Rough Volatility Models

Project outside course scope

Karl Kristian Engelund
MATH, University of Copenhagen
qvt701@ku.dk

Supervisor:
Rolf Poulsen
MATH, University of Copenhagen
rolf@math.ku.dk

January 25, 2023

Abstract

In recent years it has become apparent that volatility is rougher then a standard drift diffusion process, see [5]. In this project we look at the mathematical setup for extending the classical Bergomi model to a general Bergomi model, that encapsulates the classical and the rough Bergomi, known from [2], but also extends to more general models through a choice of kernel. We first look at the mathematical properties of so called Stochastic Volterra equations. Together with an introduction to the fractional Brownian motion this enables us to look at more general Ornstein-Uhlenbeck type processes. Next we unify the Bergomi models under the physical measure. Then we study different specifications of the market price of volatility risk, especially specifying one using the Volterra type Ornstein–Uhlenbeck process. This allows for a market price of volatility governed by so-called regimes as in [6]. Finally, we look at simulation in order to price options on the VIX. We see that the model can fit the skew shape very well for short maturities. We also consider the error on direct simulation and a common approximation, and find it to be small.

1 General results

In this section we present notation and general results that are used in the following sections. We only consider the one-dimensional case for convenience unless otherwise stated.

1.1 Volterra equations

A particular tedious stochastic integral equation that shows op in many fields is the so called Volterra equations. In general, it is an equation on the form

$$X_t = F(t) + \int_0^t K(t-s)\alpha_s ds + \int_0^t K(t-s)\beta_s dW_s \tag{1.1}$$

where W is a Brownian motion, α and β are sufficiently regular processes and $K : \mathbb{R}^+ \to \mathbb{R}^+$ sufficiently integrable function of time called the kernel. Particular kernels of interest are the constant-, the exponential- $K(x) = e^{-\theta x}$ and the fractional/power law kernel $K(x) = x^{-\gamma}$. We introduce the following notation for the convolutions

$$(K * g)_{a,b}(u) = \int_a^b K(u-s)g(s)ds, \quad (K * dY)_{a,b}(u) = \int_a^b K(u-s)\alpha_s ds + \int_a^b K(u-s)\beta_s dW_s \tag{1.2}$$

for a sufficiently integrable function g and a continuous local semi-martingale Y given by $dY_t = \alpha_t dt + \beta_t dW_t$. In particular note that $X_t = F(t) + (K * dY)_{0,t}(t)$ in equation (1.1). We may drop one or both of the subscripts when the boundary is clear. For a nice introduction to stochastic calculus of convolutions we refer to [1, Section 2]. To ensure a well-defined expression $K * dY$ we will always assume $K \in L^2_{loc}$ such that the quadratic variation inspired expression[1]

$$\langle K * dY \rangle_t := \int_0^t |K(t-s)|^2 \, d\langle Y \rangle_s \tag{1.3}$$

is finite for any $t \geq 0$. In addition to wanting local square integrability we also would like sufficient conditions for $K * dY$ to be continuous. According to [1] one such condition is assuming existence of a $\delta \in (0, 2]$ such that

$$\int_0^h K(t)^2 dt = O(h^\delta) \qquad \text{and} \qquad \int_0^T (K(t+h) - K(t))^2 dt = O(h^\delta) \tag{1.4}$$

for every $T < \infty$. We note in passing that all kernels considered in this project satisfies this property and $K \in L^2_{loc}$, see [1, Example 2.3].

Note that the convolution $(K * dY)_{a,u}(u)$ is not in general a semi-martingale, hence specifically it is not an Ito-process, so our usual results in stochastic calculus do not hold. However, instead of the usual Doob-Meyer decomposition into a martingale and a finite variation process we have the following decomposition that follows directly from linearity of integrals.

Proposition 1.1. *Let $K \in L^2_{loc}$ and let Y be an Ito-process given by $dY = \alpha dt + \beta dW$. Then for $a \leq t < u$*

$$(K * dY)_{a,u}(u) = (K * dY)_{a,t}(u) + (K * dY)_{t,u}(u). \tag{1.5}$$

The first term is an Ito-process in t while the second is a process independent of $\mathcal{F}_t$.

[1] We will just call it the quadratic variation, even though it does not have the same properties as quadratic variation of semi-martingales.

Since K is deterministic the first convolution is adapted to $(\mathcal{F}_t)_{0\le t<u}$ and markovian. Thus, we see by this decomposition that our problems lay in the second term which obviously is neither an Ito-process, markovian nor adapted to $(\mathcal{F}_t)$. In the following we circumvent this inconvenience by considering Gaussian convolutions with deterministic kernels.

1.2 Stochastic exponential

The stochastic exponential of an Ito process X is defined as

$$\mathcal{E}_t(X_.) = e^{X_t - \frac{1}{2}\langle X\rangle_t}. \tag{1.6}$$

A particular result is that if X is a martingale then so is $\mathcal{E}_t(X_.)$ and $\mathbb{E}[\mathcal{E}_t(X_.)] = 1$ if $X_0 = 0$.

Note that if the diffusion term of X is deterministic then $\langle X\rangle$ is also deterministic. Form the definition of quadratic variation $X_t^2 - \langle X\rangle_t$ is a martingale hence if $X_0 = 0$ then $\langle X\rangle_t = \mathbb{E}[X_t^2]$. So considering a Gaussian convolution with deterministic kernel, then by the decomposition 1.1 we only have problems with the $\mathcal{F}_t$ independent part. A natural extension is to define

$$\mathcal{E}((K*dW)_{t,u}(u)) = e^{(K*dW)_{t,u}(u) - \frac{1}{2}\mathrm{Var}((K*dW)_{t,u}(u))}. \tag{1.7}$$

Note that the process is centered. Usually we would use the Ito-isommetry to calculate the second moment, but the process is not a semi-martingale. However, we are in luck since $(K*dW)_{t,u}(u)$ is a centered semi-stationary Gaussian process the conclusion still holds, i.e.

$$\mathbb{E}[((K*dW)_{t,u}(u))^2] = \int_t^u K^2(u-s)ds. \tag{1.8}$$

So with our extension we can write

$$\mathcal{E}((K*dW)_{a,u}(u)) = e^{(K*dW)_{a,u}(u) - \int_a^u K^2(u-s)ds}. \tag{1.9}$$

As for the expectation we recover the result from the Ito case by recalling the moment generating function of a Gaussian, that is

$$\begin{aligned}
\mathbb{E}\left[\mathcal{E}((K*dW)_{t,u}(u))\right] =& \mathbb{E}\left[e^{(K*dW)_{t,u}(u)}\right] e^{-\frac{1}{2}\mathbb{E}\left[((K*dW)_{t,u}(u))^2\right]} && (1.10)\\
=& e^{\mathbb{E}[(K*dW)_{t,u}(u)] + \frac{1}{2}\mathbb{E}\left[((K*dW)_{t,u}(u))^2\right] - \frac{1}{2}\mathbb{E}\left[((K*dW)_{t,u}(u))^2\right]} && (1.11)\\
=& 1 && (1.12)
\end{aligned}$$

This result will be used throughout the project. Note that it also holds when conditioning on $\mathcal{F}_t$.

