\rho_{ij}}} \text{ and } \sum_i \alpha_i = 1.$$

For N=2 (two-factor Bergomi model), we find :

$$\frac{d\xi_t^u}{\xi_t^u} = \delta N_\alpha(\alpha_1 e^{-k_1(u-t)}dZ_t^1 + \alpha_2 e^{-k_2(u-t)}dZ_t^2)$$

$$N_\alpha = (\alpha_1^2 + 2\rho\alpha_1\alpha_2 + \alpha_2^2)^{\frac{-1}{2}}$$

$$\alpha_1 + \alpha_2 = 1, \quad k_1, k_2 > 0, \quad \alpha_1, \alpha_2 \in [0, 1]$$

We have a Markov representation of ξ_t^u:

$$\xi_t^u = \xi_0^u f_u(t, x_t^u, v_t(u))$$

$$x_t^u = N_\alpha \left\{\alpha_1 e^{-k_1(u-t)}X_t^1 + \alpha_2 e^{-k_2(u-t)}X_t^2\right\}$$

with :

$$
\begin{aligned}
v_t(u) &= Var(x_t^u) \\
&= N_\alpha^2(\alpha_1^2 e^{-2k_1(u-t)}v_t^1 + \alpha_2^2 e^{-2k_2(u-t)}v_t^2 \\
&\qquad + 2\alpha_1\alpha_2 e^{-(k_1+k_2)(u-t)}v_t^{1,2})
\end{aligned}
$$

where v_t^i are given by :

$$
\begin{cases} v_t^i = \frac{1-e^{-2k_i t}}{2k_i}, \\ v_t^{1,2} = \rho\frac{1-e^{-(k_1+k_2)t}}{k_1+k_2} \end{cases}
$$

and X_t^i are Ornstein Uhlenbeck processes given by :

$$
\begin{cases} dX_t^i = -k_i X_t^i dt + dZ_t^i, \\ X_0^i = 0 \end{cases}
$$

1.3 Rough Bergomi Model

We recall that for a call pricing model C (function of T maturity and strike K), the Black Scholes implied volatility is defined as the solution σ_{BS} of the following equation :

$$
C(K,T) = Call_{BS}(S_0, K, T, \sigma_{BS}(T,K), r)
$$

where $Call_{BS}$ is the Black-Scholes call price, S_0 is the initial price of the underlying asset and r is the interest rate.

We can define the implied volatility function as : $(T,K) \mapsto \sigma_{BS}(T,K)$.

We will give a brief explanation why the rough Bergomi model should be considered as is. First, we define the ATM (at the money) volatility skew as (with $\tau = T-t$ and $k = logK$) :

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

The N-factor Bergomi model generates a term structure of volatility skew that has the qualitative form :

$$
\psi(\tau) \sim \sum_i \frac{\alpha_i}{k_i\tau}\left(1 - \frac{1-e^{-k_i\tau}}{k_i\tau}\right)
$$

However, empirical data shows that $\psi(\tau) \sim \frac{1}{\tau^\alpha}$, with α small (~ 0.1). The term structure of the N-factor Bergomi model doesn't match the empirical one. But, if we consider the following dynamic for the forward variance [4]:

$$
\frac{d\xi_t^u}{\xi_t^u} = \frac{2\alpha}{(u-t)^\gamma}dW_t
$$

We find a volatility skew ATM of the form : $\psi(\tau) \sim \frac{1}{\tau^\gamma}$. To match empirical data, γ should be small too. We can see that with such a small parameter, we end up with processes that are rougher (in particular in terms of Hölder regularity) than the ones found with the previous model, which justifies the name : Rough Bergomi. From now on, we will refer to this model by rBergomi.

Chapter 2

Simulating Bergomi and Rough Bergomi

Later on, we will need ξ_t^u for a fixed t and varying u. So, we will use a discretization scheme for u. Let's suppose we already have a grid $(u_1, ..., u_N)$ and that we are looking for the values of $(\xi_t^{u_i})_{1\leq i\leq N}$.

2.1 Bergomi Model

We will only show the simulation of the two-factor Bergomi model (for the classical Bergomi models). We will mainly use the Markov property shown above to simulate the forward variance. We recall that :

$$\frac{d\xi_t^u}{\xi_t^u} = \delta N_\alpha(\alpha_1 e^{-k_1(u-t)}dZ_t^1 + \alpha_2 e^{-k_2(u-t)}dZ_t^2)$$

$$N_\alpha = (\alpha_1^2 + 2\rho\alpha_1\alpha_2 + \alpha_2^2)^{\frac{-1}{2}}$$

$$\alpha_1 + \alpha_2 = 1, \quad k_1, k_2 > 0, \quad \alpha_1, \alpha_2 \in [0,1]$$

$$Cov(dZ_t^1, dZ_t^2) = \rho dt$$

The Markov representation of ξ_t^u is given by:

$$\xi_t^u = \xi_0^u f_u(t, x_t^u, v_t(u))$$

We will specify the terms present in this formula :

- First, we start with x_t^u:

$$x_t^u = N_\alpha \left\{\alpha_1 e^{-k_1(u-t)}X_t^1 + \alpha_2 e^{-k_2(u-t)}X_t^2\right\}$$

and X_t^i are Ornstein Uhlenbeck (OU) processes given by :

$$\begin{cases} dX_t^i = -k_i X_t^i dt + dZ_t^i, \\ X_0^i = 0 \end{cases}$$

- Variance term $v_t(u)$:

$$\begin{aligned} v_t(u) &= Var(x_t^u) \\ &= N_\alpha^2(\alpha_1^2 e^{-2k_1(u-t)} v_t^1 + \alpha_2^2 e^{-2k_2(u-t)} v_t^2 \\ &\qquad + 2\alpha_1\alpha_2 e^{-(k_1+k_2)(u-t)} v_t^{1,2}) \end{aligned}$$

where v_t^i are given by :

$$\begin{cases} v_t^i = \frac{1-e^{-2k_i t}}{2k_i}, \\ v_t^{1,2} = \rho\frac{1-e^{-(k_1+k_2)t}}{k_1+k_2} \end{cases}$$

