UNIVERSITY
OF OSLO

Anton Yurchenko-Tytarenko

Stochastic Volterra volatility models

Thesis submitted for the degree of Philosophiae Doctor

Department of Mathematics
Faculty of Mathematics and Natural Sciences

2022

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2582

ISSN 1501-7710

Cover: UiO.
Print production: Graphics Center, University of Oslo.

I turn away with fright and horror from the lamentable evil of continuous functions with no derivatives.

– Charles Hermite

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisors Prof. Giulia Di Nunno and Prof. Yuliya Mishura for their continuous mentorship, valuable insights, immense support, and incredibly helpful advice. I would also like to thank my fellow PhD students,

> Dennis Schroers, my office mate with whom I had infinitely many discussions on all topics existing in the Universe (including mathematics, football, politics, pest control, and more football),
>
> Michele Giordano, who shared my passion in D&D and eventually became my co-author (perhaps those two things are related),
>
> Åsmund Sande, who, among other things, helped me *a lot* with proofreading my English.

I believe we all have become not just colleagues but good friends for life.

Finally, I would like to thank my family – my mother Marina, my father Vitaly, and my soulmate Hanna – for their constant and unconditional support.

Funding

This project was carried out within the frame and support of the ToppForsk project nr. 274410 of the Research Council of Norway with title STORM: Stochastics for Time-Space Risk Models.

Anton Yurchenko-Tytarenko
Oslo, October 2022

List of Papers

Paper I

Mishura, Yu., Yurchenko-Tytarenko, A.
"Standard and fractional reflected Ornstein-Uhlenbeck processes as the limits of square roots of Cox-Ingersoll-Ross processes".
Stochastics, 2022. DOI: 10.1080/17442508.2022.2047188.

Paper II

Di Nunno, G., Mishura, Yu., Yurchenko-Tytarenko, A.
"Sandwiched SDEs with unbounded drift driven by Hölder noises".
To appear in *Advances in Applied Probability* 55(3), 2023. arXiv: 2012.11465.

Paper III

Di Nunno, G., Mishura, Yu., Yurchenko-Tytarenko, A.
"Drift-implicit Euler scheme for sandwiched processes driven by Hölder noises".
To appear in *Numerical Algorithms*, 2022. arXiv: 2204.08827.

Paper IV

Di Nunno, G., Mishura, Yu. , Yurchenko-Tytarenko, A.
"Option pricing in Sandwiched Volterra Volatility model".
Submitted for publication. arXiv: 2209.10688.

Paper V

Di Nunno, G., Yurchenko-Tytarenko, A.
"Sandwiched Volterra Volatility model: Markovian approximations and hedging".
Submitted for publication. arXiv: 2209.13054.

Contents

Chapter 1

Introduction

1.1 The Dawn of Times: from Bachelier to Black-Scholes-Merton

In science, there is often a situation when entire areas of research flourish at the site of the fall of a single "seed" – a scientific article containing some revolutionary idea, which then goes through many stages of improvements, refinements and modifications. It is safe to say that mathematical finance is no exception and the role of such a "seed" is taken by Louis Bachelier's dissertation "*Théorie de la spéculation*" [7] published in 1900. He suggested to model stock price as some random process $S = \{S(t),\ t \in [0, T]\}$ and, analyzing empirical data on the Paris stock market, came to the conclusion that increments $S(t + \Delta) - S(t)$ have means that are close to zero and standard deviations of the order $\sqrt{\Delta}$. Of course, probability theory was still in its infancy at that time, so the reasoning in the dissertation lacked some mathematical rigour. However, if one translates it into modern mathematical language, one obtains that Bachelier's stock price had the form

$$S(t) = S(0) + \sigma W(t) \tag{1.1}$$

with $S(0)$ being initial price of the stock, $\sigma > 0$ and $W = \{W(t),\ t \in [0, T]\}$ being a standard Brownian motion. Now stochastic analysis is one of the main tools of economic science and we can truly appreciate Bachelier's revolutionary idea, but at that time his results went unnoticed. Only more than 50 years after the publication, statistician Jimmy Savage inexplicably stumbled upon this work and brought it to the attention of a number of researchers in economics. One of those researchers, Paul Samuelson, the winner of 1970 Nobel Memorial Prize in Economic Sciences, described Savage's finding in the preface to the English translation of Bachelier's dissertation [8] as follows:

> "*...Discovery or rediscovery of Louis Bachelier's 1900 Sorbonne thesis* [...] *initially involved a dozen or so postcards sent out from Yale by the late Jimmie Savage, a pioneer in bringing back into fashion statistical use of Bayesian probabilities. In paraphrase, the postcard's message said, approximately, 'Do any of you economist guys know about a 1914 French book on the theory of speculation by some French professor named Bachelier?'*
>
> *Apparently I was the only fish to respond to Savage's cast. The good MIT mathematical library did not possess Savage's 1914 reference. But it did have something better, namely Bachelier's original thesis itself.*

> *I rapidly spread the news of the Bachelier gem among early finance theorists. And when our MIT PhD Paul Cootner edited his collection of worthy finance papers, on my suggestion he included an English version of Bachelier's 1900 French text...*"

It was Samuelson who proposed[1] a simple but very important modification of Bachelier's approach: he used (1.1) to model *price logarithms* and not the *prices themselves*. That solved the most obvious problem: a Brownian motion is a Gaussian process and hence can take negative values with positive probability whereas stock prices are inherently non-negative. After a small adjustment with a linear trend, Samuelson's model took the form of a *geometric* or "*relative economic*" (the term used by Samuelson himself) Brownian motion

$$S(t) = S(0)\exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)\right\}, \quad \mu \in \mathbb{R}, \quad S(0),\ \sigma > 0, \tag{1.2}$$

or, in representation as a stochastic differential equation,

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t). \tag{1.3}$$

This log-normal process (1.2)–(1.3) subsequently became a mainstream choice for stock price models for the next couple of decades. Even now, when there are multiple arguments against the geometric Brownian motion, practitioners still use it as a benchmark model or a good "first approximation".

It is interesting to note that, in addition to the market model, Bachelier also considered the problem of option pricing and eventually derived an expression that can be called a harbinger of the now famous Black-Scholes formula. Of course, his reasoning was not based on the no-arbitrage principle and had a number of shortcomings inherent in any pioneering work. The correction of those shortcomings became the subject of a number of studies in the 60s, among which one can mention [17, 107, 110]. Samuelson himself also heavily contributed to that topic, see e.g. [99] or his paper [100] (in co-authorship with Robert Merton) where it was suggested to consider a warrant/option payoff as a function of the price of the underlying asset, and it can be said that these works were only a few steps away from the real breakthrough made by Black, Scholes and Merton just a couple of years later.

Here it is worth paying attention to the fact that the option market remained relatively illiquid until the end of the 60s. The reason for that was the lack of a consistent pricing methodology, and serious investors considered options as something from the world of gambling rather than worthy trading instruments. It is a little bit ironic that even Robert Merton himself – right in his breakthrough article [86]! – wrote the following:

> "*Because options are specialized and relatively unimportant financial securities, the amount of time and space devoted to the development of a pricing theory might be questioned...*"

[1] Samuelson himself acknowledged that the same idea was independently expressed by an astronomer M. Osborne in [91].

However, in 1968, a demand for that type of contract suddenly arose from the Chicago Board of Trade. This organization noticed a significant decrease in commodity futures trading on its exchange and therefore decided to create additional instruments for investors. They chose options, and, after overcoming some inevitable legal obstacles, the Chicago Board of Options Exchange began trading in 1973. Precisely in that year, two revolutionary papers appeared: "*The pricing of options and corporate liabilities*" [15] by Fischer Black and Myron Scholes and "*Theory of rational option pricing*" [86] by Robert Merton. As a side note, the publication of the Black and Scholes paper was far from a smooth process: in 1987 [14], Black recalled that the manuscript was rejected first by the *Journal of Political Economy* and then by the *Review of Economics and Statistics*. The paper was published only after Eugene Fama and Merton Miller personally recommended the *Journal of Political Economy* to reconsider its decision (in the meanwhile, Robert Merton showed a great deal of academic integrity by delaying the publication of his own article so that Black and Scholes would be the first). Interestingly enough, almost all the researchers mentioned in this paragraph – Scholes, Merton, Miller and Fama – received their Nobel Prizes in Economics in various years. Of course, Black would also be among them, but, unfortunately, he died just two years before the prize was awarded to Scholes and Merton (1997).

The main result of Black, Scholes and Merton can be formulated as follows: if a stock follows the model (1.2)–(1.3), then, under some assumptions, the discounted *no-arbitrage price* of a standard European call option V evolves as a function of the current time t and current price S and must satisfy a partial differential equation of the form

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{1.4}$$

with a boundary condition

$$V(T, S) = \max\{0, S - K\}, \tag{1.5}$$

where r denotes the instantaneous interest rate that is assumed to be constant, T is the maturity date of an option and K is its exercise price. Moreover, the equation (1.4)–(1.5) turns out to have an explicit solution of the form

$$V(t, S(t)) = S(t)\Phi\big(d_+(t, S(t))\big) - Ke^{-r(T-t)}\Phi\big(d_-(t, S(t))\big), \tag{1.6}$$

where

$$d_+(t, S(t)) := \frac{\log\frac{S(t)}{K} + (T-t)\left(r + \frac{\sigma^2}{2}\right)}{\sigma\sqrt{T-t}},$$

$$d_-(t, S(t)) := \frac{\log\frac{S(t)}{K} + (T-t)\left(r - \frac{\sigma^2}{2}\right)}{\sigma\sqrt{T-t}}$$

and $\Phi(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}dy$.

The ideas of Black, Scholes and Merton revolutionized mathematical finance and enjoyed empirical success: Stephen Ross, for instance, claimed in 1987 [97] that

> "*When judged by its ability to explain the empirical data, option pricing theory is the most successful theory not only in finance, but in all of economics.*"

However, this was only the beginning of a journey full of challenges – and it is ironic that the storm broke in the very year that these words of Ross were published: on October 19, 1987, the infamous "Black Monday" market crash happened.

1.2 Black Monday: before and after

The classical Black–Scholes–Merton model relies on a number of rather abstract assumptions that are not met on the real life market: specific dynamics for prices, no transaction costs, ability to buy and sell any amount of assets etc. However, being unable to reflect the reality *perfectly* is not always a big deal; after all, "*all models are wrong, but some are useful*"[2]. Black, Scholes and Merton themselves were very well aware of this: for example, [53] quotes Fisher Black on this subject:

> "*Yet that weakness is also its greatest strength. People like the model because they can easily understand its assumptions. The model is often good as a first approximation, and if you can see the holes in the assumptions you can use the model in more sophisticated ways.*"

What really mattered was a successful empirical performance of the vanilla Black–Scholes–Merton model; as it was noted by J. Wiggins, one of pioneers of continuous-time stochastic volatility modeling, "*given the elegance and tractability of the Black-Scholes formula, profitable application of alternate models requires that economically significant valuation improvements can be obtained empirically*" [111].

However, after the mentioned 1987 crash, it became crystal clear that something was very wrong with the log-normal paradigm – and something had to be done. Jackwerth & Rubinstein [75] described the problem as follows:

> "*Following the standard paradigm, assume that stock market returns are lognormally distributed with an annualized volatility of 20% (near their historical realization). On October 19, 1987, the two month S&P 500 futures price fell 29%. Under the lognormal hypothesis, this is a -27 standard deviation event with probability* 10^{-160}*. Even if one were to have lived through the entire 20 billion year life of the universe and experienced this 20 billion times (20 billion big bangs),*

[2] The aphorism is generally attributed to the statistician George Box.

that such a decline could have happened even once in this period is a virtual impossibility."

Clearly, "*virtually impossible*" price falls which left thousands of investors destitute were already a good argument to reconsider financial modeling approaches. But, except for that shock (which in principle could be branded as a single anomaly), Black Monday brought something even more annoying from the theoretical perspective: *the volatility smile.*

The volatility σ is the only parameter in the Black-Scholes formula (1.6) that is not observable. Maturity date T and exercise price K are given in specifications of the given option contract, the price $S(t)$ can be taken directly from the market – and σ has to be somehow "guessed" from the market data. In 1986, Latané & Rendelman [79] proposed an elegant method to do that. Fix some t together with the corresponding price $S(t)$ and consider the Black-Scholes option price (1.6) as a function $V_t = V_t(\tau, \kappa, \sigma)$ of the *log-moneyness* $\kappa := \log \frac{K}{S(t)}$, *time to maturity* $\tau := T - t$ and the volatility σ. Next, take the *actual* market price $\widetilde{V}_t$ of the corresponding option and notice that, since V_t is supposed to coincide with $\widetilde{V}_t$, the volatility σ can be found from the equation

$$V_t(\tau, \kappa, \sigma) - \widetilde{V}_t = 0. \tag{1.7}$$

The solution $\widehat{\sigma} = \widehat{\sigma}_t(\tau, \kappa)$ to this equation is called the *implied* volatility and, if the stock price model (1.2)–(1.3) indeed corresponds to reality well enough, $\widehat{\sigma}_t(\tau, \kappa)$ should be approximately constant for options with the same underlying asset but differing maturities T and strikes K (and hence τ and κ).

Unfortunately, this is not the case. For instance, $\widehat{\sigma}_t(\tau, \kappa)$ turns out to change with τ for fixed κ. There seems to be an easy fix of (1.2) to account for this type of variation and, in fact, Merton actually considered such a modification in his original paper [86]. Namely, if the volatility $\sigma = \sigma(t)$ is a deterministic function of time, one can obtain a version of (1.6) of the form

$$\begin{aligned} V(t, S(t)) &= S(t)\Phi\left(\frac{\log \frac{S(t)}{K} + r(T-t) + \frac{1}{2}\int_t^T \sigma^2(s)ds}{\sqrt{\int_t^T \sigma^2(s)ds}}\right) \\ &\quad - Ke^{-r(T-t)}\Phi\left(\frac{\log \frac{S(t)}{K} + r(T-t) - \frac{1}{2}\int_t^T \sigma^2(s)ds}{\sqrt{\int_t^T \sigma^2(s)ds}}\right) \\ &=: S(t)\Phi\left(\frac{-\kappa + \left(r + \frac{1}{2}\overline{\sigma}^2(t,\tau)\right)\tau}{\overline{\sigma}(t,\tau)\sqrt{\tau}}\right) \\ &\quad - Ke^{-r\tau}\Phi\left(\frac{-\kappa + \left(r - \frac{1}{2}\overline{\sigma}^2(t,\tau)\right)\tau}{\overline{\sigma}(t,\tau)\sqrt{\tau}}\right), \end{aligned}$$

where $\overline{\sigma}^2(t,\tau) := \frac{1}{\tau}\int_t^{t+\tau} \sigma^2(s)ds$. Then the counterpart of the equation (1.7) gets the form

$$V_t(\tau, \kappa, \overline{\sigma}(t,\tau)) - \widetilde{V}_t = 0$$

and its solution $\widehat{\sigma}_t(\tau, \kappa)$ is allowed to vary in τ for fixed κ. One may even argue that it is reasonable to assume that σ changes with time: as noted in [40, p. 144], "*there is nothing inconsistent about expecting high volatility this year and low volatility next year*".

As for the variation in κ for fixed τ, luckily, the implied volatility remained relatively flat (at least, the variation was subtle enough to be ignored) – exactly until the above-mentioned Black Monday crash in 1987. Since that time, investors started observing notable variability of the implied volatility in κ with very clear convex patterns (see Fig. 1.1) which were eventually called "*volatility smiles*"[3].

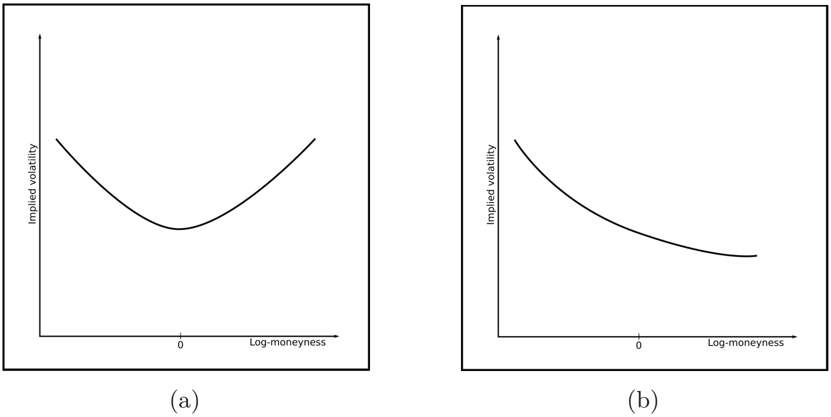


Figure 1.1: Idealised volatility smiles: (a) represents the general form of volatility smiles for foreign currency options; (b) depicts a typical implied volatility smile for equity options (see also [74]).

Such a behaviour was consistent, had a direct negative impact on empirical performance of the Black-Scholes formula and could not be explained by the price dynamics (1.2)–(1.3) – a very annoying combination for a theoretical framework. Moreover, one should not forget about the variability of implied volatility in τ: ideally, one would like a model that mimics the interplay between τ and κ, i.e. represents the behaviour of the *entire volatility surface* $(\tau, \kappa) \mapsto \widehat{\sigma}_t(\tau, \kappa)$. This creates many additional effects to reproduce:

- as noted in [26], the smile amplitude decreases very slowly as τ increases (see e.g. Fig. 1.2);
- the observed *at-the-money volatility skew* defined as

$$\Psi(\tau) := \left| \frac{\partial}{\partial \kappa} \widehat{\sigma}_t(\kappa, \tau) \right|_{\kappa=0} \tag{1.8}$$

  is known to behave as $O(\tau^{-\beta})$ when $\tau \to 0$ (see e.g. Fig. 1.3).

[3] Clearly, the shape on Fig 1.1(b) is more of a "*smirk*" rather than a "*smile*", but, as noted in [40, p. 4], "*practitioners have persisted in using the word smile to describe the relationship between implied volatilities and strikes, irrespective of the actual shape*".

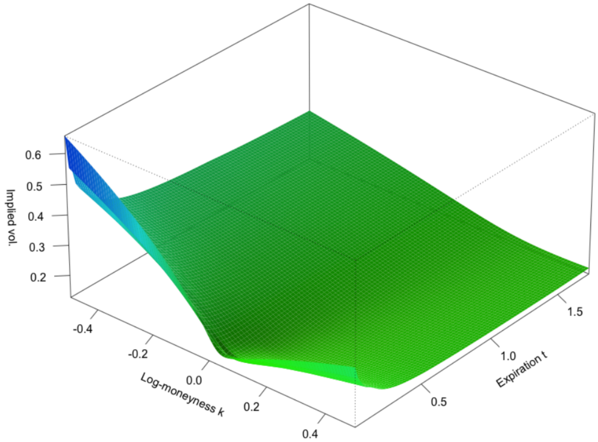


Figure 1.2: Shape of the S&P volatility surface as of June 20, 2013 volatility surface; the plot is taken from [64, Figure 1.1]. Note the gradual leveling of volatility smile as expiration time increases.

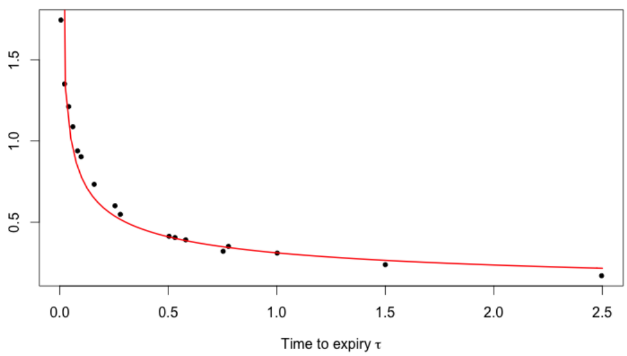


Figure 1.3: At-the-money volatility skew taken from [64, Figure 1.2]. Black dots represent estimates of the S&P volatility skews; the red curve is the power-law fit $\Psi(\tau) = C\tau^{-0.4}$.

As we will see later, it is not easy to conceive a model that reproduces jointly these two effects and the search for such a model is a a difficult puzzle for both theorists and practitioners.

1.3 Fat tails, leverage, clustering and long memory

The smile effect makes a spectacular point against the GBM dynamics (1.2)–(1.3), but it is definitely not the only argument in place. In fact, objections to log-normality of prices appeared long before 1987 – perhaps as early as the log-normal model itself. For instance, the empirical studies of Mandelbrot [82] (1963) and Fama [52] (1965) pointed out that tails of price distribution are much fatter than the ones expected from a log-normal random variable (in order to account for this, Mandelbrot suggested modeling price log-returns with α-stable distributions).

Another interesting phenomenon not grasped by the GBM is the so-called *leverage effect*: negative correlation between variance and returns of an asset. This empirical artefact, initially noticed by Black [13] and then studied in more detail by Christie [22], Cheung & Ng [21] and Duffee [48], was explained by Black himself as follows: a decrease in a stock price results in a drop of firm's equity value and hence increases its financial leverage (i.e. the company's debt rises relative to its equity). This, according to Black, makes the stock riskier and hence more volatile. Interestingly, the name "leverage effect" stuck due to this explanation although subsequent research [1, 54, 69] pointed out that this correlation may not be connected to the leverage at all. Zumbach in [115, Section 3.9.1] gives a different – and, in our opinion, quite plausible – interpretation of this negative correlation: downward moves of stock prices are usually considered as "*bad*", trigger many sales and hence increase the volatility. Upward moves, in turn, do not result in such drastic changes in investors' portfolios since most of the market participants have long positions in the first place. Zumbach [116] also studied another effect of the same nature known now as the *Zumbach effect*: pronounced *trends* in stock price movement, irrespective of sign, increase the subsequent volatility since large price moves motivate investors to modify their portfolios (unlike conditions when prices fluctuate in a smaller range).

The next empirical contradiction to the Black-Scholes-Merton framework is that any basic financial time series analysis reveals clusters of high and low volatility episodes; as noted by Mandelbrot [82], "*large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes*" (see e.g. Fig. 1.4). This clustering effect is often quantified by analyzing the autocorrelation function of absolute log-returns, i.e.

$$\operatorname{corr}(|R(t)|, |R(t+\tau)|), \tag{1.9}$$

where the log-return $R(t) := \log\left(\frac{S(t+\Delta)}{S(t)}\right)$ is defined for some given time scale Δ (which may vary between a fraction of a second for tick data to several days). According to several empirical studies such as of Ding et. al. [47], Ding & Granger [46], Bollerslev & Mikkelsen [16], Breidt et. al. [18] or Cont [28, 29, 31],

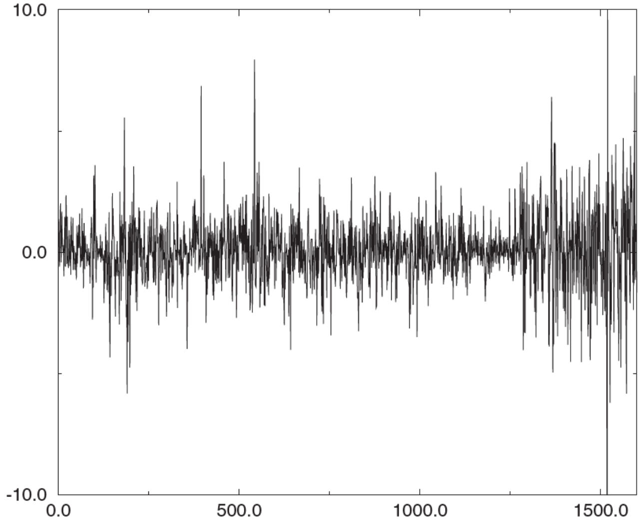


Figure 1.4: BMW daily log-returns on the Frankfurt Stock Exchange, 1992–1998; large changes in price tend to cluster together. The plot is taken from [29, Section 2.1].

the autocorrelation function (1.9) is positive and, moreover, shows signs of a slow decay of the type $O(\tau^{-\beta})$, $\tau \to \infty$, with an exponent $\beta \leq 0.5$.