1.3 Fractional Brownian Motion

A fractional Brownian motion is an extension of the classical Brownian motion, that enables us to modify the roughness using the Hurst parameter $H \in (0,1)$.

Definition 1.2 (Fractional Brownian Motion)**.** Let $H \in (0,1)$. A process $(W_t^H)_{t\in\mathbb{R}}$ is called a fractional Brownian motion(fBM) if it is an almost surely continuous, centered Gaussian process with

$$\mathbb{E}\left[W_t^H W_s^H\right] = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right) \tag{1.13}$$

for $t, s \in \mathbb{R}$.

Note in particular that the process is defined on the whole real line. We immediately see that $W^{\frac{1}{2}}$ is a two-sided standard Brownian motion, see e.g. [10, Corollary 2.7]. Furthermore, for $H < \frac{1}{2}$ the increments have a negative correlation, while the opposite is true for $H > \frac{1}{2}$. In this paper we will primarily be concerned with the case $H < \frac{1}{2}$. The negative autocorrelation produces rougher paths then the standard Brownian motion.

There are many ways to define the fBM. From [9] we have the Mandelbrott-van Ness representation[2]

$$W_t^H = C_H \left(\int_{-\infty}^{t} \frac{dW_s}{(t-s)^\gamma} - \int_{-\infty}^{0} \frac{dW_s}{(-s)^\gamma} \right) \tag{1.14}$$

where $\gamma = \frac{1}{2} - H$ and $C_H = \sqrt{\frac{2H\Gamma(3/2-H)}{\Gamma(H+1/2)\Gamma(2-2H)}}$. Note in particular that the first integral is of Volterra form with the fractional kernel, disregarding that we defined it only on $\mathbb{R}^+$. We frequently interchange between three parametrization of the Hurst index, namely

$$H = \frac{1}{2} - \gamma = \alpha - \frac{1}{2}. \tag{1.15}$$

Note in particular that when $H \in \left(0, \frac{1}{2}\right)$ then $\gamma \in \left(-\frac{1}{2}, \frac{1}{2}\right)$ and $\alpha \in \left(\frac{1}{2}, 1\right)$.

1.4 Ornstein–Uhlenbeck processes

The classical Ornstein–Uhlenbeck(OU) process X with time dependent mean reversion is the solution to the following SDE

$$dX_u = \theta(\mu_u - X_u)du + \eta dW_u \tag{1.16}$$

with some boundary condition. Here W is a standard Brownian motion, $\theta \in \mathbb{R}$, $\eta > 0$ and for now let μ be a deterministic function of time. If we assume $X_0 = 0$ then the solution is

$$X_u = \theta \int_0^u e^{-\theta(u-s)} \mu_s ds + \eta \int_0^u e^{-\theta(u-s)} dW_s. \tag{1.17}$$

By using the properties of the exponential function we can get a forward solution for $u \geq t$ on the form

$$X_u = e^{-\theta(u-t)} X_t + \theta \int_t^u e^{-\theta(u-s)} \mu_s ds + \eta \int_t^u e^{-\theta(u-s)} dW_s. \tag{1.18}$$

Note in particular that the solution involves Volterra type integrals with exponential kernels.

We will need a more general class of OU processes later. In particular, we will look at a one-dimensional OU version of an affine Volterra equation, as presented in [6],

$$X_u = F(u) + \int_0^u K(u-s)\theta(\mu_s - X_s)ds + \int_0^u K(u-s)\sigma dW_s \tag{1.19}$$

where $K \in L^p_{loc}$ for $p \geq 2$ and satisfies (1.4). Furthermore, we assume $\mu \in L^q_{loc}$ for $\frac{1}{p} + \frac{1}{q} = 1$, and for convenience we assume $\theta \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ to be constants. We will build a solution for equation (1.19). For this we need existence and uniqueness of a continuous solution. The following is from [1, Theorem 3.3] and stated in the d-dimensional case.

[2] [9] where the first to define the fBM. They did not show equivalence to definition 1.2, but we will take it as given provided the particular form of C_H, see e.g. [5].

Proposition 1.3. Let $d \in \mathbb{N}$. Assume b and σ are Lipschitz continuous functions of time with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$ and the coordinates of $K \in \mathbb{R}^{d\times d}$ are in L^2_{loc} and satisfy (1.4). Then

$$X_t = X_0 + (K * b)(t) + (K * \sigma dW)(t) \tag{1.20}$$

admits a unique continuous strong solution X for any boundary condition $X_0 \in \mathbb{R}^d$.

Remark 1.4. Using a trick of dimensionality like that of [1, Remark 3.5] it can be seen that the constant boundary X_0 of proposition 1.3 can be extended to $F(t) = \int_0^t \tilde{b}(s)ds$ for some Lipschitz continuous function $\tilde{b}$. We hypothesize that proposition 1.3 can be extended further to hold for case of

$$X_t = F(t) + (K * b)(t) + (K * \sigma dW)(t) \tag{1.21}$$

for any continuous function $F : [0,T] \to \mathbb{R}^d$.

We now return to $d = 1$ for convenience. Define R_θ as the solution to

$$K\theta * R_\theta = K\theta - R_\theta. \tag{1.22}$$

R_θ is called the resolvent of second kind of $K\theta$. Define furthermore $E_\theta = K - R_\theta * K$,

$$H_{c,d}(u) = \int_c^d \theta E_\theta(u-s)\mu_s ds \quad \text{and} \quad Y_{c,d}(u) = \int_c^d E_\theta(u-s)dW_s. \tag{1.23}$$

Theorem 1.5. *The affine Volterra equation of* (1.19) *has unique continuous strong solution on* $[0,T]$ *given by*

$$X_u = g(u) + H_{0,u}(u) + \sigma Y_{0,u}(u), \tag{1.24}$$

where

$$g(u) = F(u) - \int_0^u R_\theta(u-s)F(s)ds. \tag{1.25}$$

Proof. The proof in [6] is incorrect[3]. Let $b = \theta\mu$. Since θ is constant we have $b \in L^p_{loc}$. By assumption $K \in L^q_{loc}$. Thus, by a localized application of Hölder's inequality it can be shown, see [6, Lemma A.2], that $K * b$ is continuous. Hence, defining $\tilde{F} = F + K * b$ we have

$$X_t = \tilde{F}(t) + (-K\theta * X)(t) + (K * \sigma dW)(t) \tag{1.26}$$

which by remark 1.4 has a continuous solution X. Using [1, Lemma 2.5] we get the unique solution

$$X_t = \tilde{F}(t) - (R_\theta * \tilde{F})(t) + (E_\theta * \sigma dW)(t). \tag{1.27}$$

Then by definition of $\tilde{F}$, g, E_θ and using associativity([1, Leamma 2.1]) and distributivity of convolutions we get

$$\begin{aligned}
\tilde{F} - R_\theta * \tilde{F} &= F + K * b - R_\theta * F - R_\theta * (K * b) \\
&= g + (K - R_\theta * K) * b \\
&= g + E_\theta * b.
\end{aligned}$$

Adding the convoluted diffusion we get the result. □

[3] They validate that the proposed solution is continuous, however continuity is used to obtain the solution.

