# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Rough volatility and its application to the cryptocurrency market"

verfasst von / submitted by

Andreas Celary, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

| | |
|---|---|
| Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet: | UA 066 821 |
| Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet: | Masterstudium Mathematik |
| Betreut von / Supervisor: | Dipl.-Ing. Dr. Christa Cuchiero, Privatdoz. |

# Acknowledgements

I would like to extend deep gratitude to my advisor, Prof. Christa Cuchiero for her great support and patience throughout the process of writing this thesis. Her deep knowledge and guidance were invaluable from the beginning of my research into the field of rough volatility, to the very end. My heartfelt thanks to my family for their love and support throughout the entirety of my studies. Without them, none of this would be possible.
Finally, I wish to extend thanks to all of my professors, who nurtured my passion for mathematics during the course of my studies, as well as my friends and colleagues, whom I could always count on.

# Abstract

After Gatheral et al. suggested in [11] that market volatility shows better fits using fractional Brownian motion with Hurst exponent $H < 1/2$ as a modelling basis, the resulting framework, now dubbed “rough volatility”, gained a lot of attention in mathematical finance.
In this paper, we want to introduce some of the recent theoretical developments in the area of rough volatility. Moreover, we will examine the real world market of cryptocurrencies, in particular Bitcoin, and see if it fits into the rough framework.
In the first section, we will present some of the theory behind rough models, in particular focusing on affine Volterra processes ([1],[4]) and the rough Bergomi model ([2],[13]). We will also examine forward variance swaps using the theory of infinite-dimensional stochastic calculus ([10],[14],[6]). Finally, in the second section, we will scrutinise the Bitcoin market under the scope of rough volatility using the methodology presented in [11], and see, that it indeed fits the framework. Using this knowledge, we will make an attempt at calibrating the rough Bergomi model to simulate the price and volatility dynamics of Bitcoin.


**Keywords:** rough volatility, affine Volterra processes, rough Bergomi model, rough Heston model, forward variance curves, Bitcoin

# Abstract in Deutsch

Nachdem Gatheral et al. in [11] darauf hingewiesen hat, das die Marktvolatilität besser modelliert werden kann, wenn man eine fraktionelle Brownsche Bewegung mit Hurst Exponeneten $H < 1/2$ verwendet, hat das resultierende Rahmenkonzept, nun gennant "Rough Volatility", sehr viel Interesse im Fachgebiet der Finanzmathematik geweckt.
Wir wollen in dieser Arbeit einige der in letzter Zeit entwickelten theoretischen Neurerungen im Gebiet der Rough Volatility vorstellen. Weiters werden wir den realen Markt der Kryptowährungen, insbesondere Bitcoin, untersuchen und überprüfen, ob er in den Rahmen der Rough Volatility passt.
In dem ersten Abschnitt werden wir einen Teil der Theorie der rauen Modelle näherbringen. Wir werden uns insbesondere auf affine Volterra Prozesse ([1],[4]) konzentrieren, sowie das raue Bergomi Modell ([2],[13]). Wir werden weiters Forward Variance Swaps untersuchen, unter der Verwendung der unendlichdimensionalen stochastischen Differential- und Integralrechung ([10],[14],[6]). In dem zweiten Abschnitt werden wird dann den Bitcoinmarkt aus dem Blickwinkel der Rough Volatility analysieren, unter der Verwendung der Methodik, die in [11] präsentiert wurde. Wir werden sehen, dass er sehr wohl in den Rahmen der rauen Modelle reinpasst. Mit diesem Wissen werden wir versuchen das raue Bergomi Modell zu kalibrieren, um die Preis- und Volatilitätsdynamik des Bitcoins zu simulieren.

# Contents

# Chapter 1

# Introduction

Many stochastic models have attempted to reproduce the stock market faithfully. Few, however, have withstood the test of time. With the evolution of the theory of stochastic analysis, the rapid changes in stock market infrastructure, along with the emergence of low-timeframe algorithmic trading, mathematical finance has a lot to consider, in order to provide an acceptable modelling framework.

Markets have faced many changes over the decades, and so have the corresponding models. Basic assumptions on the market had to be reconsidered and model adjustments undertaken. Mathematical finance has evolved to tackle the problems of modelling volatility, as well as model robustness. More complicated financial instruments have emerged, which came along with an overhaul of how risk can be predicted and properly priced, especially after the American housing bubble. For many years, models used standard Brownian motion as the driving force of stochastic processes as the asset price and the volatility process. With further research, mathematicians also began to employ the fractional version of Brownian motion for modelling volatility, which does not assume independent increments. Versions of fractional Brownian motion with positive correlation were used, implying long memory on the volatility process. This, however, replicates real world trading data nowadays rather poorly, as was shown in [11]. Therein, a new model was introduced, using negative correlation. The underyling fractional Brownian motion displayed short memory, resulting in very rough looking paths. These were very similar to the observed realised variance graphs measureable from real world trading data. Consequently, the model, now dubbed "rough volatility" framework, became a source of great interest.

Various models were introduced in the rough volatility setting, and applied to the global stock and derivative market. In this paper, we shall discuss some of the work done in that regard. In particular, we will introduce two major models for simulating volatility and pricing assets: the rough Bergomi model and affine Volterra models. The goal is to provide a basic introduction of the mathematical theory behind some of the most popular stochastic models before attempting to apply them to a market which has not yet faced the same scrutiny as the global stock market: cryptocurriencies. As they were introduced only a short while ago, with the emergence of Bitcoins, many questions still demand an answer. Particularly so, after the enormous bull run at the end of the year 2017 and the subsequent crash, which resulted in a general loss of interest in the cryptocurrency market. It is our opinion that the cryptocurrency market cannot be underestimated and deserves a similar mathematical interest as other assets. There is still a process on deciding how Bitcoin will be treated and whether there should be a regulated version of the market with options, ETFs replicating its price, etc. It seems that interest in these assets is still alive and there to stay in the forseeable future. We will therefore analyse the cryptocurrency market, in particular focusing on Bitcoin, and calculate its realised variance. This will serve as an approximation of the volatility process. In the next step, we will see if it fits the rough volatility framework by estimating its Hurst parameter. In the final part, we will attempt to estimate parameters and initialise a model to replicate the Bitcoin volatility and price dynamics.

It is not entirely obvious, how well Bitcoin volatility can be analysed using the methods highlighted above. As it is a relatively new concept, it has yet to reach volumes of other, more developed markets. The market is very reactionary to global news and we believe, that it is not efficient. While trades can be traced back, information is distributed in a non-obvious way. Insider trading can happen without many repercussions, as regulations have yet to reach the level of other markets. We will see that even with certain simplifications, the rough volatility framework produces results resembling the data we see very well, giving credence to our assertion on the models viability for reproducing real world observations.

# Chapter 2

# Modelling approach

Over the years, many financial models have been developed and tested. After the emergence of the theory of rough volatility, a lot of financial mathematics revolved around asset pricing in a rough framework. We want to introduce some of the theory for the market analysis performed later. We have chosen two models for this, the affine Volterra processes and the rough Bergomi model.
None of them is Markovian, so the mathematical analysis is a bit more involved than for classical financial models. They however perform very well in real world applications and serve as a solid base for reproducing financial data. In particular, in the later part of the work, we will attempt to calibrate and fit the rough Bergomi model to the Bitcoin trading dynamics.
This chapter serves as a general introduction to commonly found results and provides a base for the theory that follows.

## 2.1 Affine Volterra processes

Throughout this section we will argue on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with the usual assumptions. We want to consider processes defined by the following equation

**Definition 2.1.** Let $d, m \in \mathbb{N}$, and $W$ be an $m$-dimensional Brownian motion. Given a kernel $K \in L^2_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^{d\times d})$ and continuous functions $b : \mathbb{R}^d \to \mathbb{R}^d, \sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$, consider the following equation

$$X_t = X_0 + \int_0^t K(t-s)b(X_s)ds + \int_0^t K(t-s)\sigma(X_s)dW_s \tag{2.1}$$

A solution to (2.1) with starting value $X_0 \in \mathbb{R}^d$ is called a *stochastic Volterra process.*

We will take the approach outlined in [1], and make use of convolution theory, beautifully outlined in [12], to deduce some important existence and regularity results.

### 2.1.1 Basics in convolution theory

We want to make use of the form of the process defined in (2.1). In order to analyse its behaviour, understanding the convolution is of paramount importance. This suggests trying to impose some basic regularity and structure on the kernel.

**Definition 2.2.** Let $I \subset \mathbb{R}$ be an interval, and let $p, q \in [1, \infty]$ satisfy $1/p + 1/q = 1$. A function $K : I^2 \to \mathbb{R}^{n \times n}$ is called a kernel of type $L^p$ on $I$ if $K$ is measurable and

$$|||K|||_{L^p(I)} := \sup_{\substack{\|g\|_{L^q(I) \le 1} \\ \|f\|_{L^p(I) \le 1}}} \int_I \int_I |g(t)K(t,s)f(s)| ds dt < \infty,$$

where the supremum is taken over all scalar functions $g \in L^q(I), f \in L^p(I)$. It is called a *Volterra kernel* if it satisfies $K(t,s) = 0$ when $s > t$ for $s, t \in I^2$.

We will now introduce the convolution of kernels and functions. It will turn out that this tool simplifies a great deal of work that will go towards showing basic existence and regularity. We want to make use of the existing theory, mainly concerning solving convolution equations with the help of resolvents. A more thorough approach is provided in [12].

**Definition 2.3.** Let $f$ be measurable and $\mu$ be a measure of locally bounded variation. The convolutions $\mu * f$ and $f * \mu$ of the function $f$ with the measure $\mu$ are defined as

$$\begin{aligned} (\mu * f)(t) &= \int_{[0,T]} f(t-s)\mu(ds), \\ (f * \mu)(t) &= \int_{[0,T]} \mu(ds) f(t-s). \end{aligned} \tag{2.2}$$

**Remark 2.4.** $f$ and $\mu$ may be matrix-valued, in which case the definitions hold for whenever the expressions are defined, i.e. when the dimensions are compatible. Note that $\mu * f$ and $f * \mu$ need not be equal in this case.

**Definition 2.5.** Let $K \in L^1_{\text{loc}}\left(\mathbb{R}_+, \mathbb{R}^{d \times d}\right)$ be a kernel.

i) The $R^{d \times d}$-valued measure $L$ on $\mathbb{R}_+$ of locally bounded variation satisfying

$$K * L = L * K \equiv id. \tag{2.3}$$

is called *resolvent of the first kind* of $K$.

ii) The kernel $R \in L^1_{\text{loc}}\left(\mathbb{R}_+, R^{d \times d}\right)$ satisfying

$$K * R = R * K = K - R. \tag{2.4}$$

is called *resolvent of the second kind* (more commonly simply resolvent) of $K$.

Resolvents allow to show a basic “variation of constants” formula to solve convolution equations.

**Lemma 2.6.** *Let $X$ be a continous process, $F : \mathbb{R}_+ \to \mathbb{R}^m$, $B \in \mathbb{R}^{d\times d}$ and $Z_t = \int_0^t bds + \int_0^t \sigma dW_s$ a continous semimartingale with $b, \sigma$ continuous. Then*

$$X = F + KB * X + K * dZ \Leftrightarrow X = F - R_B * F + E_B * dZ, \tag{2.5}$$

*where $R_B$ is the resolvent of $-KB$ and $E_B = K - R_B * K$.*

*Proof.* Assume $X = F + KB * X + K * dZ$ holds. Convolving with $R_B$ and subtracting yields

$$\begin{aligned} &X - (R_B * X) \\ &= F - R_B * F + (KB - R_B * KB) * X + (KB - R_B * KB) * dZ. \end{aligned}$$

We have $K - R_B * K = E_B$ and $KB - R_B * KB = -R_B$, since $R_B$ is the resolvent of $-KB$. This yields

$$X - (R_B * X) = F - R_B * F - R_B * X + E_B * dZ,$$

and the statement follows.
For the converse, convolve both sides with $KB$ and subtract again to get

$$\begin{aligned} X - (KB * X) = F - F * KB - (R_B * F - KB * RB * F) \\ + (E_B * dZ - KB * E_B * dZ). \end{aligned}$$

Using associativity and noting that $KB*E_B = KB*(K-R_B*K) = KB*((1-R_B)*K) = (KB - KB * R_B) * K = -R_B * K$, we have

$$\begin{aligned} X - (KB * X) = F - F * KB - (R_B * F - (KB + R_B) * F) \\ + (E_B * dZ + R_B * K * dZ) = F + K * dZ. \end{aligned}$$

This proves the claim. □

### 2.1.2 Existence of solutions and regularity results

We now want to make use of the theory provided above, and by imposing some basic regularity on the functions in (2.1) provide a result of existence of weak solutions, along with a boundedness result which proves very valuable in further questions.