The latter property should be discussed separately as it has some far-reaching consequences. Such an autocorrelation decay is known as *long range dependence* (see e.g. [11]) and, if proved to be statistically relevant, indicates *presence of memory* on the market. It should be emphasized that the long range dependence is not an easy feature to check: clearly, it manifests itself when $\tau \to \infty$ and hence one may argue that any empirically observed behaviour of autocorrelations is due to non-stationarity of financial time series over longer time periods (see e.g. discussion in [87, Section 1.4]). That being said, it must also be mentioned that there are other arguments in favour of long memory. For instance, Willinger et. al. [112] apply the so-called *rescaled range (R/S) analysis* technique to the CRSP (Center for Research in Security Prices) daily stock time series and find some weak[4] evidence of memory in the data. Lobato & Velasco [81] analyze volatility in connection to trading volumes and find that both of these financial characteristics exhibit the same degree of long memory. Another interesting point comes from the analysis of the implied volatility surface: Comte & Renault [26] noticed that the decrease of the smile amplitude as time to maturity increased was much slower than many advanced market models predicted. They argued that such an effect could be mimicked by having long memory in volatility and that claim was confirmed by e.g. a simulation study [62].

[4] As the authors write, "*...we find empirical evidence of long-range dependence in stock price returns, but because the corresponding degree of long-range dependence [...] is typically very low [...] the evidence is not absolutely conclusive*".

Of course, we cannot list all stylized facts about market behaviour contradicting the GBM dynamics due to the vast amount of material – in this regard, we refer our readers to [65, Section 2.2], [56, Section 3], well-known survey articles [27, 28] or book [115]. Two things are clear though: first, the log-normal model is way too simple and does not reflect a lot of important qualitative features of the market and, second, the behaviour of financial time series is incredibly complex and requires fairly ingenious modeling approaches. At the same time, one would want to keep the core idea of Black-Scholes-Merton – no-arbitrage pricing – since it is very intuitive and reasonable from the economic point of view. Luckily, developments in the option pricing theory subsequent to the seminal Black-Scholes and Merton papers allowed for some decent flexibility in terms of the choice of price models. Here, we refer to gradual translation of the Black-Scholes-Merton approach into the language of martingale theory which evolved in the celebrated *Fundamental Theorem of Asset Pricing* – the result which connects non-arbitrage pricing and existence of equivalent local martingale measures. In this regard, we mention the early research of Ross [96], Harrison & Kreps [67], Harrison & Pliska [68], Kreps [77] as well as subsequent seminal works of Delbaen & Schachermayer [36–38] (see also [101] for a detailed historical overview on the subject). This line of research eventually evolved into a general theory allowing for quite a broad variety of price models to choose from – hence giving researchers all the necessary tools to adjust the classical model (1.2) to account for volatility smiles and all other empirical inconsistencies. Of course, due the vast amount of different approaches developed by this time, we will not be able to cover all modeling viewpoints – for this, we refer to the overviews given in the books by Shiryaev [104] and Mariani & Florescu [84]. Here, instead, we concentrate on one particular class of models that is directly related to the framework of the present thesis: *stochastic volatility models*.

1.4 Classical stochastic volatility models

As noted in the previous sections, the “trickiest” parameter of Black-Scholes formula is the volatility σ. Empirical observations show that it varies with time, is correlated with the current price level, has clusters of low and high values and seems to have a long memory. Another important phenomenon is the so-called *excess volatility* [28, 34]: the variability in asset prices cannot be fully explained only by changes in “fundamental” economic factors. All these stylized facts together lead to an idea to modify (1.3) as

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW(t)$$

with the volatility $\{\sigma(t),\ t \ge 0\}$ being a random process that is only imperfectly correlated with the Brownian motion W. This approach can be traced back to discrete-time model of Clark (1973) [24] where asset prices were considered as subordinated stochastic processes with the time change being used to represent trading volumes and information arrival. Early contributors to continuous time stochastic volatility modeling include[5]:

- Hull & White [73] who assume that the squared volatility $\sigma^2 = \{\sigma^2(t),\ t \geq 0\}$ is itself a geometric Brownian motion, i.e. price and volatility satisfy stochastic differential equations of the form

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW(t),$$
$$d\sigma^2(t) = \theta_1\sigma^2(t)dt + \theta_2\sigma^2(t)dB(t)$$

  respectively, where B and W are two Brownian motions that are allowed to be correlated to account for the leverage effect;

- Wiggins [111] who suggests a slightly more general dynamics of the form

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW(t),$$
$$d\sigma(t) = f(\sigma(t))dt + \theta\sigma(t)dB(t);$$

- Scott [102] and Stein & Stein [108] who consider the volatility to be an Ornstein-Uhlenbeck process, i.e.

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW(t),$$
$$d\sigma(t) = \theta_1(\theta_2 - \sigma(t))dt + \theta_3 dB(t);$$

- Heston [70] who introduces the SDE of the form

$$dS(t) = \mu S(t)dt + \sqrt{\sigma(t)}S(t)dW(t),$$
$$d\sigma(t) = \theta_1(\theta_2 - \sigma(t))dt + \theta_3\sqrt{\sigma(t)}dB(t),$$

  i.e. the volatility follows the so-called *Cox-Ingersoll-Ross* or *square root* process (see also [33]) which enjoys strict positivity provided that $2\theta_1\theta_2 \geq \theta_3^2$.

Stochastic volatility models turned out to have an additional important advantage: they have an ability to reproduce, to some extent, "smiley" patterns of the implied volatility (see e.g. [93] or [55, Section 2.8.2]). However, one must acknowledge that the models listed above leave a lot of room for improvement when it comes to accuracy of grasping volatility surfaces (see e.g. [63] for a detailed overview of empirical performance of the classical stochastic volatility models). Therefore, there is no surprise that a lot of effort was made to advance the stochastic volatility framework further to account for all such inconsistencies.

At this point, let us make an important disclaimer. In this thesis, we concentrate on *continuous* stochastic volatility and present contributions in this direction. Hence we do not discuss models with jumps in this introduction and, given the popularity of the latter, we acknowledge that such a decision can be seen as a notable shortcoming. However, we still choose to omit this subject

[5] For a full coverage of stochastic volatility modeling evolution, see [103].

since it is way to broad to be covered in sufficient detail within this chapter. The reader who is interested in models with jumps is referred to the specialized books by Barndorff-Nielsen & Shepard [9], Rachev et. al. [92] or Tankov & Cont [109]. With this remark in mind, let us proceed to the next stop of our journey: *stochastic volatility models based on fractional Brownian motion.*

1.5 Overview of fractional models

1.5.1 Incorporation of long memory

The most prominent issues of classical stochastic volatility models driven by standard Brownian motion are the following:

- they cannot give a power-law behaviour of at-the-money volatility skew (1.8) (see e.g. [80, Remark 11.3.21]);
- as time to maturity increases, the smile generated by the models levels out way too fast [26].

Comte & Renault in [26] note that one of possible explanations of the second issue lies in presence of the long memory in the real life prices. To account for that, they suggest to model the volatility as follows:

$$\sigma(t) = \theta_1 \exp\left\{\theta_2 \int_0^t e^{-\theta_3(t-s)} dB^H(s)\right\},$$

where $B^H = \{B^H(t),\ t \geq 0\}$ is either a *fractional Brownian motion*[6] defined for $H \in (0,1)$ by

$$\begin{aligned} B^H(t) := &\frac{1}{\Gamma\left(H+\frac{1}{2}\right)} \int_{-\infty}^0 \left((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}\right) dB(s) \\ &+ \frac{1}{\Gamma\left(H+\frac{1}{2}\right)} \int_0^t (t-s)^{H-\frac{1}{2}} dB(s) \end{aligned} \tag{1.10}$$

or its truncated version (also known as *Riemann-Liouville fractional Brownian motion*)

$$B^H(t) := \frac{1}{\Gamma\left(H+\frac{1}{2}\right)} \int_0^t (t-s)^{H-\frac{1}{2}} dB(s),$$

where, in both cases, B is a standard Wiener process. And one must admit that the idea to use fractional Brownian motion as a driver for the volatility turned out to be extremely felicitous: the process (1.10) indeed provides the long memory (given that the so-called *Hurst index* $H \in \left(\frac{1}{2}, 1\right)$), successfully mimics the behaviour of the volatility smile amplitude mentioned above (see

[6] This process was initially considered by Kolmogorov [76] and later reintroduced by Mandelbrot and van Ness [83].

e.g. [62]) and additionally enjoys a number of other very pleasant properties. In particular, it is Gaussian, has stationary increments and is *self-similar*, i.e.

$$B^H(at) \overset{\text{Law}}{=} a^H B^H(t), \quad \forall a \geq 0. \tag{1.11}$$

The latter feature turns out to be a very convenient for testing the long memory: as noted in Section 1.3, long-range dependence is very hard to detect with statistical significance (as it requires analysis on larger time scales) whereas assuming (1.11) allows to deduce long time behavior from short time behavior; see [28] for a detailed discussion on this matter. On the top of that, it has Hölder continuous trajectories of order up to H which enables usage of pathwise calculus for stochastic integration (see e.g. [88]). This set of very pleasing mathematical properties as well as consistency with the long memory setting made fractional Brownian motion with $H > \frac{1}{2}$ a very popular choice for driving stochastic volatility; in this regard we mention e.g. the models of Rosenbaum [95], Chronopoulou and Viens [23] and Comte et. al. [25].

1.5.2 Rough revolution

However, when it comes to the power law of the implied volatility skew, the situation gets far more complicated: as noted in e.g. [5, Section 7.2.1], fractional Brownian motion with $H > \frac{1}{2}$ is not able to reproduce it. Interestingly, [5, Section 7.2.2] finds out that the required behavior of the skew can be obtained if one sacrifices long memory and takes fractional Brownian motion with $H \in \left(0, \frac{1}{2}\right)$ instead. In 2014, the latter idea obtained an additional foundation: Gatheral, Jaisson and Rosenbaum published a paper [64] with a catchy title "*Volatility is rough*". Using estimation techniques based on power variations, they came to a conclusion that the "real life" volatility has a much lower Hölder regularity than the one of fractional Brownian motion with $H > \frac{1}{2}$. Furthermore, they claim that fractional Brownian motion with $H \approx 0.1$ fits the time series much better. This quickly evolved into an enormous research area with hundreds of papers published over the years – here we refer the reader to the rough volatility literature list [98] that is regularly updated by specialists in the field. Some notable models are (in all cases, $H \in \left(0, \frac{1}{2}\right)$):

- rough Bergomi model [10] models the price as

$$\begin{aligned} dS(t) &= \sqrt{\sigma(t)} S(t) dW(t), \\ \sigma(t) &= \xi_t^t, \\ d\xi_t^u &= \theta \xi_t^u (u-t)^{H-1/2} dB(t), \quad u \geq t; \end{aligned}$$

- rough SABR model [59] where the price is given by

$$\begin{aligned} dS(t) &= \sqrt{\sigma(t)} f(S(t)) dW(t), \\ \sigma(t) &= \xi_t^t, \\ d\xi_t^u &= \theta \xi_t^u (u-t)^{H-1/2} dB(t), \quad u \geq t, \end{aligned}$$

where f is a positive continuous function;

- rough Stein-Stein model [66] with the price following the dynamics

$$dS(t) = \theta\sigma(t)S(t)dW(t),$$
$$\sigma(t) = \frac{1}{\Gamma\left(H+\frac{1}{2}\right)}\int_0^\infty (t-s)^{H-\frac{1}{2}}dB(s);$$

- rough Heston model [50]:

$$dS(t) = \sqrt{\sigma(t)}S(t)dW(t),$$
$$\sigma(t) = \sigma(0) + \int_0^t \frac{(t-s)^{H-\frac{1}{2}}}{\Gamma\left(H+\frac{1}{2}\right)}\left(\theta_1(\theta_2-\sigma(s))ds + \theta_3\sqrt{\sigma(s)}dB(s)\right).$$

Models of this type are very popular nowadays and are quite close to becoming "the new classics". That being said, we must admit that rough volatility models also have some skeletons in the cupboard. Indeed, to the best of our knowledge, there are three main points in favour of $H \in \left(0, \frac{1}{2}\right)$ in the literature:

1) direct statistical studies of high frequency volatility time series (such as the seminal paper [64] or [60]) indicate roughness of the latter;

2) rough volatility models have potency to reproduce some financial market phenomenons such as the at-the-money skew mentioned above [5] (see also [57, 58] or [114, Theorem 1]) or some form of the Zumbach effect [35, 51];

3) rough Heston model turns out to be the limit of a reasonable tick-by-tick price model based on two-dimensional Hawkes processes [49], i.e. it is possible to deduce roughness of the volatility directly from the market microstructure.

And many of these points got a fair amount of criticism. For instance, Rogers [94] and Cont and Das [30] argue that estimation techniques used in [64] are unstable and may produce very low Hurst indices even on simulated datasets driven by standard Wiener noises (although [60] seems to account for that). On the other hand, Funahashi and Kijima [61] reasonably note that, despite grasping the term structure of the at-the-money volatility skew, rough volatility models based on fractional Brownian motion with $H < 1/2$ do not give the required rate of decrease in the smile amplitude as expiration increases – and, additionally, are not consistent in general with the long memory justified by older empirical studies. In other words, we run into some sort of a "fractional modeling puzzle" (the term used by [61]): on the one hand, long memory and behaviour of the volatility smile amplitude demand $H \in \left(\frac{1}{2}, 1\right)$; on the other hand, roughness and the implied volatility skew require Hurst indices very close to 0.

1.5.3 Solving the fractional modeling puzzle

Interestingly, the way to solve the fractional puzzle lies in the very reason of that puzzle: *fractional Brownian motion itself.* After all, the contradiction between long memory and roughness appears *because of* fractional Brownian motion – in general, these two features do not depend on each other and can very well co-exist. At the same time, it would be truly wasteful to completely abandon fractional models, since they indeed manage to explain numerous empirical phenomena and are relatively simple to be used in practice. Loosely speaking, one would want a driver that, on the one hand, somehow incorporates (in some format) both long memory and roughness and, on the other hand, preserves some "nice" properties of fractional Brownian motion.

The most straightforward approach to that is to use two fractional Brownian motions with different Hurst indices instead of one. Such an attempt to "include the best of both worlds" was utilized in e.g. [61] (see also [4, Section 7.7]), where the following model was considered:

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sigma(X^1(t), X^2(t))dW(t), \\ dX^i(t) &= (\theta_1^i - \theta_2^i X^i(t))dt + \theta_3^i dB^{H_i}(t), \quad i = 1, 2, \end{aligned} \tag{1.12}$$

with B^{H_1}, B^{H_2} being two fractional Brownian motions with $H_1 > 1/2$ and $H_2 < 1/2$. The performance of this model seems to be very promising since it indeed manages to fit the implied volatility surface with the required at-the-money skew and slower decay of the smile amplitude.

Another interesting possibility – *a multifractional Brownian motion* – is advocated in Corlay et. al. [32]. There, the authors come to a conclusion that the local roughness of the volatility is heavily variable and has periods of low (≈ 0.1) and high (≈ 0.8) regularity (see Fig. 1.5). It is also important to note

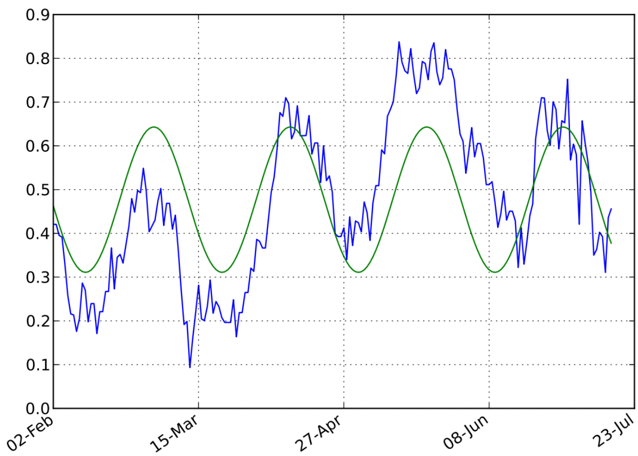


Figure 1.5: Estimated regularity of the volatility of the S&P 500 minute data (blue) and its regression (green). Source: [32, Figure 2].

that [64, Section 2.6] also reports some dependence of the volatility roughness on time, so usage of multifractional Brownian motion may indeed be a good idea.

As a convenient generalization of all approaches to the fractional puzzle mentioned above, one can also consider usage of *Gaussian Volterra processes* $Z(t) := \int_0^t \mathcal{K}(t,s)dB(s)$. The drivers on this level of generality were considered in e.g. [85] or [19] and we anticipate that processes of this type will be actively used throughout this thesis.

1.6 Technical challenges of stochastic volatility modeling

Up until now, we discussed problems related to compliance of stochastic volatility models with empirical observations and stylized facts. However, they also pose a number of challenges that have purely technical nature. For example, a common issue for stochastic volatility framework is the possibility of moment explosions in price (see e.g. [6]) – that is, $\mathbb{E}[S^r(t)]$ may be infinite for all time points t after some t_*. Moment explosions can be a notable drawback from the asset pricing perspective since, as it is noted in [6, Section 8], "*several actively traded fixed-income derivatives require at least L^2 solutions to avoid infinite model prices*".

An additional desirable property for stochastic volatility is positivity of its paths. The reason for that lies, in particular, in the procedure of measure change: densities of martingale measures in stochastic volatility models normally contain expressions of the form $\int_0^T \frac{1}{\sigma(s)}dW(s)$ and $\int_0^T \frac{1}{\sigma^2(s)}ds$ (see e.g. [12, Proposition 1.11]). Hence, if the volatility can hit zero, one may end up without a transparent description of the family of martingale measures. In practice, positivity is often achieved by modeling log-volatility and then taking exponential – but in this case one risks to get moment explosions in the price discussed above. In other cases (such as the rough Heston model), the price is modelled under the risk-neutral measure from the start which, in some sense, puts the measure change mechanism out of consideration. It is not necessarily a problem if one uses the model for option pricing, but it is still better to have a clear and coherent procedure of transition to the martingale probability if one aims to justify usage of the approach with statistical analysis of econometric time series (which is, of course, performed under the physical measure).

An additional challenge of stochastic volatility modeling comes from the numerical perspective. In many cases, it is not possible to get closed formulas neither for calibration nor for pricing, hedging and portfolio optimization and hence one must rely on numerical methods. As a result, there is always an additional layer of work related to the development of algorithms with reasonable convergence rates, computational efficiency etc. Presence of effective numerical approaches turns out to be a viable advantage for practitioners; for example, as noted in [63, p. 24], the standard Heston model was widely used despite clear empirical inconsistencies – exactly because availability of algorithms for virtually all possible applications.

1.7 Overview of the thesis

With all of the considerations described above, we are now ready to proceed to the material of this thesis where we make yet another attempt to come up with a "good" stochastic volatility model. By "*good*", we mean that it addresses the issues listed in previous sections – namely, we want the model to have the following properties:

- *flexibility in the noise*: the model should be able to accept various drivers – from fractional Brownian motions with different Hurst indices to general Hölder continuous processes – to account for different option pricing phenomenons;
- *control over the moments of the price*: we should be able to ensure existence of moments of necessary orders for the corresponding price process;
- *positivity*: the volatility process should be strictly positive and preferably have inverse moments to ensure reasonable behaviour of martingale densities;
- *existence of efficient numerics*: the model should be friendly enough to be approached by numerical algorithms for a wide range of applications such as pricing or hedging.

In the present thesis, we suggest a volatility model that comprises all the properties mentioned above. It is described by the SDE of the form

$$Y(t) = Y(0) + \int_0^t b(s, Y(s))ds + Z(t) \tag{1.13}$$

where the noise Z can potentially be *any* λ-Hölder continuous stochastic process and the drift $b = b(t, y)$ is a real function that has either

(A) an explosive growth of the type $(y - \varphi(t))^{-\gamma}$ as $y \downarrow \varphi(t)$

or

(B) an explosive growth of the type $(y - \varphi(t))^{-\gamma}$ as $y \downarrow \varphi(t)$ ***and*** an explosive decrease of the type $-(\psi(t) - y)^{-\gamma}$ as $y \uparrow \psi(t)$,

where φ and ψ are given Hölder continuous functions such that $\varphi(t) < \psi(t)$, $t \in [0, T]$. Note that we allow for an arbitrary order of Hölder regularity λ provided that

$$\gamma > \frac{1}{\lambda} - 1, \tag{1.14}$$

i.e. our driver can easily be e.g. a fractional Brownian motion (with any Hurst index H) or a multifractional Brownian motion (with no restrictions on the range of functional Hurst parameter). We prove that the SDE (1.13) has a unique strong solution whereas the chosen structure of the drift as well as (1.14) ensure that either

$$Y(t) > \varphi(t) \quad a.s., \quad t \in [0, T], \tag{1.15}$$

in case **(A)** or

$$\varphi(t) < Y(t) < \psi(t) \quad a.s., \quad t \in [0, T], \tag{1.16}$$

in case **(B)**. Having the properties (1.15)–(1.16) in mind, we call the solution to (1.13) a ***sandwiched process*** and note that the intuition behind (1.15)–(1.16) is quite simple. For example, (1.15) can be explained as follows: whenever the solution Y approaches φ, the value of $b(s, Y(s))$ becomes very big "*repelling*" Y from its lower bound while the condition (1.14) guarantees that the variation of Z is not too drastic to break this effect.

It is important to emphasize that dynamics of the type (1.13) covers some known models. As an example, one can mention

$$Y^H(t) = Y(0) + \frac{1}{2}\int_0^t \frac{k}{Y^H(s)} ds - \frac{b}{2}\int_0^t Y^H(s) ds + \frac{\sigma}{2} B^H(t) \tag{1.17}$$

considered in [89] for $H > 1/2$ (see also [72]) which can be regarded as a square root of the *fractional Cox-Ingersoll-Ross process* since $X := Y^2$ satisfies the SDE

$$X(t) = X(0) + \int_0^t (k - aX(s)) ds + \sigma \int_0^t \sqrt{X(s)} dB^H(s).$$

The opportunity we have in (1.13) to choose any Hölder continuous noise allows for both "*long memory*" and "*roughness*" and hence gives a potential to reproduce various option pricing phenomenons discussed in Sections 1.2–1.5. In turn, the lower bound in (1.15)–(1.16) allows to make the sandwiched process Y strictly positive and the upper bound in (1.16) guarantees existence of all moments for the price. Finally, we will also see that there are effective simulation schemes for (1.13) that *preserve* properties (1.15)–(1.16) and can be used for a wide range of applications such as pricing or hedging.

1.7.1 Summary of our contributions

The present thesis is based on five papers written during the PhD programme that are organized in the form of chapters, one chapter for each paper.

Paper I [90] serves as a forerunner for the main model (1.13). There, we start from a standard Brownian Cox-Ingersoll-Ross process

$$X(t) = X(0) + \int_0^t (a - bX(s))\, ds + \sigma \int_0^t \sqrt{X(s)} dW(s)$$

with $a > \frac{\sigma^2}{4}$ (i.e. we allow for $a < \frac{\sigma^2}{2}$ and hence do not exclude zero-hitting cases) and prove that its square root $Y(t) := \sqrt{X(t)}$ satisfies the SDE of the form

$$Y(t) = Y(0) + \frac{1}{2}\int_0^t \frac{k}{Y(s)}ds - \frac{b}{2}\int_0^t Y(s)ds + \frac{\sigma}{2}W(t) \tag{1.18}$$

with $k := a - \frac{\sigma^2}{4}$ (c.f. (1.17)). We investigate the SDE (1.18) from purely theoretical perspective and establish a new link to Skorokhod reflections (see e.g. [105, 106]): when $k \downarrow 0$, the square root process (1.18) converges to the *reflected Ornstein-Uhlenbeck process*

$$Y_0(t) = Y(0) - \frac{b}{2}\int_0^t Y_0(s)ds + \frac{\sigma}{2}W(t) + L_0(t), \tag{1.19}$$

where $Y_0(t) \geq 0$ and L_0 is a continuous non-decreasing process, the points of growth of which occur only when Y_0 hits zero[7]. In the second part of Paper I, we prove a similar result for the square root (1.17) of the fractional Cox-Ingersoll-Ross process.