We can see that we only need to simulate the two OU processes $X_t^i,\ i = 1, 2$ (at time t only, since t is fixed as we mentioned earlier), so that we can get the values of ξ_t^u. We can do this through a discretization scheme. For $t_n = n\Delta t$ ($\Delta = t/N$), we have

$$\begin{aligned} X_{t_{n+1}}^i &= (1 - k_i\Delta t)X_{t_n} + dZ_{t_n}^i \\ dZ_{t_n}^1 &= dB_{t_n}^1 \\ dZ_{t_n}^2 &= \rho dB_{t_n}^1 + \sqrt{1-\rho^2} dB_{t_n}^2 \end{aligned}$$

where B^1 and B^2 are two independant standard brownian motions.

Using this procedure, we get the values of X_t^i, for $i = 1, 2$. We can get the values of the forward variance for maturities $(u_1, u_2, ..., u_n)$ (ie $(\xi_t^{u_1}, \xi_t^{u_2}, ..., \xi_t^{u_n})$) by calculating $(x_t^{u_i})_i$ first, then substituting it in the previous formula :

$$x_t^{u_k} = N_\alpha \left\{ \alpha_1 e^{-k_1(u_k-t)} X_t^1 + \alpha_2 e^{-k_2(u_k-t)} X_t^2 \right\}$$

$$\xi_t^{u_k} = \xi_0^{u_k} f_{u_k}(t, x_t^{u_k}, v_t(u_k))$$

Remark : We note that the only term in the formula that we don't know up to this point is ξ_0^u. We will consider from now on that ξ_0^u is a constant so the final formula is :

$$\xi_t^{u_k} = \xi_0 f_{u_k}(t, x_t^{u_k}, v_t(u_k))$$

2.2 rBergomi Model

We know that for the rBergomi model, the forward variance can be expressed as follows :

$$\frac{d\xi_t^u}{\xi_t^u} = \frac{2\alpha}{(u-t)^\gamma} dW_t$$

We define $X_{t,u}$ by :

$$X_{t,u} = \int_0^t \frac{2\alpha}{(u-s)^\gamma} dWs - \int_0^t \frac{2\alpha^2}{(u-s)^{2\gamma}} ds, \qquad H = 0.5 - \gamma$$

Using Itô's formula on $\xi_t(u)exp(-X_{t,u})$, we find :

$$\xi_t(u) = \xi_0(u) exp\left(X_{t,u}\right)$$

Again, we will consider that $\xi_0(u)$ is independent of u, and we will denote it from now on by ξ_0.

Using Itô's Isometry Property, we find (for $u_1 < u_2$):

$$\begin{aligned} Cov(X_{t,u_1}, X_{t,u_2}) &= Cov(\int_0^t \frac{2\alpha}{(u_1-s)^\gamma} dWs, \int_0^t \frac{2\alpha}{(u_2-s)^\gamma} dWs) \\ &= \int_0^t \frac{4\alpha^2}{((u_1-s)(u_2-s))^{\frac{1}{2}-H}} ds \end{aligned}$$

As suggested in [5], we will use the hypergeometric function to calculate this quantity. We recall that the hypergeometric function is defined as follows :

$${}_2F_1(a,b;c;z) = \frac{1}{\mathrm{B}(b,c-b)} \int_0^1 x^{b-1}(1-x)^{c-b-1}(1-zx)^{-a}\, dx \qquad \Re(c) > \Re(b) > 0$$

where B is the beta function.

We can now find an easier expression of the covariance function, for $u_1 < u_2$:

$$\begin{aligned} Cov(X_{t,u_1}, X_{t,u_2}) &= 4\alpha^2 \int_0^t (u_1-s)^{H-\frac{1}{2}} (u_2-s)^{H-\frac{1}{2}} ds \\ &= 4\alpha^2 \int_{u_1-t}^{u_1} (u_2-u_1+s)^{H-\frac{1}{2}} s^{H-\frac{1}{2}} ds \\ &= 4\alpha^2 \int_0^{u_1} (u_2-u_1+s)^{H-\frac{1}{2}} s^{H-\frac{1}{2}} ds - 4\alpha^2 \int_0^{u_1-t} (u_2-u_1+s)^{H-\frac{1}{2}} s^{H-\frac{1}{2}} ds \\ &= \frac{4\alpha^2}{H+\frac{1}{2}} (u_2-u_1)^{H-\frac{1}{2}} u_1^{H+\frac{1}{2}} Hyp\left(\frac{1}{2}-H, H+\frac{1}{2}, H+\frac{3}{2}, -\frac{u_1}{u_2-u_1}\right) \\ &- \frac{4\alpha^2}{H+\frac{1}{2}} (u_2-u_1)^{H-\frac{1}{2}} (u_1-t)^{H+\frac{1}{2}} Hyp\left(\frac{1}{2}-H, H+\frac{1}{2}, H+\frac{3}{2}, -\frac{u_1-t}{u_2-u_1}\right) \end{aligned}$$

We also have :

$$m_i = \mathbb{E}(X_{t,u_i}) \quad = -\alpha^2 \frac{|u_i - t|^{2H} - u_i^{2H}}{H}$$