**Theorem 2.7.** *Let a kernel $K$ and continuous functions $\sigma$ and $b$ be given, such that the following conditions are fulfilled* *i) $K$ admits a resolvent of the first kind.*

*ii) $K \in L^2_{loc}(\mathbb{R}_+, \mathbb{R})$ and there is $\gamma \in (0, 2]$, such that $\int_0^h K^2(t)dt = O(h^\gamma)$ and $\int_0^T (K(t + h) - K(t))^2 dt = O(h^\gamma)$ for every $T < \infty$.*

*iii) $b$ and $\sigma$ satisfy the linear growth condition*

$$|b(x)| \wedge |\sigma(x)| = c_{LG}(1 + |x|), \qquad x \in \mathbb{R}^d \tag{2.6}$$

*for some constant $c_{LG}$.*

*Then (2.1) admits a continuous weak solution for any starting value $X_0 \in \mathbb{R}^d$.*

*Proof.* For the (reasonably) lengthy proof, we refer to the original paper [1]. □

**Lemma 2.8.** *Assume $b$ and $\sigma$ are continuous and satisfy the linear growth condition (2.6). Let $X$ be a continuous solution of (2.1) with starting value $X_0 \in \mathbb{R}^d$. Then, for any $p \geq 2$ and $T < \infty$, we have*

$$\sup_{t \leq T} \mathbb{E}[|X_t|^p] \leq c$$

*for some constant $c$ depending only on $|X_0|, K|_{[0,T]}, c_{LG}, p$ and $T$.*

*Proof.* Let $\tau_n := \inf\{t \geq 0 : |X_t| \geq n\} \wedge T$. We have that

$$|X_t|^p \mathbb{1}_{\{t \leq \tau_n\}} \leq \left| X_0 + \int_0^t K(t-s) b(X_t \mathbb{1}_{\{t \leq \tau_n\}}) ds + \int_0^t \sigma(X_t \mathbb{1}_{\{t \leq \tau_n\}}) dW_s \right|^p$$

Using the inequality $(a + b)^p \leq 2^p(a^p + b^p)$ two times, we get

$$\begin{aligned} |X_t|^p \mathbb{1}_{\{t \leq \tau_n\}}| &\leq 2^p |X_0|^p \\ &\quad + 4^p \underbrace{\left| \int_0^t K(t-s) b(X_s \mathbb{1}_{\{s \leq \tau_n\}}) ds \right|^p}_{\mathbf{I}} \\ &\quad + 4^p \underbrace{\left| \int_0^t K(t-s) \sigma(X_s \mathbb{1}_{\{s \leq \tau_n\}}) dW_s \right|^p}_{\mathbf{II}}. \end{aligned}$$

We shall analyse the terms one at a time. Taking expectations, using Jensen's inequality, and the linear growth assumption on $\mathbf{I}$, we obtain

$$\begin{aligned} \mathbb{E}[\mathbf{I}] &\leq c_{LG} \int_0^t |K(t-s)|^p (1 + \mathbb{E}[|X_s \mathbb{1}_{\{s \leq \tau_n\}}|])^p ds \\ &\leq 2^p c_{LG} \|K\|^{p-2}_{L^2([0,T])} \left( \|K\|^2_{L^2([0,T])} + \int_0^t |K(t-s)^2 \mathbb{E}[|X_s|^p \mathbb{1}_{\{s \leq \tau_n\}}] ds \right) \\ &\leq C_1 + C_1 |K|^2 * \mathbb{E}[|X_t|^p \mathbb{1}_{\{t \leq \tau_n\}}], \end{aligned}$$

where $C_1$ is a constant depending only on $K|_{[0,T]}$, $c_{LG}$, $p$ and T. For the second term we use the Burkholder-Davis-Gundy inequalities for the continuous local martingale $M := \int_0^t K(t-s)\sigma(X_s \mathbb{1}_{\{s\leq\tau_n\}})dW_s$ to get

$$\begin{aligned}\mathbb{E}[\mathbf{II}] &\leq C_p \mathbb{E}[\langle M\rangle_t^{p/2} \mathbb{1}_{\{t\leq\tau_n\}}] \\ &= C_p \mathbb{E}\left[\left(\int_0^t K(t-s)^2 \sigma(X_s \mathbb{1}_{\{t\leq\tau_n\}})\sigma(X_s \mathbb{1}_{\{t\leq\tau_n\}})^\top ds\right)^{p/2}\right].\end{aligned}$$

Now we can use analogous arguments as in the computations of the first term to similarly obtain

$$\mathbb{E}[\mathbf{II}] \leq C_2 + C_2|K|^2 * \mathbb{E}[|X_t|^p \mathbb{1}_{\{t\leq\tau_n\}}].$$

Putting all terms together we thus get

$$\mathbb{E}[|X_t|^p \mathbb{1}_{\{t\leq\tau_n\}}] \leq C + C|K|^2 * \mathbb{E}[|X_t|^p \mathbb{1}_{\{t\leq\tau_n\}},$$

with constant $C$ depending only on $|X_0|, K|_{[0,T]}, c_{LG}, p$ and $T$.
Consider now the kernel $K'(t,s) = C|K(t,s)|^2 \mathbb{1}_{\{s\leq t\}}$. This is a Volterra kernel by Definition 2.2. Furthermore, for any interval $[u,v] \subset \mathbb{R}_+$, Young's inequality implies

$$|||K|||_{L^p([u,v])} \leq C\|K\|_{L([0,v-u])}. \tag{2.7}$$

Thus, $-K'$ is of type $L^1$. To proceed with the proof, we want to construct a resolvent for $-K'$. To this end, we note that for $u, v$ sufficiently small, $|||K|||_{L^p([u,v])} < 1$ by (2.7). Using [12, Corollary 9.3.14] then gives us existence of a resolvent $R'$ of type $L^1$. Since $-K'$ is non-positive, so is $R'$, by [12, Proposition 9.8.1]. We may then use the generalised Gronwall lemma [12, Lemma 9.8.2] to obtain

$$\mathbb{E}[|X_t|^p \mathbb{1}_{\{t\leq\tau_n\}}] \leq C(1 + (R' * 1))(t) \leq C(1 + (R' * 1))(T) \qquad \text{for } t \in [0,T].$$

Letting $n$ tend towards infinity and using Fatou's lemma finishes the proof. □

### 2.1.3 The affine case

We now focus our attention on the particular case of affine maps in (2.1). We will show that in this case, the processes satisfy several convenient identities, particularly the affine-exponentional transform formula derived towards the end of the section.

**Definition 2.9.** For a fixed $d \in \mathbb{N}$ and a kernel $K \in L^2_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^{d\times d})$ let $a : \mathbb{R}^d \to \mathbb{S}^d(\mathbb{R})$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ be affine maps given by

$$\begin{aligned}a(x) &= A^0 + x_1 A^1 + \ldots + x_d A^d, \\ b(x) &= b^0 + x_1 b^1 + \ldots + x_d b^d\end{aligned}$$

for some $A^k \in \mathbb{S}^d(\mathbb{R})$ and $b^k \in \mathbb{R}^d$, $k = 1, \dots, d$. Furthermore, we write

$$\begin{aligned} A(u) &= \left(uA^1u^\top, \dots, uA^du^\top\right) \qquad \text{for } u \in (\mathbb{C}^d)^*, \\ B &= \left(b^1, \dots, b^d\right). \end{aligned}$$

Let $E \subseteq \mathbb{R}^d$ and $a(x)$ be positive semidefinite for each $x \in E$.
A continunous $E$-valued process $X$ is called an *affine Volterra process* (with state space $E$) if $X$ is the solution of (2.1) with $b$ as above and $a = \sigma\sigma^\top$.

The following lemma, besides being a very nice identity, will prove convenient in showing the exponential-affine transform formula.

**Lemma 2.10.** *Let $X$ be an affine Volterra process. Then for all $t < T$, we have*

$$\begin{aligned} &\mathbb{E}[X_T|\mathcal{F}_t] \\ &= \left(id - \int_0^T R_B(s)ds\right) X_0 + \left(\int_0^T E_B(s)ds\right) b^0 + \int_0^t E_B(T-s)\sigma(X_s)dW_s \end{aligned}$$

*In particular,*

$$\mathbb{E}[X_T] = \left(id - \int_0^T R_B(s)ds\right) X_0 + \left(\int_0^T E_B(s)ds\right) b^0$$

*Proof.* Since X is affine, using the conventions for $b$, it has the form $X_t = X_0 + K * (b^0dt + \sigma(X)dW_t) + KB * X$. Using Lemma 2.6, we thus have

$$X = X_0\left(id - R_B * 1\right) + E_B\left(b^0dt + \sigma(X)dW_t\right). \tag{2.8}$$

Now, consider the process $M_t = \int_0^t E_b(T-s)\sigma(X_s)dW_s$. This is a local martingale. Its quadratic variation process $\langle M \rangle_t$ satisfies

$$\begin{aligned} \mathbb{E}\left[|\langle M \rangle_T|\right] &\le \int_0^T |E_B(T-s)|^2 \mathbb{E}\left[|\sigma(X_s)|^2\right] dW_s \\ &\le \|E_B\|_{L^2([0,T])} \max_{s\in[0,T]} \mathbb{E}\left[|\sigma(X_s)|^2\right] < \infty \end{aligned}$$

by Lemma 2.8. One can now take conditional expectations on both sides of (2.8) to obtain the claim. ☐

We now come to the final result of this section, the exponential-affine transform formula, and, as a corollary, the widely used Fourier-Laplace transform of affine Volterra processes. Because of its simplicity, many stochastic models based on affine Volterra processes prove much handier for analysis than more elaborate variants.

**Theorem 2.11.** *Let $X$ be an affine Volterra process and fix some $T < \infty$, $u \in (\mathbb{C}^d)^*$, and $f \in L^1([0,T], (\mathbb{C}^d)^*)$. Assume $\psi \in L^2([0,T], (\mathbb{C}^d)^*)$ solves the Ricatti-Volterra equation*

$$\psi = uK + \left(f + \psi B + \frac{1}{2}A(\psi)\right) * K. \tag{2.9}$$

*Then, the process $Y_t, 0 \leq t \leq T$ defined by*

$$Y_t = Y_0 + \int_0^t \psi(T-s)\sigma(X_s)dW_s - \frac{1}{2}\int_0^t \psi(T-s)a(X_s)\psi(T-s)^\top ds, \tag{2.10}$$

$$Y_0 = uX_0 + \int_0^T \left(f(s)X_0 + \psi(s)b(X_0) + \frac{1}{2}\psi(s)a(X_0)\psi(s)^\top\right) ds \tag{2.11}$$

*satisfies*

$$Y_t = \mathbb{E}\left[X_T + (f*X)_T|\mathcal{F}_t\right] + \frac{1}{2}\int_t^T \psi(T-s)a(\mathbb{E}\left[X_s|\mathcal{F}_t\right])\psi(T-s)^\top ds \tag{2.12}$$

*for all $0 \leq t \leq T$. The process $\{\exp(Y_t), 0 \leq t \leq T\}$ is a local martingale and, if it is a true martingale, we have the following exponential-affine transformation formula*

$$\mathbb{E}\left[\exp\left(uX_t + (f*X)_t\right)|\mathcal{F}_t\right] = \exp(Y_t), \qquad t \leq T. \tag{2.13}$$

In order to prove the theorem, we will first need to show a prelimnary result.

**Lemma 2.12.** *The Ricatti-Volterra equation is equivalent to*

$$\psi = uE_B + \left(f + \frac{1}{2}A(\psi)\right) * E_B, \tag{2.14}$$

*where $E_B = K - R_B * K$, and $R_B$ is the resolvent of $-KB$.*

*Proof.* Assume (2.14) holds. We note that $E_B * (BK) = (K - R_B * K) * K = K * (BK) + (KB - R_B) * K = -R_B * K$. Now we can write

$$\begin{aligned}\psi - \psi * (BK) &= u(E_B + R_B * K) + (f + \frac{1}{2}A(\psi)) * E_B R_B * K) \\ &= uK + (f + \frac{1}{2}A(\psi)) * K.\end{aligned}$$

This is exactly the original equation.
For the converse, assume that (2.9) holds. Denote by $\tilde{R}_B$ the resolvent of $-BK$. Convolving on both sides and subtracting yields

$$\begin{aligned}&\psi - \psi * \tilde{R}_B \\ &= u(K - K * \tilde{R}_B) + \left(f + \frac{1}{2}A(\psi)\right) * (K - K * \tilde{R}_B) + \psi(BK - BK * \tilde{R}_B) \\ &= u(K - K * \tilde{R}_B) + \left(f + \frac{1}{2}A(\psi)\right) * (K - K * \tilde{R}_B) - \psi * \tilde{R}_B.\end{aligned}$$

Therefore, we have to show that $R_B * K = K * \tilde{R}_B$. We will show the equivalent condition, that for each $T < \infty$, there is a $\sigma > 0$, such that

$$(e^{-\sigma t} R_B) * (e^{-\sigma t} K) = (e^{-\sigma t} K) * (e^{-\sigma t} \tilde{R}_B) \qquad t \in [0, T]. \tag{2.15}$$

It follows from the definitions that $e^{-\sigma t} R_B$ is the resolvent of $-e^{-\sigma t} KB$, and that $e^{-\sigma t} \tilde{R}_B$ is the resolvent of $-e^{-\sigma t} BK$ [12, Lemma 2.3.3]. Choosing $\sigma$ large enough, so that $\|e^{-\sigma t} KB\|_{L^1([0,T])} < 1$, we construct an explicit expression for these resolvents as in the proof of [12, Theorem 2.3.1]:

$$e^{-\sigma t} R_B = -\sum_{k \geq 1} (e^{-\sigma t} KB)^{*k} \text{ and } e^{-\sigma t} \tilde{R}_B = -\sum_{k \geq 1} (e^{-\sigma t} BK)^{*k} \quad t \in [0, T],$$

where $f^{*k}$ is the $k$-fold convolution of $f$ with itself. Plugging in, we get

$$\begin{aligned} &(e^{-\sigma t} K) * (e^{-\sigma t} \tilde{R}_B) = \int_0^t e^{-\sigma(t-s)} K(t-s) \\ &\sum_{k \geq 1} \int_0^t \cdots \int_0^{s_{k-1}} \left(e^{-\sigma(s-s_1)} BK(s-s_1)\right) \cdots \left(e^{-\sigma s_k} BK(s_k)\right) ds_k \ldots ds_1 ds. \end{aligned}$$

Collecting terms and interchanging summation and integration, along with renaming variables, we get

$$\begin{aligned} &(e^{-\sigma t} K) * (e^{-\sigma t} \tilde{R}_B) = \sum_{k \geq 1} \int_0^t \cdots \int_0^{s_{k-1}} \left(e^{-\sigma(t-s_1)} K(t-s) B\right) \cdots \\ &\cdots \left(e^{-\sigma(s_k - s)} K(s_k - s) B\right) \left(e^{-\sigma s} K(s)\right) ds_k \ldots ds_1 ds. \end{aligned}$$

A symmetric computation of the left-hand side of (2.16) proves the lemma. □

*Proof of Theorem 2.11.* Denote by $\tilde{Y}_t$ the process defined in (2.12). We first show $Y_0 = \tilde{Y}_0$. To see this, we use the identity $va(x)v^\top = vA^0 v^\top + A(v)x$ to write

$$\begin{aligned} \tilde{Y}_0 &= \mathbb{E}\left[uX_T + (f * X)_T\right] + \frac{1}{2} \int_0^T \psi(T-s) A^0 \psi(T-s)^\top + A(\psi) \mathbb{E}\left[X_s\right] ds \\ &= u\mathbb{E}\left[X_T\right] + f * \mathbb{E}\left[X_T\right] + \frac{1}{2} \int_0^T \psi(T-s) A^0 \psi(T-s)^\top + A(\psi) \mathbb{E}\left[X_s\right] ds, \end{aligned}$$

where in the last step we use the stochastic Fubini theorem. Subtracting and collecting terms yields

$$\begin{aligned} \tilde{Y}_0 - Y_0 = &\, u\mathbb{E}\left[X_T - X_0\right] + f * \mathbb{E}\left[X_T - X_0\right] \\ &- \psi * (b^0 + BX_0) + \frac{1}{2} A(\psi) * \mathbb{E}[X - X_0]. \end{aligned}$$