It can easily be seen that the classical OU process (1.16) is equivalent to (1.19) with a constant kernel. By noting that the resolvent of second kind to the constant kernel times θ is $R_\theta(t) = \theta e^{-\theta t}$, and likewise, $E_\theta(t) = e^{-\theta t}$ we could extend the solution to our classical OU process by allowing μ to be stochastic, and having a non-zero boundary condition.

For the constant kernel we could write up a forward solution, however this is not possible in general. Yet we note that the measurability pattern in (1.18) and that of proposition 1.1 are similar.

2 Models under $\mathbb{P}$

In this section we will introduce the models to be studied under the physical measure $\mathbb{P}$. We assume existence of two independent standard Brownian motions W and $\bar{W}$. We model the price of an asset S with the dynamics

$$\frac{dS_u}{S_u} = \zeta_u du + \sqrt{V_u} dB_u \tag{2.1}$$

with $S_0 = s_0$ and $B = \rho W_t + \sqrt{1-\rho^2}\bar{W}_t$ being a standard Brownian motion. The ζ function is assumed to be deterministic while the V process will be stochastic. Note in particular that S and V are one dimensional.

2.1 The classical Bergomi model

We first present a classical stochastic volatility model, namely the one-factor Bergomi model. We model a latent variable X as a standard OU process by (1.16) with $\mu = 0$ and $\theta, \eta > 0$. The solution is given by (1.18)

$$X_u = X_t e^{-\theta(u-t)} + \eta (K * dW)_t(u) \tag{2.2}$$

for $t \leq u$ where K is the exponential kernel with parameter θ and $(K * dW)_t(u) = (K * dW)_{t,u}(u)$. We then model the squared volatility by defining a $\mathcal{F}_t$ measurable $\overline{A}$ and

$$V_u = \overline{A}_t(u) \mathcal{E}_u (X_\cdot) \,. \tag{2.3}$$

For later convenience we rewrite as

$$V_u = A_t(u) \mathcal{E} \left(\eta (K * dW)_t(u) \right) \tag{2.4}$$

where we have moved all the $\mathcal{F}_t$ measurable quantities into A, hence A is a known function at time t. Note that $\mathbb{E}_t \left[\mathcal{E} \left(\eta (K * dW)_t(u) \right) \right] = 1$ so taking conditional expectation we get $A_t(u) = \mathbb{E} \left[V_u \mid \mathcal{F}_t \right]$. Hence, we arrive at

$$V_u = \mathbb{E} \left[V_u \mid \mathcal{F}_t \right] \mathcal{E} \left(\eta (K * dW)_t(u) \right) . \tag{2.5}$$

2.2 Rough Bergomi model

We can roughen the model by using a fractional Brownian motion. Like [2], consider the relation

$$\log V_u - \log V_t = 2\nu (W_u^H - W_t^H) \tag{2.6}$$

for $u \geq t$. Using the Mandelbrott-van Ness representation, see equation (1.14), we have

$$\log V_u - \log V_t = 2\nu C_H \left(\int_{-\infty}^{u} \frac{dW_s^{\mathbb{P}}}{(u-s)^{\gamma}} - \int_{-\infty}^{t} \frac{dW_s^{\mathbb{P}}}{(t-s)^{\gamma}} \right) \tag{2.7}$$

$$= 2\nu C_H \left(\underbrace{\int_{t}^{u} \frac{dW_s^{\mathbb{P}}}{(u-s)^{\gamma}}}_{=(K*dW)_t(u)} \right) + \underbrace{\int_{-\infty}^{t} \frac{1}{(u-s)^{\gamma}} - \frac{1}{(t-s)^{\gamma}} dW_s^{\mathbb{P}}}_{=Z_t(u)} \Bigg) \tag{2.8}$$

by linearity of integrals where K is the fractional kernel with parameter γ. Note that $K * dW$ is independent of $\mathcal{F}_t$ while Z is $\mathcal{F}_t$ measurable. Let $\eta = 2\nu C_H$. By taking exponential, multiplying V_t and taking conditional expectation we thus get

$$\mathbb{E}\left[V_u \mid \mathcal{F}_t\right] = V_t \mathbb{E}\left[\exp\left(\eta (K * dW)_t(u)\right)\right] \exp\left(\eta Z_t(u)\right). \tag{2.9}$$

By the results from section 1.2 and integrating the fractional kernel we get

$$\mathbb{E}\left[\exp\left(\eta (K * dW)_t(u)\right)\right] = \exp\left(\frac{\eta^2}{2} \frac{(u-t)^{2H}}{2H} \right). \tag{2.10}$$

Hence, by multiplying and dividing by the MGF we may write

$$V_u = V_t \exp\left[\eta (K * dW)_t(u) + \eta Z_t(u)\right] \tag{2.11}$$

$$= V_t \exp\left[\eta (K * dW)_t(u) - \frac{\eta^2}{2} \mathbb{E}\left[((K * dW)_t(u))^2 \right] \right] \exp\left[\frac{\eta^2}{2} \frac{(u-t)^{2H}}{2H} + \eta Z_t(u) \right] \tag{2.12}$$

$$= \mathbb{E}\left[V_u \mid \mathcal{F}_t\right] \mathcal{E}(\eta (K * dW)_t(u)). \tag{2.13}$$

The last equality follows from (2.9).

2.3 Generalized Bergomi model

As we have seen in the preceding sections we arrive at the same model form for the classical Bergomi model and the rough Bergomi model. Hence, as in [6] we propose a generalized Bergomi model on the form

$$\frac{dS_u}{S_u} = \zeta_u du + \sqrt{V_u} dB_u, \tag{2.14}$$

$$V_u = \mathbb{E}\left[V_u \mid \mathcal{F}_t\right] \mathcal{E}(\eta (K * dW)_t(u)) \tag{2.15}$$

with $S_t = s$, $B_u = \rho W_u + \sqrt{1-\rho^2} \bar{W}_u$ and suitable kernel $K \in L^2_{loc}$ satisfying (1.4).

3 Measure Change

We need to consider the model under a risk neutral measure $\mathbb{Q}$ making the discounted asset a martingale. For the sake of convenience we assume a zero interest risk-free rate. We use a superscript $\mathbb{P}$ to denote expectations under $\mathbb{P}$ and that $W^{\mathbb{P}}$ is a Brownian motion under $\mathbb{P}$. Under $\mathbb{Q}$ we leave out the superscript.