Paper II [44] studies the SDE (1.13) in full generality: we prove existence and uniqueness of solution for general Hölder continuous noises as well as properties (1.15) and (1.16). Additionally, we study the moments of sandwiched processes (1.13) and verify that for all $r > 0$

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y(t)|^r\right] < \infty, \quad \mathbb{E}\left[\sup_{t\in[0,T]} |Y(t) - \varphi(t)|^{-r}\right] < \infty \tag{1.20}$$

in case **(A)** and

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y(t) - \varphi(t)|^{-r}\right] < \infty, \quad \mathbb{E}\left[\sup_{t\in[0,T]} |\psi(t) - Y(t)|^{-r}\right] < \infty \tag{1.21}$$

in case **(B)** provided that the Hölder constant

$$\Lambda := \sup_{0\leq s<t\leq T} \frac{|Z(t) - Z(s)|}{|t-s|^\lambda} \tag{1.22}$$

of Z satisfies some mild moment assumptions. We regard (1.20)–(1.21) as one of the most important technical results of the thesis: it allows to control the behaviour of sandwiched processes near the bounds and it is a crucial tool for the analysis of e.g. numerical schemes. We also discuss the connection of (1.13) to Skorokhod reflections and generalize the results of Paper I. Namely, we prove that reflected processes appear as limits of sandwiched processes and obtain new representations of the corresponding regulators. Finally, we suggest two stochastic volatility models which we regard as generalizations of the Cox-Ingersoll-Ross and Chan–Karolyi–Longstaff–Sanders (see e.g. [20]) processes. The results are illustrated with simulations.

[7]In the literature, such process L_0 is called a *reflection function* or *regulator* of the Skorokhod reflection problem (1.19).

Paper III [42] is devoted to a simulation method for (1.13). We suggest an algorithm in the spirit of the *drift-implicit* Euler scheme constructed for the classical Cox-Ingersoll-Ross process in [2, 3, 39] and extended to the case of the fractional Brownian motion with $H > \frac{1}{2}$ in [71, 78, 113]: according to this approach, the approximation $\widehat{Y}(t_{k+1})$ of $Y(t_{k+1})$ is obtained from the equation

$$\widehat{Y}(t_{k+1}) = \widehat{Y}(t_k) + b(t_{k+1}, \widehat{Y}(t_{k+1}))(t_{k+1} - t_k) + (Z(t_{k+1}) - Z(t_k)). \quad (1.23)$$

Under some mild assumptions on the drift and Λ from (1.22), we prove $L^r(\Omega; L^\infty([0,T]))$-convergence of this scheme. Additionally, we show that for all points t_k of the partition

$$\widehat{Y}(t_k) > \varphi(t_k)$$

in the setting **(A)** and

$$\varphi(t_k) < \widehat{Y}(t_k) < \psi(t_k)$$

in the case **(B)**. This *sandwich-preserving* property is a major improvement in comparison to the scheme used in Paper II as it allows to reproduce the behaviour of sandwiched processes more accurately. In particular, we prove that $(\widehat{Y} - \varphi)^{-1}$ and $(\psi - \widehat{Y})^{-1}$ converge to $(Y - \varphi)^{-1}$ and $(\psi - Y)^{-1}$ correspondingly which is crucial for e.g. approximating martingale densities in stochastic volatility models.

Paper IV [43] introduces the market model which we call the *Sandwiched Volterra Volatility (SVV)* model. Namely, we consider

$$S_i(t) = S_i(0) + \int_0^t \mu_i(s) S_i(s) ds + \int_0^t Y_i(s) S_i(s) dB_i^S(s), \quad (1.24)$$

$$Y_i(t) = Y_i(0) + \int_0^t b_i(s, Y_i(s)) ds + \int_0^t \mathcal{K}_i(t,s) dB_i^Y(s), \quad (1.25)$$

$i = 1, ..., d$, where S_i and Y_i are price and volatility processes respectively, μ_i are deterministic continuous functions, b_i are drifts in the setting **(B)** with strictly positive lower bounds, B_i^S, B_i^Y are correlated standard Brownian motions and $\mathcal{K}_i$ are arbitrary square integrable kernels such that each process $Z_i(t) := \int_0^t \mathcal{K}_i(t,s) dB_i^Y(s)$ is Hölder continuous up to the order $H_i \in (0,1)$. We

- prove that the prices S_i have moments of all orders,
- obtain an exhaustive description of the set of equivalent local martingale measures for (1.24)–(1.25),
- show Malliavin differentiability of prices and volatilities and
- use Malliavin calculus to develop an efficient algorithm for numerical pricing of options with discontinuous payoffs.

Note that the procedure of measure change is often left out of consideration in rough volatility models (such as the rough Heston model) and hence the transparency of the SVV model in this regard turns out to be its substantial advantage.

Paper V [45] is devoted to Markovian approximations of the SVV model

$$
\begin{aligned}
X(t) &= X(0) + \int_0^t Y(s)X(s)\left(\rho dB_1(s) + \sqrt{1-\rho^2}dB_2(s)\right),\\
Y(t) &= Y(0) + \int_0^t b(s, Y(s))ds + \int_0^t \mathcal{K}(t,s)dB_1(s),
\end{aligned}
$$

where X represents the discounted price. Our core idea is to approximate $\mathcal{K}$ with a degenerate kernel $\mathcal{K}_m(t,s) = \sum_{i=1}^m e_{m,i}(t)f_{m,i}(s)$ in such a way that the noise $Z_m(t) = \sum_{i=1}^m e_{m,i}(t)\int_0^t f_{m,i}(s)dB_1(s)$ has the same Hölder regularity as $Z(t) := \int_0^t \mathcal{K}(t,s)dB_1(s)$ and then take the $(m+2)$-dimensional Markov process $(X_m, Y_m, U_{m,1}, ..., U_{m,m})$ of the form

$$
\begin{aligned}
X_m(t) &= X(0) + \int_0^t Y_m(s)X_m(s)\left(\rho dB_1(s) + \sqrt{1-\rho^2}dB_2(s)\right),\\
Y_m(t) &= Y(0) + \int_0^t b(s, Y_m(s))ds + \sum_{i=1}^m e_{m,i}(t)U_{m,i}(t),\\
U_{m,1}(t) &= \int_0^t f_{m,1}(s)dB_1(s),\\
&\vdots\\
U_{m,m}(t) &= \int_0^t f_{m,m}(s)dB_1(s)
\end{aligned}
$$

as approximation. We prove convergence of (X_m, Y_m) to (X, Y) and provide explicit form of such approximations for rough fractional and general Hölder continuous kernels. In the second part of the Chapter, we use this technique to develop numerical algorithms for the quadratic hedging problem. Namely, we take the explicit representation of the optimal hedge in form of the *non-anticipating derivative* (see e.g. [41]) and replace the original X and Y with their Markovian approximations. The described strategy eliminates dependence of X and Y on the past and allows for numerical approximation of conditional expectations that appear in the non-anticipating derivative. As a result, we suggest two algorithms to compute the optimal hedging strategy: *Nested Monte Carlo* and *Least Squares Monte Carlo*.

References

[1] Aït-Sahalia, Y., Fan, J., and Li, Y. "The leverage effect puzzle: Disentangling sources of bias at high frequency". In: *Journal of financial economics* vol. 109, no. 1 (2013), pp. 224–249.

[2] Alfonsi, A. "On the discretization schemes for the CIR (and Bessel squared) processes". In: *Monte Carlo Methods Appl.* vol. 11, no. 4 (2005), pp. 355–384.

[3] Alfonsi, A. "Strong order one convergence of a drift implicit Euler scheme: application to the CIR process". In: *Statist. Probab. Lett.* vol. 83, no. 2 (2013), pp. 602–607.

[4] Alos, E. and Garcia Lorite, D. *Malliavin calculus in finance: Theory and practice*. London, England: CRC Press, 2021.

[5] Alòs, E., León, J. A., and Vives, J. "On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility". In: *Finance and stochastics* vol. 11, no. 4 (2007), pp. 571–589.

[6] Andersen, L. B. G. and Piterbarg, V. V. "Moment explosions in stochastic volatility models". In: *Finance and Stochastics* vol. 11, no. 1 (Sept. 2006), pp. 29–50.

[7] Bachelier, L. "Théorie de la spéculation". In: *Annales Scientifiques de l'Ecole Normale Superieure. Quatrieme Serie* vol. 17 (1900), pp. 21–86.

[8] Bachelier, L. *Louis Bachelier's theory of speculation: The origins of modern finance*. Ed. by Davis, M. and Etheridge, A. Princeton, NJ: Princeton University Press, 2006.

[9] Barndorff-Nielsen, O. E. and Shephard, N. *Lévy driven volatility models*. 2011.

[10] Bayer, C., Friz, P., and Gatheral, J. "Pricing under rough volatility". In: *Quantitative finance* vol. 16, no. 6 (2016), pp. 887–904.

[11] Beran, J. *Statistics for Long-Memory Processes*. Chapman and Hall/CRC, 1994.

[12] Biagini, F., Guasoni, P., and Pratelli, M. "Mean-Variance Hedging for Stochastic Volatility Models". In: *Mathematical Finance* vol. 10, no. 2 (Apr. 2000), pp. 109–123.

[13] Black, F. "Studies of stock price volatility changes". In: *Proceedings of the Business and Economics Section of the American Statistical Association*. 1976, pp. 177–181.

[14] Black, F. "Citation Classic – The Pricing of Options and Corporate Liabilities". In: *Citation Classic*, no. 33 (1987), p. 16.

[15] Black, F. and Scholes, M. "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* vol. 81, no. 3 (1973), pp. 637–654.

[16] Bollerslev, T. and Ole Mikkelsen, H. “Modeling and pricing long memory in stock market volatility”. In: *Journal of Econometrics* vol. 73, no. 1 (1996), pp. 151–184.

[17] Boness, A. J. “Elements of a Theory of Stock-Option Value”. In: *Journal of Political Economy* vol. 72, no. 2 (1964), pp. 163–175.

[18] Breidt, F. J., Crato, N., and de Lima, P. “The detection and estimation of long memory in stochastic volatility”. In: *Journal of Econometrics* vol. 83, no. 1 (1998), pp. 325–348.

[19] Catalini, G. and Pacchiarotti, B. “Asymptotics for multifactor Volterra type stochastic volatility models”. In: *Stochastic analysis and applications* (2022), pp. 1–31.

[20] Chan, K. C. et al. “An empirical comparison of alternative models of the short-term interest rate”. In: *The journal of finance* vol. 47, no. 3 (1992), pp. 1209–1227.

[21] Cheung, Y.-W. and Ng, L. K. “Stock price dynamics and firm size: An empirical investigation”. In: *The journal of finance* vol. 47, no. 5 (1992), pp. 1985–1997.

[22] Christie, A. “The stochastic behavior of common stock variances: Value, leverage and interest rate effects”. In: *Journal of financial economics* vol. 10, no. 4 (1982), pp. 407–432.

[23] Chronopoulou, A. and Viens, F. G. “Estimation and pricing under long-memory stochastic volatility”. In: *Annals of finance* vol. 8, no. 2–3 (2012), pp. 379–403.

[24] Clark, P. K. “A subordinated stochastic process model with finite variance for speculative prices”. In: *Econometrica: journal of the Econometric Society* vol. 41, no. 1 (1973), p. 135.

[25] Comte, F., Coutin, L., and Renault, E. “Affine fractional stochastic volatility models”. In: *Annals of finance* vol. 8, no. 2–3 (2012), pp. 337–378.

[26] Comte, F. and Renault, E. “Long memory in continuous-time stochastic volatility models”. In: *Mathematical Finance. An International Journal of Mathematics, Statistcs and Financial Economics* vol. 8, no. 4 (1998), pp. 291–323.

[27] Cont, R. “Empirical properties of asset returns: stylized facts and statistical issues”. In: *Quantitative finance* vol. 1, no. 2 (2001), pp. 223–236.

[28] Cont, R. “Long range dependence in financial markets”. In: *Fractals in Engineering*. London: Springer-Verlag, 2005, pp. 159–179.

[29] Cont, R. “Volatility clustering in financial markets: Empirical facts and agent-based models”. In: *Long Memory in Economics*. Springer Berlin Heidelberg, 2006, pp. 289–309.

[30] Cont, R. and Das, P. “Rough volatility: fact or artefact?” 2022.

[31] Cont, R., Potters, M., and Bouchaud, J.-P. “Scaling in stock market data: Stable laws and beyond”. In: *Scale Invariance and Beyond*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, pp. 75–85.

[32] Corlay, S., Lebovits, J., and Lévy Véhel, J. “Multifractional stochastic volatility models”. In: *Mathematical Finance. An International Journal of Mathematics, Statistcs and Financial Economics* vol. 24, no. 2 (2014), pp. 364–402.

[33] Cox, J. C., Ingersoll, J. E., and Ross, S. A. “A theory of the term structure of interest rates”. In: *Econometrica: Journal of the Econometric Society* vol. 53, no. 2 (1985), pp. 385–407.

[34] Cutler, D., Poterba, J., and Summers, L. “What moves stock prices?” In: *Journal of Portfolio Management* (1989), pp. 4–12.

[35] Dandapani, A., Jusselin, P., and Rosenbaum, M. “From quadratic Hawkes processes to super-Heston rough volatility models with Zumbach effect”. In: *Quantitative finance* vol. 21, no. 8 (2021), pp. 1235–1247.

[36] Delbaen, F. and Schachermayer, W. “The fundamental theorem of asset pricing for unbounded stochastic processes”. In: *Mathematische annalen* vol. 312, no. 2 (1998), pp. 215–250.

[37] Delbaen, F. and Schachermayer, W. “A general version of the fundamental theorem of asset pricing”. In: *Mathematische annalen* vol. 300, no. 1 (1994), pp. 463–520.

[38] Delbaen, F. and Schachermayer, W. *Mathematics of arbitrage*. New York, NY: Springer, 2006.

[39] Dereich, S., Neuenkirch, A., and Szpruch, L. “An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process”. In: *Proceedings of the Royal Society A. Mathematical, physical, and engineering sciences* vol. 468, no. 2140 (2012), pp. 1105–1115.

[40] Derman, E., Miller, M. B., and Park, D. *The volatility smile*. Nashville, TN: John Wiley & Sons, 2016.

[41] Di Nunno, G. “Stochastic integral representations, stochastic derivatives and minimal variance hedging”. In: *Stoch. Stoch. Rep.* vol. 73, no. 1-2 (2002), pp. 181–198.

[42] Di Nunno, G., Mishura, Y., and Yurchenko-Tytarenko, A. “Drift-implicit Euler scheme for sandwiched processes driven by Hölder noises”. In: *Numerical Algorithms (to appear)* (2022).

[43] Di Nunno, G., Mishura, Y., and Yurchenko-Tytarenko, A. “Option pricing in Volterra Sandwiched Volatility model”. In: *arXiv:2209.10688* (2022).

[44] Di Nunno, G., Mishura, Y., and Yurchenko-Tytarenko, A. “Sandwiched SDEs with unbounded drift driven by Hölder noises”. In: *Advances in Applied Probability (to appear)* vol. 55, no. 3 (2023).

[45] Di Nunno, G., Mishura, Y., and Yurchenko-Tytarenko, A. "Volterra Sandwiched Volatility model: Markovian approximations and hedging". In: *arXiv:2209.13054* (2022).

[46] Ding, Z. and Granger, C. W. J. "Modeling volatility persistence of speculative returns: A new approach". In: *Journal of Econometrics* vol. 73, no. 1 (1996), pp. 185–215.

[47] Ding, Z., Granger, C. W. J., and Engle, R. F. "A long memory property of stock market returns and a new model". In: *Journal of empirical finance* vol. 1, no. 1 (1993), pp. 83–106.

[48] Duffee, G. "Stock returns and volatility a firm-level analysis". In: *Journal of financial economics* vol. 37, no. 3 (1995), pp. 399–420.

[49] El Euch, O., Fukasawa, M., and Rosenbaum, M. "The microstructural foundations of leverage effect and rough volatility". In: *Finance and stochastics* vol. 22, no. 2 (2018), pp. 241–280.

[50] El Euch, O. and Rosenbaum, M. "The characteristic function of rough Heston models". In: *Mathematical Finance. An International Journal of Mathematics, Statistcs and Financial Economics* vol. 29, no. 1 (2019), pp. 3–38.

[51] El Euch, O. et al. "The Zumbach effect under rough Heston". In: *Quantitative finance* vol. 20, no. 2 (2020), pp. 235–241.

[52] Fama, E. F. "The Behavior of Stock-Market Prices". In: *The Journal of Business* vol. 38, no. 1 (1965), pp. 34–105.

[53] Fengler, M. *Semiparametric modeling of implied volatility*. 2005th ed. Berlin, Germany: Springer, 2005.

[54] Figlewski, S. and Wang, X. "Is the "leverage effect" a leverage effect?" In: *SSRN Electronic Journal* (2001).

[55] Fouque, J.-P., Papanicolaou, G., and Sircar, K. R. *Derivatives in financial markets with stochastic volatility*. Cambridge, England: Cambridge University Press, 2000.

[56] Frey, R. "Derivative Asset Analysis in Models with Level-Dependent and Stochastic Volatility". In: *Research Papers in Economics* (1997).

[57] Fukasawa, M. "Asymptotic analysis for stochastic volatility: martingale expansion". In: *Finance and stochastics* vol. 15, no. 4 (2011), pp. 635–654.

[58] Fukasawa, M. "Volatility has to be rough". In: *Quantitative finance* vol. 21, no. 1 (2021), pp. 1–8.

[59] Fukasawa, M. and Gatheral, J. "A rough SABR formula". In: *Frontiers of Mathematical Finance* vol. 1, no. 1 (2022), p. 81.

[60] Fukasawa, M., Takabatake, T., and Westphal, R. "Is volatility rough?" 2019.

[61] Funahashi, H. and Kijima, M. “A solution to the time-scale fractional puzzle in the implied volatility”. In: *Fractal and fractional* vol. 1, no. 1 (2017), p. 14.

[62] Funahashi, H. and Kijima, M. “Does the Hurst index matter for option prices under fractional volatility?” In: *Annals of finance* vol. 13, no. 1 (2017), pp. 55–74.

[63] Gatheral, J. *The volatility surface: A practitioner’s guide.* Ed. by Gatheral, J. Nashville, TN: John Wiley & Sons, 2006.

[64] Gatheral, J., Jaisson, T., and Rosenbaum, M. “Volatility is rough”. In: *SSRN Electronic Journal* (2014).

[65] Ghysels, E., Harvey, A. C., and Renault, E. “Stochastic volatility”. In: *Handbook of Statistics.* Elsevier, 1996, pp. 119–191.

[66] Harms, P. and Stefanovits, D. “Affine representations of fractional processes with applications in mathematical finance”. In: *Stochastic processes and their applications* vol. 129, no. 4 (2019), pp. 1185–1228.

[67] Harrison, J. M. and Kreps, D. M. “Martingales and arbitrage in multiperiod securities markets”. In: *Journal of Economic Theory* vol. 20, no. 3 (1979), pp. 381–408.

[68] Harrison, J. M. and Pliska, S. R. “Martingales and stochastic integrals in the theory of continuous trading”. In: *Stochastic processes and their applications* vol. 11, no. 3 (1981), pp. 215–260.

[69] Hasanhodzic, J. and Lo, A. W. “On Black’s leverage effect in firms with no leverage”. In: *The Journal of Portfolio Management* vol. 46, no. 1 (2019), pp. 106–122.

[70] Heston, S. L. “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. In: *Review of Financial Studies* vol. 6, no. 2 (1993), pp. 327–43.

[71] Hong, J. et al. “Optimal strong convergence rate of a backward Euler type scheme for the Cox–Ingersoll–Ross model driven by fractional Brownian motion”. In: *Stochastic Processes and their Applications* vol. 130, no. 5 (2020), pp. 2675–2692.

[72] Hu, Y., Nualart, D., and Song, X. “A singular stochastic differential equation driven by fractional Brownian motion”. In: *Statistics & probability letters* vol. 78, no. 14 (2008), pp. 2075–2085.

[73] Hull, J. and White, A. “The pricing of options on assets with stochastic volatilities”. In: *The journal of finance* vol. 42, no. 2 (1987), pp. 281–300.

[74] Hull, J. C. *Options, Futures, and Other Derivatives.* 11th ed. New York: Pearson, 2022.

[75] Jackwerth, J. C. and Rubinstein, M. “Recovering probability distributions from option prices”. In: *The journal of finance* vol. 51, no. 5 (1996), p. 1611.

[76] Kolmogorov, A. "Wiener Spirals and some other interesting curves in Hilbert space (In Russian)". In: *Doklady AN SSSR* vol. 26, no. 2 (1940).

[77] Kreps, D. M. "Arbitrage and equilibrium in economies with infinitely many commodities". In: *Journal of mathematical economics* vol. 8, no. 1 (1981), pp. 15–35.

[78] Kubilius, K. and Medžiūnas, A. "Positive solutions of the fractional SDEs with non-Lipschitz diffusion coefficient". In: *Mathematics* vol. 9, no. 1 (2020).

[79] Latané, H. A. and Rendleman Richard J., J. "Standard deviations of stock price ratios implied in option prices". In: *The journal of finance* vol. 31, no. 2 (1976), pp. 369–381.

[80] Lee, R. W. "Implied volatility: Statics, dynamics, and probabilistic interpretation". In: *Recent Advances in Applied Probability.* Boston: Kluwer Academic Publishers, 2006, pp. 241–268.

[81] Lobato, I. N. and Velasco, C. "Long memory in stock-market trading volume". In: *Journal of business & economic statistics: a publication of the American Statistical Association* vol. 18, no. 4 (2000), pp. 410–427.

[82] Mandelbrot, B. "The Variation of Certain Speculative Prices". In: *The Journal of Business* vol. 36, no. 4 (1963), pp. 394–419.

[83] Mandelbrot, B. B. and Van Ness, J. W. "Fractional Brownian motions, fractional noises and applications". In: *SIAM review* vol. 10, no. 4 (1968), pp. 422–437.

[84] Mariani, M. C. and Florescu, I. *Quantitative Finance.* Nashville, TN: John Wiley & Sons, 2020.

[85] Merino, R. et al. "Decomposition formula for rough Volterra stochastic volatility models". In: *International journal of theoretical and applied finance* vol. 24, no. 02 (2021), p. 2150008.

[86] Merton, R. C. "Theory of rational option pricing". In: *The Bell journal of economics and management science* vol. 4, no. 1 (1973), p. 141.

[87] Mikosch, T. and Starica, C. "Is it really long memory we see in financial returns?" In: *Extremes and Integrated Risk Management.* Risk Books, 2000, pp. 149–168.

[88] Mishura, Y. *Stochastic calculus for fractional Brownian motion and related processes.* 2008th ed. Berlin, Germany: Springer, 2007.

[89] Mishura, Y. and Yurchenko-Tytarenko, A. "Fractional Cox–Ingersoll–Ross process with non-zero "mean"". In: *Modern Stochastics: Theory and Applications* vol. 5, no. 1 (2018), pp. 99–111.

[90] Mishura, Y. and Yurchenko-Tytarenko, A. "Standard and fractional reflected Ornstein–Uhlenbeck processes as the limits of square roots of Cox–Ingersoll–Ross processes". In: *Stochastics* (2022), pp. 1–19.

[91] Osborne, M. F. M. “Brownian Motion in the Stock Market”. In: *Operations Research* vol. 7, no. 2 (1959), pp. 145–173.

[92] Rachev, S. T. et al. *Financial models with Lévy processes and volatility clustering*. Chichester, England: John Wiley & Sons, 2011.

[93] Renault, E. and Touzi, N. “Option hedging and implied volatilities in a stochastic volatility model”. In: *Mathematical Finance. An International Journal of Mathematics, Statistcs and Financial Economics* vol. 6, no. 3 (1996), pp. 279–302.

[94] Rogers, L. C. G. *Things we think we know*. 2019.

[95] Rosenbaum, M. “Estimation of the volatility persistence in a discretely observed diffusion model”. In: *Stochastic processes and their applications* vol. 118, no. 8 (2008), pp. 1434–1462.

[96] Ross, S. A. “A Simple Approach to the Valuation of Risky Streams”. In: *The Journal of Business* vol. 51, no. 3 (1978), pp. 453–475.

[97] Ross, S. A. “Finance”. In: *The New Palgrave Dictionary of Economics*. Ed. by Eatwell, J., Milgate, M., and Newman, P. Vol. 2. London: Macmillan, 1987, pp. 322–336.

[98] *Rough Volatility Literature*. https://sites.google.com/site/roughvol/home/risks-1. Accessed: 2022-09-15.