Now, observe that from the definition of the process $X_t$, we have

$$\mathbb{E}\left[X_T - X_0\right] = K * \mathbb{E}\left[b(X) + \sigma(X)dW\right] = K * \left(b^0 + B\mathbb{E}\left[X\right]\right).$$

And, furthermore,

$$\begin{aligned}
&\frac{1}{2}A(\psi) * \mathbb{E}[X - X_0] \\
&= \frac{1}{2}A(\psi) * K * \left(b^0 + B\mathbb{E}[X]\right) = (\psi - uK - (f + \psi B) * K) * (b^0 + B\mathbb{E}[X]) \\
&= \psi * \left(b^0 + B\mathbb{E}[X]\right) - u\mathbb{E}[X - X_0] - (f + \psi B) * \mathbb{E}[X - X_0].
\end{aligned}$$

Using Lemma 2.12., and substituting into the original equation, we finally get

$$\begin{aligned}
\tilde{Y}_0 - Y_0 &= u\mathbb{E}[X_T - X_0] + f * \mathbb{E}[X_T - X_0] - \psi * (b^0 + BX_0) \\
&\quad + \psi * (b^0 + B\mathbb{E}[X_T] - u\mathbb{E}[X_T - X_0] - (f + \psi B) * \mathbb{E}[X_T - X_0] \\
&= 0.
\end{aligned}$$

For the second part, we again make use of the identity $va(x)v^\top = vA^0v^\top + A(v)x$ and the stochastic Fubini theorem to obtain

$$\begin{aligned}
\tilde{Y}_t &= u\mathbb{E}[X_T|\mathcal{F}_t] + \int_0^T f(T-s)\mathbb{E}[X_s|\mathcal{F}_t]ds \\
&+ \frac{1}{2}\int_t^T \psi(T-s)A^0\psi(T-s)^\top + A(\psi)\mathbb{E}[X_s|\mathcal{F}_t]ds \\
&= u\mathbb{E}[X_T|\mathcal{F}_t] + \int_0^T f(T-s)\mathbb{E}[X_s|\mathcal{F}_t]ds \\
&+ \frac{1}{2}\int_0^T \psi(T-s)A^0\psi(T-s)^\top + A(\psi)\mathbb{E}[X_s|\mathcal{F}_t]ds \\
&- \frac{1}{2}\int_0^t \psi(T-s)a(\mathbb{E}[X_s|\mathcal{F}_t])\psi(T-s)^\top ds \\
&= u\mathbb{E}[X_T|\mathcal{F}_t] + \int_0^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)\mathbb{E}[X_s|\mathcal{F}_t] \\
&- \frac{1}{2}\int_0^t \psi(T-s)a(X_s)\psi(T-s)^\top ds + \underbrace{\frac{1}{2}\int_0^T \psi(T-s)A^0\psi(T-s)^\top ds}_{=:C_1},
\end{aligned}$$

where, in the last equality we used measurability of $X_s$ with respect to $\mathcal{F}_t$ for $s \leq T$ and $C_1$ is independent of $t$ and is $\mathcal{F}_0$-measurable.

For the next step, we use Lemma 2.10. and , for notational convenience, denote the $t$-independent component by $C_2$, i.e.

$$\mathbb{E}[X_T|\mathcal{F}_t] = C_2 + \int_0^t E_B(T-s)\sigma(X_s)dW_s$$

and further compute

$$
\begin{aligned}
&\int_0^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)\mathbb{E}[X_s|\mathcal{F}_t]ds \\
&= \underbrace{\int_0^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)\left(\left(id - \int_0^T R_B(r)dr\right)X_0 + \left(\int_0^T E_B(r)dr\right)b^0\right)ds}_{=:C_3} \\
&+ \int_0^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)\int_0^s E_B(T-r)\sigma(T-r)dW_r ds \\
&= C_3 + \int_0^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)\int_0^t \mathbb{1}_{\{r<s\}}E_B(s-r)\sigma(X_r)dW_r ds \\
&= C_3 + \int_0^t \int_r^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)E_B(s-r)ds\sigma(X_r)dW_r \\
&= C_3 + \int_0^t \left(f + \frac{1}{2}A(\psi)\right) * E_B(T-r)\sigma(X_r)dW_r.
\end{aligned}
$$

By grouping up terms and using Lemma 2.12., we thus get

$$
\begin{aligned}
\tilde{Y}_t = &\underbrace{C_1 + uC_2 + C_3}_{=:C} + u\int_0^t E_B(T-r)\sigma(X_r)dW_r \\
&+ \int_0^T \left(f + \frac{1}{2}A(\sigma)\right) * E_B(T-r)\sigma(X_r)dW_r \\
&- \frac{1}{2}\int_0^t \psi(T-s)a(X_s)\psi(T-s)^\top ds \\
&= C + \int_0^t \sigma(X_r)\left(uE_B(T-r) + \left(f + \frac{1}{2}A(\psi)\right) * E_B(T-r)\right)dW_r \\
&- \frac{1}{2}\int_0^t \psi(T-s)a(X_s)\psi(T-s)^\top ds \\
&= C + \int_0^t \sigma(X_r)\psi(T-r)dW_r - \frac{1}{2}\int_0^t \psi(T-s)a(X_s)\psi(T-s)^\top ds
\end{aligned}
$$

Now, what is left to be shown is that $C = \tilde{Y}_0$. We plug in the definitions to obtain

$$\begin{aligned}
C &= \frac{1}{2}\int_0^T \psi(T-s)A^0\psi(T-s)^\top ds \\
&+ u\left(id - \int_0^T R_B(s)ds\right)X_0 + u\left(\int_0^T E_B(s)ds\right)b^0 \\
&+ \int_0^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)\left(\left(id - \int_0^T R_B(r)dr\right)X_0 + \left(\int_0^T E_B(r)dr\right)b^0\right)ds \\
&= b^0\int_0^T uE_B(s) + \left(f + \frac{1}{2}A(\psi)\right) * E_B(s)ds + \frac{1}{2}\int_0^T \psi(T-s)A^0\psi(T-s)^\top \\
&+ \left(f + \frac{1}{2}A(\psi)\right)(T-s)\left(\left(id - \int_0^T R_B(r)dr\right)X_0\right)ds \\
&+ uX_0\left(id - \int_0^T R_B(s)ds\right) \\
&= b^0\int_0^T \psi(T-s)ds + \frac{1}{2}\int_0^T \psi(T-s)A^0\psi(T-s)^\top + \frac{1}{2}A(\psi)X_0 ds \\
&+ uX_0 - X_0\int_0^T \left(f + \frac{1}{2}A(\psi)\right) * R_B(s) + uR_B(s)ds + X_0\int_0^T f(T-s)ds.
\end{aligned}$$

We observe that, since $E_B * B = (K - R_B * K) * B = KB - R_B * (KB) = KB - (KB - R_B) = R_B$, we have

$$\begin{aligned}
\psi B &= uE_B * B + \left(f + \frac{1}{2}A(\psi)\right) * E_B * B = \\
&= uR_B + \left(f + \frac{1}{2}A(\psi)\right) * R_B.
\end{aligned}$$

Plugging in this identity yields

$$\begin{aligned}
C &= b^0\int_0^T \psi(T-s) + \frac{1}{2}\psi(T-s)a(X_0)\psi(T-s)^\top ds \\
&+ uX_0 - X_0\int_0^T B\psi(T-s)ds \\
&= uX_0 + \int_0^T f(s)X_0 + \psi(s)b(X_0) + \frac{1}{2}\psi(s)a(X_0)\psi(s)^\top ds = Y_0.
\end{aligned}$$

This finishes the proof of the initial statement of the theorem.
To prove that $\exp(Y_t)$ is a local martingale, we note that by (2.9),

$$Y_t - \langle Y \rangle_t = Y_0 + \int_0^t \psi(T-s)\sigma(X_s)dW_s.$$

is a local martingale. Applying Îto's formula to $S_t := \exp(Y_t)$ swiftly shows

$$dS_t = S_t \left( dY_t + \frac{1}{2} d\langle Y \rangle_t \right),$$

so $S_t$ is a local martingale as well. Finally, in the true martingale case, we note that by (2.11),

$$\exp(Y_t) = \mathbb{E}[\exp(Y_T)|\mathcal{F}_t] = \mathbb{E}[\exp(uX_T + (f * X_T)|\mathcal{F}_t].$$

This completes the proof. □

**Remark 2.13.** In the particular case $f \equiv 0$ and $t = 0$, Theorem 2.11. yields the following expressions for the Fourier-Laplace transform of X:

$$\begin{aligned} \mathbb{E}[e^{uX_T}] &= \exp\left( \mathbb{E}[uX_T] + \frac{1}{2} \int_0^T \psi(T-s) a(\mathbb{E}[X_s]) \psi(T-s)^\top ds \right) \\ &= \exp\left(\phi(T) + \chi(T) X_0\right), \end{aligned}$$

where the functions $\phi$ and $\chi$ are defined by

$$\begin{aligned} \phi'(t) &= \psi(t) b^0 + \frac{1}{2} \psi(t) A^0 \psi(t)^\top, & \phi(0) &= 0, \\ \chi'(t) &= \psi(t) B + \frac{1}{2} A(\psi)(t), & \chi(0) &= u. \end{aligned}$$

### 2.1.4 Examples: the Volterra Heston model

In this section we want to consider an important example from the world of mathematical finance, the Heston model for stock prices. We will attempt to analyse this from the angle of Volterra processes and then consider the case of the rough Heston Model.
Take $d = 2$ and consider the process $X_t = (\log S_t, V_t)$, where $S$ is the price process and $V$ its variance, given by

$$\frac{dS_t}{S_t} = \sqrt{V_t} \left( \sqrt{1-\rho^2} dW_s^1 + \rho dW_s^2 \right) \tag{2.16}$$

$$V_t = V_0 + \int_0^t K(t-s) \left( \kappa(\theta - V_s) ds + \sigma \sqrt{V_s} dW_s^2 \right), \tag{2.17}$$

where $K \in L^2_{\text{loc}}(\mathbb{R}_+, \mathbb{R})$ is a kernel, $W = (W^1, W^2)$ is a two-dimensional standard Brownian motion, and $V_0, \kappa, \theta, \sigma \in \mathbb{R}_+, \rho \in [-1, 1]$ are parameters.
The process $S_t$ satisfies

$$\log S_t = \log S_0 - \frac{1}{2} \int_0^t V_s ds + \int_0^t \sqrt{V_s} \left( \sqrt{1-\rho^2} dW_s^1 + \rho dW_s^2 \right).$$

We therefore have for $X_t$

$$X_t = \begin{pmatrix} \log S_0 \\ V_0 \end{pmatrix} + \int_0^t \begin{pmatrix} \frac{1}{2}V_s \\ K(t-s)\kappa(\theta - V_s) \end{pmatrix} ds + \int_0^t \begin{pmatrix} \sqrt{V_s}\sqrt{1-\rho^2} & \rho \\ 0 & K(t-s)\sigma\sqrt{V_s} \end{pmatrix} dW_s.$$

This shows the process $X_t$ is indeed an affine Volterra process with kernel $K' = \begin{pmatrix} 1 & 0 \\ 0 & K \end{pmatrix}$, and $a$ and $b$ given by

$$A^0 = A^1 = 0, \quad A^2 = \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix},$$
$$b^0 = \begin{pmatrix} 0 \\ \kappa\theta \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -\frac{1}{2} \\ 0 & -\kappa \end{pmatrix}.$$

We will make use of Remark 2.13 to look at the Fourier-Laplace transform. To do this, we compute

$$A(\psi) = \left(0, \psi \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix} \psi^\top \right)^\top = \left(0, \psi_1^2 + 2\rho\sigma\psi_1\psi_2 + \sigma^2\psi_2^2\right)^\top.$$

This yields the following for the components of the Ricatti-Volterra equation

$$\psi_1 = u_1 + 1 * f_1, \tag{2.18}$$
$$\psi_2 = u_2 K + K * \left( f_2 + \frac{1}{2}(\psi_1^2 - \psi_1) - \kappa\psi_2 + \frac{1}{2}(\sigma^2\psi_2^2 + 2\rho\sigma\psi_1\psi_2) \right). \tag{2.19}$$

We want to consider the special case of the Heston model with a fractional kernel $K(t) = \frac{t^{1-\alpha}}{\Gamma(\alpha}$ for $\alpha \in (1/2, 1]$. This is the rough Heston model introduced in [8] and [9].
Using that $L(dt) = \frac{t^{-\alpha}}{\Gamma(1-\alpha)} dt$ is a resolvent of $K$, along with Lemma 2.12 and convolving with $L$, we get

$$\begin{aligned} \psi * L &= uE_B * L + \left( f + \frac{1}{2}A(\psi) \right) * E_B * L \\ &= u(id - R_B * id) + \left( f + \frac{1}{2}A(\psi) \right) * (id - R_B * id) \\ &= u + \left( f + \frac{1}{2}A(\psi) \right) * id - \left( uR_B + \left( f + \frac{1}{2}A(\psi) \right) * R_B \right) * id \\ &= u + \left( f + \frac{1}{2}A(\psi) \right) * id - \psi B * id. \end{aligned}$$

By Remark 2.13. it is easy to see that $\chi = \psi * L$ whenever $L$ is a resolvent of $K$. In the particular case of the fractional kernel we have $\psi_2 * L = I^{1-\alpha}\psi$. We therefore obtain

$$\chi = \left(\psi_1, I^{1-\alpha}\psi_2\right).$$

This yields the Fourier-Laplace transform

$$\mathbb{E}\left[e^{(u_1 \log S_T + u_2 V_T + (f_1 * \log S)_T + (f_2 * V)_T}\right] = \exp\left(\phi(T) - \psi_1(T)\log S_0 + I^{1-\alpha}\psi_2(T)V_0\right),$$

where $\psi_1$ is given by (2.18) and $\psi_2$ and $\phi$ are defined by

$$\begin{aligned} \phi' &= \kappa\theta\psi_2 & \phi(0) &= 0 \\ D^\alpha\psi_2 &= f_2 + \frac{1}{2}\left(\psi_1^2 - \psi_1\right) + (\rho\sigma\psi_1 - \kappa)\psi_2 + \frac{1}{2}\sigma^2\psi_2^2 & I^{1-\alpha}\psi_2 &= u_2. \end{aligned}$$

Here $D^\alpha = \frac{d}{dt}I^{1-\alpha}$ and $I^{1-\alpha}$ denote the fractional derivative and fractional integral, respectively.
For the more complete treatment of this case, along with existence and regularity results, we refer to the original paper [1].