Hence, if we have a discrete grid for u, for example $(u_1, ..., u_N)$, we can get the values of $(\xi_t^{u_i})_{1\leq i\leq N}$ through the Variance-Covariance matrix of $(X_{t,u_i})_{1\leq i\leq N}$ and its mean. In fact, if we denote Σ the Variance-Covariance matrix of $(X_{t,u_i})_{1\leq i\leq N}$ (which we can compute through the hypergeometric function as shown above) and its mean m, we know that if we can find a square root of Σ (ie S st $SS^T = \Sigma$) , we have :

$$\begin{pmatrix} \xi_t^{u_1} \\ \vdots \\ \xi_t^{u_N} \end{pmatrix} = \xi_0 \; exp\left(m + SZ\right), \text{ where } Z \sim \mathcal{N}(0, I_N)$$

The choice of the method that we should use to find S, is crucial and could depend on how we calculate the hypergeometric function. Since some the eigenvalues of Σ are zero, and since the precision of the hypergeometric function doesn't exceed 10^{-14}, we might end up with a Σ that has slightly negative eigenvalues . To avoid this, we used the diagonalization technique and before we take the square root of the eigenvalues, we remove those that are slightly negative (-10^{-14}).

In other terms, if we have $\Sigma = PDP^T$, $S = P\sqrt{D1_{D>0}}P^T$.

Chapter 3

Volatility options pricing

3.1 Introduction to VIX

The Volatility Index, or VIX, was introduced in 1993 by the Chicago Board Options Exchange (CBOE for short). It is defined as the annualized square root of the price of a basic option with payoff equal to $log(S_{t+\Theta}/S_t)$ with S representing the SPX 500 index and the maturity Θ equal to 30 days. It can be formally written via risk-neutral expectation under the form :

$$VIX_t = \sqrt{\mathbb{E}[log(\frac{S_{t+\Theta}}{S_t})]} \times 100$$

We can also express the VIX at time t using the forward variance through the continuous-time monitoring formula :

$$VIX_t = \sqrt{\frac{1}{\Theta}\int_t^{t+\Theta} \xi_t^u du}$$

3.2 Pricing VIX options under rBergomi Model

We are interested here in pricing an option with pay-off at time t given by:

$$f(\frac{1}{\Theta}\int_t^{t+\Theta} \xi_t^u du)$$

with Θ being one month, the Call option on the VIX for instance corresponds to $f(x) = (\sqrt{x} - K)^+$.
This option price seen at time t=0 is :

$$P_0 = \mathbb{E}[f(\frac{1}{\Theta}\int_t^{t+\Theta} \xi_t^u du)]$$

(we're considering the interest rate r to be equal to 0).

To compute this option price through Monte Carlo, we shall use a discrete version of the integral $\frac{1}{\Theta}\int_t^{t+\Theta} \xi_t^u du$. We shall consider two discretisation schemes in the next section.

3.2.1 Integral Discretization method

Rectangular scheme

The first approach to estimate the integral $\int_t^{t+\Theta} \xi_t^u du$, is to use the rectangular discretization scheme. We substitute ξ_t^u for $u \in [u_i, u_{i+1}]$, by $\xi_t^{u_i}$.

$$\begin{cases} \frac{1}{\Theta}\int_t^{t+\Theta} \xi_t^u du \approx \frac{1}{\Theta}\sum_{i=0}^{N-1} \xi_t^{u_i}(u_{i+1}-u_i) \\ \qquad\qquad\qquad = \frac{1}{N}\sum_{i=0}^{N-1} \xi_t^{u_i}, \\ u_i = t + \Theta\frac{i}{N} \end{cases}$$

The error committed with this discrete approximation scheme is bounded by $1/N$. In other terms, we have for f (Lipschitz) :

$$|f\left(\frac{1}{N}\sum_{i=0}^{N-1}\xi_t^{u_i}\right) - f\left(\frac{1}{\Theta}\int_t^{t+\Theta}\xi_t^u du\right)| = \mathcal{O}(\frac{1}{N})$$

To summarize this method, if we want to sample the VIX value at time t, we can get it through the following formula:

$$VIX_t^{R_N} = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}\xi_t^{u_i}}$$

Trapezoidal scheme

The second approach to estimate the integral $\int_t^{t+\Theta} \xi_t^u du$ is to use the trapezoidal discretization scheme. We substitute ξ_t^u for $u \in [u_i, u_{i+1}]$ by $\frac{u_{i+1}-u}{u_{i+1}-u_i}\xi_t^{u_i} + \frac{u-u_i}{u_{i+1}-u_i}\xi_t^{u_{i+1}}$.

$$\begin{cases} \frac{1}{\Theta}\int_t^{t+\Theta} \xi_t^u du \approx \frac{1}{\Theta}\sum_{i=0}^{N-1}\int_{u_i}^{u_{i+1}}\left(\frac{u_{i+1}-u}{u_{i+1}-u_i}\xi_t^{u_i} + \frac{u-u_i}{u_{i+1}-u_i}\xi_t^{u_{i+1}}\right)du, \\ \qquad\qquad\qquad = \frac{1}{\Theta}\sum_{i=0}^{N-1}(u_{i+1}-u_i)\frac{\xi_t^{u_i}+\xi_t^{u_{i+1}}}{2}, \\ \qquad\qquad\qquad = \frac{1}{N^2}\sum_{i=0}^{N-1}(2i+1)\frac{\xi_t^{u_i}+\xi_t^{u_{i+1}}}{2}, \\ u_i = t + \Theta\left(\frac{i}{N}\right)^2 \end{cases}$$

The error committed with this discrete approximation scheme is bounded by $1/N^2$. In other terms, we have for f (Lipschitz) :

$$|f\left(\frac{1}{N^2}\sum_{i=0}^{N-1}(2i+1)\frac{\xi_t^{u_i}+\xi_t^{u_{i+1}}}{2}\right) - f\left(\frac{1}{\Theta}\int_t^{t+\Theta}\xi_t^u du\right)| = \mathcal{O}(\frac{1}{N^2})$$

If we want to use for example the trapezoidal scheme to sample the VIX value at time t, we can get it through the following formula:

$$VIX_t^{T_N} = \sqrt{\frac{1}{N^2}\sum_{i=0}^{N-1}(2i+1)\frac{\xi_t^{u_i}+\xi_t^{u_{i+1}}}{2}}$$

3.2.2 Control Variate

Since we are only estimating the VIX through a Monte Carlo method, it would be preferable to reduce the variance of our estimator. One efficient way to achieve this, is to use a control variate.