As of now the problem is inherently two-dimensional since we have the two sources of risk $W^{\mathbb{P}}$ and $\overline{W}^{\mathbb{P}}$. However, when we consider a stochastic measure change we will allow for the stochasticity to be dependent on only part of the noise in the volatility process. Hence, we consider a three-dimensional standard Brownian motion $\boldsymbol{W}^{\mathbb{P}} = (\bar{W}^{\mathbb{P}}, Z^{\mathbb{P}}, \bar{Z}^{\mathbb{P}})^\top$, where $W^{\mathbb{P}} = \nu Z^{\mathbb{P}} + \bar{\nu}\bar{Z}^{\mathbb{P}}$ for a scalar $\nu \in (-1, 1)$ and use the notation $\bar{\nu} = \sqrt{1-\nu^2}$. Consider the process

$$\phi_t = \left(\frac{1}{\bar{\rho}} \left(\frac{\zeta_t}{\sqrt{V_t}} - \rho\lambda(t) \right),\ 0\ ,\ \frac{1}{\bar{\nu}}\lambda(t) \right)^\top \tag{3.1}$$

where λ is a process to be determined. Assuming ϕ satisfies the Novikov condition then using Girsanov we have

$$d\boldsymbol{W}_t^{\mathbb{P}} = d\boldsymbol{W}_t + \phi_t dt \tag{3.2}$$

where $\boldsymbol{W} = (W, Z, \bar{Z})^\top$ is a three-dimensional standard Brownian motion under $\mathbb{Q} \sim \mathbb{P}$ with independent coordinates. Define $B_t = \rho W_t + \bar{\rho}\bar{W}_t$ and $W = \nu Z + \bar{\nu}\bar{Z}$. Then by multiplying $(\overline{\rho}, \rho\nu, \rho\overline{\nu})$ and $(0, \nu, \overline{\nu})$ we get

$$dB_t^{\mathbb{P}} = dB_t + \frac{\zeta_t}{\sqrt{V_t}}dt, \tag{3.3}$$

$$dW_t^{\mathbb{P}} = dW_t + \lambda(t)dt \tag{3.4}$$

respectively. Hence, we see that

$$dS_t = \sqrt{V_t}dB_t \tag{3.5}$$

so S is a martingale since V_t is $\mathcal{F}_t$ measurable. We call λ the market price of volatility risk(MPVR). It is free as long as the conditions of Girsanov are met by ϕ. Note that

$$||\phi_t||^2 = \frac{1}{\bar{\rho}^2}\left(\frac{\zeta_t}{\sqrt{V_t}} - \rho\lambda(t) \right)^2 + \frac{1}{\bar{\nu}^2}\lambda(t)^2 \leq \frac{2}{\bar{\rho}^2}\left(\frac{\zeta_t^2}{V_t} + \rho^2\lambda(t)^2 \right) + \frac{1}{\bar{\nu}^2}\lambda(t)^2. \tag{3.6}$$

So we only need to check Novikov on $\frac{\zeta_t}{\sqrt{V_t}}$ and $\lambda(t)$. By general theory we assume that the condition holds for the first process. Before specifying the MPVR note that

$$\begin{aligned} V_u =& \mathbb{E}^{\mathbb{P}}\left[V_u \mid \mathcal{F}_t\right] \mathcal{E}(\eta(K * dW^{\mathbb{P}})_t(u)) && (3.7)\\ =& \mathbb{E}^{\mathbb{P}}\left[V_u \mid \mathcal{F}_t\right] \mathcal{E}(\eta(K * dW)_t(u)) \exp\left(\eta(K*\lambda)_t(u)\right) && (3.8) \end{aligned}$$

since $\langle (K * dW)(u) \rangle = \langle (K * dW^{\mathbb{P}})(u) \rangle$. We now turn to look at two specific specifications.

3.1 Deterministic MPVR

Letting λ be a sufficiently integrable deterministic function of time obviously obeys the Novikov condition. However, the distribution of V continues to be lognormal under $\mathbb{Q}$ witch does not conform with empirical data. Taking conditional expectation and using the result from section 1.2 we see that the forward variance

$$\begin{aligned} \xi_t(u) := \mathbb{E}[V_u \mid \mathcal{F}_t] &= \mathbb{E}\left[\mathbb{E}^{\mathbb{P}}\left[V_u \mid \mathcal{F}_t\right] \mathcal{E}(\eta(K * dW)_t(u)) \exp\left(\eta(K*\lambda)_t(u)\right) \mid \mathcal{F}_t\right] && (3.9)\\ &= A(u) \exp\left(\eta(K*\lambda)_t(u)\right) && (3.10) \end{aligned}$$

where $A(u) = \mathbb{E}^{\mathbb{P}}[V_u \mid \mathcal{F}_t]$ is a $\mathcal{F}_t$ measurable variable that we could consider a parameter.

3.2 Stochastic MPVR

A general observation in markets is that market price of risks is stochastic. Following the idea from [6] we assume that the MPVR has so-called regimes governed a CTMC μ that is independent of the Brownian noise. That is, we essentially assume a non-market driven process that determines the agents collective view on volatility risk. In particular, we let the MPVR be

$$\lambda_s = \theta(\mu_s - X_s) \tag{3.11}$$

where X is a OU type Volterra process, (1.19), with Z as driving Brownian motion and $\sigma = \nu$. In the case of constant kernel we know that, given μ, X is mean reverting towards μ, hence the MPVR will revert towards 0. So we assume that the market adapts and mean reverts after a shock is produced by jumps in μ that are independent of the market.

As for Novikov we obviously have that ϕ is adapted, and we assume a finite state space for μ hence it is bounded. Note

$$||\lambda(t)||^2 = \theta^2(\mu_t - X_t)^2 \leq 2\theta^2(\mu_t^2 + X_t^2) \tag{3.12}$$

so we only need Novikov on X. Using the solution form theorem 1.5 we see that we can use the bound on μ to get that we only need to verify the Novikov condition on the Y process, which is a well-behaved Gaussian process. For the general technical details see [1], and for this particular ϕ, modulo some sign changes, see [6, Appendix B].

Let $t = 0$ and drop the index for convenience. For this particular MPVR we note that by (1.19) we have

$$\log\left\{\mathcal{E}(\eta(K * dW)(u))\exp\left(\eta(K * \lambda)(u)\right)\right\} \tag{3.13}$$

$$= \eta(K * dW)(u) - \frac{\eta^2}{2}\mathbb{E}\left[((K * dW)(u))^2\right] + \eta\left(X_u - F(u) - \nu(K * dZ)(u)\right) \tag{3.14}$$

$$= \eta(K * dW)(u) - \frac{\eta^2}{2}\mathbb{E}\left[((K * dW)(u))^2\right] \tag{3.15}$$

$$+ \eta\left(X_u - F(u) + \overline{\nu}(K * d\overline{Z})(u) - (\overline{\nu}(K * d\overline{Z})(u) + \nu(K * dZ)(u))\right) \tag{3.16}$$

$$= -\frac{\eta^2}{2}\mathbb{E}\left[\left((K * dW(u))^2\right)\right] + \eta\left(X_u - F(u) + \overline{\nu}M(u)\right) \tag{3.17}$$

where we have added and subtracted $\overline{\nu}M(u) := \overline{\nu}M_{0,u}(u)$ where $M_{c,d}(u) = \int_c^d K(u-s)d\overline{Z}_s$ in the third equality. Note that the variance term is deterministic so inserting back into equation (3.8) we get

$$V_u = A(u)\exp\left(\eta\left(X_u - F(u) + \overline{\nu}M(u)\right)\right) \tag{3.18}$$

where A is a $\mathcal{F}_t$ measurable variable which includes the variance term.