## 2.2 The Rough Bergomi Model

We want to consider another model for pricing assets which aims to replicate the dynamics of the underlying volatility. Throughout this short exposition, we will follow [2]. We refer to the original paper for the more thorough treatment, along with some arguments for and against the model.
To begin with, it has been noted that at any reasonable time scales, the time series of realised variance $\sigma_t$ behaves in accordance with

$$\log\sigma_{t+\Delta} - \log\sigma_t = \nu(W^H_{t+\Delta} - W^H_t), \tag{2.20}$$

for $\Delta > 0$, where $\nu$ is a positive parameter and $W^H$ is a fractional Brownian motion with Hurst parameter H. This is the rough fractional stochastic volatility (RFSV) model introduced in [11]. We note that affine models do not satisfy this identity.
As Mandelbrot and Van Ness have shown in [16], it is possible to represent fractional Brownian Motion in terms of Wiener integrals

$$W^H_t = C_H\left(\int_{-\infty}^{t}(t-s)^{-\gamma}\,dW_s - \int_{-\infty}^{0}(-s)^{-\gamma}\,dW_s\right),$$

where $\gamma = 1/2 - H$ and $C_H = \sqrt{\frac{2H\Gamma(3/2-H)}{\Gamma(H+1/2)\Gamma(2-2H)}}$ is chosen, so that

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

one of the defining features of fractional Brownian motion.
Denote $V_t = \sigma_t^2$ and use the Mandelbrot-Van Ness representation for (2.20), to write

$$\begin{aligned} \log V_u - \log V_t &= 2\nu C_H \left( \int_{-\infty}^{u} \frac{1}{(u-s)^{\gamma}} dW_s - \int_{-\infty}^{t} \frac{1}{(t-s)^{\gamma}} dW_s \right) \\ &= 2\nu C_H \Bigg( \underbrace{\int_{t}^{u} \frac{1}{(u-s)^{\gamma}} dW_s}_{=: M_t(u)} + \underbrace{\int_{-\infty}^{t} \left( \frac{1}{(u-s)^{\gamma}} - \frac{1}{(t-s)^{\gamma}} \right) dW_s}_{=: Z_t(u)} \Bigg), \end{aligned} \tag{2.21}$$

where we have split the terms in such a way that $Z_t(u)$ is $\mathcal{F}_t$-measurable, and $M_t(u)$ is independent of $\mathcal{F}_t$ and Gaussian with mean zero and variance $\frac{1}{2H}(u-t)^{2H}$. Introduce

$$\tilde{W}_t(u) = \sqrt{2H} \int_{t}^{u} \frac{1}{(u-s)^{\gamma}} dW_s, \tag{2.22}$$

$$\eta = \frac{2\nu C_H}{\sqrt{2H}}, \tag{2.23}$$

so that $\tilde{W}_t(u)$ is a rescaled version of $M_t(u)$ with variance $(u-t)^{2H}$, and $2\nu C_H M_t(u) = \eta \tilde{W}_t(u)$. Rewriting and taking conditional expectations in (2.21), we get

$$\mathbb{E}\left[V_u | \mathcal{F}_t\right] = V_t \mathbb{E}\left[\exp\left(2\nu C_H Z_t(u) + \eta \tilde{W}_t(u)\right) | \mathcal{F}_t\right].$$

Using the $\mathcal{F}_t$-measurability of $Z_t(u)$ and the independence of $\tilde{W}_t(u)$ of $\mathcal{F}_t$, we can write

$$\mathbb{E}\left[V_u | \mathcal{F}_t\right] = V_t \exp\left(2\nu C_H Z_t(u)\right) \mathbb{E}\left[\exp\left(\eta \tilde{W}_t(u)\right)\right].$$

Finally, using the moment generating formula for the Gaussian random variable $\tilde{W}_t(u)$ with mean and variance as described above, we obtain

$$\mathbb{E}\left[V_u | \mathcal{F}_t\right] = V_t \exp\left(2\nu C_H Z_t(u) + \frac{1}{2}\eta^2 (u-t)^{2H}\right).$$

Recall the definition of the stochastic Wick exponential of a zero mean Gaussian random variable $X$:

$$\mathcal{E}(X) = \exp\left(X - \frac{1}{2}\mathbb{E}[|X|^2]\right).$$

Using this, we can write

$$\mathcal{E}(\eta \tilde{W}_t(u)) = \exp\left(\eta \tilde{W}_t(u) - \frac{1}{2}\eta^2 (u-t)^{2H}\right),$$

and rewrite the volatility as

$$\begin{aligned} V_u &= V_t \exp\left(\eta \tilde{W}_t(u) + 2\nu C_H Z_t(u)\right) \\ &= \mathbb{E}[V_u|\mathcal{F}_t]\mathcal{E}\left(\eta \tilde{W}_t(u)\right). \end{aligned} \tag{2.24}$$

We can thus write down the model for the price of an asset with an underlying volatility process

$$\begin{aligned} \frac{dS_u}{S_u} &= \mu_u du + \sqrt{V_u} dZ_u \\ V_u &= V_t \exp\left(\eta \tilde{W}_t(u) + 2\nu C_H Z_t(u)\right), \end{aligned} \tag{2.25}$$

where $\mu_u$ is a suitable drift term, and $Z_u$ is a standard Brownian motion correlated with $W_u$ used to define $\tilde{W}_t(u)$ and $Z_t(u)$, i.e.

$$dZ_t = \rho dW_t + \sqrt{1-\rho^2} dW_t^{\perp},$$

with $W_t^{\perp}$ a Brownian motion independent of $W_t$, and $\rho$ a suitable correlation coefficient.

We note here, that all calculations were assuming the underlying physical measure $\mathbb{P}$. In the next step, we want to change from $\mathbb{P}$ to an equivalent martingale measure. Henceforth, all processes will be equipped with a superscript for their corresponding measure. We will also use

$$\tilde{W}_t^{\cdot}(u) = \sqrt{2H} \int_t^u \frac{1}{(u-s)^{\gamma}} dW_s^{\cdot}$$

for the fractional process defined above with the corresponding measure-dependent Brownian motion.
Assume now that the two Brownian motions $Z^{\mathbb{P}}$ and $W^{\mathbb{P}}$ are correlated with constant $\rho$. We now want to write down the model in terms of the equivalent martingale measure $\mathbb{Q} \sim \mathbb{P}$ using this information. Under $\mathbb{Q}$, the asset price process is a martingale, that is

$$\frac{dS_u}{S_u} = \sqrt{V_u} dZ_u^{\mathbb{Q}}.$$

By Girsanov's theorem, we have

$$dZ_u^{\mathbb{Q}} = dZ_u^{\mathbb{P}} + \frac{\mu_u}{\sqrt{V_u}} du.$$

Since $Z^{\mathbb{P}}$ and $W^{\mathbb{P}}$ are of constant correlation $\rho$, they fulfill

$$W^{\mathbb{P}} = \rho Z^{\mathbb{P}} + \bar{\rho} \bar{Z}^{\mathbb{P}}, \qquad \rho^2 + \bar{\rho}^2 = 1,$$

where $\bar{Z}^{\mathbb{P}}$ is a standard Brownian motion independent of $Z^{\mathbb{P}}$. The change of measure for this second component then has the form

$$d\bar{Z}_u^{\mathbb{Q}} = d\bar{Z}_u^{\mathbb{P}} + \gamma_u du,$$

with a suitable adapted process $\gamma_u$. We can now write down the formula for $W^{\mathbb{Q}}$ under the change of measure:

$$\begin{aligned} W^{\mathbb{Q}} &:= \rho Z^{\mathbb{Q}} + \bar{\rho}\bar{Z}^{\mathbb{Q}}, \\ dW_u^{\mathbb{P}} &= dW_u^{\mathbb{Q}} - \underbrace{\left(\frac{\rho\mu_u}{\sqrt{V_u}} + \bar{\rho}\gamma_u\right)}_{=:-\lambda_u} du. \end{aligned} \tag{2.26}$$

Thus, after the change of measure, (2.24) becomes

$$\begin{aligned} V_u &= \mathbb{E}^{\mathbb{P}}[V_u|\mathcal{F}_t] \exp\left(\eta\sqrt{2H}\int_t^u \frac{1}{(u-s)^\gamma} dW_s^{\mathbb{P}} - \frac{\eta^2}{2}(u-t)^{2H}\right) \\ &= \mathbb{E}^{\mathbb{P}}[V_u|\mathcal{F}_t] \exp\left(\eta\sqrt{2H}\left(\int_t^u \frac{1}{(u-s)^\gamma} dW_s^{\mathbb{Q}} + \int_t^u \frac{\lambda_s}{(u-s)^\gamma} ds\right) - \frac{\eta^2}{2}(u-t)^{2H}\right) \\ &= \mathbb{E}^{\mathbb{P}}[V_u|\mathcal{F}_t] \mathcal{E}\left(\eta\tilde{W}_t^{\mathbb{Q}}(u)\right) \exp\left(\eta\sqrt{2H}\int_t^u \frac{\lambda_s}{(u-s)^\gamma} ds\right). \end{aligned} \tag{2.27}$$

For the rough Bergomi model, we then consider the simplest case of measure change

$$dW_u^{\mathbb{P}} = dW_u^{\mathbb{Q}} + \lambda(u) du,$$

with $\lambda(u)$ a deterministic function of $u$. Plugging into (2.27), we get

$$V_u = \xi_t(u)\mathcal{E}\left(\eta\tilde{W}_t^{\mathbb{Q}}(u)\right),$$

where

$$\xi_t(u) = \mathbb{E}^{\mathbb{Q}}[V_u|\mathcal{F}_t] = \mathbb{E}^{\mathbb{P}}[V_u|\mathcal{F}_t] \exp\left(\eta\sqrt{2H}\int_t^u \frac{\lambda(s)}{(u-s)^\gamma} ds\right)$$

is the so called *forward variance curve*. It is the product of a term depending on the history of the Brownian motion ($\mathbb{E}^{\mathbb{P}}[V_u|\mathcal{F}_t]$), and the other term depending on the market price of risk $\lambda(s)$. The model is non-Markovian. In case of a non-deterministic $\lambda_u$, although $V_u$ has a lognormal conditional distribution under $\mathbb{P}$, it will not necessarily do so under $\mathbb{Q}$.
After this theoretical derivation, we want to include another representation of the rough Bergomi Model which will serve as an extension of the one derived so far, following [13]. In particular, we consider a volatility process of the form

$$\sigma_t = \Xi(t)\exp(X_t) \qquad \text{with} \qquad X_t = \int_0^t g(t,s)^\top dW_s^{\mathbb{Q}}, \tag{2.28}$$

where $W^{\mathbb{Q}}$ is a d-dimensional standard Brownian motion under the equivalent martingale measure $\mathbb{Q}$, and $g$ is a kernel satisfying the square integrability conditon

$$\int_0^t \|g(t,s)\|^2 ds < \infty, \qquad \text{for all } t \geq 0.$$

The function $\Xi(t)$ is deterministic and locally square integrable. It calibrates the initial variance curve. We can immediately recover the initial formulation of the rough Bergomi model by setting for the kernel $g(t,s) = \alpha(t-s)^{-\gamma}$ with $\alpha = 2\nu C_H$. It of course satisfies the square integrability condition. In the preceding exposition, we made use of the stochastic Wick exponential instead of the ordinary one. Both variants are of course equivalent by letting the deterministic part of the expression be absorbed by $\Xi(t)$. Thus, we see the rough Bergomi model defined in the preceding paragraphs is a special case of (2.28).
We can immediately write down the dynamics of the forward variance curve $\xi_t(u) = \mathbb{E}^{\mathbb{Q}}[\sigma_u^2|\mathcal{F}_t]$ under the equivalent martingale measure $\mathbb{Q}$:

$$d\xi_t(u) = 2\xi_t(u)g(u,t)^{\top}dW_t^{\mathbb{Q}}.$$

Because of the small number of parameters to initialise the model, we will attempt to fit it to the Bitcoin market data in the continuation of the paper.

## 2.3 Forward variance

In this section we want to examine the forward variance curve $\xi_t(u)$ a little closer. In particular, we want to approach the problem from a different angle and consider the forward rate process as a function of time to maturity of the underlying derivative. In particular, we want to consider the function

$$x \mapsto \lambda_t(x) := \xi_t(t+x) = \mathbb{E}[V_{t+x}|\mathcal{F}_t].$$

This is the so-called Musiela parametrisation. This gives us the following dynamics

$$d\lambda_t(x) = \frac{d}{dx}\lambda_t(x)dt + K(x)\lambda_t(x)dW_t, \tag{2.29}$$

where $K(x) = 2\alpha x^{-\gamma}$, which is just $g(t+x,t)$ from the formulation in the previous section. Now, considering this process, we note that it has been lifted to an infinite dimensional setting, considering curves as functions of time to maturity $x$. In case of (2.29), we are therefore dealing with an infinite-dimensional stochastic differential equation with the unbounded generator $A = \frac{d}{dx}$. In order to properly analyse (2.29), we need to provide the right setting.
We will introduce an appropriate (Hilbert) space for the process $\lambda_t$ to apply the methods of infinite-dimensional stochastic calculus. In particular, we want to derive a Feynman-Kac

formula for a contingent claim on the underlying forward variance curve. We follow here the work of Filipović in [10]. As our reference to the general theory of infinite-dimensional calculus, we will mostly be using [7].