[99] Samuelson, P. A. “Rational theory of warrant pricing”. In: *Industrial Management Review* vol. 6, no. 2 (1959), pp. 13–39.

[100] Samuelson, P. A. and Merton, R. C. “A Complete Model of Warrant Pricing that Maximizes Utility”. In: *Industrial Management Review* vol. 10 (1972), pp. 17–46.

[101] Schachermayer, W. “The fundamental theorem of asset pricing”. In: *Handbook of the Fundamentals of Financial Decision Making*. World Scientific, 2013, pp. 31–48.

[102] Scott, L. O. “Option pricing when the variance changes randomly: Theory, estimation, and an application”. In: *Journal of financial and quantitative analysis* vol. 22, no. 4 (1987), p. 419.

[103] Shephard, N. and Andersen, T. G. “Stochastic volatility: Origins and overview”. In: *Handbook of Financial Time Series*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 233–254.

[104] Shiryaev, A. N. *Essentials of stochastic finance: Facts, models, theory*. Singapore: World Scientific Publishing, 1999.

[105] Skorokhod, A. V. “Stochastic Equations for Diffusion Processes in a Bounded Region”. In: *Theory of Probability & Its Applications* vol. 6, no. 3 (Jan. 1961), pp. 264–274.

[106] Skorokhod, A. V. “Stochastic equations for diffusion processes in a bounded region. II”. In: *Theory of probability and its applications* vol. 7, no. 1 (1962), pp. 3–23.

[107] Sprenkle, C. M. “Warrant prices as indicators of expectations and preferences”. In: *Yale economic essays* vol. 1, no. 2 (1959), pp. 179–231.

[108] Stein, E. M. and Stein, J. C. “Stock Price Distributions with Stochastic Volatility: An Analytic Approach”. In: *Review of Financial Studies* vol. 4 (1991), pp. 727–752.

[109] Tankov, P. and Cont, R. *Financial modelling with jump processes*. Philadelphia, PA: Chapman & Hall/CRC, 2003.

[110] Thorp, E. O. and Kassouf, S. T. *Beat the Market: A Scientific Stock Market System*. New York: Random House, 1967.

[111] Wiggins, J. B. “Option values under stochastic volatility: Theory and empirical estimates”. In: *Journal of financial economics* vol. 19, no. 2 (1987), pp. 351–372.

[112] Willinger, W., Taqqu, M. S., and Teverovsky, V. “Stock market prices and long-range dependence”. In: *Finance and stochastics* vol. 3, no. 1 (1999), pp. 1–13.

[113] Zhang, S.-Q. and Yuan, C. “Stochastic differential equations driven by fractional Brownian motion with locally Lipschitz drift and their implicit Euler approximation”. In: *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* (2020), pp. 1–27.

[114] Zhu, Q. et al. “Markovian approximation of the rough Bergomi model for Monte Carlo option pricing”. In: *Mathematics* vol. 9, no. 5 (2021), p. 528.

[115] Zumbach, G. *Discrete time series, processes, and applications in finance*. 2013th ed. Berlin, Germany: Springer, 2012.

[116] Zumbach, G. “Volatility conditional on price trends”. In: *Quantitative finance* vol. 10, no. 4 (2010), pp. 431–442.

Papers

Paper I

Standard and fractional reflected Ornstein-Uhlenbeck processes as the limits of square roots of Cox-Ingersoll-Ross processes

Yuliya Mishura, Anton Yurchenko-Tytarenko

Published in *Stochastics* (2022).
DOI: 10.1080/17442508.2022.2047188.

Abstract

In this paper, we establish a new connection between Cox-Ingersoll-Ross (CIR) and reflected Ornstein-Uhlenbeck (ROU) models driven by either a standard Wiener process or a fractional Brownian motion with $H > \frac{1}{2}$. We prove that, with probability 1, the square root of the CIR process converges uniformly on compacts to the ROU process as the mean reversion parameter tends to either $\sigma^2/4$ (in the standard case) or to 0 (in the fractional case). This also allows to obtain a new representation of the reflection function of the ROU as the limit of integral functionals of the CIR processes. The results of the paper are illustrated by simulations.

Contents

The present research is carried out within the frame and support of the ToppForsk project nr. 274410 of the Research Council of Norway with title STORM: Stochastics for Time-Space Risk Models. The first author is supported by the National Research Fund of Ukraine under grant 2020.02/0026.

I.1 Introduction

Both the reflected Ornstein-Uhlenbeck (ROU) and the Cox-Ingersoll-Ross (CIR) processes are extremely popular models in a variety of fields. Without attempting to give a complete overview of possible applications due to the large amount of literature on the topic, we only mention that the ROU process is widely used in queueing theory [9, 29–31], in population dynamics modeling [24, 37], in economics and finance for modeling regulated markets [2, 3, 16, 32], interest rates [10] and stochastic volatility [26] (see also [8, 19] and references therein for more details on applications of the ROU in various fields) while the most notable usages of the CIR process are related to representing the dynamics of interest rates [4–6] and stochastic volatility in the Heston model [11].

It is well-known [20, 27] that the CIR process has strong links with the standard OU dynamics; in particular, if $B = (B_1, ..., B_d)$ is a d-dimensional Brownian motion and $U = (U_1, ..., U_d)$ is a standard d-dimensional OU process given by

$$U_i(t) = U_i(0) - \frac{b}{2}\int_0^t U_i(s)ds + \frac{\sigma}{2}B_i(t), \quad t \geq 0, \quad i = 1, ..., d,$$

then it is easy to see via Itô's formula that the process $\sum_{i=1}^d U_i^2(t)$, $t \geq 0$, is the CIR process of the form

$$X(t) = X(0) + \int_0^t (a - bX(s))\, ds + \sigma \int_0^t \sqrt{X(s)}dW(s), \quad t \geq 0, \tag{I.1}$$

with $a = \frac{d\sigma^2}{4}$ and $W(t) := \sum_{i=1}^d \int_0^t \frac{U_i(s)}{\sqrt{\sum_{j=1}^d U_j^2(s)}}dB_i(s)$ (which is a standard Brownian motion by Levy's characterization). The value $d = \frac{4a}{\sigma^2}$ is sometimes referred to as a *dimension* or a *number of degrees of freedom* of the CIR process (see e.g. [20] and references therein) and thus, in this terminology, a square of a standard one-dimensional OU process turns out to be a CIR process with one degree of freedom w.r.t. another Brownian motion.

In this paper, we investigate a connection between the CIR and the ROU processes that is in some sense related to the one described above. Namely, in the first part we prove that the ROU process

$$Y(t) = Y(0) - \frac{b}{2}\int_0^t Y(s)ds + \frac{\sigma}{2}W(t) + L(t), \quad t \geq 0, \tag{I.2}$$

where W is a standard Brownian motion and L is a continuous non-decreasing process that can have points of growth only at zeros of Y, coincides with the square root of the CIR process of the type (I.1) with $a = \frac{\sigma^2}{4}$ (i.e. with one degree of freedom) driven by the same Brownian motion W. Moreover, if $\{\varepsilon_n,\ n \geq 1\}$ is a sequence of positive numbers such that $\varepsilon_n \downarrow 0$ as $n \to \infty$, then, with probability 1, for all $T > 0$

$$\sup_{t\in[0,T]} \left| L(t) - \frac{1}{2}\int_0^t \frac{\varepsilon_n}{\sqrt{X_{\varepsilon_n}(s)}}ds \right| \to 0, \quad n \to \infty, \tag{I.3}$$

where X_{ε_n} is the CIR process of the form

$$X_{\varepsilon_n}(t) = X(0) + \int_0^t \left(\frac{\sigma^2}{4} + \varepsilon_n - bX_{\varepsilon_n}(s) \right) ds + \sigma \int_0^t \sqrt{X_{\varepsilon_n}(s)} dW(s).$$

The second part of the paper discusses the connection between fractional counterparts of equations (I.1) and (I.2) driven by fractional Brownian motion $\{B^H(t),\ t \geq 0\}$ with Hurst index $H > \frac{1}{2}$. Namely, we consider a fractional Cox-Ingersoll-Ross process

$$X_\varepsilon^H(t) = X(0) + \int_0^t \left(\varepsilon - bX_\varepsilon^H(s) \right) ds + \sigma \int_0^t \sqrt{X_\varepsilon^H(s)} dB^H(s), \quad t \geq 0,$$

where the integral $\int_0^t \sqrt{X^H(s)} dB^H(s)$ is understood as the pathwise limit of Riemann-Stieltjes integral sums (see [22] or [7, Subsection 4.1]) and prove that with probability 1 the paths of $\{\sqrt{X_\varepsilon^H(t)},\ t \geq 0\}$ a.s. converge to the reflected fractional Ornstein-Uhlenbeck (RFOU) process uniformly on each compact $[0, T]$ as $\varepsilon \downarrow 0$. Moreover, an analogue of the representation (I.3) also takes place: if L^H is a reflection function of the RFOU process, then, with probability 1, for each $T > 0$

$$\sup_{t \in [0,T]} \left| L^H(t) - \frac{1}{2} \int_0^t \frac{\varepsilon}{\sqrt{X_\varepsilon^H(s)}} ds \right| \to 0, \quad \varepsilon \downarrow 0.$$

The paper is organised as follows. In section I.2, we consider the link between the CIR and the ROU processes in the standard Wiener case. Section I.3 is devoted to the fractional setting. Section I.4 contains simulations that illustrate our results.

I.2 Classical reflected Ornstein-Uhlenbeck and Cox-Ingersoll-Ross processes

The main goal of this section is to establish a connection between Cox-Ingersoll-Ross (CIR) and reflected Ornstein-Uhlenbeck (ROU) processes in the standard Brownian setting. We shall start from the definition of a *reflection function* following the one given in the classical work [28].

Definition I.2.1. Let $\xi = \{\xi(t),\ t \geq 0\}$ be some stochastic process. The process $\zeta = \{\zeta(t),\ t \geq 0\}$ is called *a reflection function* for ξ, if ζ is, with probability 1, a continuous non-decreasing process such that $\zeta(0) = 0$ and the points of growth of ζ can occur only at zeros of ξ.

Definition I.2.2. Stochastic process $\widetilde{Y} = \{\widetilde{Y}(t),\ t \geq 0\}$ is called a *reflected Ornstein-Uhlenbeck (ROU) process* if it satisfies a stochastic differential equation of the form

$$\widetilde{Y}(t) = Y(0) - \widetilde{b} \int_0^t \widetilde{Y}(s) ds + \widetilde{\sigma} W(t) + \widetilde{L}(t), \quad t \geq 0, \tag{I.4}$$

where $Y(0)$, $\widetilde{b}$ and $\widetilde{\sigma}$ are positive constants, $W = \{W(t),\ t \geq 0\}$ is a standard Brownian motion, $\{\widetilde{L}(t),\ t \geq 0\}$ is a reflection function for $\widetilde{Y}$ and $\widetilde{Y} \geq 0$ a.s.

Remark I.2.3. The ROU process is well-known and studied in the literature, see e.g. [31] and references therein. Note also that, despite (I.4) has two unknown functions $\widetilde{Y}$ and $\widetilde{L}$, the solution is still unique. Indeed, let $\widetilde{Y}$ and $\widehat{Y}$ be two stochastic processes satisfying

$$\widetilde{Y}(t) = Y(0) - \widetilde{b}\int_0^t \widetilde{Y}(s)ds + \widetilde{\sigma}W(t) + \widetilde{L}(t)$$

and

$$\widehat{Y}(t) = Y(0) - \widetilde{b}\int_0^t \widehat{Y}(s)ds + \widetilde{\sigma}W(t) + \widehat{L}(t),$$

where $\widetilde{L}$ and $\widehat{L}$ are the corresponding reflection functions. Assume that on some $\omega \in \Omega$ such that both $\widetilde{Y}$ and $\widehat{Y}$ are continuous

$$\widetilde{Y}(t) - \widehat{Y}(t) > 0 \tag{I.5}$$

and consider $\tau(t) := \sup\{s \in [0,t) :\ \widetilde{Y}(t) - \widehat{Y}(t) = 0\}$. Then $\widetilde{Y}(u) - \widehat{Y}(u) > 0$ for all $u \in (\tau(t), t]$; moreover, $\widetilde{Y}(u) > 0$ for $u \in (\tau(t), t]$, therefore $\widetilde{L}$ is non-increasing on $(\tau(t), t]$. It means that the difference $\widetilde{Y}(u) - \widehat{Y}(u)$ is also non-increasing on $(\tau(t), t]$ since

$$\widetilde{Y}(u) - \widehat{Y}(u) = -\widetilde{b}\int_{\tau(t)}^u (\widetilde{Y}(s) - \widehat{Y}(s))ds + \left(\widetilde{L}(u) - \widehat{L}(u)\right) - \left(\widetilde{L}(\tau(t)) - \widehat{L}(\tau(t))\right)$$

and the right-hand side is non-increasing w.r.t. u. Whence, taking into account that $\widetilde{Y}(\tau(t)) - \widehat{Y}(\tau(t)) = 0$ due to the definition of $\tau(t)$ and continuity of both $\widetilde{Y}$ and $\widehat{Y}$, the difference $\widetilde{Y}(u) - \widehat{Y}(u)$ cannot be positive for any $u \in (\tau(t), t]$ which contradicts (I.5). Interchanging the roles of $\widetilde{Y}$ and $\widehat{Y}$, one can easily verify that $\widetilde{Y}(t) - \widehat{Y}(t)$ cannot be negative either and whence $\widehat{Y} = \widetilde{Y}$, $\widehat{L} = \widetilde{L}$.

Now, consider a standard CIR process defined as a continuous modification of the unique solution to the equation

$$X(t) = X(0) + \int_0^t (a - bX(s))\, ds + \sigma\int_0^t \sqrt{X(s)}dW(s), \quad t \geq 0, \tag{I.6}$$

where $X(0), a, b, \sigma > 0$ and $W = \{W(t),\ t \geq 0\}$ is a classical Wiener process. It is well-known (see e.g. [14, Example 8.2]) that for $a > 0$ the solution $\{X(t),\ t \geq 0\}$ is non-negative a.s. for any $t \geq 0$; moreover, the solution is strictly positive a.s. provided that $a \geq \frac{\sigma^2}{2}$, see e.g. [15, Chapter 5]. Therefore, if $a > 0$, the square-root process $Y = \{Y(t),\ t \in [0,T]\} := \{\sqrt{X(t)},\ t \in [0,T]\}$ is well-defined.

For an arbitrary $\varepsilon > 0$, consider a stochastic process $\{\sqrt{X(t)+\varepsilon},\ t \in [0,T]\}$. By Itô's formula, for any $t \geq 0$

$$\sqrt{X(t)+\varepsilon} = \sqrt{X(0)+\varepsilon} + \frac{1}{2}\int_0^t \left(\frac{a}{\sqrt{X(s)+\varepsilon}} - \frac{\sigma^2}{4}\frac{X(s)}{(X(s)+\varepsilon)^{\frac{3}{2}}} \right) ds$$
$$- \frac{1}{2}\int_0^t \frac{bX(s)}{\sqrt{X(s)+\varepsilon}} ds + \frac{\sigma}{2}\int_0^t \frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon}} dW(s) \tag{I.7}$$

and, since the left-hand side of (I.7) converges to $\sqrt{X(t)} = Y(t)$ a.s. as $\varepsilon \to 0$, moving $\varepsilon \to 0$ in the right-hand side would give us the dynamics of Y.

First, it is clear that for any $t \geq 0$

$$\sqrt{X(0)+\varepsilon} \to Y(0) \quad a.s. \tag{I.8}$$

and

$$\int_0^t \frac{X(s)}{\sqrt{X(s)+\varepsilon}} ds \to \int_0^t Y(s) ds \quad a.s. \tag{I.9}$$

as $\varepsilon \to 0$. Further, by the monotone convergence,

$$\int_0^t \frac{1}{\sqrt{X(s)+\varepsilon}} ds \to \int_0^t \frac{1}{Y(s)} ds \in [0,\infty) \cup \{\infty\} \quad a.s.,$$
$$\int_0^t \frac{X(s)}{(X(s)+\varepsilon)^{\frac{3}{2}}} \to \int_0^t \frac{1}{Y(s)} ds \in [0,\infty) \cup \{\infty\} \quad a.s. \tag{I.10}$$

as $\varepsilon \to 0$. Finally, by Burkholder-Davis-Gundy inequality and dominated convergence theorem, for any $T > 0$

$$\mathbb{E}\left(\sup_{t\in[0,T]} \left| \int_0^t \frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon}} dW(s) - W(t) \right| \right)^2 \leq 4\mathbb{E}\int_0^T \left(\frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon}} - 1 \right)^2 ds$$
$$= 4\mathbb{E}\int_0^T \left(\frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon}} - 1 \right)^2 \mathbb{1}_{\{X(s)>0\}} ds + 4\mathbb{E}\int_0^T \mathbb{1}_{\{X(s)=0\}} ds$$
$$= 4\mathbb{E}\int_0^T \left(\frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon}} - 1 \right)^2 \mathbb{1}_{\{X(s)>0\}} ds \to 0, \quad \varepsilon \to 0,$$

where we used continuity of the distribution of $X(s)$ for each $s > 0$ to state that $4\mathbb{E}\int_0^T \mathbb{1}_{\{X(s)=0\}} ds = 0$ (see e.g. [20] and references therein). This implies that

$$\sup_{t\in[0,T]} \left| \int_0^t \frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon}} dW(s) - W(t) \right| \xrightarrow{L^2(\Omega)} 0, \qquad \varepsilon \to 0. \tag{I.11}$$

By (I.11), it is evident that there exists a sequence $\{\varepsilon_n,\ n \geq 1\}$ which depends on T such that

$$\sup_{t\in[0,T]} \left| \int_0^t \frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon_n}} dW(s) - W(t) \right| \to 0 \quad a.s., \quad n \to \infty \tag{I.12}$$

and along this sequence

$$\lim_{n\to\infty} \frac{1}{2} \int_0^t \left(\frac{a}{\sqrt{X(s)+\varepsilon_n}} - \frac{\sigma^2}{4} \frac{X(s)}{(X(s)+\varepsilon_n)^{\frac{3}{2}}} \right) ds < \infty, \quad t \in [0,T], \tag{I.13}$$

a.s. because all other limits in (I.7) as $\varepsilon_n \to 0$ are finite a.s. However, the integral $\int_0^t \frac{1}{Y(s)} ds$ which arises in (I.10) may be infinite and thus the explicit form of the limit above for now remains obscure. This issue as well as the connection of Y to the ROU process is addressed in the next theorem.

Theorem I.2.4. *Let $Y = \{Y(t), t \ge 0\} = \{\sqrt{X(t)},\ t \ge 0\}$ be the square root process, where X is the CIR process defined by* (I.6)*.*

$$\tau := \inf\{t \ge 0 : X(t) = 0\} = \inf\{t \ge 0 : Y(t) = 0\}.$$

(a) *If $a > \frac{\sigma^2}{4}$, then for any $t \ge 0$*

$$\int_0^t \frac{1}{Y(s)} ds < \infty \quad a.s.$$

Moreover, the square root process Y a.s. satisfies the SDE of the form

$$Y(t) = Y(0) + \frac{1}{2}\left(a - \frac{\sigma^2}{4}\right) \int_0^t \frac{1}{Y(s)} ds - \frac{b}{2}\int_0^t Y(s) ds + \frac{\sigma}{2} W(t), \tag{I.14}$$

$Y(0) = \sqrt{X(0)}$, and the solution to this equation is unique among non-negative stochastic processes.

(b) *If $a = \frac{\sigma^2}{4}$, then*

$$\int_0^\tau \frac{1}{Y(s)} ds < \infty \quad a.s.$$

while

$$\int_0^{\tau+\gamma} \frac{1}{Y(s)} ds = \infty \quad a.s.$$

for any $\gamma > 0$. Moreover, the square root process Y satisfies the SDE of the form

$$Y(t) = Y(0) - \frac{b}{2}\int_0^t Y(s) ds + \frac{\sigma}{2} W(t) + L(t), \tag{I.15}$$

where the process L from (I.15) *is a continuous nondecreasing process the points of growth of which can occur only at zeros of Y, i.e. Y is a reflected Ornstein-Uhlenbeck process.*

Proof. Case (a): $a > \frac{\sigma^2}{4}$. Denote $p := a - \frac{\sigma^2}{4} > 0$. Our goal is to prove that the integral

$$\int_0^t \frac{1}{\sqrt{X(s)}} ds = \int_0^t \frac{1}{Y(s)} ds$$

is finite a.s. Define $A(t) := \left\{ \omega \in \Omega : \int_0^t \frac{1}{\sqrt{X(s)}} ds = +\infty \right\}$ and assume that for some $t > 0$: $\mathbb{P}(A(t)) > 0$. Fix $T > t$, the corresponding sequence $\{\varepsilon_n, \ n \geq 1\}$ such that convergence (I.12) holds and an arbitrary $\omega \in A(t) \cap \Omega'$, where $\Omega' \subset \Omega$, $\mathbb{P}(\Omega') = 1$, is the set where (I.12) takes place (in what follows, ω in brackets will be omitted). Then

$$\int_0^t \left(\frac{a}{\sqrt{X(s) + \varepsilon_n}} - \frac{\sigma^2}{4} \frac{X(s)}{(X(s) + \varepsilon_n)^{\frac{3}{2}}} \right) ds$$
$$= \frac{\sigma^2}{4} \int_0^t \left(\frac{1}{\sqrt{X(s) + \varepsilon_n}} - \frac{X(s)}{(X(s) + \varepsilon_n)^{\frac{3}{2}}} \right) ds + p \int_0^t \frac{1}{\sqrt{X(s) + \varepsilon_n}} ds.$$

Obviously, for all $s \in [0, t]$

$$\frac{1}{\sqrt{X(s) + \varepsilon_n}} \geq \frac{X(s)}{(X(s) + \varepsilon_n)^{\frac{3}{2}}} \quad a.s.,$$

so, for $\omega \in A(t) \cap \Omega'$

$$\int_0^t \left(\frac{a}{\sqrt{X(s) + \varepsilon_n}} - \frac{\sigma^2}{4} \frac{X(s)}{(X(s) + \varepsilon_n)^{\frac{3}{2}}} \right) ds \to \infty \quad a.s., \quad n \to \infty,$$

whence, taking into account (I.7)–(I.9), we obtain that

$$\sqrt{X(t)} - \sqrt{X(0)} + \frac{b}{2} \int_0^t \sqrt{X(s)} ds - \frac{\sigma}{2} W(t) = \infty,$$

which is impossible a.s. We get a contradiction, whence $\mathbb{P}(A(t)) = 0$ for all $t \geq 0$ and $\int_0^t \frac{1}{\sqrt{X(s)}} ds = \int_0^t \frac{1}{Y(s)} ds < \infty$ a.s. By going to the limit in (I.7), we immediately get (I.14).