In our model, the squared VIX index $VIX_t^2 = \frac{1}{\Theta}\int_t^{t+\Theta} \xi_t^u du$ is an integral over a family of lognormal random variables. Mimicking Kemna and Vorst 's control variate trick [6](originally proposed in the context of Asian options in lognormal models), it is natural to roughly approximate this integral by the exponential of an integral of a corresponding family of Gaussian random variables:

$$\overline{VIX}_t^2 := exp\left(\frac{1}{\Theta}\int_t^{t+\Theta} log(\xi_t^u)du\right)$$

The interesting property of this quantity ($\overline{VIX}_t^2$) is the fact that it is a lognormal process, meaning that its mean can be calculated through the Black Scholes call price formula (we have an explicit formula for its mean).
(In fact, since ξ_t^u is lognormal, we know that $\int_t^{t+\Theta} log(\xi_t^u)du$ is normal, which makes $exp\left(\frac{1}{\Theta}\int_t^{t+\Theta} log(\xi_t^u)du\right)$ lognormal).

Using the Black Scholes Formula, for Y Gaussian with mean m_Y and variance σ_Y^2 we have that :

$$\bar{P}_0 = \mathbb{E}\left[\left(e^{\frac{1}{2}Y} - K\right)^+\right] = \bar{Y}\mathcal{N}\left(\frac{\log\frac{\bar{Y}}{K} + \frac{1}{8}\sigma_Y^2}{\sigma_Y/2}\right) - K\mathcal{N}\left(\frac{\log\frac{\bar{Y}}{K} - \frac{1}{8}\sigma_Y^2}{\sigma_Y/2}\right)$$

with $\bar{Y} = \frac{1}{2}m_Y + \frac{1}{8}\sigma_Y^2$.

We will use the same discretization scheme used to estimate the VIX option price to estimate this control variate, and in order to avoid any additional bias due to this discretization, we will consider the Black Scholes formula for the discretized variate, i.e: we would consider $\bar{P}_0^N$ instead of $\bar{P}_0$, where

$$\bar{P}_0^N = \mathbb{E}(\text{Discrete version of the control } \overline{VIX}_t^N)$$

Rectangular scheme

Under the rectangular discretization scheme, we have $Y_{N,R_N} = \frac{1}{N}\sum_{i=0}^{N-1} log(\xi_t^{u_i})$ with :

- the mean:

$$m_{N,R_N} = \frac{1}{N}\sum_{i=0}^{N-1}\mathbb{E}(X_{t,u_i}) = -\frac{1}{N}\sum_{i=0}^{N-1}\alpha^2\frac{|u_i - t|^{2H} - u_i^{2H}}{H}$$

- the variance:

$$\sigma_{N,R_N}^2 = \begin{pmatrix} 1/N \\ \vdots \\ 1/N \end{pmatrix}^T .Cov(X_{t,u_1}, .., X_{t,u_N}). \begin{pmatrix} 1/N \\ \vdots \\ 1/N \end{pmatrix}$$

Trapezoidal scheme

Under the trapezoidal discretization scheme, we have :

$Y_{N,T_N} = \frac{1}{2N^2}\big(\xi_t^{u_0} + (2N+1)\xi_t^{u_N} + \sum_{i=1}^{N-1} 4i\xi_t^{u_i}\big)$
To compute the mean and the variance we need to define first, the vector V_{N,T_N} :

$$V_{N,T_N} := \frac{1}{2N^2} \begin{pmatrix} 1 \\ 4 \\ \vdots \\ 4i \\ \vdots \\ 4(N-1) \\ 2N+1 \end{pmatrix}$$

- the mean:

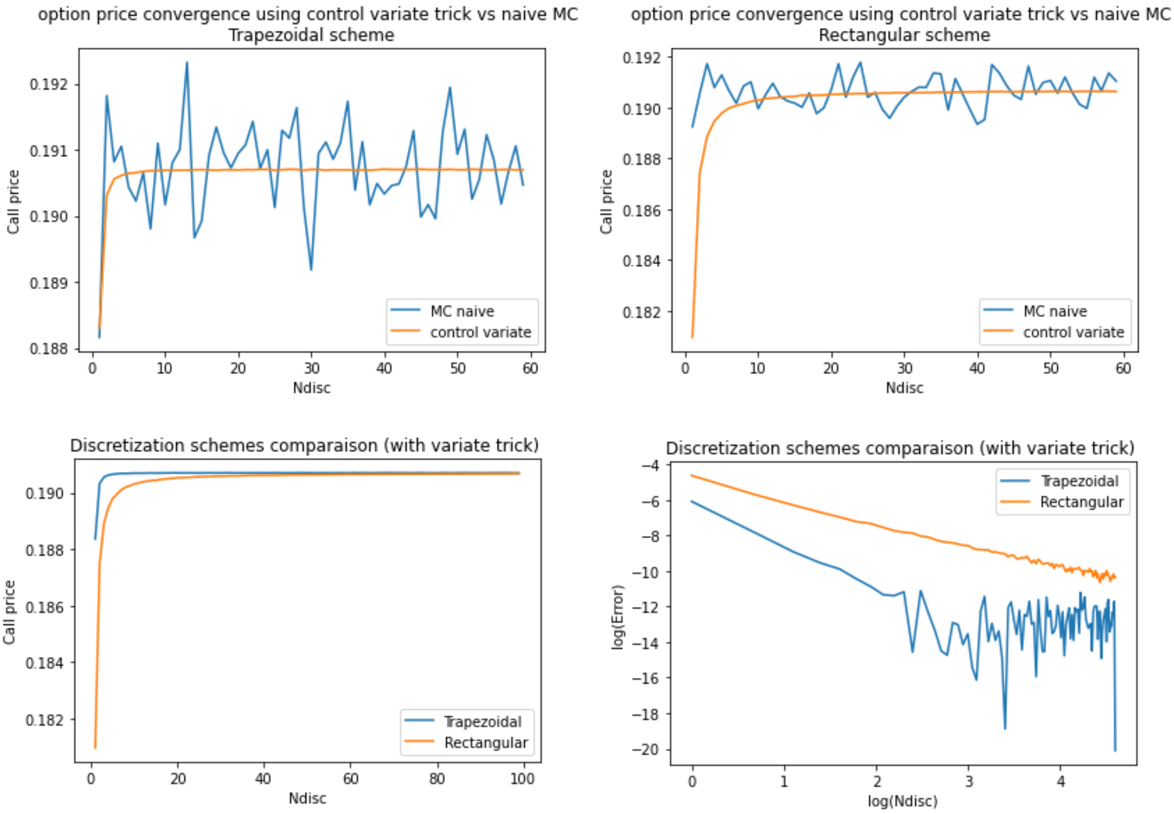
$$m_{N,T_N} = V_{N,T_N}^T . \begin{pmatrix} \mathbb{E}(X_{T,u_0}) \\ \vdots \\ \mathbb{E}(X_{T,u_N}) \end{pmatrix}$$

- the variance:

$$\sigma_{N,T_N}^2 = V_{N,T_N}^T . Cov(X_{T,u_1}, .., X_{T,u_N}).V_{N,T_N}$$

3.2.3 Numerical Results and Comparison between the different procedures

Figure 3.1: 'Comparison between trapezoidal and rectangular schemes, and between MC naive and control variate estimation'



Values of the parameters used for our numerical implementation :

- T=1 maturity , K=0 for strike, $\Theta = 0.1$ (1 month),$M = 50000$ paths (Monte Carlo Samples), $\xi_0^2 = 20\%$ (flat forward variance) and $\alpha = 0.2$

Figure 3.1 illustrates the fast convergence of trapezoidal scheme with the control variate procedure. That's why in order to estimate VIX options prices and compute the underlying implied volatility, we will be using this method.

We can see that the log-error committed by the trapezoidal scheme (right bottom graph) is double that committed by the rectangular scheme. This is reminiscent of the fact that the trapezoidal scheme is bounded by $1/N^2$ whereas the rectangular scheme is only bounded by $1/N$.

Chapter 4

VIX implied volatility surface and extension of the rBergomi model

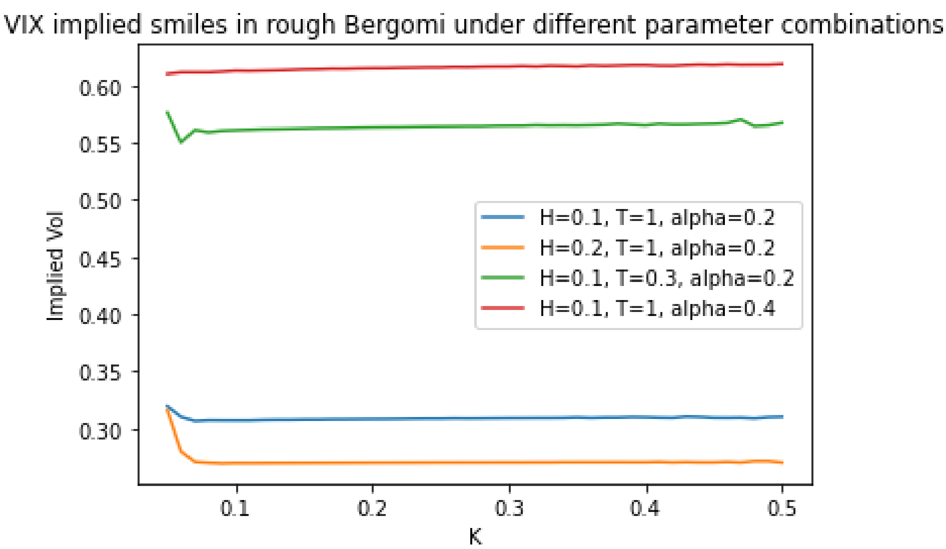
4.1 Implied Volatility surface using rBergomi

Using the previously described methodology, we saw that we can calculate (through a Monte Carlo estimation) the volatility call option price. If we note C(T,K) the price of a call of maturity T and strike K on the VIX, we know that the Black Scholes implied volatility is defined as the solution $\sigma_B S$ of the following equation :