**Remark 2.14.** Although we base our considerations around the (rough) Bergomi model, the infinite-dimensional approach is not particularly exclusive to variance curves based on that framework. We can use an analogous methodology to look at (rough) Volterra-Heston models and more. For some of the many applications, see for instance [21],[14], or [6], as well as [4].

As a beginning, we note that the forward variance curve need not be differentiable in $x$. We therefore want to relax the notions of a "solution" to (2.29) in order to work with the operator $A$ in a meaningful way.
To prepare our abstract setting, we require the following

i) An infinite-dimensional Wiener process $W$ on a Hilbert space $U$ with covariance operator $Q$.

ii) A strongly continuous semigroup $\{S(t)|t \in \mathbb{R}_+\}$ on $H$ with infinitesmial generator $A$.

iii) Measurable functions $F(t, \omega, h) : (\mathbb{R}_+ \times \Omega \times H, \mathcal{B}(\mathbb{R}_+) \times \mathcal{F} \times \mathcal{B}(H)) \to (H, \mathcal{B}(H))$ and $\sigma(t, \omega, h) = (\sigma_j(t, \omega, h))_{j\in\mathbb{N}} : (\mathbb{R}_+ \times \Omega \times H, \mathcal{B}(\mathbb{R}_+) \times \mathcal{F} \times \mathcal{B}(H)) \to (L_2^0(H), \mathcal{B}(L_2^0(H)))$, where $L_2^0(H)$ is defined as the space of Hilbert-Schmidt operators from $Q^{1/2}(U)$ into $H$.

iv) A non-random starting value $h_0 \in H$.

In order not to delve too deep into the theoretical exposition, we refer to the orginial works [10] and [7] for a more thorough treatment, along with the properly rigorous definitions of the components.
Consider the following stochastic differential equation of a process $X$ in a Hilbert space $H$

$$\begin{aligned} dX_t &= (AX_t + F(t, X_t))dt + \sigma(t, X_t)dW_t, \\ X_0 &= h_0. \end{aligned} \tag{2.30}$$

We will introduce the following solution concepts from [10, Definition 1.19] to proceed with in the continuation of this section

**Definition 2.15.** Let $X$ be a predictable process with values in $H$ satisfying for $\tau > 0$ a stopping time the following equation

$$\mathbb{P}\left[\int_0^{t\wedge\tau} \left(\|X_s\|_H + \|F(s, X_s)\|_H + \|\sigma(s, X_s)\|^2_{L_2^0(H)}\right) ds \leq \infty\right] = 1, \quad \forall t \in \mathbb{R}_+.$$

We call $X$

i) a *local mild solution* to (2.30), if the following variation of constants formula holds

$$X_t = S(t \wedge \tau)h_0 + \int_0^{t\wedge\tau} S((t \wedge \tau) - s)F(s, X_s)ds$$
$$+ \int_0^{t\wedge\tau} S((t \wedge \tau) - s)\sigma(s, X_s)dW_s, \qquad \mathbb{P}\text{-a.s.} \qquad \forall t \in \mathbb{R}_+,$$

where S denotes again the semigroup generated by A.

ii) a *local weak solution* to (2.30), if for any $\xi \in D(A^*)$ it holds

$$\langle \xi, X_t \rangle_H = \langle \xi, h_0 \rangle_H + \int_0^{t\wedge\tau} (\langle A^*\xi, X_s \rangle_H + \langle \xi, F(s, X_s) \rangle_H)\, ds$$
$$+ \int_0^{t\wedge\tau} \langle \xi, \sigma(s, X_s) \rangle_H dW_s, \qquad \mathbb{P}\text{-a.s.} \qquad \forall t \in \mathbb{R}_+.$$

iii) a *local strong solution* to (2.30), if $X \in D(A), dt \otimes d\mathbb{P}$-a.s.

$$\mathbb{P}\left[\int_0^{t\wedge\tau} \|AX_s\|_H ds \leq \infty\right] = 1, \qquad t \in \mathbb{R}_+,$$

and the following integral equation holds

$$X_t = h_0 + \int_0^{t\wedge\tau} (AX_s + F(s, X_s))\, ds$$
$$+ \int_0^{t\wedge\tau} \sigma(s, X_s)dW_s, \quad \mathbb{P}\text{-a.s.} \quad \forall t \in \mathbb{R}_+.$$

If $\tau = \infty$, we just say $X$ is a *mild*, *weak*, or *strong solution*.

We see immediately that by considering weak solutions, we can properly deal with the unbounded operator $A$. There are however some points we must address. By transforming (2.30) into the weak formulation, we can use the adjoint of $A$ to circumvent the problem of differentiability of the forward curve. We need to work with a space that will allow us to work with these operators. In particular, we will use a Hilbert space which will make the semigroup of left shifts strongly continuous, along with other desirable properties which will prove very convenient for further computation. The space in question was introduced by Filipović, see [10].

**Definition 2.16.** Let $w : \mathbb{R}_+ \to [1, \infty)$ be a non-decreasing $C^1$-function such that $w^{-\frac{1}{3}} \in L^1(\mathbb{R}_+)$. We introduce the scalar product

$$\langle f, g \rangle_w := f(0)g(0) + \int_{\mathbb{R}_+} f'(x)g'(x)w(x)dx,$$

along with the induced norm

$$\|h\|_w^2 := |h(0)|^2 + \int_{\mathbb{R}_+} |h'(x)|^2 w(x) dx,$$

and define the *forward curve space* $H_w$ as

$$H_w := \{h \in L^1_{\text{loc}}(\mathbb{R}_+) | \exists h' \in L^1_{\text{loc}}(\mathbb{R}_+) \text{ and } \|h\|_w < \infty\}.$$

**Theorem 2.17.** *The space $H_w$ equipped with $\|\cdot\|_w$ forms a separable Hilbert space satisfying the following conditions:*

*i) The functions $h \in H_w$ are continuous, and the pointwise evaluation $ev_x(h) := h(x)$ is a continuous linear functional on $H_w$ for each $x \in \mathbb{R}_+$.*

*ii) The semigroup $\{S(t) | t \in \mathbb{R}_+\}$ of left shifts, i.e. $S(t)f(x) = f(x+t)$, is strongly continuous in $H_w$ with infinitesimal generator $A = \frac{d}{dx}$.*

*iii) Define, for $f \in H_w$ and $x \in \mathbb{R}_+$,*

$$(\mathcal{T}h)(x) := h(x) \int_0^x h(\xi) d\xi.$$

*Then there exists a constant $K$, such that*

$$\|\mathcal{T}h\|_w \leq K \|h\|_w^2$$

*for all $h \in H_w$ with $\mathcal{T}h \in H_w$.*

*Proof.* We refer to the original work [10]. □

We will drop the suffix in the norm and the scalar product in the continuation of the section whenever there is no risk of confusion.
As we will be working with a functional space, and consider differential equations thereon, we want to recall the definition of a suitable notion of a functional derivative.

**Definition 2.18.** Let $(H, \|\cdot\|)$ be a Hilbert space and $f : H \to \mathbb{R}$ be a map. We say $f$ is *Fréchet differentiable* at $x \in H$ if there exists a functional $\partial f(x) \in H$, such that

$$\lim_{\|h\| \to 0} \frac{|f(x+h) - f(x) - \langle \partial f(x), h \rangle|}{\|h\|} = 0.$$

In that case, we call $\partial f(x)$ the Fréchet derivative of $f$ at $x$. In case $f$ is in that way differentiable at all points in $H$, $\partial f$ defines a function $\partial f : H \to H$ which we again can check for differentiablility. Proceeding iteratively, we thus obtain the $k$-th order Fréchet derivative of $f$ as a function

$$\partial^k f : H \to \bigotimes_{j=1}^{k} H.$$

We will follow the approach in [7, Section 9.3.2.] in the proof of the Feynman-Kac formula. In particular, assuming a measurable function $F : H_w \to \mathbb{R}$, which will serve as our contingent claim, we want to use approximations of the the process $F(\lambda_t)$ via cylindrical functions using the Filipović space. That way, we can ensure semimartingale properties and show via the infinite-dimensional Itô formula that in the limit, the process satisfies the Kolmogorov equation with the unbounded operator $A$, in complete analogy to the finite-dimensional Feynman-Kac formula. We will need additonal regularity assumptions to use the results of existence and uniqueness in the theory infinite-dimensional stochastic differential equations, however. We have to be particularly careful in case of a fractional version of the Kernel $K$, as it explodes at 0.

We first recall the definition of a cylindrical function

**Definition 2.19.** Let $H$ be a Hilbert space. A function $\phi : H \to \mathbb{R}$ is called *k-cylindrical* if there are functionals $h_1, \dots, h_k \in H$, and $\psi : \mathbb{R}^k \to \mathbb{R}$, such that $\phi(x) = \psi(h_1(x), \dots, h_k(x))$ for all $x \in H$.

There are many useful results we will be using to conduct our analysis. In particular, $C^\infty$-cylindrical functions on $H$ form a dense subspace of $L^p(H)$ for any $1 \leq p \leq \infty$. For a more thorough exposition along with proofs, we refer, for instance, to [7].

To work with cylindrical functions, we have to choose appropriate functionals for which the pairings $\langle h, A\lambda\rangle$ and $\langle h, K\lambda\rangle$ are actually defined. We therefore set

$$\begin{aligned} \mathrm{dom}(A) &= \{h \in H_w : Ah \in H_w\} \\ \mathrm{dom}(A^*) &= \{h \in H_w : \exists C \geq 0 \text{ with } |\langle h, Ax\rangle| \leq C\|x\|_w \text{ for all } x \in \mathrm{dom}(A)\} \end{aligned}.$$

The adjoint is then defined as the unique linear operator $A^* : \mathrm{dom}(A^*) \to H_w$ determined by $\langle h, A\lambda\rangle = \langle A^* h, \lambda\rangle$.

**Remark 2.20.** Set $G_t(x) = \int_0^t K(x)\lambda_s(x)dW_s$. The kernel in the fractional setting has a singularity at 0, therefore $G_t$ does not belong to the Filipović space. As observed in [6], there are two methods to accomodate this problem.

i) Restate the problem using the concept of mild solutions to obtain a *weakly mild* formulation of (2.29):

$$\langle h, \lambda_t\rangle = \langle h, S(t)\lambda_0\rangle + \int_0^t \langle h, S(t-s)K(t-s+\cdot)\lambda_s(t-s+\cdot)\rangle dW_s, \tag{2.31}$$

where $S(t)$ again denotes the shift operator of the semigroup generated by $A$.

ii) Restrict the domains $\mathrm{dom}(A)$ (and therefore $\mathrm{dom}(A^*)$) in such a way that $\langle h, G_t\rangle$ is defined. In particular, consider the set

$$H_w^z := \{h \in H_w : S(z)h \in H_w\},$$

where $S(z)$ is again the shift operator. Set $B := \bigcap_{z>0} H_w^z$. Then, for $h \in H_w, K \in B$, define $\langle\langle h, K\rangle\rangle := \lim_{z\to\infty}\langle h, K(\cdot + z)\rangle$ whenever the limit is defined and $\langle\langle h, K\rangle\rangle := \infty$ otherwise. Then $\langle\langle g, h\rangle\rangle = \langle g, h\rangle$ for each $h \in H_w$.

We note that weakly mild solutions then are weak solutions if we restrict the domain of $A^*$ to a subset $D \subseteq \mathrm{dom}(A^*)$ where $D = \{h \in \mathrm{dom}(A^*) : \langle\langle h, G_t\rangle\rangle \text{ is well-defined.}\}$.

We will try to determine the infinitesimal generator L of the process $\lambda$. For this we will work with the space of (twice) Fréchet differentiable functions. We will also need to restrict ourselves again to functions of at most polynomial growth. Let then

$$\begin{aligned} C_m^k(H_w) :=\{&F : H_w \to \mathbb{R} : \text{F is k-Fréchet differentiable and fulfills} \\ &|F(\xi)| < C\left(|1 + \|\xi\|^m\right) \text{ for some } C > 0 \text{ and all } \xi \in H_w\}. \end{aligned}$$

Let now $F \in C_m^k(H_w)$. While the process $F(\lambda_t)$ is in general not a semimartingale, we can approximate it using clindrical functions. As weak formulation, we have

$$\langle h, \lambda_t\rangle = \int_0^t \langle A^* h, \lambda_s\rangle ds + \int_0^t \langle h, K\lambda_s\rangle dW_s$$

for $h \in D$. The process $\langle h, \lambda_t\rangle$ is a semimartingale. Because of that, by choosing functionals from our appropriately restricted domain D, the n-dimensional process $F_n(\langle h_1, \lambda_t\rangle, \dots, \langle h_n, \lambda_t\rangle)$, for a twice (Fréchet) differentiable function $F_n : \mathbb{R}^n \to \mathbb{R}$ is also a semimartingale. The generator then takes the form

$$\begin{aligned} LF_n(\langle h_1, \lambda_t\rangle, \dots, \langle h_n \lambda_t\rangle) = &\sum_{i=1}^n \partial_i F_n(\langle h_1, \lambda_t\rangle, \dots, \langle h_n, \lambda_t\rangle)\langle A^* h_i, \lambda_t\rangle \\ &+ \frac{1}{2}\sum_{i,j=1}^n \partial_{ij} F_n(\langle h_1, \lambda_t\rangle, \dots, \langle h_n, \lambda_t\rangle)\langle h_i, K\lambda_t\rangle\langle h_j, K\lambda_t\rangle \end{aligned}$$

for all $n$. Because of the aforementioned density of the cylindrical functions, the generator acts on the function $F \in C_m^2(H_w)$ by

$$LF(\lambda) = \langle A^* \partial_\lambda F, \lambda\rangle + \frac{1}{2}\langle \partial_{\lambda\lambda} F, K\lambda \otimes K\lambda\rangle,$$

where, by a slight abuse of notation, $\langle\ ,\ \rangle$ denotes the multilinear pairing of the functionals $\partial_\lambda F$ and $\partial_{\lambda\lambda} F$. We also require that $\partial_\lambda F \in \mathrm{dom}(A^*)$, as well as $\partial_{\lambda\lambda F}$ is such, that $\langle \partial_{\lambda\lambda} F, K\lambda \otimes K\lambda\rangle$ is well-defined.