Concerning the uniqueness of solution to (I.14), let $\widetilde{Y}(t)$ be any of its non-negative solutions. Then, by Itô's formula,

$$\widetilde{Y}^2(t) = X(0) + \int_0^t \left(a - b\widetilde{Y}^2(s) \right) ds + \sigma \int_0^t \widetilde{Y}(s) dW(s)$$

so $\widetilde{Y}$ satisfies equation (I.6) and thus coincides with X. Therefore

$$\widetilde{Y}(t) = \sqrt{X(t)} = Y(t) \quad a.s., \quad t \geq 0.$$

Case (b): $a = \frac{\sigma^2}{4}$. Fix $T > 0$ and take the corresponding sequence

$\{\varepsilon_n,\ n \geq 1\}$ such that (I.12) holds. By (I.7), for any $t \in [0, T]$

$$\sqrt{X(t)+\varepsilon_n} = \sqrt{X(0)+\varepsilon_n} + \frac{1}{2}\int_0^t \left(\frac{a}{\sqrt{X(s)+\varepsilon_n}} - \frac{\sigma^2}{4}\frac{X(s)}{(X(s)+\varepsilon_n)^{\frac{3}{2}}} \right) ds$$
$$- \frac{1}{2}\int_0^t \frac{bX(s)}{\sqrt{X(s)+\varepsilon_n}} ds + \frac{\sigma}{2}\int_0^t \frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon_n}} dW(s)$$
$$= \sqrt{X(0)+\varepsilon_n} + \frac{\sigma^2}{8}\int_0^t \frac{\varepsilon_n}{(X(s)+\varepsilon_n)^{\frac{3}{2}}} ds$$
$$- \frac{1}{2}\int_0^t \frac{bX(s)}{\sqrt{X(s)+\varepsilon_n}} ds + \frac{\sigma}{2}\int_0^t \frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon_n}} dW(s),$$

and (I.13) implies that there exists $\Omega' \subset \Omega$, $\mathbb{P}(\Omega') = 1$, such that for all $\omega \in \Omega'$ the limit

$$L(t) := \lim_{n\to\infty} \frac{\sigma^2}{8}\int_0^t \frac{\varepsilon_n}{(X(s)+\varepsilon_n)^{\frac{3}{2}}} ds$$

is well-defined and finite for all $t \in [0, T]$. It is evident that $L(0) = 0$ a.s. due to continuity of X and the fact that $X(0) > 0$. Moreover, since a.s.

$$L(t) = \sqrt{X(t)} - \sqrt{X(0)} + \frac{b}{2}\int_0^t \sqrt{X(s)} ds - \frac{\sigma}{2} W(t) \quad t \in [0, T],$$

L is continuous in t. Furthermore,

$$\int_0^{t_1} \frac{\varepsilon_n}{(X(s)+\varepsilon_n)^{\frac{3}{2}}} ds \leq \int_0^{t_2} \frac{\varepsilon_n}{(X(s)+\varepsilon_n)^{\frac{3}{2}}} ds$$

for all $t_1 < t_2$, and whence L is non-decreasing in t a.s. Finally, if $X(t) = x > 0$, there exists an interval $[t_1, t_2]$ containing t such that $X(s) > \frac{x}{2}$ for all $s \in [t_1, t_2]$ and thus

$$L(t_1) - L(t_2) = \lim_{n\to\infty} \frac{\sigma^2}{8}\int_{t_1}^{t_2} \frac{\varepsilon_n}{(X(s)+\varepsilon_n)^{\frac{3}{2}}} ds \to 0, \quad n \to \infty,$$

i.e. L can increase only at points of zero hitting of X that coincide with the ones of Y. Taking into the account all of the above as well as an arbitrary choice of T, L is the reflection function for Y and the latter is indeed a ROU process.

Now, let us prove that $\int_0^\tau \frac{1}{Y_s} ds < \infty$ a.s. Consider a standard Ornstein-Uhlenbeck process $U = \{U(t),\ t \geq 0\}$ of the form

$$U(t) = \sqrt{X(0)} - \frac{b}{2}\int_0^t U(s) ds + \frac{\sigma}{2} W(t), \tag{I.16}$$

with W being the same Brownian motion that drives X. It is evident that Y coincides with U until τ a.s. and thus it is sufficient to prove that $\int_0^\tau \frac{1}{U(s)} ds < \infty$ a.s. For any $\varepsilon > 0$ consider

$$\frac{\sigma^2}{4}\int_0^\tau \frac{1}{U(s)} \mathbb{1}_{\{\varepsilon < U(s) < 1\}} ds = \int_\varepsilon^1 \frac{L_U(\tau, x)}{x} dx,$$

where L_U denotes the local time of U, and observe that

$$\frac{\sigma^2}{4}\int_0^\tau \frac{1}{U(s)}ds \le \lim_{\varepsilon\downarrow 0}\int_0^\tau \frac{1}{U(s)}\mathbb{1}_{\{\varepsilon<U(s)<1\}}ds = \int_0^1 \frac{L_U(\tau,x)}{x}dx.$$

Computations similar to the ones in [25, Section IV.44] indicate that the local time $L_U(t,x)$ of U is Hölder continuous in x up to order $\frac{1}{2}$ over bounded time intervals and thus $\int_0^\tau \frac{1}{U(s)}ds = \int_0^\tau \frac{1}{Y(s)}ds < \infty$ a.s.

Finally, assume that for some $\gamma > 0$,

$$\int_0^{\tau+\gamma} \frac{1}{Y(s)}ds < \infty$$

with positive probability. On $\omega \in \Omega$ where this property holds, we have that

$$\begin{aligned}&\int_0^{\tau+\gamma}\frac{ds}{\sqrt{X(s)+\varepsilon}} - \int_0^{\tau+\gamma}\frac{X(s)}{(X(s)+\varepsilon)^{\frac{3}{2}}}ds\\ &\quad\to \int_0^{\tau+\gamma}\frac{1}{Y(s)}ds - \int_0^{\tau+\gamma}\frac{1}{Y(s)}ds = 0, \quad \varepsilon\to 0.\end{aligned}$$

Therefore, for such ω, Y satisfies the equation of the form

$$Y(t) = Y(0) - \frac{b}{2}\int_0^t Y(s)ds + \frac{\sigma}{2}W(t)$$

on the interval $[0,\tau+\gamma]$, i.e. such paths of Y coincide with the corresponding paths of the Ornstein-Uhlenbeck process U defined by (I.16) up until $\tau+\gamma$. This implies that $U(\tau) = 0$ and U is nonnegative on the interval $[\tau,\tau+\gamma]$ for such ω, which is impossible due to the non-tangent property of Gaussian processes stated by [33], see also [23]. ■

Remark I.2.5. Since the integral $\int_0^t \frac{1}{\sqrt{X(s)}}ds$ is finite a.s. for $a > \frac{\sigma^2}{4}$,

$$\begin{aligned}&\sqrt{X(t)+\varepsilon} - \sqrt{X(0)+\varepsilon} - \frac{1}{2}\int_0^t\left(\frac{a}{\sqrt{X(s)+\varepsilon}} - \frac{\sigma^2}{4}\frac{X(s)}{(X(s)+\varepsilon)^{\frac{3}{2}}}\right)ds\\ &\qquad + \frac{1}{2}\int_0^t \frac{bX(s)}{\sqrt{X(s)+\varepsilon}}ds\\ &\xrightarrow{a.s.} \sqrt{X(t)} - \sqrt{X(0)} - \frac{1}{2}\left(a-\frac{\sigma^2}{4}\right)\int_0^t\frac{1}{\sqrt{X(s)}}ds + \frac{b}{2}\int_0^t\sqrt{X(s)}ds < \infty\end{aligned}$$

as $\varepsilon \to 0$. Therefore, taking into account (I.7),

$$\int_0^t \frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon}}dW(s) \xrightarrow{a.s.} W(t), \quad \varepsilon\to 0.$$

As a corollary of Theorem I.2.4, we have a representation of the reflection function of the ROU process as the limit of integral functionals of the CIR processes. It is interesting that the reflection function is singular w.r.t. the Lebesgue measure (see Remark I.2.7) while the processes that converge to it are absolutely continuous a.s.

Theorem I.2.6. *Let $\{W(t),\ t \ge 0\}\}$ be a continuous modification of a standard Brownian motion, $Y(0)$, b, $\sigma > 0$ be given constants and $\{\varepsilon_n,\ n \ge 1\}$ be an arbitrary sequence such that $\varepsilon_n \downarrow 0$, $n \to \infty$. For any ε_n from this sequence, consider the CIR process $X_{\varepsilon_n} = \{X_{\varepsilon_n}(t),\ t \ge 0\}$ given by*

$$X_{\varepsilon_n}(t) = X(0) + \int_0^t \left(\frac{\sigma^2}{4} + \varepsilon_n - bX_{\varepsilon_n}(s)\right) ds + \sigma \int_0^t \sqrt{X_{\varepsilon_n}(s)} dW(s)$$

and denote its square root by $Y_{\varepsilon_n}(t) := \sqrt{X_{\varepsilon_n}(t)}$. Then, with probability 1,

1) *the limit $\lim_{n\to\infty} Y_{\varepsilon_n}(t) =: Y(t)$ is well-defined, finite and non-negative for any $t \ge 0$;*

2) *the limit process $Y = \{Y(t),\ t \ge 0\}$ is a ROU process satisfying the equation of the form*

$$Y(t) = Y(0) - \frac{b}{2}\int_0^t Y(s)ds + \frac{\sigma}{2}W(t) + L(t), \quad t \ge 0,$$

with $Y(0) = \sqrt{X(0)} > 0$ and L being the reflection function for Y;

3) *for any $T > 0$*

$$\sup_{t\in[0,T]} |Y(t) - Y_{\varepsilon_n}(t)| \to 0, \quad n \to \infty, \tag{I.17}$$

and

$$\sup_{t\in[0,T]} \left| L(t) - \frac{1}{2}\int_0^t \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)} ds \right| \to 0, \quad n \to \infty. \tag{I.18}$$

Proof. Denote by X the CIR process of the form

$$X(t) = X(0) + \int_0^t \left(\frac{\sigma^2}{4} - bX(s)\right) ds + \sigma \int_0^t \sqrt{X(s)} dW(s).$$

By Theorem I.2.4, there exists $\Omega' \subset \Omega$, $\mathbb{P}(\Omega') = 1$, such that for all $\omega \in \Omega'$ X and each Y_{ε_n}, $n \ge 1$, are continuous and the latter satisfy equations of the form

$$Y_{\varepsilon_n}(t) = Y(0) + \frac{1}{2}\int_0^t \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)} ds - \frac{b}{2}\int_0^t Y(s)ds + \frac{\sigma}{2}W(t), \quad t \ge 0,$$

with the integral $\int_0^t \frac{1}{Y_{\varepsilon_n}(s)} ds < \infty$. Furthermore, since each $X_{\varepsilon_n} = Y_{\varepsilon_n}^2$ is a CIR process that satisfies conditions of the comparison theorem from [13], this Ω' can be chosen such that for all $\omega \in \Omega'$

$$Y_{\varepsilon_n}(\omega, t) \ge Y_{\varepsilon_{n+1}}(\omega, t) \ge \sqrt{X(t)} \ge 0, \quad t \ge 0, \quad n \ge 1. \tag{I.19}$$

Fix $\omega \in \Omega'$ (in what follows, we will omit ω in brackets for notational simplicity). Since the sequence $\{Y_{\varepsilon_n}(t),\ n \ge 1\}$ is non-increasing for each $t \ge 0$, there exists a pointwise limit $Y(t) := \lim_{n\to\infty} Y_{\varepsilon_n}(t) \in [0,\infty)$. Moreover, it is evident that $\lim_{n\to\infty} \int_0^t Y_{\varepsilon_n}(s)ds = \int_0^t Y(s)ds$ and since

$$\begin{aligned} Y(t) &= \lim_{n\to\infty} Y_{\varepsilon_n}(t) \\ &= Y(0) - \lim_{n\to\infty} \frac{b}{2}\int_0^t Y_{\varepsilon_n}(s)ds + \frac{\sigma}{2}W(t) + \lim_{n\to\infty}\frac{1}{2}\int_0^t \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)}ds \\ &= Y(0) - \frac{b}{2}\int_0^t Y(s)ds + \frac{\sigma}{2}W(t) + \lim_{n\to\infty}\frac{1}{2}\int_0^t \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)}ds, \end{aligned}$$

the limit $L(t) := \lim_{n\to\infty} \frac{1}{2}\int_0^t \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)}ds$ is well-defined, nonnegative and finite.

In order to obtain the claim of the theorem, it is sufficient to check that the function L defined above is indeed a reflection function for Y, i.e. is continuous and nondecreasing process that starts at zero and the points of growth of which occur only at zeros of Y. Note that continuity of L would also imply the uniform convergences (I.17) and (I.18) on each compact $[0,T]$. Indeed, since $Y_{\varepsilon_n}(t) \ge Y_{\varepsilon_{n+1}}(t)$ for all $t \ge 0$, $n \ge 1$ and continuity of L would imply continuity of Y, Dini's theorem guarantees (I.17). The same argument applies to (I.18): the right-hand side of

$$\frac{1}{2}\int_0^t \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)}ds = Y_{\varepsilon_n}(t) - Y(0) + \frac{b}{2}\int_0^t Y_{\varepsilon_n}(s)ds - \frac{\sigma}{2}W(t)$$

is non-increasing w.r.t. t, therefore for each $t \ge 0$ and $n \ge 1$

$$\frac{1}{2}\int_0^t \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)}ds \ge \frac{1}{2}\int_0^t \frac{\varepsilon_{n+1}}{Y_{\varepsilon_{n+1}}(s)}ds$$

and Dini's theorem implies (I.18) as well.

By (I.19), continuity of X and the fact that $X(0) > 0$, there exists an interval $[0,t_0)$ such that for all $t \in [0,t_0)$ and $n \ge 1$ $Y_{\varepsilon_n}(t) \ge \frac{Y(0)}{2}$. Thus for any $t \in [0,t_0)$

$$L(t) = \lim_{n\to\infty}\frac{1}{2}\int_0^t \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)}ds \le \lim_{n\to\infty}\frac{t_0\varepsilon_n}{Y(0)} = 0,$$

i.e. $L(t) = 0$ for all $t \in [0,t_0]$.

For the reader's convenience, we will split the further proof into four steps.

Step 1: L in non-decreasing. Monotonicity of L is obvious since for any fixed $n \ge 1$ and $t_1 < t_2$

$$\int_0^{t_1} \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)}ds \le \int_0^{t_2} \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)}ds.$$

Step 2: right-continuity. Let us show that L is continuous from the right. For any fixed $t \ge 0$, denote $L(t+) := \lim_{\delta\downarrow 0} L(t+\delta)$ (the right limit exists since L is non-decreasing) and assume that $L(t+) - L(t) = \alpha > 0$. Due to the monotonicity of L, this implies that for all $\delta > 0$

$$L(t+\delta) - L(t) \geq \alpha > 0. \tag{I.20}$$

Now, take n_0 such that for all $n \geq n_0$

$$\frac{1}{2}\int_0^t \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)} ds \in \Big[L(t), L(t) + \frac{\alpha}{4}\Big)$$

and $\delta_0 > 0$ such that

$$\frac{1}{2}\int_0^{t+\delta_0} \frac{\varepsilon_{n_0}}{Y_{\varepsilon_{n_0}}(s)} ds \in \Big[L(t), L(t) + \frac{\alpha}{2}\Big).$$

As it was noted previously, for each $s \geq 0$ the values of $\frac{1}{2}\int_0^s \frac{\varepsilon_n}{Y_{\varepsilon_n}(u)} du$ are non-increasing when $n \to \infty$. Thus for any $n \geq n_0$

$$\frac{1}{2}\int_0^{t+\delta_0} \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)} ds \leq \frac{1}{2}\int_0^{t+\delta_0} \frac{\varepsilon_{n_0}}{Y_{\varepsilon_{n_0}}(s)} ds \leq L(t) + \frac{\alpha}{2},$$

i.e.

$$L(t+) \leq \lim_{n\to\infty} \frac{1}{2}\int_0^{t+\delta_0} \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)} ds < L(t) + \frac{\alpha}{2},$$

which contradicts (I.20). Therefore, $L(t+) - L(t) = 0$, i.e. L is right-continuous.

Step 3: left-continuity. Now, let us show that L is continuous from the left. Assume that it is not true and there exists $t > 0$ such that $L(t) - L(t-) > 0$ (note that $L(t-) = \lim_{\delta\downarrow 0} L(t-\delta)$ is well-defined due to the monotonicity of L). Since L may have only positive jumps, so does Y and, moreover, the points of jumps of L and Y coincide. This implies that $Y(t) - Y(t-) > 0$ and we now consider two cases.

Case 1: $Y(t-) = y > 0$. Then $Y(t) = Y(t+) > y$ (note that Y is right-continuous by Step 2) and there exists an interval $[t-\delta, t+\delta]$ such that $Y_{\varepsilon_n}(s) \geq Y(s) > \frac{y}{2}$ for all $s \in [t-\delta, t+\delta]$. This implies that

$$L(t+\delta) - L(t-\delta) = \lim_{n\to\infty} \frac{1}{2}\int_{t-\delta}^{t+\delta} \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)} ds \leq \lim_{n\to\infty} \frac{2\delta\varepsilon_n}{y} = 0,$$

i.e. L cannot have a jump at t. This means that Y cannot have a jump at point t either and we obtain a contradiction.

Case 2: $Y(t-) = 0$ *and* $Y(t+) = Y(t) = y > 0$. Fix $T > t$, $\lambda \in \left(0, \frac{1}{2}\right)$ and let Λ be a random variable such that for all $t_1, t_2 \in [0,T]$

$$|W(t_1) - W(t_2)| \leq \Lambda |t_1 - t_2|^\lambda.$$

Take $n_1 \geq 1$ and $\delta_1 > 0$ such that $\varepsilon_{n_1} + \delta_1 + \frac{\sigma\Lambda}{2}\delta_1^\lambda < y$ and note that there exists $\delta_2 < \delta_1$ such that $Y(t-\delta_2) < \varepsilon_{n_1}$. Since $Y_{\varepsilon_n}(t-\delta_2) \downarrow Y(t-\delta_2)$ as $n \to \infty$, there exists $n_2 > n_1$ such that $Y_{\varepsilon_{n_2}}(t-\delta_2) < \varepsilon_{n_1} < y$. Moreover, $Y_{\varepsilon_{n_2}}(t) \geq Y(t) = y$ thus one can define

$$\tau := \sup\{s \in (t-\delta_2, t),\ Y_{\varepsilon_{n_2}}(s) = \varepsilon_{n_1}\}.$$

Observe that $Y_{\varepsilon_{n_2}}(\tau) = \varepsilon_{n_1}$ and $Y_{\varepsilon_{n_2}}(s) \geq \varepsilon_{n_1}$ for all $s \in [\tau, t]$, whence

$$\begin{aligned} Y_{\varepsilon_{n_2}}(t) &= Y_{\varepsilon_{n_2}}(\tau) + \frac{1}{2}\int_\tau^t \frac{\varepsilon_{n_2}}{Y_{\varepsilon_{n_2}}(s)}ds - \frac{b}{2}\int_\tau^t Y_{\varepsilon_{n_2}}(s)ds + \frac{\sigma}{2}(W(t) - W(\tau)) \\ &\leq \varepsilon_{n_1} + \frac{\varepsilon_{n_2}}{2\varepsilon_{n_1}}(t-\tau) + \frac{\sigma\Lambda}{2}(t-\tau)^\lambda \\ &\leq \varepsilon_{n_1} + \delta_1 + \frac{\sigma\Lambda}{2}\delta_1^\lambda < y, \end{aligned}$$

which contradicts the assumption that $Y_{\varepsilon_{n_2}}(t) \geq y$. This contradiction together with all of the above implies that Y (and thus L) is continuous at each point $t \geq 0$.

Step 4: points of growth. Now, let us prove that the points of growth of L may occur only at zeros of Y. Indeed, let $t > 0$ be such that $Y(t) = y > 0$. Since Y is continuous, there exists $\delta_3 > 0$ such that for any $s \in (t-\delta_3, t+\delta_3)$

$$Y(s) > \frac{y}{2} > 0.$$

This, in turn, implies that for all $s \in (t-\delta_3, t+\delta_3)$ and $n \geq 1$

$$Y_{\varepsilon_n}(s) \geq Y(s) > \frac{y}{2} > 0$$

and thus for any $\delta \in [0, \delta_3)$

$$L(t+\delta) - L(t-\delta) = \lim_{n\to\infty} \frac{1}{2}\int_{t-\delta}^{t+\delta} \frac{\varepsilon_n}{Y_{\varepsilon_n}(s)}ds \leq \lim_{n\to\infty} \frac{2\delta}{y}\varepsilon_n = 0.$$

Therefore $L(t+\delta) - L(t-\delta) = 0$ and L does not grow in some neighbourhood of t. ■

Remark I.2.7. It is well-known (see e.g. [1, Appendix A] or [35, Subsection 3.3.1]) that the absolute value of OU and ROU processes with non-zero mean reversion levels do not coincide. In turn, in the "symmetric" case with zero mean reversion parameter, absolute value of the OU process and ROU process have the same distribution but do not coincide pathwisely. Theorem I.2.4 allows to clarify this subtle difference in the following manner.

Let $B = \{B(t),\ t \geq 0\}$ be some standard Brownian motion and

$$U(t) = U(0) - \frac{b}{2}\int_0^t U(s)ds + \frac{\sigma}{2}B(t), \quad t \geq 0,$$

be a standard Ornstein-Uhlenbeck process with non-random positive initial value $U(0) > 0$. By Itô's formula,

$$\begin{aligned} U^2(t) &= U^2(0) + \int_0^t \left(\frac{\sigma^2}{4} - bU^2(s) \right) ds + \sigma \int_0^t U(s) dB(s) \\ &= U^2(0) + \int_0^t \left(\frac{\sigma^2}{4} - bU^2(s) \right) ds + \sigma \int_0^t |U(s)| \operatorname{sign}(U(s)) dB(s) \\ &= U^2(0) + \int_0^t \left(\frac{\sigma^2}{4} - bU^2(s) \right) ds + \sigma \int_0^t |U(s)| dW(s), \end{aligned}$$

where $W(t) := \int_0^t \operatorname{sign}(U(s)) dB(s)$ is a standard Brownian motion (which can be easily verified by Levy's characterization). Thus, the process $X(t) := U^2(t)$, $t \geq 0$, is a CIR process w.r.t. W. By Theorem I.2.4, the square root process $Y(t) := \sqrt{X(t)}$, $t \geq 0$, is a reflected Ornstein-Uhlenbeck process with respect to W satisfying the SDE of the form

$$Y(t) = U(0) - \frac{b}{2} \int_0^t Y(s) ds + \frac{\sigma}{2} W(t) + L(t), \quad t \geq 0, \tag{I.21}$$

with L being the reflection function for Y. Since $Y(t) = \sqrt{X(t)} = |U(t)|$, by Tanaka's formula

$$\begin{aligned} Y(t) &= U(0) + \int_0^t \operatorname{sign}(U(s)) dU(s) + L_U(t) \\ &= U(0) - \frac{b}{2} \int_0^t \operatorname{sign}(U(s)) U(s) ds + \frac{\sigma}{2} \int_0^t \operatorname{sign}(U(s)) dB(s) + L_U(t) \quad \text{(I.22)} \\ &= U(0) - \frac{b}{2} \int_0^t Y(s) ds + \frac{\sigma}{2} W(t) + L_U(t), \end{aligned}$$

with L_U being the local time of U at zero. Comparing (I.21) and (I.22), we obtain that $L(t) = L_U(t)$, i.e. the reflection function of the ROU process Y coincides with local time at zero of the OU process U.

I.3 Fractional Cox-Ingersoll-Ross and fractional reflected Ornstein-Uhlenbeck processes

Let now $\{B^H(t),\ t \geq 0\}$ be a continuous modification of a fractional Brownian motion with Hurst index $H > \frac{1}{2}$. Consider a stochastic differential equation of the form

$$Y^H(t) = Y(0) + \frac{1}{2} \int_0^t \left(\frac{a}{Y^H(s)} - bY^H(s) \right) ds + \frac{\sigma}{2} dB^H(t), t \geq 0, \tag{I.23}$$

where $Y(0) > 0$ is a given constant, a, b, $\sigma > 0$. According to [22] (see also [7]), the SDE (I.23) a.s. has a unique pathwise solution $\{Y^H(t),\ t \geq 0\}$ such that

$Y^H(t) > 0$ for all $t \geq 0$, and the subset of Ω where this solution exists does not depend on $Y(0)$, a, b or σ (in fact, the solution exists for all $\omega \in \Omega$ such that $B^H(\omega, t)$ is locally Hölder continuous in t). Moreover, it can be shown (see [22, Theorem 1] or [7, Subsection 4.1]) that the process $X^H(t) = (Y^H(t))^2$, $t \geq 0$, satisfies the SDE of the form

$$X^H(t) = X(0) + \int_0^t (a - bX^H(s))ds + \sigma \int_0^t \sqrt{X^H(s)}dB^H(s), \quad t \geq 0, \quad \text{(I.24)}$$

where $X(0) = Y^2(0)$ and the integral with respect to the fractional Brownian motion exists as the pathwise limit of the corresponding Riemann-Stieltjes integral sums. Taking into account the form of (I.24), the process $\{X^H(t),\ t \geq 0\}$ can be interpreted as a natural fractional generalisation of the Cox-Ingersoll-Ross process with $\{Y^H(t),\ t \geq 0\}$ being its square root.