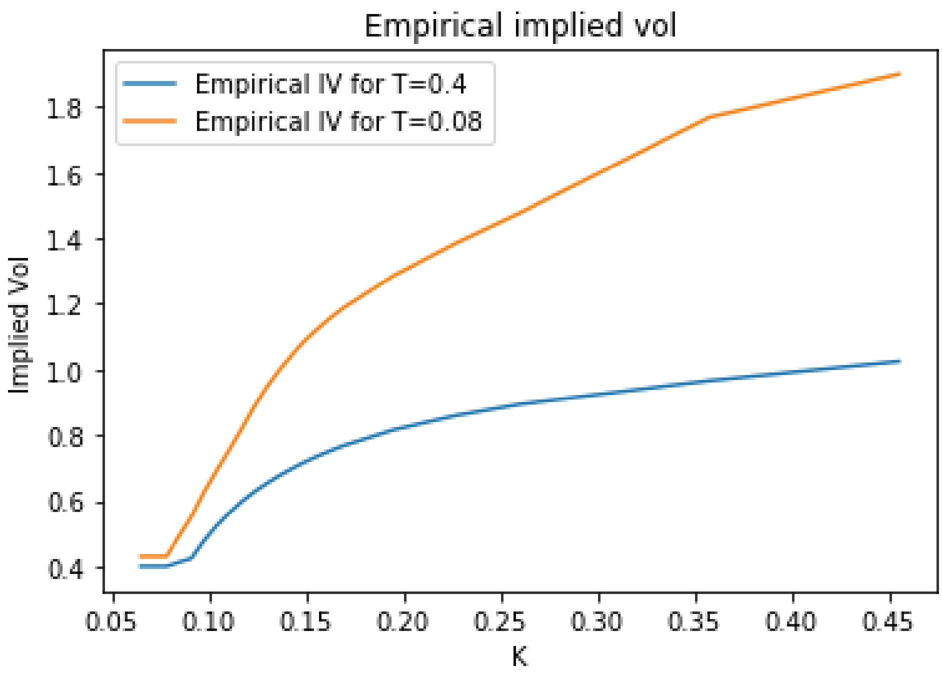
$$C(K,T) = Call_{BS}(S_0, K, T, \sigma_{BS}(T,K), r)$$

We use the bisection method to extract σ_{BS}. We didn't use Newton's method because it was unstable particularly because of a "lack of convexity" in this particular problem.

We found the following results using the rBergomi model :



However, when we look at the empirical implied volatility (below), we can clearly see that this model produced a shape of IV that doesn't match that of real data.



Nonetheless, we note that although the rBergomi model fails to capture the upside implied volatility skew for the VIX, it reproduces one important property : looking at the data we can see that we should expect a bigger IV value for calls that have a short maturity. This is what we see in the first graph (blue and green graphs).

4.2 Extension of the rBergomi model

Since the rBergomi model fails to capture the upside implied volatility skew, an extension was proposed through the works of Pr. De Marco.

In a similar fashion to the initial 2-Factor Bergomi model, we will consider a new model with two factors :

$$\xi_t^u = \xi_0^u \left((1-\delta)\, exp\left(\int_0^t \frac{2\alpha_1}{(u-s)^\gamma} dWs - \int_0^t \frac{2\alpha_1^2}{(u-s)^{2\gamma}} ds \right) + \delta\ exp\left(\int_0^t \frac{2\alpha_2}{(u-s)^\gamma} dWs - \int_0^t \frac{2\alpha_2^2}{(u-s)^{2\gamma}} ds \right) \right.$$

If we denote $V_t^u = \int_0^t \frac{1}{(u-s)^\gamma} dWs$

We have :

$$\xi_t^u = \xi_0^u \left(\delta\ exp\left(2\alpha_1 V_t^u - 2\alpha_1^2 Var(V_t^u) \right) + (1-\delta)\, exp\left(2\alpha_2 V_t^u - 2\alpha_2^2 Var(V_t^u) \right) \right)$$

We can simulate V_t^u through the same method we used for X_t^u (in fact V is a special case of X for which $\alpha = 0.5$). The only problem we have in this case, is with the control variate since : $log(\xi_t^u)$ in this case is no longer normally distributed. Thus we need to find another control variat

4.2.1 New control Variate trick

We propose to compare 2 control variate tricks:

Control Variate N1 :

Mimicking the control variate trick used in the rBergomi Monte Carlo simulations, we can consider a new exponential of a family of Gaussian random variables:

$$\overline{VIX}_{T,1}^2 := \xi_0^u \ exp\left(\frac{1}{\Theta}\int_t^{t+\Theta} \delta \ log(\xi_{1,t}^u) + (1-\delta) \ \ log(\xi_{2,t}^u)du\right)$$

with: $\xi_{k,T}^u = exp\left(2\alpha_k V_t^u - 2\alpha_k^2 Var(V_t^u)\right)$ for $k \in$\{1,2\}

We can now proceed in the same fashion as described in section (3.2.2), to define a new Black Scholes price formula for our new variate.

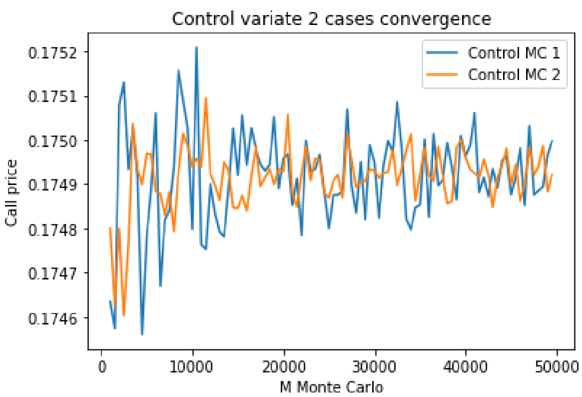
Control Variate N2 :

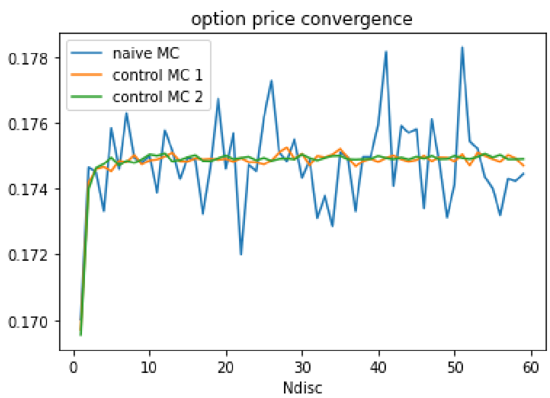
In this second approach, our control variate is the convex sum of the two control variates for each factor in our extended rBergomi Model(each factor characterized by its alpha parameter). In other words, we will define our new control variate as :

$$\overline{VIX}_{T,2} := \delta \ \overline{VIX}_T(\alpha_1) + (1-\delta) \, \overline{VIX}_T(\alpha_2)$$

In this case we will be using the exact same Black Scholes formulas as in section (3.2.2).