We want to obtain a Feynman-Kac formula following these considerations, as mentioned in the beginning of the section. Consider a general stochastic differential equation for some process $X_t$ with values on a Hilbert space $H$ of the form

$$dX_t = (AX_t + b(t, X_t))\, dt + \sigma(t, X_t)dW_t. \tag{2.32}$$

Following the general theory concerning existence and uniqueness of solutions presented in [7], we impose the following regularity conditions.

R1) $|b(t,x) - b(t,y)| + \|\sigma(t,x) - \sigma(t,y)\|_{L_2^0} < C|x-y|$ for some $C > 0$ and for all $x, y \in H,\ t \in [0,T]$

R2) $|b(t,x)|^2 + \|\sigma(t,x)\|_{L_2^0}^2 < C\left(1 + |x|^2\right)$ for some $C > 0$ and for all $x \in H,\ t \in [0,T]$

This is needed to prove the existence and uniqueness of the solution to (2.32). See [7, Hypothesis 7.1 and Theorem 9.25].
In our case, $b \equiv 0$ and $\sigma(t, \xi) = \sigma(\xi) = K(\cdot)\xi$. In case of a fractional kernel, we have to allow additional restrictions. Following Remark 2.20, we therefore want consider kernels $K \in L^2_{\text{loc}} \cap B$, with $B$ defined as in Remark 2.20(ii), such that

$$\int_0^T \sup_{t\in[0,T]} K(x+t)^2 dx < \infty$$

for each T. This is to ensure that the weakly mild formulation is well-defined. We then consider the following space.

$$\mathcal{D} = \{h \in \text{dom}(A^*) : \exists C \geq 0, \text{ so that } |\langle\langle h, Kg\rangle\rangle| < C\|g\|_w \text{ for all } g \in H_w\}$$

By considering weakly mild solutions, we can therefore apply the same reasoning as in [7, Section 9.3.2.] to obtain a function defined by the transition semigroup

$$P_t\phi(\lambda) = \mathbb{E}\left[\phi(\langle h, \lambda_T\rangle)|\lambda_t = \lambda\right] = f(t, \lambda),$$

where $f$ solves the infinite-dimensional PDE

$$\begin{aligned} \partial_t f(t,\lambda) &= Lf(t,\lambda) \\ f(T,\lambda) &= \phi(\langle h, \lambda\rangle). \end{aligned} \tag{2.33}$$

Using the same reasoning, we can also evaluate claims on a security $S$, given by the dynamics

$$dS_t = \mu S_t dt + S_t\sigma(\lambda_t)dZ_t, \tag{2.34}$$

where $\sigma : H_w \to \mathbb{R}$ and $Z$ is a Brownian motion correlated with $W$ with correlation coefficient $\rho$. The generator for the cylindrical functions then has the form

$$
\begin{aligned}
&LF_n(\langle h_1, \lambda\rangle, \ldots, \langle h_n, \lambda\rangle, s) \\
&= \sum_{i=1}^{n} \partial_i F_n(\langle h_1, \lambda\rangle, \ldots, \langle h_n, \lambda\rangle, s)\langle A^* h_i, \lambda\rangle \\
&+ \partial_{n+1} F_n(\langle h_1, \lambda\rangle, \ldots, \langle h_n, \lambda\rangle, s)\mu s \\
&+ \frac{1}{2}\sum_{i,j=1}^{n} \partial_{ij} F_n(\langle h_1, \lambda\rangle, \ldots, \langle h_n, \lambda\rangle, s)\langle h_i, K\lambda\rangle\langle h_j, K\lambda\rangle \\
&+ \sum_{i=1}^{n} \partial_{in+1} F_n(\langle h_1, \lambda\rangle, \ldots, \langle h_n, \lambda\rangle, s)\langle h_i, K\lambda\rangle s\sigma(\lambda)\rho \\
&+ \frac{1}{2}\partial_{n+1}^2 F_n(\langle h_1, \lambda\rangle, \ldots, \langle h_n, \lambda\rangle, s)s^2\sigma(\lambda)^2,
\end{aligned}
$$

where, again, we consider elements $\{h_1, \cdots, h_n\} \in H_w$. For the corresponding Feynman-Kac formula we then have the function

$$\mathbb{E}\left[\phi(\langle h, \lambda_T\rangle, S_T)|\lambda_t = \lambda, S_t = x\right] = f(t, \lambda, x),$$

where $f$ solves

$$
\begin{aligned}
\partial_t f &= \langle A^* \partial_\lambda f, \lambda\rangle + \mu x \partial_x f + x\sigma(\lambda)\rho\langle \partial_{x\lambda} f, K\lambda\rangle \\
&\quad + \frac{1}{2}\langle \partial_{\lambda\lambda} f, K\lambda \otimes K\lambda\rangle + \frac{1}{2}x^2\sigma(\lambda)^2\partial_{xx} f \qquad (2.35)\\
f(T, \lambda, x) &= \phi(\langle h, \lambda\rangle, x).
\end{aligned}
$$

# Chapter 3

# Applications to the Bitcoin market

We shall now attempt to use the rough volatility framework to analyse a real world market. We will take a look at the cryptocurrency market which rose to prominence over the last few years. As it still is relatively novel, it has not yet been treated with the same scrutiny as the security markets.
Our approach throughout this chapter will be to extract price data on the arguably most popular and widespread cryptocurrency, the Bitcoin (henceforth referred to as BTC, for brevity), as it is the largest traded asset in the class by volume. We will then compute the daily volatility and put it under scrutiny to check if it fits under the scope of the rough volatility framework. We will analyse the empirical distribution of the increments and compute the Hurst parameter to see if the process driving the volatility can be modelled as a fractional Brownian motion. This will allow us to check whether we can deduce parameters and fit a model which reproduces the stochastic dynamics of the volatility process. In the final section, we will derive a forecasting variance formula. We will see that it provides a reasonable estimator for predicting short range volatility, up to a scaling factor.
The methodology employed throughout this chapter is a version of the approach used in [11] and [2], adapted and tailored to be used on the BTC market data. We also note [20] as a paper that conducts a similar analysis on the BTC market, using a slightly different method of estimating the Hurst parameter and showing the applicability of the rough volatility framework.

## 3.1 Data description

We want to look at the price evolution of BTC. For this, we consider a fixed time horizon. We choose data from January 1st, 2014 to October 3rd, 2019. This was done to skip the time period of very small volume and make observations easier.
Time series of the BTC price were extracted from various free data mining resources. In

particular, we have the following price chart in the considered time period due to hourly data from the freely accessible BTC price data bank on Coindesk [5].

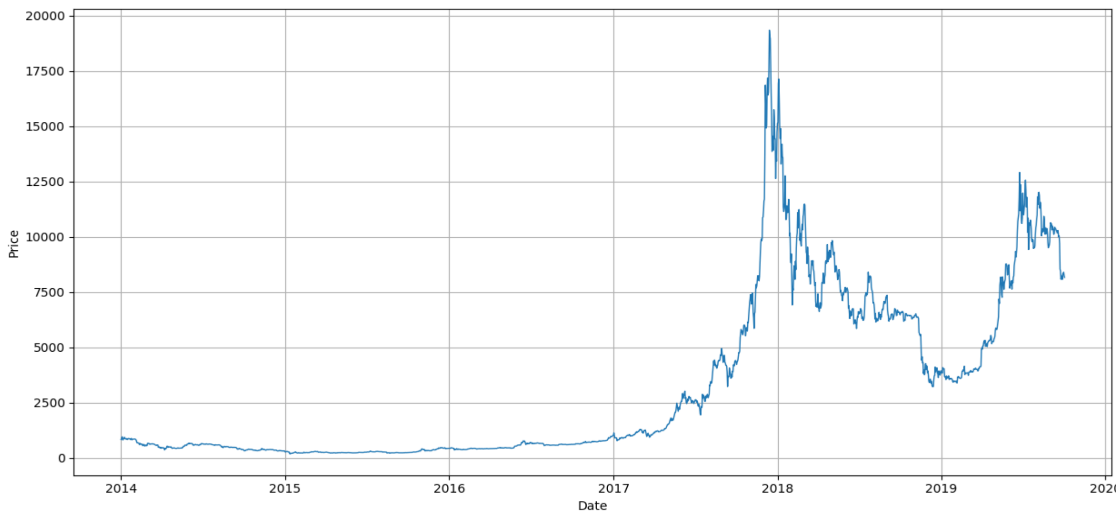



Figure 3.1: BTC price chart

## 3.2 Volatility estimation

In order to estimate the stochastic volatility, we shall be using a different set of data than the one above. We want to base our calculations on minute trading data and compute the daily volatility. This is done by computing the realised variance over all data in a given day with a predetermined time increment $\Delta$ (in this case 1 minute). Let $\{S_\tau\}$ be the sequence of minute price measurements. Denote by $\{S_t\} := \{S_{\tau_l}\}$ the subsequence of daily measurements at the start of the day, and $K_t$ be the number of minute price measurements in a given day. Then the daily empirical variance is calculated via the formula

$$\sigma_t^2 = \frac{1}{K_t} \sum_{k=0}^{K_t-1} \left(\log(S_{\tau+(k+1)\Delta}) - \log(S_{\tau+k\Delta})\right)^2 . \tag{3.1}$$

As the problem of missing trading entries seems to be a somewhat ubiquitous issue with BTC time series for small time frames, several steps have been taken to ensure an acceptable standard of quality for the data used in all future calculations. Several data sets from the free data bank service Kaggle, see [15],[17], were glued together to account for most missing entries and to fill bigger gaps of information. We use the combined data from the major online cryptocurrency exchanges Bitfinex, Coinbase, and Bitstamp. Whenever gaps of missing data were small and still left uncovered, realised variance was calculated over the remainig measurements. We thus have $K_t \approx 1313$ as a mean value for the number of minute observations in a given day as opposed to 1440. Using the fact that single entries do not cause strong deviations in sums over all minute measurements in a given day, we make the assumption that no considerable impact was made on the calculations of the empirical variance. With the mentioned steps, we produce the following daily realised variance chart for BTC.

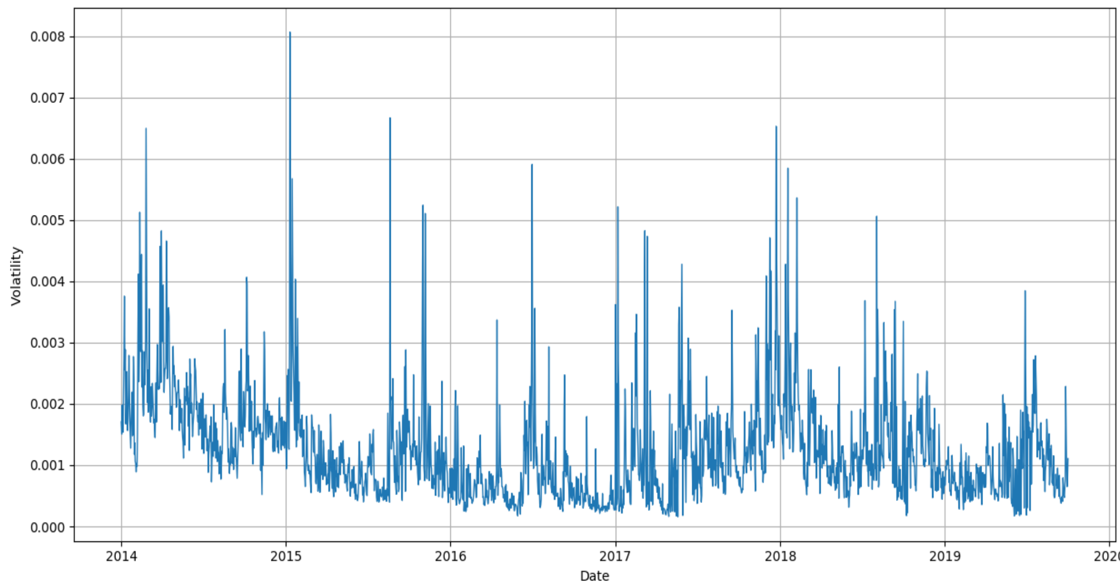



Figure 3.2: Realised volatility

Note that a fractional Brownian motion $(W_t^H)_{t\in\mathbb{R}}$ with Hurst parameter $H \in (0,1)$ satisfies for any $t \in \mathbb{R}, \Delta \geq 0, q > 0$

$$\mathbb{E}[|W_{t+\Delta}^H - W_t^H|^q] = C_q \Delta^{qH}, \tag{3.2}$$

where $C_q = \mathbb{E}[|W_1^H|^q]$. We will base our analysis around this observation and attempt to derive an estimation for the Hurst parameter of the fractional Brownian motion driving the volatility process and see if it fits the rough volatility model.
Assume we have the volatility process $\sigma_t$ on a time horizon $[0,T]$. Fix $N = \lfloor T/\Delta \rfloor$. Define

$$m(q,\Delta) = \frac{1}{N}\sum_{k=0}^{N-1} |\log(\sigma_{(k+1)\Delta}) - \log(\sigma_{k\Delta})|^q. \tag{3.3}$$

This will serve as our estimator for $\mathbb{E}[|\log(\sigma_\Delta) - \log(\sigma_0)|^q]$. We will regress the logarithm of the function $m(q,\Delta)$ against $\log\Delta$ and see, if we can find some regularity. The following

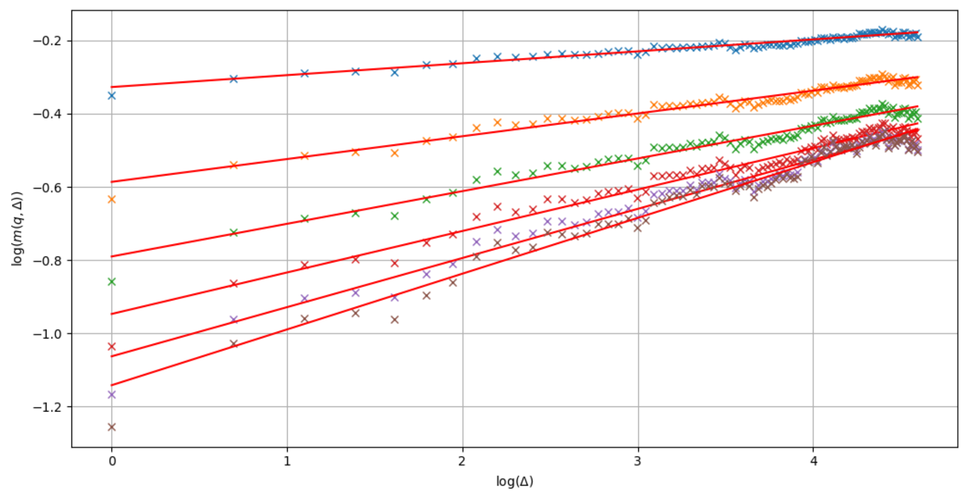



Figure 3.3: $m(q, \Delta)$ fitted with a power law for smaller values of $q$

charts present our results.