To get an expression for the forward variance we need to take conditional expectation in (3.18). We first look at the conditional moment generating functions of X and M. Using proposition 1.1 we can take out the $\mathcal{F}_t$ measurable part. Since M is a centered Gaussian we have that

$$\mathbb{E}\left[e^{w\overline{\nu}M(u)} \mid \mathcal{F}_t\right] = \exp\left\{w\overline{\nu}M_{0,t}(u)\right\}\exp\left\{\frac{w^2\overline{\nu}^2}{2}\mathbb{E}\left[M_{t,u}^2(u) \mid \mathcal{F}_t\right]\right\} \tag{3.19}$$

$$= \exp\left\{w\overline{\nu}M_{0,t}(u) + w^2\underbrace{\frac{\overline{\nu}^2}{2}\int_t^u K^2(u-s)ds}_{m_t(u)}\right\}. \tag{3.20}$$

For the cMGF of X we have by theorem 1.5 that the solution is $X_u = g(u) + H_u + \nu Y_u$. In particular the H process depends on μ. Let $\mathcal{G}_T = \sigma(\mu_s \mid 0 \leq s \leq T)$ then by the Tower Property we have that

$$\mathbb{E}\left[e^{wX_u} \mid \mathcal{F}_t\right] = \mathbb{E}\left[\mathbb{E}\left[e^{wX_u} \mid \mathcal{F}_t \vee \mathcal{G}_T\right] \mid \mathcal{F}_t\right]. \tag{3.21}$$

Since $g(u)$ and $H_{0,u}(u)$ are $\mathcal{F}_t \vee \mathcal{G}_T$ measurable and

$$\mathbb{E}\left[e^{w\nu Y_u} \mid \mathcal{F}_t \vee \mathcal{G}_T\right] = \exp\left\{w\nu Y_{0,t}(u) + w^2 \underbrace{\frac{\nu^2}{2}\int_t^u E_\theta^2(u-s)ds}_{e_t(u)}\right\} \tag{3.22}$$

then by linearity

$$\mathbb{E}\left[e^{wX_u} \mid \mathcal{F}_t\right] = \mathbb{E}\left[e^{w(g(u)+H_{0,t}(u)+H_{t,u}(u)+\nu Y_{0,t}(u))+w^2 e_t(u)} \mid \mathcal{F}_t\right] \tag{3.23}$$

$$= e^{w(g(u)+H_{0,t}(u)+\nu Y_{0,t}(u))+w^2 e_t(u)}\mathbb{E}\left[e^{wH_{t,u}(u)} \mid \mathcal{F}_t\right]. \tag{3.24}$$

By the Markov property of μ and since $H_{t,u}(u)$ is $\sigma(\mu_s \mid s \geq t)$ measurable we get

$$\mathbb{E}\left[e^{wH_{t,u}(u)} \mid \mathcal{F}_t\right] = \mathbb{E}\left[e^{wH_{t,u}(u)} \mid \mu_t\right] \tag{3.25}$$

$$= \mathbb{E}\left[\exp\left\{w\int_t^u E_\theta(u-x)\theta\mu_x dx\right\} \mid \mu_t\right] \tag{3.26}$$

$$= \mathbb{E}\left[\exp\left\{w\int_0^\tau E_\theta(\tau-s)\theta\mu_s ds\right\} \mid \mu_0 = z\right]\Bigg|_{\substack{\tau=u-t\\ z=\mu_t}} \tag{3.27}$$

$$=: G(w,\tau,z)\Big|_{\substack{\tau=u-t\\ z=\mu_t}} \tag{3.28}$$

where we have made the change of variable with $x = s - t$. In the following we only need $w = \eta$ hence drop the w in the G notation. We summarize our findings in the following proposition.

Proposition 3.1. For a generalized Bergomi model with MPVR defined by (3.11), then the forward variance is given by

$$\xi_t(u) = \xi_0(u)\frac{G(u-t,\mu_t)}{G(u,\mu_0)}\exp\left(\eta\Lambda(t,u) + \eta^2\lambda(t,u)\right) \tag{3.29}$$

where

$$\Lambda(t,u) = H_{0,t}(u) + \nu Y_{0,t}(u) + \overline{\nu} M_{0,t}(u) \tag{3.30}$$

and

$$\lambda(t,u) = e_t(u) - e_0(u) + m_t(u) - m_0(u). \tag{3.31}$$

Proof. Note that since $Z, \overline{Z}$ and μ are independent, thus by taking conditional expectation in equation (3.18) we get

$$\xi_t(u) = \mathbb{E}\left[V_u \mid \mathcal{F}_t\right] = A(u)e^{-\eta F(u)}\mathbb{E}\left[e^{\eta X_u} \mid \mathcal{F}_t\right]\mathbb{E}\left[e^{\eta\overline{\nu}M(u)} \mid \mathcal{F}_t\right] \tag{3.32}$$

$$= A(u)e^{\eta(-F(u)+g(u)+H_{0,t}(u)+\nu Y_{0,t}(u)+\overline{\nu}M_{0,t}(u))+\eta^2(e_t(u)+m_t(u))}G(u-t,\mu_t). \tag{3.33}$$

Specifically for $t = 0$ we get

$$\xi_0(u) = A(u)\exp\left\{\eta\left(-F(u)+g(u)\right) + \eta^2\left(e_0(u)+m_0(u)\right)\right\}G(u,\mu_0). \tag{3.34}$$

Solving for $A(u)$ and inserting into $\xi_t(u)$ we get the wanted. □

So now we have expressions for the forward variance in both cases of MPVR for a generalized Bergomi model.

4 Simulation and calibration

In the present section we will study simulation of the VIX. We will only consider the fractional case with stochastic market price of risk.

As is common in the literature we take the definition of the VIX to be[4]

$$VIX_t = \sqrt{\frac{1}{\Delta}\int_t^{t+\Delta} \xi_t(u)du} \tag{4.1}$$

where t is the time to maturity and Δ is one month.

The code is implemented in Python 3.10 using the `numpy` library and can be found on `https://github.com/Karl99Kristian/RoughVolatility`.

4.1 Direct simulation

We want to simulate directly from (4.1), so we need to simulate $\xi_t(u)$ for fixed t and $u \in (t, t+\Delta)$. Note that $\xi_t(u)$ is $\mathcal{F}_t$-measurable for all $u \geq t$ so using proposition 3.1 we need to simulate $Z, \overline{Z}$ and μ on $[0,t]$. Consider a fixed width $\Delta t > 0$ and the grid $(i \cdot \Delta t)_{i\in\mathbb{N}_0}$. We denote the indices of a sub-grid by

$$\pi_A = \{i \in \mathbb{N} \mid s_i := i \cdot \Delta t \in A\} \tag{4.2}$$

for a set $A \subseteq [0, t+\Delta]$.