Remark I.3.1. It is evident that the solution to (I.24) is unique in the class of non-negative stochastic processes with paths that are Hölder-continuous up to the order H. Indeed, by the fractional pathwise counterpart of the Itô's formula (see e.g. [36, Theorem 4.3.1]) the square root of the solution must satisfy the equation (I.23) until the first moment of zero hitting. However, as it was noted above, the solution to (I.23) is unique and strictly positive a.s., i.e. never hits zero.

Now, let us recall the definition of the reflected fractional Ornstein-Uhlenbeck (RFOU) process.

Definition I.3.2. Stochastic process $\widetilde{Y}^H = \{\widetilde{Y}^H(t),\ t \geq 0\}$ is called a *fractional reflected Ornshein-Uhlenbeck (RFOU) process* if it satisfies a stochastic differential equation of the form

$$\widetilde{Y}^H(t) = Y(0) - \widetilde{b}\int_0^t \widetilde{Y}^H(s)ds + \widetilde{\sigma}B^H(t) + \widetilde{L}^H(t), \quad t \geq 0, \quad \text{(I.25)}$$

where $Y(0)$, $\widetilde{b}$ and $\widetilde{\sigma}$ are positive constants, $B^H = \{B^H(t),\ t \geq 0\}$ is a fractional Brownian motion, $\{\widetilde{L}^H(t),\ t \geq 0\}$ is a reflection function for $\widetilde{Y}^H$ in the sense of Definition I.2.1 and $\widetilde{Y}^H \geq 0$ a.s.

Remark I.3.3. For more details on properties of the RFOU process see e.g. [18] and references therein. Note that, by the argument similar to the one stated in Remark I.2.3, the solution (Y^H, L^H) to the equation (I.25) is unique.

When it comes to the connection between FCIR and RFOU processes, there is a notable difference from the standard Brownian case discussed in section I.2: in the standard case the ROU process turned out to coincide with the square root of the CIR process with $a = \frac{\sigma^2}{4}$ which is not true for the fractional case. More precisely, if $a > 0$, X^H is strictly positive a.s. and thus $\sqrt{X^H}$ cannot coincide with the RFOU process. Furthermore, for $a = 0$ [21, Theorem 6] claims existence and uniqueness of solution to (I.24) when $H \in \left(\frac{2}{3}, 1\right)$, and this solution turns out to stay in zero after hitting it, i.e. its square root is also different from the RFOU process. However, it is still possible to establish a clear connection between FCIR and RFOU processes highlighted in the next theorem.

Theorem I.3.4. *Let $\{B^H(t),\ t \geq 0\}$ be a continuous modification of a fractional Brownian motion with Hurst index $H > \frac{1}{2}$, $Y(0)$, b, $\sigma > 0$ be given constants. For any $\varepsilon > 0$, consider a square root process $Y_\varepsilon^H = \{Y_\varepsilon^H(t),\ t \geq 0\}$ given by*

$$Y_\varepsilon^H(t) = Y(0) + \frac{1}{2}\int_0^t \left(\frac{\varepsilon}{Y_\varepsilon^H(s)} - bY_\varepsilon^H(s)\right) ds + \frac{\sigma}{2}B^H(t), \quad t \geq 0.$$

Then, with probability 1,

1) *the limit $\lim_{\varepsilon\downarrow 0} Y_\varepsilon^H(t) =: Y^H(t)$ is well-defined, finite and non-negative for any $t \geq 0$;*

2) *the limit process $Y^H = \{Y^H(t),\ t \geq 0\}$ is a RFOU process satisfying the equation of the form*

$$Y^H(t) = Y(0) - \frac{b}{2}\int_0^t Y^H(s)ds + \frac{\sigma}{2}B^H(t) + L^H(t), \quad t \geq 0; \tag{I.26}$$

3) *for any $T > 0$*

$$\sup_{t\in[0,T]} |Y^H(t) - Y_\varepsilon^H(t)| \to 0, \quad \varepsilon \downarrow 0,$$

and

$$\sup_{t\in[0,T]} \left| L^H(t) - \frac{1}{2}\int_0^t \frac{\varepsilon}{Y_\varepsilon^H(s)}ds \right| \to 0, \quad \varepsilon \downarrow 0.$$

Proof. Let $\omega \in \Omega$ such that $B^H(\omega, t)$ is locally Hölder continuous in t be fixed (for notational simplicity, we will omit it in brackets). As it was noted above, for such ω all Y_ε^H are well-defined and strictly positive. Moreover, by the comparison theorem (see e.g. [22, Lemma 1] or [7, Lemma A.1]), for all $t \geq 0$ and $\varepsilon_1 > \varepsilon_2$

$$Y_{\varepsilon_1}^H(t) > Y_{\varepsilon_2}^H(t) > 0.$$

This implies that for any fixed $t \geq 0$ the limits $\lim_{\varepsilon\downarrow 0} Y_\varepsilon^H(t) = Y^H(t)$ and $\lim_{\varepsilon\downarrow 0} \frac{1}{2}\int_0^t \frac{\varepsilon}{Y_\varepsilon^H(s)}ds =: L^H(t)$ are well-defined, non-negative and finite. Furthermore, by comparison theorem, each Y_ε^H exceeds the fractional Ornstein-Uhlenbeck of the form

$$U^H(t) = Y(0) - \frac{b}{2}\int_0^t U^H(s)ds + \frac{\sigma}{2}B^H(t), \quad t \geq 0,$$

and hence there exists an interval $[0, t_0)$ such that for all $t \in [0, t_0)$ and $\varepsilon > 0$ it holds that $Y_\varepsilon^H(t) \geq \frac{Y(0)}{2}$. Thus for any $t \in [0, t_0)$

$$L^H(t) = \lim_{\varepsilon\downarrow 0} \frac{1}{2}\int_0^t \frac{\varepsilon}{Y_\varepsilon^H(s)}ds \leq \lim_{\varepsilon\downarrow 0} \frac{t_0\varepsilon}{Y(0)} = 0,$$

i.e. $L^H(t) = 0$ for all $t \in [0, t_0)$.

The remaining part of the proof is identical to the one of Theorem I.2.6. ■

Remark I.3.5. Theorem I.3.4 and the preceding remark highlight that the FCIR process (I.24) is not continuous at zero w.r.t. the mean-reversion parameter a.

I.4 Simulations

Let us illustrate the results with simulations. On Fig. I.1, the black paths depict simulated trajectories of the square root $\{Y_\varepsilon^H(t),\ t \ge 0\}$ of the FCIR process given by an equation of the form

$$Y_\varepsilon^H(t) = Y(0) + \frac{1}{2}\int_0^t \frac{\varepsilon}{Y_\varepsilon^H(s)}ds - \frac{b}{2}\int_0^t Y_\varepsilon^H(s)ds + \frac{\sigma}{2}B^H(t)$$

with $Y(0) = 0.25$, $b = 1$, $\sigma = 1$, $\varepsilon = 0.0001$ and different Hurst indices H; the red lines are the corresponding integrals $\frac{1}{2}\int_0^t \frac{\varepsilon}{Y_\varepsilon^H(s)}ds$. In order to simulate Y_ε^H, the backward Euler approximation technique from [17] was used, see also [12, 34].

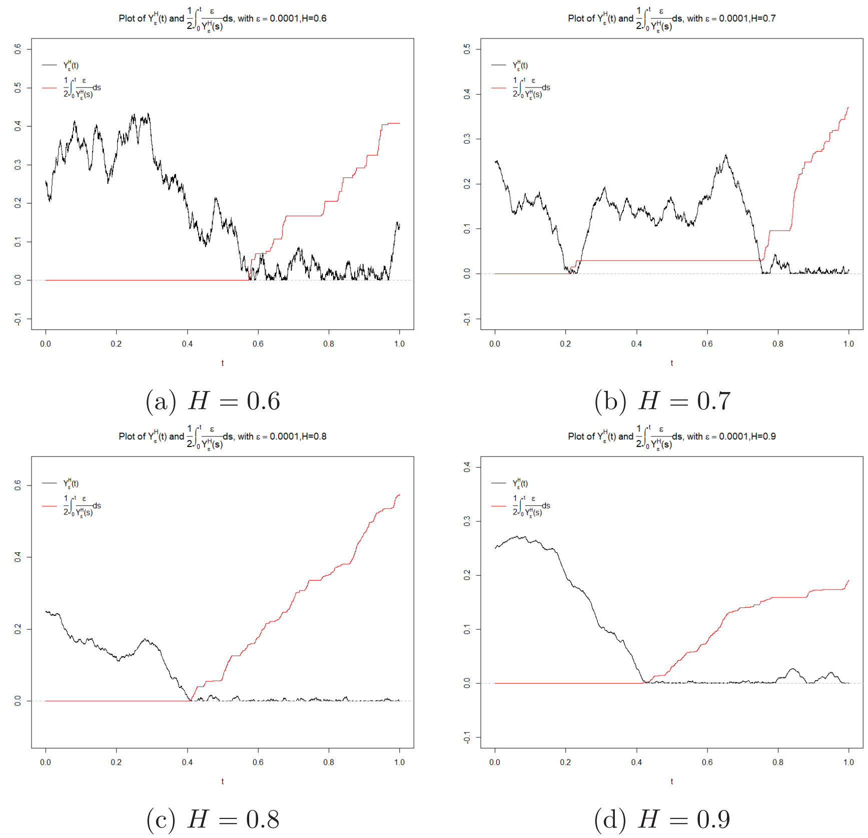
Theorem I.3.4 states that the red line approximates the reflection function L^H of the RFOU process and it can be clearly seen that the plot is well agreed with the theory: the integral $\frac{1}{2}\int_0^t \frac{\varepsilon}{Y_\varepsilon^H(s)}ds$ shows notable growth only when the corresponding path of Y_ε^H is very close to zero.

Fig. I.2 illustrates the uniform convergence of paths of Y_ε^H to the path of RFOU process as $\varepsilon \downarrow 0$. On the picture, $H = 0.6$, $Y(0) = 0.25$, $b = 1$, $\sigma = 1$ and the path of the FROU process Y^H was simulated using the Euler-type method:

$$Y^H(0) = Y(0),$$

$$Y^H(t_{n+1}) = \left(Y^H(t_n) - \frac{b}{2}Y^H(t_n)(t_{n+1} - t_n) + \frac{\sigma}{2}\left(B^H(t_{n+1}) - B^H(t_n)\right)\right) \vee 0.$$

When $\varepsilon = 0.0001$, the path of Y_ε^H (purple) is so close to the corresponding path of the ROU process (bold black) that they are not distinguishable on the plot.



(a) $H = 0.6$ (b) $H = 0.7$

(c) $H = 0.8$ (d) $H = 0.9$

Figure I.1: Sample paths of $Y_\varepsilon^H(t)$ (black) and $\frac{1}{2}\int_0^t \frac{\varepsilon}{Y_\varepsilon^H(s)}ds$ (red) for $\varepsilon = 0.0001$ and different Hurst indices H.

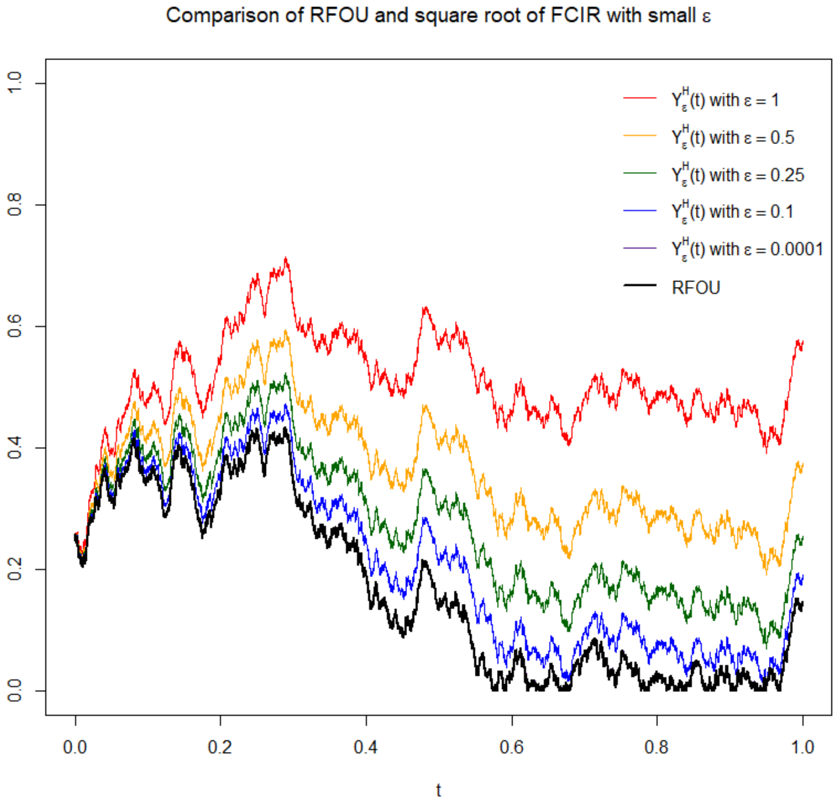


Figure I.2: Comparison of the Y_ε^H with $\varepsilon = 1$ (red), $\varepsilon = 0.5$ (orange), $\varepsilon = 0.25$ (green) $\varepsilon = 0.1$ (blue), $\varepsilon = 0.0001$ (purple) and the RFOU process (bold black). Note that the purple path ($\varepsilon = 0.0001$) is not visible on the plot since it almost completely coincides with the bold black trajectory of the RFOU process.

References

[1] Ball, C. A. and Roma, A. “Stochastic volatility option pricing”. In: *Journal of financial and quantitative analysis* vol. 29, no. 4 (1994), pp. 589–607.

[2] Bo, L., Wang, Y., and Yang, X. “Some integral functionals of reflected SDEs and their applications in finance”. In: *Quantitative finance* vol. 11, no. 3 (2011), pp. 343–348.

[3] Bo, L. et al. “On the conditional default probability in a regulated market: a structural approach”. In: *Quantitative finance* vol. 11, no. 12 (2011), pp. 1695–1702.

[4] Cox, J. C., Ingersoll, J. E., and Ross, S. A. “A re-examination of traditional hypotheses about the term structure of interest rates”. In: *The journal of finance* vol. 36, no. 4 (1981), pp. 769–799.

[5] Cox, J. C., Ingersoll, J. E., and Ross, S. A. “A theory of the term structure of interest rates”. In: *Econometrica: journal of the Econometric Society* vol. 53, no. 2 (1985), p. 385.

[6] Cox, J. C., Ingersoll, J. E., and Ross, S. A. “An intertemporal general equilibrium model of asset prices”. In: *Econometrica: journal of the Econometric Society* vol. 53, no. 2 (1985), p. 363.

[7] Di Nunno, G., Mishura, Y., and Yurchenko-Tytarenko, A. “Sandwiched SDEs with unbounded drift driven by Hölder noises”. In: *Advances in Applied Probability (to appear)* vol. 55, no. 3 (2023).

[8] Giorno, V., Nobile, A. G., and Cesare, R. di. “On the reflected Ornstein–Uhlenbeck process with catastrophes”. In: *Applied mathematics and computation* vol. 218, no. 23 (2012), pp. 11570–11582.

[9] Giorno, V., Nobile, A. G., and Ricciardi, L. M. “On some diffusion approximations to queueing systems”. In: *Advances in applied probability* vol. 18, no. 4 (1986), pp. 991–1014.

[10] Goldstein, R. S. and Keirstead, W. P. “On the term structure of interest rates in the presence of reflecting and absorbing boundaries”. In: *SSRN Electronic Journal* (1997).

[11] Heston, S. L. “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. In: *Review of Financial Studies* vol. 6, no. 2 (1993), pp. 327–43.

[12] Hong, J. et al. “Optimal strong convergence rate of a backward Euler type scheme for the Cox–Ingersoll–Ross model driven by fractional Brownian motion”. In: *Stochastic processes and their applications* vol. 130, no. 5 (2020), pp. 2675–2692.

[13] Ikeda, N. and Watanabe, S. “A comparison theorem for solutions of stochastic differential equations and its applications”. In: *Osaka Journal of Mathematics* vol. 14, no. 3 (1977), pp. 619–633.

[14] Ikeda, N. and Watanabe, S. *Stochastic Differential Equations and Diffusion Processes*. Elsevier, 1981.

[15] Karatzas, I. and Shreve, S. E. *Brownian Motion and Stochastic Calculus*. Springer New York, 1998.

[16] Krugman, P. R. "Target zones and exchange rate dynamics". In: *The Quarterly Journal of Economics* vol. 106, no. 3 (1991), pp. 669–682.

[17] Kubilius, K. "Estimation of the Hurst index of the solutions of fractional SDE with locally Lipschitz drift". In: *Lithuanian Association of Nonlinear Analysts (LANA). Nonlinear Analysis. Modelling and Control* vol. 25, no. 6 (2020), pp. 1059–1078.

[18] Lee, C. and Song, J. "On drift parameter estimation for reflected fractional Ornstein–Uhlenbeck processes". In: *Stochastics. An International Journal of Probability and Stochastic Processes* vol. 88, no. 5 (2016), pp. 751–778.

[19] Linetsky, V. "On the transition densities for reflected diffusions". In: *Advances in applied probability* vol. 37, no. 2 (2005), pp. 435–460.

[20] Maghsoodi, Y. "Solution of the extended CIR term structure and bond option valuation". In: *Mathematical Finance. An International Journal of Mathematics, Statistcs and Financial Economics* vol. 6, no. 1 (1996), pp. 89–109.

[21] Melnikov, A., Mishura, Y., and Shevchenko, G. "Stochastic viability and comparison theorems for mixed stochastic differential equations". In: *Methodology and computing in applied probability* vol. 17, no. 1 (2015), pp. 169–188.

[22] Mishura, Y. and Yurchenko-Tytarenko, A. "Fractional Cox–Ingersoll–Ross process with non-zero "mean"". In: *Modern Stochastics: Theory and Applications* vol. 5, no. 1 (2018), pp. 99–111.

[23] Piterbarg, V. I. *Twenty Lectures About Gaussian Processes*. Atlantic Financial Press, 2015.

[24] Ricciardi, L. M. "Stochastic population theory: Diffusion processes". In: *Mathematical Ecology*. Springer Berlin Heidelberg, 1986, pp. 191–238.

[25] Rogers, L. C. G. and Williams, D. *Diffusions, Markov Processes and Martingales*. Cambridge University Press, 2000.

[26] Schöbel, R. and Zhu, J. "Stochastic volatility with an Ornstein–Uhlenbeck process: An extension". In: *Review of finance* vol. 3, no. 1 (1999), pp. 23–46.

[27] Shiga, T. and Watanabe, S. "Bessel diffusions as a one-parameter family of diffusion processes". In: *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete* vol. 27, no. 1 (1973), pp. 37–46.

[28] Skorokhod, A. V. "Stochastic Equations for Diffusion Processes in a Bounded Region". In: *Theory of Probability & Its Applications* vol. 6, no. 3 (Jan. 1961), pp. 264–274.

[29] Ward, A. R. and Glynn, P. W. "A diffusion approximation for a GI/GI/1 queue with balking or reneging". In: *Queueing systems* vol. 50, no. 4 (2005), pp. 371–400.

[30] Ward, A. R. and Glynn, P. W. "A Diffusion Approximation for a Markovian Queue with Reneging". In: *Queueing systems* vol. 43 (2003), pp. 103–128.

[31] Ward, A. R. and Glynn, P. W. "Properties of the Reflected Ornstein–Uhlenbeck Process". In: *Queueing systems* vol. 44 (2003), pp. 109–123.

[32] Yang, X. et al. "Modeling the exchange rates in a target zone by reflected Ornstein-Uhlenbeck process". In: *SSRN Electronic Journal* (2012).

[33] Ylvisaker, D. "A Note on the Absence of Tangencies in Gaussian Sample Paths". In: *The Annals of Mathematical Statistics* vol. 39, no. 1 (1968), pp. 261–262.

[34] Zhang, S.-Q. and Yuan, C. "Stochastic differential equations driven by fractional Brownian motion with locally Lipschitz drift and their implicit Euler approximation". In: *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* (2020), pp. 1–27.

[35] Zhu, J. *Applications of Fourier transform to smile modeling*. Springer Berlin Heidelberg, 2010.

[36] Zähle, M. "Integration with respect to fractal functions and stochastic calculus. I". In: *Probability theory and related fields* vol. 111, no. 3 (1998), pp. 333–374.

[37] Aalen, O. O. and Gjessing, H. K. "Survival models based on the Ornstein-Uhlenbeck process". In: *Lifetime data analysis* vol. 10, no. 4 (2004), pp. 407–423.

Authors' addresses

Yuliya Mishura Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, 64/13, Volodymyrska St., Kyiv 01601, Ukraine, myus@univ.kiev.ua

Anton Yurchenko-Tytarenko Department of Mathematics, University of Oslo, Moltke Moes vei 35, 0851 Oslo, Norway, antony@math.uio.no

Paper III

Drift-implicit Euler scheme for sandwiched processes driven by Hölder noises

Giulia Di Nunno, Yuliya Mishura, Anton Yurchenko-Tytarenko

To appear in *Numerical Algorithms* (2022).
arXiv: 2204.08827.

III

Abstract

In this paper, we analyze the drift-implicit (or backward) Euler numerical scheme for a class of stochastic differential equations with unbounded drift driven by an arbitrary λ-Hölder continuous process, $\lambda \in (0,1)$. We prove that, under some mild moment assumptions on the Hölder constant of the noise, the $L^r(\Omega; L^\infty([0,T]))$-approximation error converges to 0 as $O(\Delta^\lambda)$, $\Delta \to 0$. To exemplify, we consider numerical schemes for the generalized Cox–Ingersoll–Ross and Tsallis–Stariolo–Borland models. The results are illustrated by simulations.

Contents

The present research is carried out within the frame and support of the ToppForsk project nr. 274410 of the Research Council of Norway with title STORM: Stochastics for Time-Space Risk Models. The second author was supported by Japan Science and Technology Agency CREST JPMJCR14D7, CREST Project reference number JPMJCR2115.

III.1 Introduction

We analyze the *drift-implicit* (also known as *backward*) Euler numerical scheme for stochastic differential equations (SDEs) of the form

$$Y(t) = Y(0) + \int_0^t b(s, Y(s))ds + Z(t), \quad t \in [0, T], \tag{III.1}$$

where Z is a general λ-Hölder continuous noise, $\lambda \in (0, 1)$, and the drift b is unbounded and has one of the following two properties:

(A) $b(t, y)$ has an explosive growth of the type $(y - \varphi(t))^{-\gamma}$ as $y \downarrow \varphi(t)$, where φ is a given Hölder continuous function of the same order λ as Z and $\gamma > \frac{1}{\lambda} - 1$;

(B) $b(t, y)$ has an explosive growth of the type $(y - \varphi(t))^{-\gamma}$ as $y \downarrow \varphi(t)$ *and* an explosive decrease of the type $-(\psi(t) - y)^{-\gamma}$ as $y \uparrow \psi(t)$, where φ and ψ are given Hölder continuous functions of the same order λ as Z such that $\varphi(t) < \psi(t)$, $t \in [0, T]$, and $\gamma > \frac{1}{\lambda} - 1$.

The SDEs of this type were extensively studied in [15]. It was shown that the properties **(A)** or **(B)**, along with some relatively weak additional assumptions, ensure that the solution to (III.1) is bounded from below (*one-sided sandwich case*) by the function φ in the setting **(A)**, i.e.

$$Y(t) > \varphi(t), \quad t \in [0, T], \tag{III.2}$$

or stays between φ and ψ (*two-sided sandwich case*) in the setting **(B)**, i.e.

$$\varphi(t) < Y(t) < \psi(t), \quad t \in [0, T]. \tag{III.3}$$

We emphasize that the SDE type (III.1) includes and generalizes several widespread stochastic models. For example, the process given by

$$Y(t) = Y(0) - \int_0^t \frac{\kappa Y(s)}{1 - Y^2(s)} ds + Z(t), \quad t \in [0, T],$$

where Z is λ-Hölder continuous with $\lambda > \frac{1}{2}$, fits into the setting **(B)** and can be regarded as a natural extension of the Tsallis–Stariolo–Borland (TSB) model employed in biophysics (for more details on the standard Brownian TSB model see e.g. [16, Subsection 2.3] or [13, Chapter 3 and Chapter 8]). Another important example is

$$Y(t) = Y(0) + \int_0^t \left(\frac{\kappa_1}{Y^\gamma(s)} - \kappa_2 Y(s) \right) ds + Z(t), \quad t \in [0, T], \tag{III.4}$$

where Z is λ-Hölder continuous, $\lambda \in (0, 1)$, and $\gamma > \frac{1}{\lambda} - 1$. It can be shown (see [15, Subsection 4.2]) that, if $\lambda > \frac{1}{2}$, stochastic process $X(t) := Y^{1+\gamma}(t)$ satisfies the SDE

$$X(t) = X(0) + (1+\gamma) \int_0^t (\kappa_1 - \kappa_2 X(s))\, ds + \int_0^t X^\alpha(s) dZ(s), \quad t \in [0, T], \tag{III.5}$$

where $\alpha := \frac{\gamma}{1+\gamma} \in (0,1)$ and the integral w.r.t. Z exists as a pathwise limit of Riemann-Stieltjes integral sums. Equations of the type (III.5) are used in finance in the standard Brownian setting and are called *Chan—Karolyi—Longstaff—Sanders* (CKLS) or *constant elasticity of variance* (CEV) model (see, e.g., [4, 8, 9]). If $\alpha = \frac{1}{2}$, the equation (III.5) is also known as the *Cox–Ingersoll–Ross* (CIR) equation, see , e.g., [10–12].