For small values of $q$, it is easy to see that the points mostly lie on a straight line. We also include the results for larger values of $q$. We present the results in two different charts for the sake of better clarity.

Note that for larger values of $q$, the lines seem to spread out with growing $\log(\Delta)$. The fit also appears to be less obvious. We still notice a decent approximation however. Assuming stationarity of the increments and that the law of large numbers can be applied, the linear fit implies that, in expectation, the log-volatility increments satisfy

$$\mathbb{E}[|\log(\sigma_\Delta) - \log(\sigma_0)|^q] = C_q \Delta^{\zeta_q}, \tag{3.4}$$

where $\zeta_q > 0$ is the slope of the line associated with $q$. Plotting $\zeta_q$ against $q$, we notice that, roughly, $\zeta_q \sim qH$ for a small enough range of $q$, which supports the results in the two charts above. As is the case in the original paper, concerning estimating the Hurst parameter for various indices, we observe a slight concavity in the curve $\zeta_q$. Evidently, approximations for greater values of $q$ get worse. It was noted that the same effect is observed in estimating the Hurst parameter for fractional Brownian motion replacing the log-volatility process for the same number of points. This effect is therefore attributed to the finite sample size, rather than inadequacies in the model. Restricting the range of $q$ lets us estimate the Hurst parameter relatively well.
A different approach may be used, however, as is noted in [20]. There, a generalised Hurst

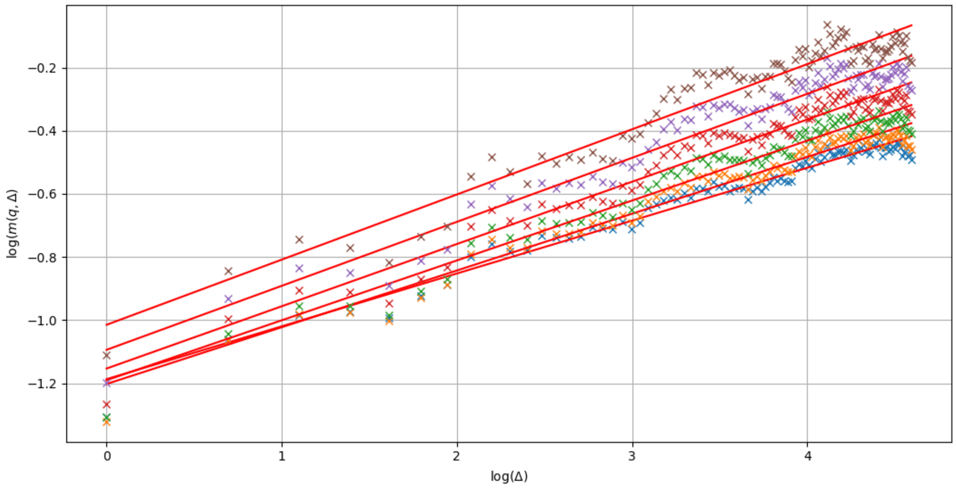
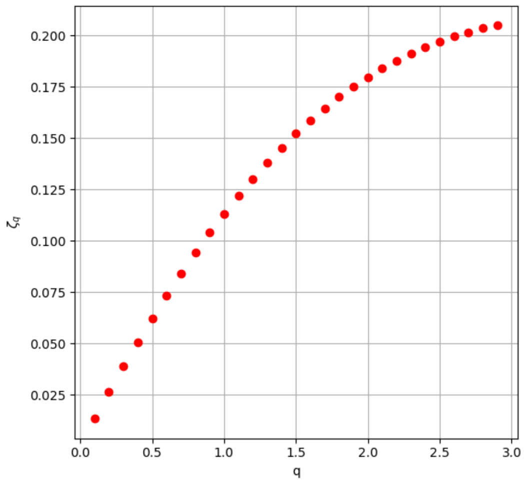


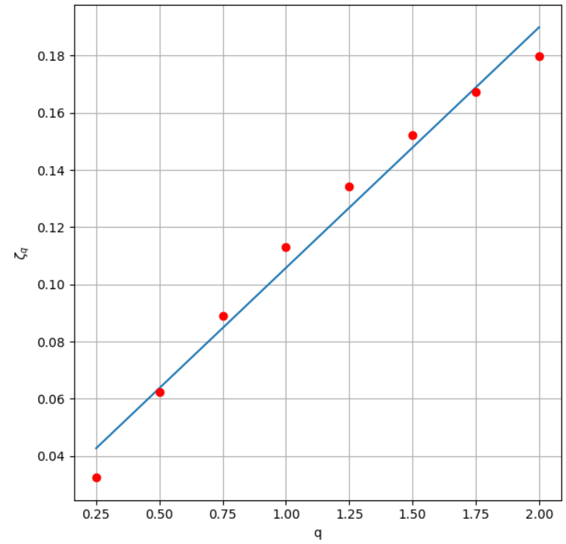

Figure 3.4: $m(q, \Delta)$ fitted with a power law for larger values of $q$

exponent is calculated using the method of multifractal detrended fluctuation analysis (MF-DFA). This accomodates the dependence of the Hurst exponent on $q$, which is consistent with our observations. The results suggest a multifractal nature of the log-volatility process, rather than a monofractal one, which a constant Hurst exponent would be evidence of. We shall not investigate those findings further, and instead refer to the original paper.
With this, the volatility of the BTC market seems to belong to the category of rough models with a Hurst parameter $H \approx 0.084$. The evident roughness is also noted in [20], albeit with a slightly higher Hurst exponent. Given the slight dependence on $q$, this may be the result of choosing our ranges differently. Slight differences may also be due to our using 1 minute trading data, as opposed to 5 minute data used in the paper. It is safe to say, that given finite sample sizes and the fact that the volatility process can only be approximated, estimators will accumulate error terms in the computations, resulting in deviations from the ideal model. In addition, we have to expect noise accumulating from measuring and calculating the realised variance in (3.1).
In the next step, we want to look at the increments $\log(\sigma_{t+\Delta}) - \log(\sigma_t)$ for varying lags $\Delta$ and determine the underyling distribution to see if fractional Brownian motion can actually be employed as the underlying stochastic process for the log-volatlity. We follow the method outlined in [11].

The histograms depict the fitted distributions for the increments, with orange being the empirical fit, red the normal fit and blue the normal fit scaled by $\Delta^H$. We can confirm that the increments fit the normal distribution, in particular with the scaling factor. This




(a) $\zeta_q$ plotted against $q$




(b) A linear polynomial fitted by restricting the range of $q$

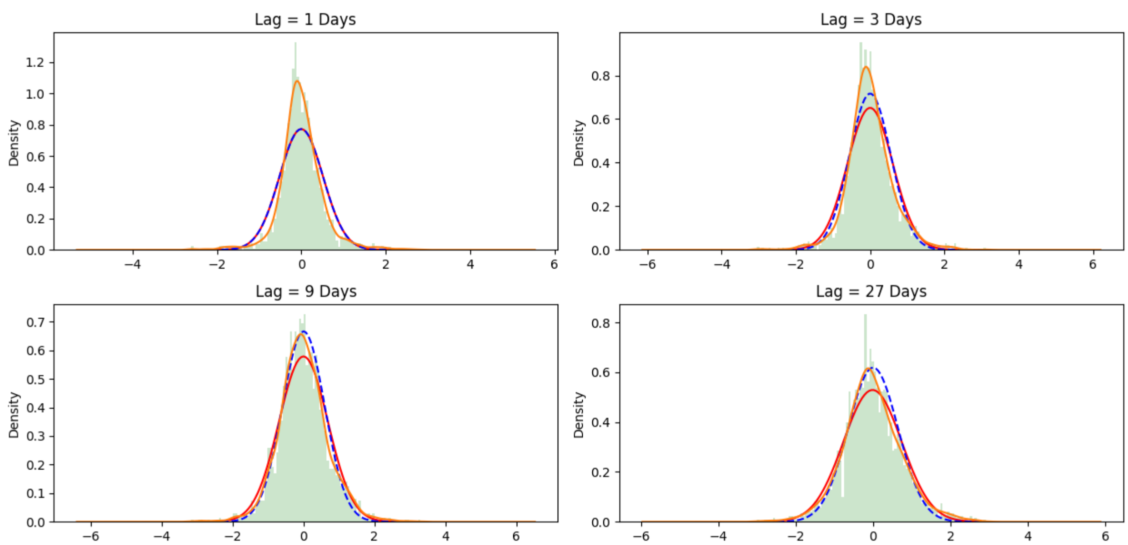



Figure 3.6: Histograms of the increments

is consistent with the findings in other markets, where a similar analysis was conducted. Putting all of the information together, we can observe the log-volatility seems to fit the fractional Brownian motion setting to an acceptable level, as noted at the start of Section

3.2. Taking into account the various possible shortcomings in the gathering of the samples, and calculations with finite precision, the rough volatility framework still appears to reflect the available data better than most classical models.

## 3.3 Model fitting

Now that we have laid our theoretical basis and gathered our data, we shall attempt to fit a model that simulates our market. As mentioned before, we will be using the rough Bergomi model in our attempts to simulate the real world BTC market and see how well it mirrors the observed volatility process, if at all. This choice was made because of the small number of parameters that have to be estimated. As the BTC market is stil very young, there are still various pieces of information we cannot extract. We have mentioned, for instance, that the rough Bergomi model depends on the forward variance curve. In an ordinary stock and derivatives framework this can be estimated by gathering data on the price of variance swap contracts and the "volatility of volatility". By gathering data on the VIX and VVIX, along with prices of the traded variance swaps, one can estimate the forward variance curve needed to initialise the model. In case of the BTC market, a reasonable derivative market is yet to be established. Thus, we cannot calibrate the forward variance curve. We have to sacrifice some precision and simplify the calibration. We will see, however, that this still provides remarkably well-fitting simulations.
For our fitting attempt, we shall make use of the form given in (2.25). We therefore outline the following process of calibrating the rough Bergomi model:

- Using volatility data, estimate the Hurst coefficient $H$.
- Estimate the parameter $\nu$ by using (2.20).
- Calibrate the model without the forward variance curve by using the form given in (2.25) and setting the origin $t = 0$.
- Guess the correlation coefficient $\rho$. As noted in [2], this can usually be done by analysing the ATM volatility skew. In particular, the product $\rho\nu$ sets the level of the ATM skew for longer expirations of options thereon. As we do not have access to such data, due to the absence of a regulated options market, we borrow the values from the SPX skew used in the paper mentioned above.

As we already have determined the Hurst parameter in our previous calculations, we can turn to estimating the volatility parameter $\nu$ of the log-volatility process. To this end, we will need the following

**Definition 3.1.** Let $X = (X_t)_{t\geq 0}$ be a stochastic process with continuous trajectories and

$q > 0$ be fixed. Define the *q-th variation* of $X$ on an interval $[0, T]$ as

$$v_q(X, [0,T]) = \lim_{n\to\infty} \sum_{k=1}^{n} \left| X_{k\frac{T}{n}} - X_{(k-1)\frac{T}{n}} \right|^q,$$

where the limit is to be understood as a limit in probability.

If the the limit exists, and it is nonzero a.s., then for any $p > q$ we have $v_p(X, [0,T]) = 0$ and for any $r < q$, $v_r(X, [0,T]) = \infty$. It is a generalisation of the quadratic variation, which is just the 2-variation. We are interested in the $1/H$-variation of a fractional Brownian motion $W^H$. In particular, as has been noted in [18], the self-similarity property of fractional Brownian motion implies that

$$\sum_{k=1}^{n} \left| X_{k\frac{T}{n}} - X_{(k-1)\frac{T}{n}} \right|^{1/H}$$

has the same distribution as

$$\frac{T}{n} \sum_{k=1}^{n} \left| X_k - X_{(k-1)} \right|^{1/H}.$$

Now, by the Ergodic Theorem (see, for instance, [3, Sections 24 and 36], this converges in $L^1$ a.s. to $c_H T$, where $c_H = \mathbb{E}\left[|W_1^H|^{1/H}\right]$. Applying this to (2.20), we get

$$\sum_{k=1}^{n} |\log \sigma_{t_k} - \log \sigma_{t_{k-1}}|^{1/H} = \nu \sum_{k=1}^{n} |W_{t_k}^H - W_{t_{k-1}}^H|^{1/H} \to \nu c_H T, \tag{3.5}$$

where $(t_k)_{k=0}^n$ is a partition of the interval $[0, T]$ which becomes finer as $n \to \infty$.
We can therefore estimate our parameter $\nu$ by computing the partial sum $v_{1/H}^{est}$ of the $1/H$-variation of our log-volitility process approximated by realised variance up to an acceptable precision. Then $\nu \approx \frac{v_{1/H}^{est}}{c_H T}$, all of which are known constants. In our case, after implementing the method in Python, we obtain the value $\nu \approx 0.043$ for the log-volatility data obtained from the BTC market.
We can now begin the calibration of our model. As stated above, we will, for simplicity, assume $t = 0$ and ignore the calibration of the forward variance curve by setting it to a constant $V_0$. We therefore want to implement the following model (see equations (2.22) and (2.25) for the corresponding definitions of the processes):

$$\begin{aligned} dS_u &= S_0 \left( \mu_u du + \sqrt{V_u} dZ_u \right) \\ V_u &= V_0 \exp\left( \eta \tilde{W}_0(u) + 2\nu C_H Z_0(u) \right), \end{aligned} \tag{3.6}$$

with an underlying regular Brownian motion $W_u$, and $Z_u$ a correlated Brownian motion with correlation parameter $\rho$. Note that $Z_u$ is not related to the process $Z_0(u)$. To generate the random paths, we use a direct approach and make use of the definitions of the processes $Z_0(u)$ and $\tilde{W}_0(u)$, and compute discretised versions of the respective Îto integrals. In order to compute $Z_0(u)$ numerically, we cut off any value below a certain arbitrary threshold. Considering all the concessions we made, claims on the precision of the simulation were kept modest. We come to conclusion, however, that even then, the model resembles the real world data remarkably well.