To simulate the Brownian motions we simulate increments of Z and $\overline{Z}$ as a 2 dimensional uncorrelated normal i.e.

$$\left(\begin{pmatrix}\Delta Z_i \\ \Delta \overline{Z}_i\end{pmatrix}\right)_{i\in\pi_{[0,t)}}, \text{ where } \frac{1}{\sqrt{\Delta t}}\begin{pmatrix}\Delta Z_i \\ \Delta \overline{Z}_i\end{pmatrix} \sim \mathcal{N}(0, I). \tag{4.3}$$

For the CTMC μ we approximate by a discrete time Markov chain $(\mu_i)_{i\in\pi_{[0,t]}}$. We specify the state space to be $\mathcal{S} = \{\mu^0, \mu^1\}$. Let Q be the intensity matrix of μ with diagonal $-(q_1, q_2)$. To simulate μ we calculate the transition probability matrix P by

$$P = e^{Q\Delta t} \tag{4.4}$$

where e is the matrix-exponential. We then simulate $\mu_{s_{i+1}}$ by evaluating a Bernoulli random variable with probabilities given by the row of P corresponding to the state index of μ_{s_i}. We assume for simplicity that $\mu_0 = \mu^0$.

Consider the particular quantities in the expression for the forward variance. We set the $\xi_0(u) = \xi_0$ and consider it a parameter. The integrals M and Y will be approximated by a trapezoidal method, i.e.

$$Y_{0,t}(u) = \sum_{i\in\pi_{[0,t)}} \frac{E_\theta(u-s_i) + E_\theta(u-s_{i+1})}{2}\Delta Z_i, \tag{4.5}$$

$$M_{0,t}(u) = \sum_{i\in\pi_{[0,t)}} \frac{K(u-s_i) + K(u-s_{i+1})}{2}\Delta \overline{Z}_i, \tag{4.6}$$

[4] See [4] for the actual definition. In continuous time our definition is sufficiently accurate.

where $K(x) = \frac{1}{x^{\frac{1}{2}-H}}$ is the fractional kernel and thus by [6, Lemma A3]

$$E_\theta(x) = x^{\alpha-1}\Gamma(\alpha)E_{\alpha,\alpha}(-\theta\Gamma(\alpha)x^\alpha) \tag{4.7}$$

where $E_{\alpha,\beta} = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n+\beta)}$ is the Mittag-Leffler function[5].

For the μ integrals note that μ is piecewise constant hence,

$$H_{0,t}(u) = \sum_{i\in\pi_{[0,t)}} \mu_{s_i} I_i(u) \tag{4.8}$$

where

$$I_i(u) = \int_{s_i}^{s_{i+1}} \theta E_\theta(u-s)ds. \tag{4.9}$$

We can determine the I integral explicitly up to a specification of the Mittag-Leffler function since by [6, Lemma A3+A5] we have for $\alpha \in (1/2, 1)$ that

$$\int_s^u \theta E_\theta(u-t)dt = \int_0^{u-s} \theta E_\theta(t)dt = 1 - E_{\alpha,1}(-\theta\Gamma(\alpha)(u-s)^\alpha) \tag{4.10}$$

so

$$I_i(u) = \int_{s_i}^u \theta E_\theta(u-t)dt - \int_{s_{i+1}}^u \theta E_\theta(u-t)dt \tag{4.11}$$

$$= E_{\alpha,1}(-\theta\Gamma(\alpha)(u-s_{i+1})^\alpha) - E_{\alpha,1}(-\theta\Gamma(\alpha)(u-s_i)^\alpha). \tag{4.12}$$

For the G functions we also need to approximate an expectation. Since we assumed a finite state space for μ we can estimate the conditional expectation by simulating M paths of μ for each initial condition in $\mathcal{S}$. Calling the j'th path $\tilde{\mu}^{(k,j)}$ with $\tilde{\mu}_0^{(k,j)} = \mu^k$ for all j and $k \in \{0,1\}$ we can utilize the law of large numbers to get

$$G(\tau,\mu^k) = \frac{1}{M}\sum_{j=1}^M \exp\left(\eta\int_0^\tau \theta E_\theta(\tau-s)\tilde{\mu}_s^{(k,j)}ds\right) \tag{4.13}$$

$$= \frac{1}{M}\sum_{j=1}^M \exp\left(\eta\sum_{i\in\pi_{[0,\tau)}} \tilde{\mu}_{s_i}^{(k,j)} I_i(\tau)ds\right). \tag{4.14}$$

For the functions in λ note that since we are in the fractional case

$$m_t(u) = \frac{\overline{\nu}}{2}\int_t^u K^2(u-s)ds = \frac{\overline{\nu}}{2}\int_t^u \frac{1}{(u-s)^{1-2H}}ds = \frac{\overline{\nu}}{2}\frac{(u-t)^{2H}}{2H} \tag{4.15}$$

since $H \in (0, 1/2)$. For the e_t functions note that

$$\frac{2}{\eta^2}e_t(u) = \int_t^u E_\theta^2(u-s)ds = \int_0^{u-t} E_\theta^2(s)ds = \int_\varepsilon^{u-t} E_\theta^2(s)ds + \int_0^\varepsilon E_\theta^2(s)ds \tag{4.16}$$

for $\varepsilon > 0$. Note that $E_\theta^2(s) = s^{2H-1}\psi(s)^2$ and since $H \in (0, 1/2)$ then s^{2H-1} may explode around 0 in a trapezoidal scheme. However, recalling the particular form of the Mittag-Leffler function it can easily be shown that ψ is slowly varying at 0, hence also ψ^2. Thus, we can utilize the idea

[5] The Mittag-Leffler function is implemented from the package [7].

form [3] to approximate the integral around 0 by letting ψ^2 be constant, namely the mean value $\frac{1}{2}(\psi^2(0)+\psi^2(\varepsilon))$, and integrate the other function analytically, such that

$$\int_0^{\varepsilon} E_\theta^2(s)ds = \frac{\psi^2(0)+\psi^2(\varepsilon)}{2}\int_0^{\varepsilon} s^{2H-1}ds = \frac{\psi^2(0)+\psi^2(\varepsilon)}{2}\frac{\varepsilon^{2H}}{2H} =: \overline{e}_\varepsilon. \tag{4.17}$$

So to approximate $e_t(u)$ we make a trapezoidal scheme away form 0 and take $\varepsilon = \Delta t$, i.e.

$$e_t(u) = \frac{\eta^2}{2}\left(\overline{e}_\varepsilon + \sum_{i\in\pi_{(0,u-t)}} \frac{E_\theta^2(s_i)+E_\theta^2(s_{i+1})}{2}\Delta t\right). \tag{4.18}$$

The simulation procedure is thus as follows:

Procedure

1. Simulate $\tilde{\mu}^{(k,j)}$ on $\pi_{[0,t+\Delta]}$ for each $k\in\{0,1\}$ and $0<j\leq M$.
2. Simulate increments of the Brownian motions on $\pi_{[0,t)}$.
3. Simulate μ on $\pi_{[0,t]}$ with $\mu_0=\mu^0$.
4. For each $i\in\pi_{[t,t+\Delta]}$ let $u=s_i$ and
   - Calculate $Y_{0,t}(u)$ and $M_{0,t}(u)$ by trapezoidal scheme,
   - Calculate $m_s(u)$ analytically and $e_s(u)$ by the hybrid scheme for $s=t$ and $s=0$,
   - Calculate $G(u,s_0)$ using LLN,
   - Based on $\mu^k=\mu_u$ calculate $G(u-t,\mu^k)$ using LLN,
   - Calculate $\xi_t(u)$ from proposition 3.1.
5. Calculate VIX_t by a trapezoidal scheme of 4.1.