In this work, we develop a numerical approximation (both pathwise and in $L^r(\Omega; L^\infty([0,T])))$ for sandwiched processes (III.1) which is similar to the *drift-implicit* (also known as *backward*) Euler scheme constructed for the classical Cox-Ingersoll-Ross process in [2, 3, 14] and extended to the case of the fractional Brownian motion with $H > \frac{1}{2}$ in [18, 21, 22]. In this drift-implicit scheme, in order to generate $\widehat{Y}(t_{k+1})$, one has to solve the equation of the type

$$\widehat{Y}(t_{k+1}) = \widehat{Y}(t_k) + b(t_{k+1}, \widehat{Y}(t_{k+1}))\Delta_N + (Z(t_{k+1}) - Z(t_k)) \qquad \text{(III.6)}$$

with respect to $\widehat{Y}(t_{k+1})$ which is in general a more computationally heavy problem in comparison to the standard Euler-type techniques (see e.g. [15, Appendix A]). However, this drift-implicit numerical method also has a substantial advantage: the approximation $\widehat{Y}$ maintains the property of being sandwiched, i.e., for all points t_k of the partition

$$\widehat{Y}(t_k) > \varphi(t_k)$$

in the setting **(A)** and

$$\varphi(t_k) < \widehat{Y}(t_k) < \psi(t_k)$$

in the case **(B)**. Having this in mind, we shall say that the drift-implicit scheme is *sandwich preserving.*

We note that a similar approximation scheme was studied in [21] and [18, 22] for processes of the type (III.4) driven by a fractional Brownian motion with $H > 1/2$. Our work can be seen as an extension of those. However, we emphasize that our results have several elements of novelty. In particular, the paper [21] discusses only pathwise convergence and not convergence in $L^r(\Omega; L^\infty([0,T]))$. The approach of [18] and [22] is very noise specific as both use Malliavin calculus techniques in the spirit of [19, Proposition 3.4] to estimate inverse moments of the considered process (which turns out to be crucial to control explosive growth of the drift). As a result, two limitations appear: a restrictive condition involving the time horizon T (see e.g. [18, Eq. (8) and Remark 3.1]) and sensitivity to the choice of the noise, i.e. their method cannot be applied directly for drivers other than fBm with $H > 1/2$. This lack of flexibility in terms of the choice of the noise is a crucial disadvantage in e.g. finance where modern empirical studies justify the use of fBm with extremely low Hurst index ($H < 0.1$) [7] or even drivers with time-varying roughness [1]. Our approach makes use of [15, Theorem 3.2] based on the pathwise calculus and allows us to obtain strong convergence with no limitations on T for a substantially larger class of noises. In fact, we require only Hölder continuity of the noise and some moment condition on the corresponding Hölder coefficient which is often satisfied and shared by e.g. *all* Hölder continuous Gaussian processes.

The paper is organized as follows. Section III.2 describes the setting in detail and contains some necessary statements on the properties of the sandwiched processes. In Section III.3, we give the convergence results in the setting **(B)** which turns out to be a bit simpler than **(A)** due to boundedness of the process. Section III.4 extends the scheme to the setting **(A)**. In Section III.5, we give some examples and simulations; in particular we show that in some cases (e.g. for the generalized TSB and CIR models) equations (III.6) can be solved explicitly which drastically improves the computational efficiency of the algorithm.

III.2 Preliminaries and assumptions

Fix $T > 0$ and define

$$\begin{aligned} \mathcal{D}_{a_1} &:= \{(t,y) \in [0,T] \times \mathbb{R}_+,\ y \in (\varphi(t) + a_1, \infty)\}, \\ \mathcal{D}_{a_1,a_2} &:= \{(t,y) \in [0,T] \times \mathbb{R}_+,\ y \in (\varphi(t) + a_1, \psi(t) - a_2)\} \end{aligned} \tag{III.7}$$

with $a_1 \geq 0$ and $a_1, a_2 \in \left[0, \frac{1}{2}\|\psi - \varphi\|_\infty\right)$ respectively and φ, $\psi \in C([0,T])$ being such that $\varphi(t) < \psi(t)$, $t \in [0,T]$.
Throughout the paper, we will be dealing with a stochastic differential equation of the form

$$Y(t) = Y(0) + \int_0^t b(s, Y(s))ds + Z(t), \quad t \in [0,T]. \tag{III.8}$$

The noise $Z = \{Z(t),\ t \in [0,T]\}$ is always assumed to satisfy the following conditions:

(Z1) $Z(0) = 0$ a.s.;

(Z2) Z has a.s. λ-Hölder continuous paths for some $\lambda \in (0,1)$, i.e. there exists a positive random variable Λ such that

$$|Z(t) - Z(s)| \leq \Lambda |t-s|^\lambda, \quad s,t \in [0,T], \quad a.s.$$

Given the noise Z satisfying **(Z1)**–**(Z2)**, the initial value $Y(0)$ and the drift b satisfy *one of the two* assumptions given below.

Assumption III.2.1. (*One-sided sandwich case*) There exists a λ-Hölder continuous function φ: $[0,T] \to \mathbb{R}$ with λ being the same as in **(Z2)** such that

(A1) $Y(0)$ is deterministic and $Y(0) > \varphi(0)$,

(A2) b: $\mathcal{D}_0 \to \mathbb{R}$ is continuous and for any $\varepsilon \in (0,1)$

$$|b(t_1,y_1) - b(t_2,y_2)| \leq \frac{c_1}{\varepsilon^p}\left(|y_1 - y_2| + |t_1 - t_2|^\lambda\right), \quad (t_1,y_1),(t_2,y_2) \in \mathcal{D}_\varepsilon,$$

where $c_1 > 0$ and $p > 1$ are some given constants and λ is from **(Z2)**,

(A3)

$$b(t,y) \geq \frac{c_2}{(y-\varphi(t))^\gamma}, \quad (t,y) \in \mathcal{D}_0 \setminus \mathcal{D}_{y_*},$$

where y_*, $c_2 > 0$ are some given constants and $\gamma > \frac{1}{\lambda} - 1$ with λ being from **(Z2)**,

(A4) the partial derivative $\frac{\partial b}{\partial y}$ with respect to the spacial variable exists, is continuous and bounded from above, i.e.

$$\frac{\partial b}{\partial y}(t,y) < c_3, \quad (t,y) \in \mathcal{D}_0,$$

for some $c_3 > 0$.

Assumption III.2.2. (*Two-sided sandwich case*) There exist λ-Hölder continuous functions φ, ψ: $[0,T] \to \mathbb{R}$, $\varphi(t) < \psi(t)$, $t \in [0,T]$, with λ being the same as in **(Z2)** such that

(B1) $Y(0)$ is deterministic and $\varphi(0) < Y(0) < \psi(0)$,

(B2) b: $\mathcal{D}_{0,0} \to \mathbb{R}$ is continuous and for any $\varepsilon \in \left(0, \min\left\{1, \frac{1}{2}\|\psi - \varphi\|_\infty\right\}\right)$

$$|b(t_1,y_1) - b(t_2,y_2)| \leq \frac{c_1}{\varepsilon^p}\left(|y_1 - y_2| + |t_1 - t_2|^\lambda\right), \quad (t_1,y_1),(t_2,y_2) \in \mathcal{D}_{\varepsilon,\varepsilon},$$

where $c_1 > 0$ and $p > 1$ are some given constants and λ is from **(Z2)**,

(B3)

$$b(t,y) \geq \frac{c_2}{(y-\varphi(t))^\gamma}, \quad (t,y) \in \mathcal{D}_{0,0} \setminus \mathcal{D}_{y_*,0},$$

$$b(t,y) \leq -\frac{c_2}{(\psi(t)-y)^\gamma}, \quad (t,y) \in \mathcal{D}_{0,0} \setminus \mathcal{D}_{0,y_*},$$

where y_*, $c_2 > 0$ are some given constants and $\gamma > \frac{1}{\lambda} - 1$ with λ being from **(Z2)**,

(B4) the partial derivative $\frac{\partial b}{\partial y}$ with respect to the spacial variable exists, is continuous and bounded from above, i.e.

$$\frac{\partial b}{\partial y}(t,y) < c_3, \quad (t,y) \in \mathcal{D}_{0,0},$$

for some $c_3 > 0$.

Both Assumptions III.2.1 and III.2.2 along with **(Z1)**–**(Z2)** ensure that the SDE (III.8) has a unique solution. In the theorem below, we provide some relevant results related to sandwiched processes (see [15, Theorems 2.3, 2.5, 2.6, 3.1 and 3.2]).

Theorem III.2.3. *Let $Z = \{Z(t),\ t \in [0,T]\}$ be a stochastic process satisfying **(Z1)**–**(Z2)**.*

1) *If the initial value $Y(0)$ and the drift b satisfy assumptions **(A1)–(A3)**, then the SDE has a unique strong pathwise solution such that for all $t \in [0,T]$*

$$Y(t) > \varphi(t) \quad a.s. \tag{III.9}$$

Moreover, there exist deterministic constants L_1, L_2, L_3 and $L_4 > 0$ depending only on $Y(0)$, the shape of b and λ, such that for all $t \in [0,T]$ the estimate (III.9) *can be refined as follows:*

$$\varphi(t) + \frac{L_1}{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \le Y(t) \le L_3 + L_4\Lambda \quad a.s., \tag{III.10}$$

*where Λ is from **(Z2)** and γ is from **(A3)**. In particular, if Λ is such that*

$$\mathbb{E}\left[\Lambda^{\frac{r}{\gamma\lambda+\lambda-1}}\right] < \infty \tag{III.11}$$

for some $r > 0$, then

$$\mathbb{E}\left[\sup_{t\in[0,T]} \frac{1}{(Y(t)-\varphi(t))^r}\right] < \infty,$$

and, if

$$\mathbb{E}\Lambda^r < \infty \tag{III.12}$$

for some $r > 0$, then

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y(t)|^r\right] < \infty.$$

2) *If the initial value $Y(0)$ and the drift b satisfy assumptions **(B1)–(B3)**, then the SDE has a unique strong pathwise solution such that for all $t \in [0,T]$*

$$\varphi(t) < Y(t) < \psi(t) \quad a.s. \tag{III.13}$$

Moreover, there exist deterministic constants L_1 and $L_2 > 0$ depending only on $Y(0)$, the shape of b and λ, such that for all $t \in [0,T]$ the estimate (III.13) *can be refined as follows:*

$$\varphi(t) + \frac{L_1}{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \le Y(t) \le \psi(t) - \frac{L_1}{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \quad a.s., \tag{III.14}$$

*where Λ is from **(Z2)** and γ is from **(B3)**. In particular, if Λ can be chosen in such a way that*

$$\mathbb{E}\left[\Lambda^{\frac{r}{\gamma\lambda+\lambda-1}}\right] < \infty \tag{III.15}$$

for some $r > 0$, then

$$\mathbb{E}\left[\sup_{t\in[0,T]} \frac{1}{(Y(t)-\varphi(t))^r}\right] < \infty, \quad \mathbb{E}\left[\sup_{t\in[0,T]} \frac{1}{(\psi(t)-Y(t))^r}\right] < \infty.$$

Remark III.2.4. Properties (III.9)–(III.10) and (III.13)–(III.14) hold on each $\omega \in \Omega$ such that $Z(\omega; t)$, $t \in [0, T]$, is Hölder continuous and we always consider only such $\omega \in \Omega$ in all proofs with pathwise arguments. For notational simplicity, we will also omit ω in brackets.

Remark III.2.5. Due to the property (III.13), the setting described in Assumption III.2.2 will be referred to as the **two-sided sandwich case** since the solution is "sandwiched" between φ and ψ a.s. Similarly, the property (III.9) justifies the name **one-sided sandwich case** for the setting corresponding to Assumption III.2.1. In both cases III.2.1 and III.2.2, the solution to (III.8) will be referred to as a **sandwiched process**.

Remark III.2.6. Note that assumptions **(A4)** and **(B4)** are not required for Theorem III.2.3 to hold and will be used later on.

In what follows, conditions (III.11), (III.12) and (III.15) will play an important role since the $L^r(\Omega; L^\infty([0,T]))$-convergence of the approximation scheme will directly follow from the integrability of Λ. However it should be noted that these conditions are not very restricting as indicated in the following example.

Example III.2.7. *(Hölder Gaussian noises)* Let $Z = \{Z(t),\ t \in [0,T]\}$ be an arbitrary Hölder continuous Gaussian process satisfying **(Z1)**–**(Z2)**, e.g. standard or fractional Brownian motion. In this case, by [6], the random variable Λ from **(Z2)** can be chosen to have moments of *all* orders.

We now complete the Section with some examples of the sandwiched processes.

Example III.2.8. (*Generalized CIR and CKLS/CEV models*) Let $\varphi \equiv 0$, Z satisfy **(Z1)**–**(Z2)** with $\lambda \in (0,1)$ and $Y(0)$, κ_1, $\kappa_2 > 0$, $\gamma > \frac{1}{\lambda} - 1$ be given. Then, by Theorem III.2.3, 1), the SDE of the form

$$Y(t) = Y(0) + \int_0^t \left(\frac{\kappa_1}{Y^\gamma(s)} - \kappa_2 Y(s) \right) ds + Z(t), \tag{III.16}$$

$t \in [0,T]$, has a unique positive solution. Moreover, it can be shown (see [15, Subsection 4.2]) that, if $\lambda > \frac{1}{2}$, stochastic process $X(t) := Y^{1+\gamma}(t)$, $t \in [0,T]$, a.s. satisfies the SDE of the form

$$X(t) = X(0) + (1+\gamma) \int_0^t (\kappa_1 - \kappa_2 X(s))\, ds + (1+\gamma) \int_0^t X^\alpha(s) dZ(s), \tag{III.17}$$

$t \in [0,T]$, where $\alpha := \frac{\gamma}{1+\gamma} \in (0,1)$ and the integral w.r.t. Z exists a.s. as a pathwise limit of Riemann-Stieltjes integral sums. As mentioned already, the (III.17) appears in finance in the standard Brownian setting and is called *Chan–Karolyi–Longstaff–Sanders* (CKLS) or *constant elasticity of variance* (CEV) model (see e.g. [4, 8, 9]). If $\alpha = \frac{1}{2}$ (i.e. when $\gamma = 1$), the equation (III.17) is also known as the *Cox-Ingersoll-Ross* (CIR) equation [10–12].

Remark III.2.9. *(Connection with the classical Brownian CIR/CKLS models)*

1) If $\gamma = 1$ in (III.16) (CIR case), Assumption **(A3)** demands Z to be Hölder continuous of order $\lambda > \frac{1}{2}$. That means that Example III.2.8 does not cover the classical Brownian CIR model since the continuous modification of a standard Brownian motion has paths that are Hölder continuous only up to (but not including) the order 1/2. However, it is still possible to establish a clear connection between our setting and the classical CIR model. Indeed, let $\{W(t),\ t \in [0,T]\}$ be the continuous modification of a standard Brownian motion. Consider the CIR process $X = \{X(t),\ t \in [0,T]\}$ defined by

$$dX(t) = a(b - X(t))dt + \sigma\sqrt{X(t)}dW(t), \quad X_0 > 0,$$

where a, b, $\sigma > 0$ and $2ab > \sigma^2$. The latter condition ensures that X has positive paths a.s. and hence one can define $Y := \sqrt{X}$. By Itô's formula, Y satisfies the SDE

$$dY(t) = \left(\frac{\kappa_1}{Y(t)} - \kappa_2 Y(t)\right)dt + \frac{\sigma}{2}dW(t), \quad Y_0 = \sqrt{X_0} > 0, \qquad \text{(III.18)}$$

with $\kappa_1 := \frac{4ab-\sigma^2}{8}$ and $\kappa_2 := \frac{a}{2}$, which has a type very similar to (III.16). The SDE (III.18) can then be used to define a drift-implicit Euler scheme of the form (III.6) which turns out to converge to the original process (III.18). For more details on the drift-implicit Euler scheme for the classical Brownian CIR process, see e.g. [14].

2) If $\gamma > 1$ in (III.16), Assumptions **(Z1)**–**(Z2)** and **(A1)**–**(A4)** allow Z to be a standard Brownian motion. However, in this case one cannot use pathwise calculus to obtain (III.17) whereas the standard Itô's formula shows that $X := Y^{1+\gamma}$ does not coincide with the standard CKLS process. In order to cover the standard CKLS model, we have to modify the drift in (III.16) to compensate for the second order term in Itô's formula as follows:

$$dY(t) = \left(\frac{\kappa_1}{Y^\gamma(t)} - \frac{\gamma\sigma^2}{2Y(t)} - \kappa_2 Y(t)\right)dt + \sigma dW(t). \qquad \text{(III.19)}$$

The SDE (III.19) satisfies Assumption III.2.1 and $X := Y^{1+\gamma}$ is the solution to the SDE

$$X(t) = X(0) + (1+\gamma)\int_0^t (\kappa_1 - \kappa_2 X(s))\,ds + (1+\gamma)\sigma\int_0^t X^\alpha(s)dW(s), \quad ,$$

with $\alpha = \frac{\gamma}{1+\gamma}$, i.e. $X := Y^{1+\gamma}$ is the classical CKLS process.

Example III.2.10. (*Generalized TSB model*) Let $\varphi \equiv -1$, $\psi \equiv 1$, $Y(0) \in (-1,1)$, Z satisfy **(Z1)**–**(Z2)** with $\lambda > \frac{1}{2}$ and $\kappa > 0$. Then, by Theorem III.2.3, 2), the SDE of the form

$$Y(t) = Y(0) - \int_0^t \frac{\kappa Y(s)}{1 - Y^2(s)}ds + Z(t), \qquad \text{(III.20)}$$

$t \in [0,T]$, has a unique solution such that $-1 < Y(t) < 1$ for all $t \in [0,T]$ a.s. In the standard Brownian setting, the SDE of the type (III.20) is known as the Tsallis–Stariolo–Borland (TSB) model and is used in biophysics (for more details, see e.g. [16, Subsection 2.3] or [13, Chapter 3 and Chapter 8]).

Example III.2.11. For the given Z satisfying **(Z1)**–**(Z2)** with $\lambda \in (0,1)$, λ-Hölder continuous functions φ, ψ, $\varphi(t) < \psi(t)$, $t \in [0,T]$, and $Y(0) \in (\varphi(0), \psi(0))$ consider the SDE of the form

$$Y(t) = Y(0) + \int_0^t \left(\frac{\kappa_1}{(Y(s) - \varphi(s))^\gamma} - \frac{\kappa_2}{(\psi(s) - Y(s))^\gamma} - \kappa_3 Y(s) \right) ds + Z(t),$$

$t \in [0,T]$, where κ_1, $\kappa_2 > 0$, $\kappa_3 \in \mathbb{R}$ and $\gamma > \frac{1}{\lambda} - 1$. By Theorem III.2.3, 2), this SDE has a unique solution such that $\varphi(t) < Y(t) < \psi(t)$ a.s. Note that the TSB drift from (III.20) also has this shape with $\varphi \equiv -1$, $\psi \equiv 1$, $\gamma = 1$, $\kappa_1 = \kappa_2 = \frac{\kappa}{2}$ and $\kappa_3 = 0$ since

$$-\frac{\kappa y}{1 - y^2} = \frac{\kappa}{2} \left(\frac{1}{y+1} - \frac{1}{1-y} \right).$$

Notation III.2.12. In what follows, C denotes any positive deterministic constant that does not depend on the partition and the exact value of which is not relevant. Note that C may change from line to line (or even within one line).

III.3 The approximation scheme for the two-sided sandwich

We will start by considering the numerical scheme for the *two-sided sandwich case* which turns out to be slightly simpler due to boundedness of Y. Let the noise Z satisfy **(Z1)**–**(Z2)**, $Y(0)$ and b satisfy Assumption III.2.2 and $Y = \{Y(t),\ t \in [0,T]\}$ be the unique solution of the SDE (III.8). Consider a uniform partition $\{0 = t_0 < t_1 < ... < t_N = T\}$ of $[0,T]$, $t_k := \frac{Tk}{N}$, $k = 0, 1, ..., N$, with the mesh $\Delta_N := \frac{T}{N}$ such that

$$c_3 \Delta_N < 1, \tag{III.21}$$

where c_3 is an upper bound for $\frac{\partial b}{\partial y}$ from **(B4)**. Let us define $\widehat{Y}(t)$ as follows:

$$\begin{aligned} \widehat{Y}(0) &= Y(0), \\ \widehat{Y}(t_{k+1}) &= \widehat{Y}(t_k) + b(t_{k+1}, \widehat{Y}(t_{k+1}))\Delta_N + (Z(t_{k+1}) - Z(t_k)), \\ \widehat{Y}(t) &= \widehat{Y}(t_k), \quad t \in [t_k, t_{k+1}), \end{aligned} \tag{III.22}$$

where the second expression is considered as an equation with respect to $\widehat{Y}(t_{k+1})$.

Remark III.3.1. Equation with respect to $\widehat{Y}(t_{k+1})$ from (III.22) has a unique solution such that $\widehat{Y}(t_{k+1}) \in (\varphi(t_{k+1}), \psi(t_{k+1}))$. Indeed, for any fixed $t \in [0,T]$ and any $z \in \mathbb{R}$, consider the equation

$$y - b(t,y)\Delta_N = z \tag{III.23}$$

w.r.t. y. Assumption **(B4)** together with condition (III.21) imply that $(y - b(t,y)\Delta_N)'_y > 0$ and, by **(B3)**,

$$y - b(t,y)\Delta_N \to -\infty, \quad y \to \varphi(t)+,$$
$$y - b(t,y)\Delta_N \to \infty, \quad y \to \psi(t)-.$$

Thus there exists a unique $y \in (\varphi(t), \psi(t))$ satisfying (III.23).

Remark III.3.2. The value of $\widehat{Y}(t)$ for $t \in [0,T] \setminus \{t_0, ..., t_N\}$ can also be defined via linear interpolation for $t \in [t_k, t_{k+1})$, $k = 0, ..., N-1$, as

$$\widehat{Y}(t) = \frac{1}{\Delta_N}\left((t_{k+1} - t)\widehat{Y}(t_k) + (t - t_k)\widehat{Y}(t_{k+1})\right).$$

In such case all results of this section hold with almost no changes in the proofs.

Remark III.3.3. The algorithms of the type (III.22) are sometimes called the **drift-implicit** [2, 3, 14] or **backward** [18] Euler approximation schemes.

Before presenting the main results of this section, we require some auxiliary lemmas. First of all, we note that the values $\widehat{Y}(t_n)$, $n = 0, 1, ..., N$, of the discretized process are bounded away from both φ and ψ by random variables that do not depend on the partition. Namely, we have the following result that can be regarded as a discrete modification of arguments in [15, Theorem 3.2].