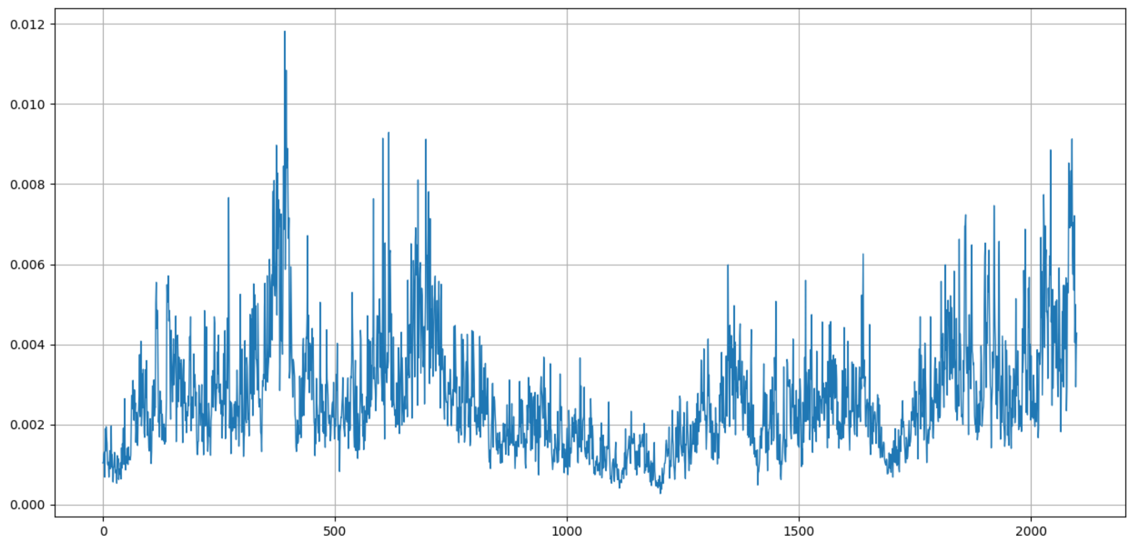



Figure 3.7: Sample path for the volatility process

While in the simulated volatility process we notice similar rough paths as in the observed data, with the same anti-persistence encountered in the framework of fractional Brownian motion. We thus conclude the model can generate very realistic volatility simulations, even with skipping the calibration of the forward variance curve.
As for the price sample paths generated using our model (3.6), we simulate some paths using the parameters, $\rho = -0.9$, $S_0 = 700$ and an arbitrary drift term $\mu_u = \frac{1}{2000}\exp(u/2000)$. We encounter frequently the same boom-bust patterns as in the observed price action in the cryptocurrency markets. We also note that the anti-persistent volatility process seems to generate rather persistent price processes, which seems to be a common occurence, as is also noted in [11].
The model appears to consistently generate very plausible paths and looks promising for further inspection, particularly in a situation, where possible derivates on BTC and other cryptocurrencies might emerge. Should such deravites ever appear, it is our conclusion that the model provides a reasonable framework for pricing and risk forecasting.

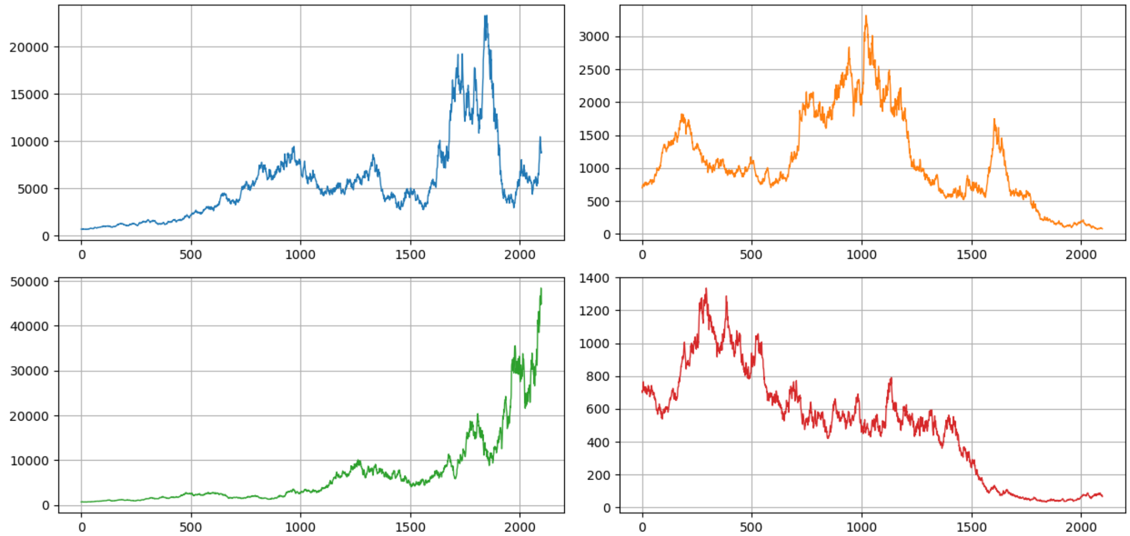



Figure 3.8: Some samples paths for the price process

## 3.4 Variance forecast

In this last section, we want to consider a forecasting method for the variance process derived from the results presented in [19]. In particular, Nuzman and Poor have shown that the "forward" fractional Brownian motion $W^H_{t+\Delta}$ is conditionally Gaussian with the following formulas for the conditional expectation and conditional variance

$$\mathbb{E}[W^H_{t+\Delta}|\mathcal{F}_t] = \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\int_{-\infty}^{t}\frac{W^H_s}{(t+\Delta-s)(t-s)^{H+1/2}}ds$$
$$\mathrm{Var}[W^H_{t+\Delta}|\mathcal{F}_t] = C_H\Delta^{2H},$$

where $C_H = \frac{\Gamma(3/2-H)}{\Gamma(H+1/2)\Gamma(2-2H)}$.
Following [2], by using the Laplace transform, we can therefore derive the following variance forecast formula

$$\mathbb{E}[V_{t+\Delta}|\mathcal{F}_t] = \exp\left(\mathbb{E}[\log(V_{t+\Delta})|\mathcal{F}_t] + 2C_H\nu^2\Delta^{2H}\right), \tag{3.7}$$

where

$$\mathbb{E}[\log(V_{t+\Delta})|\mathcal{F}_t] = \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\int_{-\infty}^{t}\frac{\log(V_s)}{(t+\Delta-s)(t-s)^{H+1/2}}ds. \tag{3.8}$$

We can implement the method in Python and compute the values $\mathbb{E}[V_{t+\Delta}|\mathcal{F}_t]$ as our predictions for the volatility. To calculate the integral in (3.8) we use Riemannian sums with step length 1/2. Since we cannot obtain the entire past history of the volatility process towards minus infinity, we instead set the arbitrary starting point at the 500th day of the observed volatility data obtained in (3.1). We make rolling predictions for every day afterwards using $\Delta = 1$, i.e. one day in advance. Denoting the predicted volatilities by $\bar{V}$, we thus calculate the following values $\bar{V}_{t+1} \approx \mathbb{E}[V_{t+1}|\mathcal{F}_t]$. We visualise the results in a the form of a scatter chart of predicted volatilities against our empirical observations.

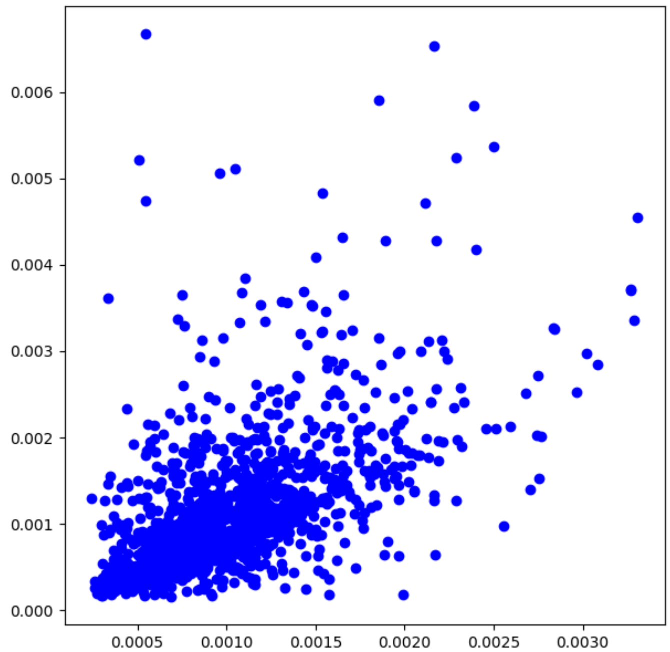



Figure 3.9: Predicted volatility against estimated observed volatility

The true strength of the forecasting formula can better be shown by superimposing the two graphs of the predicted and realised volatilites.
We see that the predicted volatily values fit our observed graph reasonably well. We notice a slight scaling discrepancy which again may be due to estimation errors and noise. Appropriate compensation may however result in even better fits.

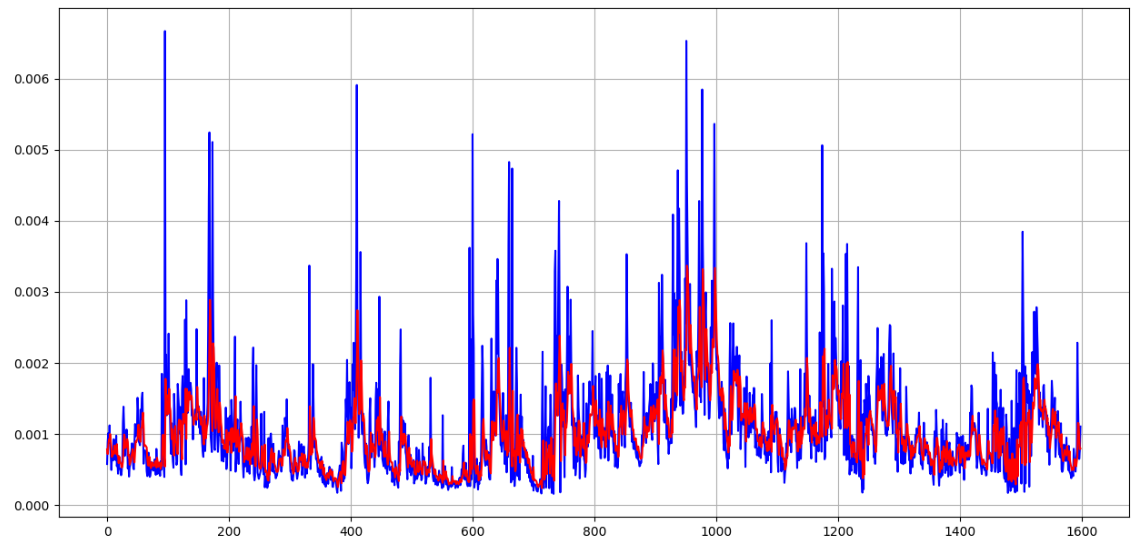



Figure 3.10: Predicted volatility in red, estimated observed volatility in blue

# Bibliography


[1] Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine Volterra processes. *arXiv e-prints*, page arXiv:1708.08796v3, Oct 2019.

[2] Christian Bayer, Peter Friz, and Jim Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16:1–18, 11 2015.

[3] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Statistics. Wiley, 2012.

[4] Christa Cuchiero and Josef Teichmann. Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case. *arXiv e-prints*, 2019.

[5] Coindesk. Bitcoin price index. `https://www.coindesk.com/price/bitcoin`, last accessed on 10.10.2019.

[6] Christa Cuchiero and Sara Svaluto-Ferro. Infinite dimensional polynomial processes. *arXiv e-prints*, page arXiv:1911.02614, 2019.

[7] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2014.

[8] Omar El Euch and Mathieu Rosenbaum. Perfect hedging in rough Heston models. *arXiv e-prints*, page arXiv:1703.05049, Mar 2017.

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

[10] Damir Filipović. Consistency Problems for Heath-Jarrow-Morton Interest Rate Models. 1760, 01 2001.

[11] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.

[12] G. Gripenberg, S. O. Londen, and O. Staffans. *Volterra Integral and Functional Equations*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1990.

[13] Blanka Horvath, Antoine Jacquier, and Peter Tankov. Volatility options in rough volatility models. *arXiv e-prints*, 2018.

[14] Antoine Jacquier and Mugad Oumgari. Deep PPDEs for rough local stochastic volatility. *arXiv e-prints*, 2019.

[15] Carsten Klein. Historical crypto currency data from the bitfinex exchange including bitcoin, version 5. `https://www.kaggle.com/tencars/392-crypto-currency-pairs-at-minute-resolution`, last accessed 11.10.2019.

[16] Benoit B. Mandelbrot and John W. Van Ness. Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review*, 10(4):422–437, 1968.

[17] Zielinski M.C. Bitcoin data at 1-min intervals from select exchanges, jan 2012 to august 2019, version 17. `https://www.kaggle.com/mczielinski/bitcoin-historical-data`, last accessed 17.10.2019.

[18] David Nualart. Fractional Brownian motion: stochastic calculus and applications. In *Proceedings of the international congress of mathematicians (ICM), Madrid, Spain, August 22–30, 2006. Volume III: Invited lectures*, pages 1541–1562. Zürich: European Mathematical Society (EMS), 2006.

[19] Carl J. Nuzman and H. Vincent Poor. Linear estimation of self-similar processes via Lamperti's transformation. *Journal of Applied Probability*, 37(2):429–452, 2000.

[20] Tetsuya Takaishi. Rough volatility of Bitcoin. *Finance Research Letters*, page 101379, Nov 2019.

[21] Frederi Viens and Jianfeng Zhang. A Martingale Approach for Fractional Brownian Motions and Related Path Dependent PDEs. *arXiv e-prints*, 2017.