In order to price options on the VIX we reproduce step 2-5 N times giving $(VIX_t^{(j)})_{j=1,\dots,N}$. Since we assume zero interest rates we can utilize the law of large numbers to get

$$Call(VIX_0,K,t) = \mathbb{E}^{\mathbb{Q}}\left[(VIX_t-K)^+\right] \approx \frac{1}{N}\sum_{j=1}^{N}(VIX_t^{(j)}-K)^+ \tag{4.19}$$

for a range of strikes K.

4.2 Data and calibration

We gather real prices of European call options on the VIX from Yahoo Finance[6] spot date is Jan. 17th 2023 for one day to maturity and Jan. 20th for 3 days, 1 month and 3 months. Strikes that have not been traded on the observation date are removed, furthermore we remove strikes that make the price highly non-convex. We use a daycount convention of 252 days. Latest trade price is used to calculate the implied volatility assuming 0 interest rate. Given parameters Θ we can run the direct simulation as described above and calculate the implied volatility $\overline{\sigma}_K(\Theta)$ for each present strike K in the data $\mathcal{K}$. We seek to minimize the mean squared error

$$MSE(\Theta) = \frac{1}{n_K}\sum_{K\in\mathcal{K}}(\sigma_K-\overline{\sigma}_K(\Theta))^2 \tag{4.20}$$

[6] https://finance.yahoo.com/quote/%5EVIX/options?p=%5EVIX.

We found the method of trust constraints[7] to be particularly effective. Since efficiency is not of particular importance in this project, the model evaluation can be very timeconsuming dependent on choice of N and Δt. Thus, we vary Δt with the time to maturity. The calibrated parameters are reported in table 1.

	min	max	ttm: 1 day	ttm: 3 days	ttm: 1 month	ttm: 3 months
H	0.07	0.13	0.101	0.0961	0.128	0.13
η	0.02	0.99	0.540	0.530	0.519	0.490
θ	0.1	10.0	1.313	5.844	5.760	0.509
ν	0.01	0.99	0.446	0.271	0.548	0.896
μ^0	0.0	1.0	0.466	0.622	0.197	0.0565
μ^1	0.0	20.0	10.510	4.701	4.900	20.000
q_0	0.0	2.0	1.001	0.651	0.893	0.913
q_1	0.0	15.0	5.953	3.426	3.713	4.344
ξ_0	0.001	0.25	0.0393	0.0426	0.0545	0.0635
Grid Δt:			5 pr day	3 pr day	2 pr day	1 pr day
$MSE \cdot 10^3$			6.8396	8.7334	1.4805	0.42667

Table 1: Constraints and calibrated parameters for fractional model with stochastic MPVR.

We can also plot the VIX smile. This is done in figure 1. We see that for the short maturities the fit is pretty good. For the longer maturities the fit is not as good. However, there are multiple reasons for this. Firstly note that the scales on the plots are not uniform, just looking at the MSE we see that these are no worse fits then for the short maturities. Furthermore, for the one-month maturity there is a liquid option in the money with very large IV giving the VIX-skew a smile. Finally, we see that for the three-months maturity we hit the constraints in H and μ^1. This indicates that on longer maturities the roughness decreases, but the jumps in MPVR can be rather large. We hypothesize that adding more states to μ might make the model better at fitting deeper ITM options and longer maturities for reasonable parameters.

5 VIX approximation

As we saw in the previous section simulation of the VIX can be quite cumbersome and time-consuming. It is common in the literature to approximate the arithmetic average of our VIX definition with a geometric average, i.e.

$$VIX_t^2 = \frac{1}{\Delta}\int_t^{t+\Delta} \xi_t(u)du \approx \exp\left(\frac{1}{\Delta}\int_t^{t+\Delta} \log \xi_t(u)du\right) =: \exp\left(N_t\right). \tag{5.1}$$

It is e.g. done in [2] to get expressions for the VVIX and in [6] and [8] as an approximate control variate. In the classical control variate technique we would need equality between $\mathbb{E}[VIX_t]$ and $\mathbb{E}[\exp(N_t/2)]$. However by an application of Jensen's inequality we have that the arithmetic average is greater than the geometric. In this section we will compare simulations of the VIX by a direct simulation as in the previous section with simulations using this approximation.

[7]This is provided by the `SciPy` package.

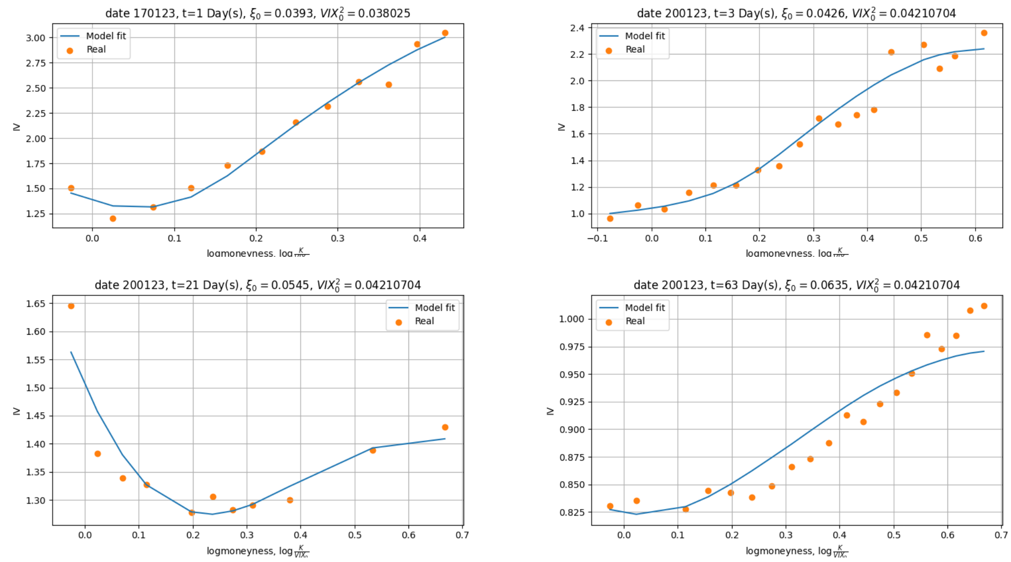

Figure 1: Calibrated VIX IV's for different maturities.