Comparison between these two control variate tricks :





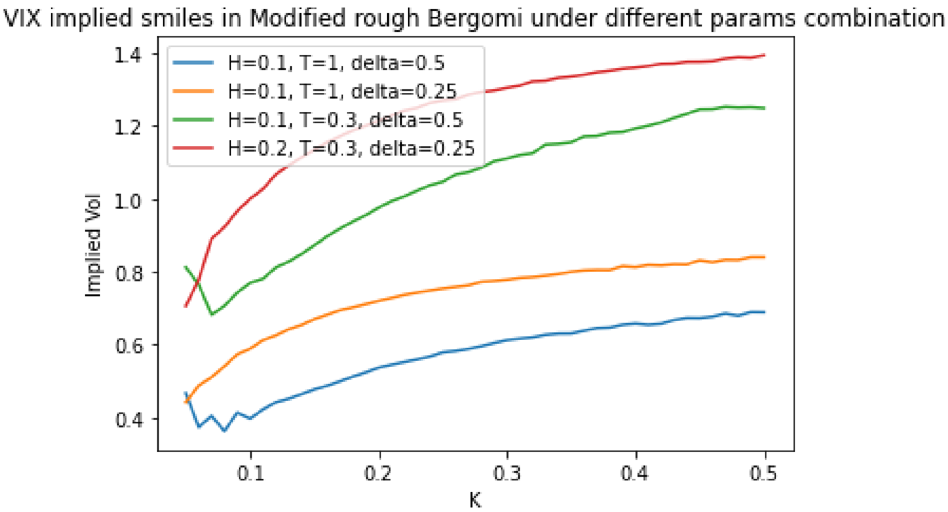
values of parameters:

- T=1 maturity , K=0 for strike, $\Theta = 0.1$ (1 month),$M = 10000$ paths (Monte Carlo Samples), $\xi_0^2 = 20\%$ (flat forward variance), $\alpha_1 = 0.2$, $\alpha_2 = 0.5$ and $\delta = 0.5$ convexity parameter

Since the second trick is slightly better at estimating the call price we will consider it in order to compute the implied volatility below.

4.3 Implied Volatility surface

Once again, we look at the shape of implied volatility :



We can see in this case that the extended rBergomi model produces the shape we are looking for. We need to examine the degree to which this model would fit the data. We are not interested in just reproducing the overall shape, but rather have a model that fits the data decently.

Chapter 5

Calibration

We fitted both rBergomi model and its extended version on real data. We chose the maturities $T = \frac{21}{252}, \frac{100}{252}, \frac{188}{252}$. We fitted the parameters using the Powell method.

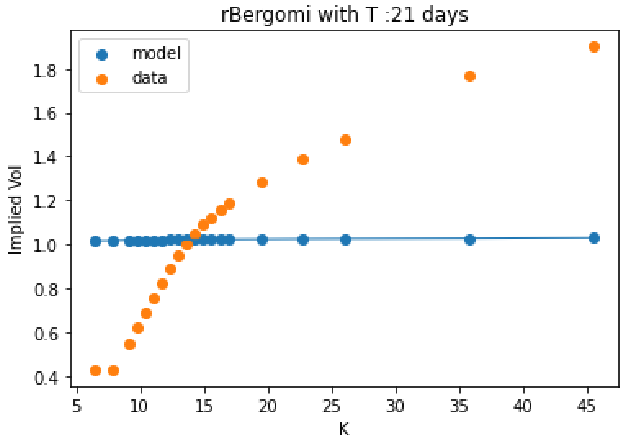
5.1 rBergomi Model

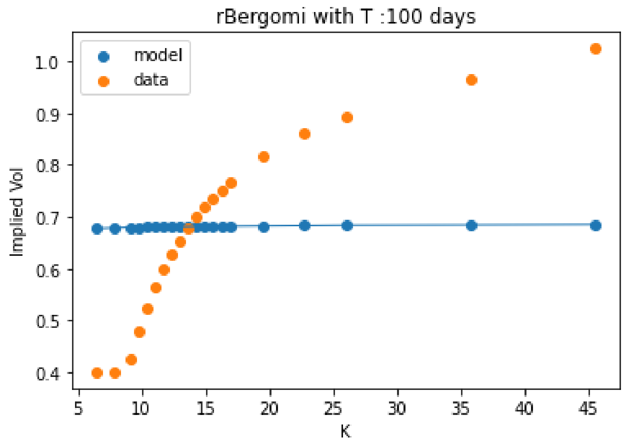
We show once again that even if we try to calibrate the rBergomi model, the best result we can find is that of a line passing in the middle (the horizontal line that minimizes the MSE). As we see in the graphs below, no matter what the value of the maturity, this model always fails to capture the upside skew.

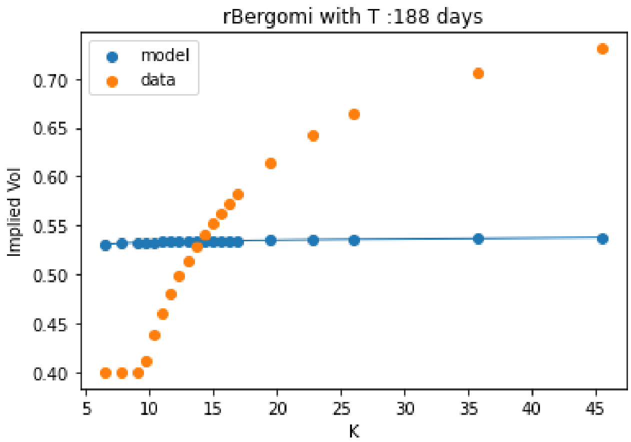
We chose the following bounds to fit the model :

- H : [0.1,0.3]
- ξ_0 : [0.01,0.4]
- α : [0.05,0.9]

(we used the function minimize from **scipy.optimize** a python library in order to minimize the mean square error (MSE) between our model's Implied Volatility and the empirical one).



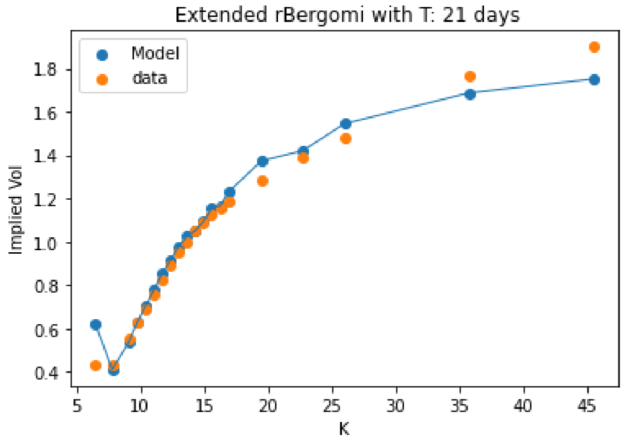


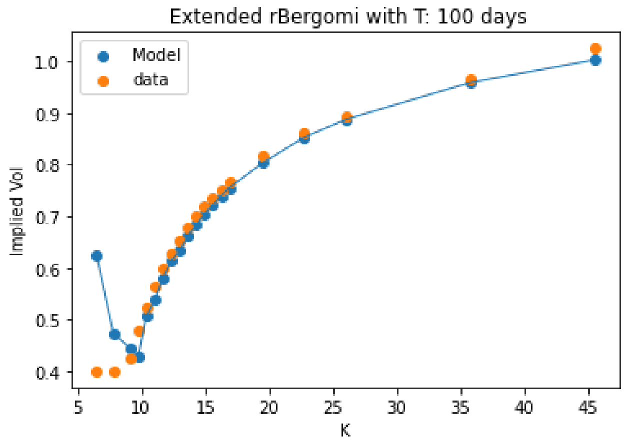


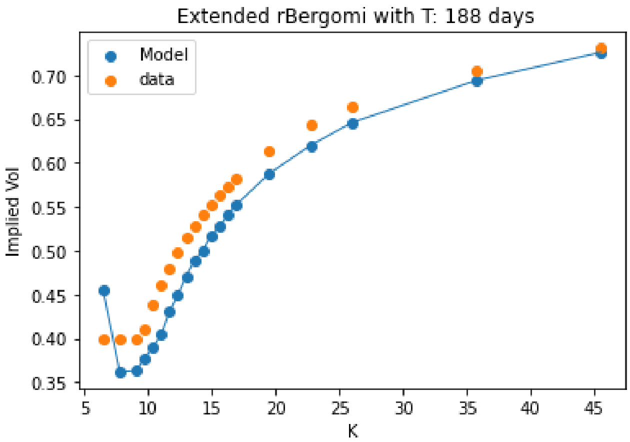
5.2 Extended rBergomi Model

We saw earlier that the extended model produces the shape we are looking for of the the implied volatility, but we still need to examine if that shape decently fits the real data. So we tried to fit the parameters on our data. The bounds we chose are :

- H : [0.1,0.5]
- ξ_0 : [0.01,0.4]
- α_1 : [0.05,0.9]
- α_2 : [0.05,0.9]
- δ : [0,1]







We can see that the extended rBergomi model produces an astonishing fit of the data.

Important remarks : We note that the value of H was big for the shortest maturity $(T = 21/252)$: it was around 0.3 whereas in the other two cases it was at around 0.15 . The opposite effect was observed for ξ_0. This suggests the presence of structural differences between long term and short term derivatives.

5.3 Deep Calibration

The calibration process described above is very expensive time-wise, this is particularly due to the fact that the function we are trying to minimize isn't very smooth. To solve this issue, we can approximate the pricing function C with a smoother function, then try to work with this approximation for which it will be easier to calculate the gradient or Hessian. One effective way to approximate C is to use neural networks.

We set a grid for T and K : $(T_i, K_j)_{1\leq i\leq n, 1\leq j\leq m}$. The architecture of our NN is as follows:

$$\text{Input : Models parameters } (H, \xi_0, \alpha_1, ...)$$
$$\text{Output : } C(T_i, K_j) \text{ for } 1 \leq i \leq n, 1 \leq j \leq m$$

However, due to a lack of time, we couldn't finish this part.

Bibliography

[1] Gatheral J. Jaisson T. and Rosenbaum M. *Volatility is rough.* 2014. URL: https://arxiv.org/pdf/1410.3394.pdf (page 1).

[2] Lorenzo Bergomi. *Smile Dynamics II*. 2005. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1493302 (page 1).

[3] Lorenzo Bergomi. *Smile Dynamics IV*. 2009. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1520443 (page 1).

[4] Friz Bayer and Gatheral. *Pricing Under Rough Volatility*. 2015. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2554754 (page 3).

[5] Jacquier Horvath and Tankov. *Volatility options in rough volatility models*. 2019. URL: https://arxiv.org/pdf/1802.01641.pdf (page 7).

[6] A.G.Z Kemna and A.C.F. Vorst. *A pricing method for options based on average asset values. Journal of Banking and Finance, 14(1): 113-129,* 1990 (page 11).