5.1 Simulation of approximation

In the following we fix a $t \geq 0$ and consider $\mathcal{G}_t = \sigma(\mu_s | s \leq t)$. By proposition 3.1 we can write

$$\log \xi_t(u) = \overline{m}(u) + \eta \left(\nu Y_{0,t}(u) + \overline{\nu} M_{0,t}(u)\right) \tag{5.2}$$

where $\overline{m}(u) = \log \xi_0(u) + \eta^2 \lambda(t,u) + \log \frac{G(u-t,\mu_t)}{G(u,\mu_0)} + \eta H_{0,t}(u)$. Note that Y and M are independent centered Gaussian, specifically also when conditioning on $\mathcal{G}$. Hence,

$$\mu_N := \mathbb{E}[N_t \mid \mathcal{G}_t] = \frac{1}{\Delta} \int_t^{t+\Delta} \overline{m}(u) du \tag{5.3}$$

and

$$\sigma_N^2 := \mathrm{Var}\left(N_t \mid \mathcal{G}_t\right) = \frac{\eta^2}{\Delta^2} \left(\nu^2 \mathrm{Var}\left(\int_t^{t+\Delta} Y_{0,t}(u) du \right) + \overline{\nu}^2 \mathrm{Var}\left(\int_t^{t+\Delta} M_{0,t}(u) du \right) \right). \tag{5.4}$$

Note in particular that the variance does not depend on μ. Using stochastic Fubini, Ito-isometry and [6, Lemma A.5] we get

$$\mathrm{Var}\left(\int_t^{t+\Delta} Y_{0,t}(u) du \right) = \frac{1}{\theta^2} \int_0^t \left(E_{\alpha,1}(-\theta\Gamma(\alpha)(t-s)^\alpha) - E_{\alpha,1}(-\theta\Gamma(\alpha)(t+\Delta-s)^\alpha) \right)^2 ds \tag{5.5}$$

and likewise by brute force integration we get

$$\mathrm{Var}\left(\int_t^{t+\Delta} M_{0,t}(u) du \right) = \frac{1}{\alpha} \left(\frac{1}{\beta} \left((t+\Delta)^\beta - \Delta^\beta + t^\beta \right) - 2 \int_0^t (s^2 + s\Delta)^\alpha ds \right) \tag{5.6}$$

with $\beta = 2\alpha + 1$. So we have

$$\log \xi_t(u)\Big|_{\mathcal{G}_t} \sim \mathcal{N}(\mu_N, \sigma_N^2). \tag{5.7}$$

So to approximate the VIX we may simulate μ to calculate $\overline{m}$ and then

$$\overline{VIX}_t = \exp\left\{\frac{1}{2}\left(\frac{1}{\Delta}\int_t^{t+\Delta} \overline{m}(u)du + \overline{N}\right)\right\} \tag{5.8}$$

where $\overline{N} \sim \mathcal{N}(0, \sigma_N^2)$ is simulated independently of μ. In the implementation we have calculated all integrals using a trapezoidal quadrature as explained above.

5.2 Results

We have simulated the VIX using the parameters from 3 days to maturity with tri-daily updates. For the direct simulation we use $N = 10,000$. We simulate the approximation across different N in a range of 100 to $1,000,000$. We then computed the first three central moments and compared with the direct simulation. The absolute error is plotted in the upper panel of figure 2. We see that the moments are not too far off. However, there is a bias as expected.

In the case of deterministic MPVR we know that the forward variance is lognormal. This can be captured in the stochastic MPVR by setting μ_s constant since then $\frac{G(u-t,\mu_t)}{G(u,\mu_0)}\exp(\eta H_{0,t}(u)) = 1$. When ξ is lognormal then the VIX approximation is also exactly lognomral. However, the distribution of the true VIX from (4.1) is unknown. Using the same parameters except setting $\mu^k = \lim_{s\to\infty} \mathbb{E}\mu_s \approx 1.27$ and running the test we get the errors in the lower panel of figure 2. We see that the error is even smaller hence the approximation in [2] is justified empirically.

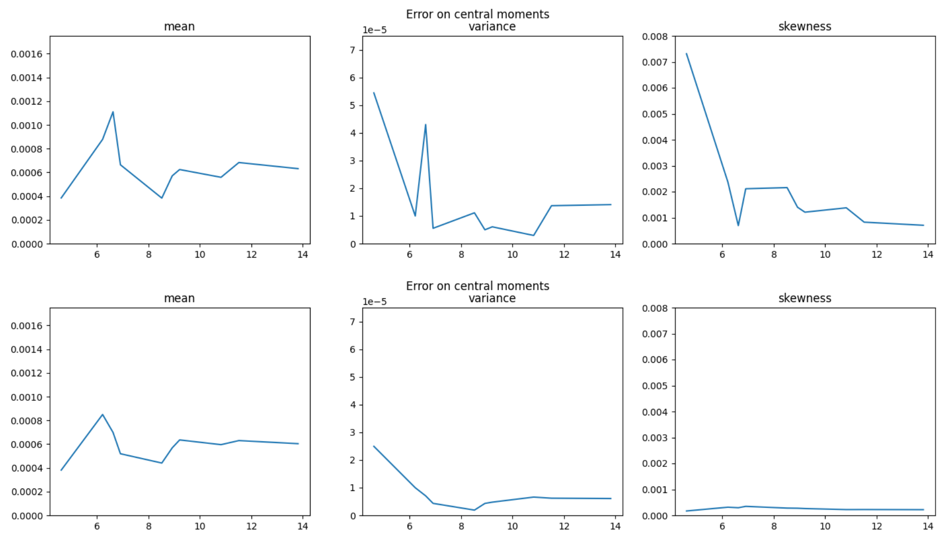

Figure 2: Error on empirical moments between direct and approximate simulation. Lower panel is lognormal case. N is on log scale.

References

[1] Abi Jaber, E.; Larsson, M.; Pulido, S. (2019), "Affine Volterra processes", https://arxiv.org/pdf/1708.08796.pdf

[2] Bayer, C.; Friz, P.; Gatheral, J. (2015), "Pricing under rough volatility", https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2554754

[3] Bennedsen, M.; Lunde, A.; Pakkanen, M. S. (2017), "Hybrid scheme for Brownian semistationary processes", https://arxiv.org/pdf/1507.03004v4.pdf

[4] Chicago Board Options Exchange (2019), "White Paper, Cboe Volatility index", https://www.sfu.ca/~poitras/419_VIX.pdf

[5] Gatheral, J., Jaisson, T. and Rosenbaum, M. (2014). "Volatility is rough", https://arxiv.org/pdf/1410.3394.pdf.

[6] Guerreiro, H. and Guerra, J. (2022). "VIX pricing in the rBergomi model under a regime switching change of measure", https://arxiv.org/pdf/2201.10391.pdf.

[7] Hinsen, K. (2017). "The Mittag-Leffler function in Python", https://github.com/khinsen/mittag-leffler.

[8] Horvath, B., Jacquier, A. and Tankov, P. (2019). "Volatility options in rough volatility models", https://arxiv.org/pdf/1802.01641v2.pdf.

[9] Mandelbrot, B.; van Ness, J.W. (1968), "Fractional Brownian motions, fractional noises and applications", SIAM Review, 10 (4): 422–437.

[10] Schilling, R. L., Partzsch, L. (2010), "Brownian Motion", (second edition), De Gruyter Graduate.