Lemma III.3.4. *Let Z satisfy **(Z1)**–**(Z2)**, Assumption III.2.2 hold and the mesh of the partition Δ_N satisfy* (III.21)*. Then there exist deterministic constants L_1 and $L_2 > 0$ depending only on $Y(0)$, the shape of the drift b and λ, such that with probability 1*

$$\varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}} \le \widehat{Y}(t_n) \le \psi(t_n) - \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}}, \quad n = 0, 1, ..., N,$$

*where Λ is from **(Z2)** and γ is from **(B3)**.*

Proof. We will prove that

$$\varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}} \le \widehat{Y}(t_n), \quad n = 0, 1, ..., N, \quad a.s. \tag{III.24}$$

by using the pathwise argument (see Remark III.2.4). The other inequality can be derived in a similar manner. Recall that, by Assumption III.2.2, φ and ψ are λ-Hölder continuous, i.e. there exists $K > 0$ such that

$$|\varphi(t) - \varphi(s)| + |\psi(t) - \psi(s)| \le K|t - s|^\lambda, \quad t, s \in [0, T].$$

Denote also

$$\beta := \frac{\lambda^{\frac{\lambda}{1-\lambda}} - \lambda^{\frac{1}{1-\lambda}}}{c_2^{\frac{\lambda}{1-\lambda}}} > 0,$$

where c_2 is from **(B3)**,

$$L_2 := K + (2\beta)^{\lambda-1}\left(\frac{(Y(0)-\varphi(0))\wedge y_*\wedge(\psi(0)-Y(0))}{2}\right)^{1-\lambda-\gamma\lambda} > 0,$$

with the constants y_* and γ also from **(B3)**, and

$$\varepsilon := \frac{1}{(2\beta)^{\frac{1-\lambda}{\gamma\lambda+\lambda-1}}(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}.$$

Note that, with probability 1,

$$|\varphi(t)-\varphi(s)| + |\psi(t)-\psi(s)| + |Z(t)-Z(s)| \leq (L_2+\Lambda)|t-s|^\lambda, \quad t,s\in[0,T],$$

and, furthermore, it is easy to check that $\varepsilon < Y(0)-\varphi(0)$, $\varepsilon < \psi(0)-Y(0)$ and $\varepsilon < y_*$.

If $\widehat{Y}(t_n) \geq \varphi(t_n)+\varepsilon$ for a particular $n = 0,1,...,N$, then, by definition of ε, the bound of the type (III.24) holds automatically. Suppose that there exists $n = 1,...,N$ such that $\widehat{Y}(t_n) < \varphi(t_n)+\varepsilon$. Denote by $\kappa(n)$ the last point of the partition before t_n on which $\widehat{Y}$ stays above ε, i.e.

$$\kappa(n) := \max\{k = 0,...,n-1 \mid \widehat{Y}_{t_k} \geq \varphi(t_k)+\varepsilon\}$$

(note that such point exists since $\widehat{Y}(t_0)-\varphi(0) = Y(0)-\varphi(0) > \varepsilon$). Then, for all $k = \kappa(n)+1,...,n$ we have that $\widehat{Y}(t_k) < \varepsilon < y_*$ and therefore, using **(B3)**, we obtain that, with probability 1,

$$\begin{aligned}
&\widehat{Y}(t_n)-\varphi(t_n)\\
&\quad= \widehat{Y}(t_{\kappa(n)})-\varphi(t_n)+\Delta_N\sum_{k=\kappa(n)+1}^{n} b(t_k,\widehat{Y}(t_k)) + Z(t_n)-Z(t_{\kappa(n)})\\
&\quad\geq \varepsilon+\varphi(t_{\kappa(n)})-\varphi(t_n)+\frac{c_2}{\varepsilon^\gamma}(t_n-t_{\kappa(n)})+Z(t_n)-Z(t_{\kappa(n)})\\
&\quad\geq \varepsilon+\frac{c_2}{\varepsilon^\gamma}(t_n-t_{\kappa(n)})-(L_2+\Lambda)(t_n-t_{\kappa(n)})^\lambda.
\end{aligned}$$

Consider a function $F_\varepsilon : \mathbb{R}_+ \to \mathbb{R}$ such that

$$F_\varepsilon(t) = \varepsilon + \frac{c_2}{\varepsilon^\gamma}t - (L_2+\Lambda)t^\lambda.$$

It is straightforward to verify that F_ε attains its minimum at

$$t_* := \left(\frac{\lambda}{c_2}\right)^{\frac{1}{1-\lambda}}\varepsilon^{\frac{\gamma}{1-\lambda}}(L_2+\Lambda)^{\frac{1}{1-\lambda}}$$

and, taking into account the explicit form of ε,

$$\begin{aligned} F_\varepsilon(t_*) &= \varepsilon + \frac{\lambda^{\frac{1}{1-\lambda}}}{c_2^{\frac{\lambda}{1-\lambda}}}\varepsilon^{\frac{\gamma\lambda}{1-\lambda}}(L_2+\Lambda)^{\frac{1}{1-\lambda}} - \frac{\lambda^{\frac{\lambda}{1-\lambda}}}{c_2^{\frac{\lambda}{1-\lambda}}}\varepsilon^{\frac{\gamma\lambda}{1-\lambda}}(L_2+\Lambda)^{\frac{1}{1-\lambda}} \\ &= \varepsilon - \beta\varepsilon^{\frac{\gamma\lambda}{1-\lambda}}(L_2+\Lambda)^{\frac{1}{1-\lambda}} = \frac{1}{2^{\frac{\gamma\lambda}{\gamma\lambda+\lambda-1}}\beta^{\frac{1-\lambda}{\gamma\lambda+\lambda-1}}(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \\ &= \frac{\varepsilon}{2}. \end{aligned}$$

Namely, even if $\widehat{Y}(t_n) < \varphi(t_n) + \varepsilon$, we still have that, with probability 1,

$$\widehat{Y}(t_n) - \varphi(t_n) \ge F_\varepsilon(t_n - t_{\kappa(n)}) \ge F_\varepsilon(t_*) = \frac{\varepsilon}{2},$$

and thus, with probability 1, for any $n = 0, 1, ..., N$

$$\begin{aligned} \widehat{Y}(t_n) \ge \varphi(t_n) + \frac{\varepsilon}{2} &= \varphi(t_n) + \frac{1}{2^{\frac{\gamma\lambda}{\gamma\lambda+\lambda-1}}\beta^{\frac{1-\lambda}{\gamma\lambda+\lambda-1}}(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \\ &=: \varphi(t_n) + \frac{L_1}{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}, \end{aligned}$$

where $L_1 := \frac{1}{2^{\frac{\gamma\lambda}{\gamma\lambda+\lambda-1}}\beta^{\frac{1-\lambda}{\gamma\lambda+\lambda-1}}}$. ■

Remark III.3.5. It is clear that constants L_1 and L_2 in Lemma III.3.4 can be chosen jointly for Y and $\widehat{Y}$, so that the inequalities

$$\varphi(t) + \frac{L_1}{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \le Y(t) \le \psi(t) - \frac{L_1}{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}, \quad t \in [0,T],$$

and

$$\varphi(t_n) + \frac{L_1}{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \le \widehat{Y}(t_n) \le \psi(t_n) - \frac{L_1}{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}, \quad n = 0, 1, ..., N,$$

hold simultaneously with probability 1.

Next, we proceed with a simple property of the sandwiched process Y in (III.8).

Lemma III.3.6. *Let Z satisfy **(Z1)**–**(Z2)** and assumptions **(B1)**–**(B3)** hold.*

1) *There exists a positive random variable Υ such that, with probability 1,*

$$|Y(t) - Y(s)| \le \Upsilon|t-s|^\lambda, \quad t, s \in [0,T].$$

2) *If, for some $r \ge 1$,*

$$\mathbb{E}\left[\Lambda^{\frac{r\max\{p,\gamma\lambda+\lambda-1\}}{\gamma\lambda+\lambda-1}}\right] < \infty, \tag{III.25}$$

*where λ and Λ are from **(Z2)**, p is from **(B2)** and γ is from **(B3)**, then one can choose Υ such that*

$$\mathbb{E}[\Upsilon^r] < \infty.$$

Proof. Denote $\phi(t) := \frac{1}{2}(\psi(t) + \varphi(t))$, $t \in [0, T]$. By (III.14),

$$\varphi(t) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \leq Y(t) \leq \psi(t) - \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}, \quad t \in [0, T], \quad a.s.,$$

i.e. with probability 1 $(t, Y(t)) \in \mathcal{D}_{\frac{1}{\xi}, \frac{1}{\xi}}$, $t \in [0, T]$, where

$$\xi := \frac{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}{L_1} \tag{III.26}$$

and $\mathcal{D}_{\frac{1}{\xi}, \frac{1}{\xi}}$ is defined by (III.7). It is evident that $(t, \phi(t)) \in \mathcal{D}_{\frac{1}{\xi}, \frac{1}{\xi}}$, $t \in [0, T]$, therefore, using **(Z2)**, **(B2)** and (III.13), we can write that, with probability 1, for all $0 \leq s < t \leq T$:

$$\begin{aligned} |Y(t) - Y(s)| &\leq \int_s^t |b(u, Y(u))| du + |Z(t) - Z(s)| \\ &\leq \int_s^t |b(u, Y(u)) - b(u, \phi(u))| du + \int_s^t |b(u, \phi(u))| du + \Lambda(t - s)^\lambda \\ &\leq c_1 \xi^p \int_s^t |Y(u) - \phi(u)| du + \max_{u \in [0,T]} |b(u, \phi(u))|(t - s) + \Lambda(t - s)^\lambda \\ &\leq \left(c_1 \xi^p \|\psi - \varphi\|_\infty + \max_{u \in [0,T]} |b(u, \phi(u))| \right) (t - s) + \Lambda(t - s)^\lambda \\ &\leq C(\xi^p + \Lambda + 1)(t - s)^\lambda, \end{aligned} \tag{III.27}$$

where C is a positive constant. Now one can put

$$\Upsilon := C(\xi^p + \Lambda + 1) \tag{III.28}$$

and observe that the definition of Υ, (III.25) and (III.26) imply that

$$\mathbb{E}[\Upsilon^r] < \infty.$$

■

Next, using Lemma III.3.4 and following the proof of Lemma III.3.6, it is easy to obtain the following result.

Corollary III.3.7. *Let **(Z1)**–**(Z2)** and Assumption III.2.2 hold. Then there exists a random variable Υ independent of the partition such that with probability 1*

$$|\widehat{Y}(t_k) - \widehat{Y}(t_n)| \leq \Upsilon |t_k - t_n|^\lambda, \quad k, n = 0, ..., N. \tag{III.29}$$

Furthermore, if (III.25) *holds for some for $r \geq 1$, then*

$$\mathbb{E}[\Upsilon^r] < \infty.$$

Finally, Υ can be chosen jointly for Y and $\widehat{Y}$, so that

$$|Y(t) - Y(s)| \leq \Upsilon |t - s|^\lambda, \quad t, s \in [0, T],$$

holds simultaneously with (III.29) *with probability 1.*

Lemma III.3.8. *Let Z satisfy **(Z1)**–**(Z2)**, Assumption III.2.2 hold and the mesh of the partition Δ_N satisfy* (III.21)*. Then*

1) *for any $r \geq 1$, there exists a positive random variable $\mathcal{C}_1$ that does not depend on the partition such that*

$$\sup_{k=0,1,\dots,N} |Y(t_k) - \widehat{Y}(t_k)|^r \leq \mathcal{C}_1 \Delta_N^{\lambda r} \quad a.s.;$$

2) *if, additionally,*

$$\mathbb{E}\left[\Lambda^{\frac{r(p+\max\{p,\gamma\lambda+\lambda-1\})}{\gamma\lambda+\lambda-1}}\right] < \infty, \tag{III.30}$$

*where λ and Λ are from **(Z2)**, p is from **(B2)** and γ is from **(B3)**, then one can choose $\mathcal{C}_1$ such that $\mathbb{E}[\mathcal{C}_1] < \infty$, i.e. there exists a deterministic constant C that does not depend on the partition such that*

$$\mathbb{E}\left[\sup_{k=0,1,\dots,N} |Y(t_k) - \widehat{Y}(t_k)|^r\right] \leq C\Delta_N^{\lambda r}.$$

Proof. Fix $\omega \in \Omega$ such that $Z(\omega, t)$, $t \in [0,T]$, is Hölder continuous (for simplicity of notation, we will omit ω in the brackets). Denote $e_n := Y(t_n) - \widehat{Y}(t_n)$, $\Delta Z_n := Z(t_n) - Z(t_{n-1})$. Then

$$\begin{aligned} e_n &= Y(t_{n-1}) + \int_{t_{n-1}}^{t_n} b(s, Y(s))ds + \Delta Z_n \\ &\quad - \widehat{Y}(t_{n-1}) - b(t_n, \widehat{Y}(t_n))\Delta_N - \Delta Z_n \\ &= e_{n-1} + \left(b(t_n, Y(t_n)) - b(t_n, \widehat{Y}(t_n))\right)\Delta_N \\ &\quad + \int_{t_{n-1}}^{t_n} (b(s, Y(s)) - b(t_n, Y(t_n)))ds. \end{aligned} \tag{III.31}$$

By the mean value theorem,

$$\begin{aligned} \left(b(t_n, Y(t_n)) - b(t_n, \widehat{Y}(t_n))\right)\Delta_N &= \frac{\partial b}{\partial y}(t_n, \Theta_n)\Delta_N(Y(t_n) - \widehat{Y}(t_n)) \\ &= \frac{\partial b}{\partial y}(t_n, \Theta_n)\Delta_N e_n \end{aligned}$$

with $\Theta_n \in (Y(t_n) \wedge \widehat{Y}(t_n), Y(t_n) \vee \widehat{Y}(t_n))$. Using this, we can rewrite (III.31) as follows:

$$\left(1 - \frac{\partial b}{\partial y}(t_n, \Theta_n)\Delta_N\right) e_n = e_{n-1} + \int_{t_{n-1}}^{t_n} (b(s, Y(s)) - b(t_n, Y(t_n)))ds, \tag{III.32}$$

where

$$1 - \frac{\partial b}{\partial y}(t_n, \Theta_n)\Delta_N > 1 - c_3\Delta_N > 0$$

by **(B4)** and (III.21).

Next, denote

$$\zeta_0 := 1, \quad \zeta_n := \prod_{i=1}^{n} \left(1 - \frac{\partial b}{\partial y}(t_i, \Theta_i)\Delta_N\right)$$

and define $\tilde{e}_n := \zeta_n e_n$. By multiplying both sides of (III.32) by ζ_{n-1}, we obtain that

$$\tilde{e}_n = \tilde{e}_{n-1} + \zeta_{n-1} \int_{t_{n-1}}^{t_n} (b(s, Y(s)) - b(t_n, Y(t_n)))ds \tag{III.33}$$

and, expanding the terms $\tilde{e}_{i-1}$ in (III.33) one by one, $i = n, n-1, ..., 1$, and taking into account that $\tilde{e}_0 = 0$, we obtain that

$$\tilde{e}_n = \sum_{i=1}^{n} \zeta_{i-1} \int_{t_{i-1}}^{t_i} (b(s, Y(s)) - b(t_i, Y(t_i)))ds.$$

Therefore

$$e_n = \sum_{i=1}^{n} \frac{\zeta_{i-1}}{\zeta_n} \int_{t_{i-1}}^{t_i} (b(s, Y(s)) - b(t_i, Y(t_i)))ds.$$

Observe that, by assumption **(B4)** and (III.21), for any $i, n \in \mathbb{N}$, $i < n$,

$$\begin{aligned}
\frac{\zeta_k}{\zeta_n} &= \prod_{i=k+1}^{n} \left(1 - \frac{\partial b}{\partial y}(t_i, \Theta_i)\Delta_N\right)^{-1} \le \prod_{i=k+1}^{N} (1 - c_3\Delta_N)^{-1} \\
&\le (1 - c_3\Delta_N)^{-N} = \left(1 - \frac{c_3 T}{N}\right)^{-N} \to e^{c_3 T}, \quad N \to \infty,
\end{aligned}$$

whence there exists a constant C that does not depend on i, n or N such that

$$\frac{\zeta_k}{\zeta_n} \le C.$$

Using this, one can deduce that

$$\begin{aligned}
|e_n|^r &\le C \left| \sum_{i=1}^{n} \frac{\zeta_{i-1}}{\zeta_n} \int_{t_{i-1}}^{t_i} (b(s, Y(s)) - b(t_i, Y(t_i)))ds \right|^r \\
&\le C \left(\sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} |b(s, Y(s)) - b(t_i, Y(t_i))| \, ds \right)^r.
\end{aligned}$$

Note that $(t, Y(t)) \in \mathcal{D}_{\frac{1}{\xi},\frac{1}{\xi}}$, where ξ is defined by (III.26) and $\mathcal{D}_{\frac{1}{\xi},\frac{1}{\xi}}$ is defined via (III.7), hence, by **(B2)** as well as Lemma III.3.6, we can deduce that

$$\begin{aligned}
&\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |b(s,Y(s)) - b(t_i,Y(t_i))|\,ds\right)^r \\
&\le C\xi^{pr}\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |s-t_i|^\lambda\,ds\right)^r + C\xi^{pr}\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |Y(s)-Y(t_i)|\,ds\right)^r \\
&\le C\xi^{pr}\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |s-t_i|^\lambda\,ds\right)^r + C\xi^{pr}\Upsilon^r\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |s-t_i|^\lambda\,ds\right)^r \\
&= C\xi^{pr}(1+\Upsilon^r)\left(\sum_{i=1}^{n}\frac{1}{(1+\lambda)}\Delta_N^{1+\lambda}\right)^r \\
&\le C\xi^{pr}(1+\Upsilon^r)\Delta_N^{\lambda r}.
\end{aligned}$$

In other words, there exists a constant C that does not depend on the partition such that

$$|e_n|^r = |Y(t_n) - \widehat{Y}(t_n)|^r \le C\xi^{pr}(1+\Upsilon^r)\Delta_N^{\lambda r}$$

and, since the right-hand side of the relation above does not depend on n or N, we have

$$\sup_{n=0,\dots,N} |Y(t_n) - \widehat{Y}(t_n)|^r \le C\xi^{pr}(1+\Upsilon^r)\Delta_N^{\lambda r} =: \mathcal{C}_1\Delta_N^{\lambda r}. \tag{III.34}$$

It remains to notice that, by (III.26) and (III.28),

$$\mathbb{E}\left[\xi^{pr}(1+\Upsilon^r)\right] < \infty$$

whenever (III.30) holds, which finally implies

$$\mathbb{E}\left[\sup_{n=0,\dots,N} |Y(t_n) - \widehat{Y}(t_n)|^r\right] \le \mathbb{E}[\mathcal{C}_1]\Delta_N^{\lambda r} =: C\Delta_N^{\lambda r}.$$

■

Now we are ready to proceed to the main results of this subsection.

Theorem III.3.9. *Let Z satisfy **(Z1)**–**(Z2)**, Assumption III.2.2 hold and the mesh of the partition Δ_N satisfy* (III.21). *Then*

1) *for any $r \ge 1$, there exists a random variable $\mathcal{C}_2$ that does not depend on the partition such that*

$$\sup_{t\in[0,T]} |Y(t) - \widehat{Y}(t)|^r \le \mathcal{C}_2\Delta_N^{\lambda r} \quad a.s.;$$

2) *if, additionally,*

$$\mathbb{E}\left[\Lambda^{\frac{r(p+\max\{p,\gamma\lambda+\lambda-1\})}{\gamma\lambda+\lambda-1}}\right] < \infty,$$

where λ *and* Λ *are from* ***(Z2)****,* p *is from* ***(B2)*** *and* γ *is from* ***(B3)****, then one can choose* $\mathcal{C}_2$ *such that* $\mathbb{E}[\mathcal{C}_2] < \infty$*, i.e. there exists a deterministic constant* C *that does not depend on the partition such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y(t) - \widehat{Y}(t)|^r\right] \le C\Delta_N^{\lambda r}.$$

Proof. Fix $\omega \in \Omega$ such that $Z(\omega, t)$, $t \in [0, T]$, is Hölder continuous (for simplicity of notation, we again omit ω in the brackets) and consider an arbitrary $t \in [0, T]$. Denote

$$n(t) := \max\{n = 0, 1, ..., N \mid t \ge t_n\},$$

i.e. $t \in [t_{n(t)}, t_{n(t)+1})$. Then

$$\begin{aligned} |Y(t) - \widehat{Y}(t)|^r &\le C\left(|Y(t) - Y(t_{n(t)})|^r + |Y(t_{n(t)}) - \widehat{Y}(t_{n(t)})|^r\right) \\ &\le C\Upsilon^r (t - t_{n(t)})^{\lambda r} + C(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda+\lambda-1}}(1 + \Upsilon^r)\Delta_N^{\lambda r} \\ &\le C\left(\Upsilon^r + (1 + \Upsilon^r)(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda+\lambda-1}}\right)\Delta_N^{\lambda r}, \end{aligned}$$

where we used Lemma III.3.6 to estimate $|Y(t) - Y(t_{n(t)})|^r$ and bound (III.34) to estimate $|Y(t_{n(t)}) - \widehat{Y}(t_{n(t)})|^r$. Therefore

$$\sup_{t\in[0,T]} |Y(t) - \widehat{Y}(t)|^r \le C\left(\Upsilon^r + (1 + \Upsilon^r)(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda+\lambda-1}}\right)\Delta_N^{\lambda r} =: \mathcal{C}_2\Delta_N^{\lambda r}.$$

Finally, using the same arguments as in Lemma III.3.6 and Lemma III.3.8, one can easily show that the condition

$$\mathbb{E}\left[\Lambda^{\frac{r(p+\max\{p,\gamma\lambda+\lambda-1\})}{\gamma\lambda+\lambda-1}}\right] < \infty$$

implies that

$$\mathbb{E}\left[\Upsilon^r + (1 + \Upsilon^r)(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda+\lambda-1}}\right] < \infty,$$

therefore

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y(t) - \widehat{Y}(t)|^r\right] \le C\left(\Upsilon^r + (1 + \Upsilon^r)(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda+\lambda-1}}\right)\Delta_N^{\lambda r}$$

for some constant $C > 0$ that does not depend on the partition. ■

Theorem III.3.10.

1) *Let Z satisfy **(Z1)**–**(Z2)**, Assumption III.2.2 hold and the mesh of the partition Δ_N satisfy* (III.21). *Then, for any $r \ge 1$, there exists a random variable $\mathcal{C}_3$ that does not depend on the partition such that*

$$\sup_{n=0,1,\dots,N} \left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r \le \mathcal{C}_3 \Delta_N^{\lambda r} \quad a.s.$$

and

$$\sup_{n=0,1,\dots,N} \left| \frac{1}{\psi(t_n) - Y(t_n)} - \frac{1}{\psi(t_n) - \widehat{Y}(t_n)} \right|^r \le \mathcal{C}_3 \Delta_N^{\lambda r} \quad a.s.$$

2) *If, additionally,*

$$\mathbb{E}\left[\Lambda^{\frac{r(2+p+\max\{p, \gamma\lambda+\lambda-1\})}{\gamma\lambda+\lambda-1}}\right] < \infty, \tag{III.35}$$

*where λ and Λ are from **(Z2)**, p is from **(B2)** and γ is from **(B3)**, then one can choose $\mathcal{C}_3$ such that $\mathbb{E}[\mathcal{C}_3] < \infty$, i.e. there exists a deterministic constant C that does not depend on the partition such that*

$$\mathbb{E}\left[\sup_{n=0,1,\dots,N} \left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r\right] \le C \Delta_N^{\lambda r}$$

and

$$\mathbb{E}\left[\sup_{n=0,1,\dots,N} \left| \frac{1}{\psi(t_n) - Y(t_n)} - \frac{1}{\psi(t_n) - \widehat{Y}(t_n)} \right|^r\right] \le C \Delta_N^{\lambda r}.$$

Proof. By Remark III.3.5 and estimate (III.34), with probability 1 for any $n = 0, \dots, N$:

$$\begin{aligned}
&\left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r \\
&\quad = \frac{|Y(t_n) - \widehat{Y}(t_n)|^r}{(Y_{t_n} - \varphi(t_n))^r (\widehat{Y}_{t_n} - \varphi(t_n))^r} \\
&\quad \le \frac{(L_2 + \Lambda)^{\frac{2r}{\gamma\lambda+\lambda-1}}}{L_1^{2r}} \sup_{n=0,1,\dots,N} |Y(t_n) - \widehat{Y}(t_n)|^r \\
&\quad \le C (L_2 + \Lambda)^{\frac{2r}{\gamma\lambda+\lambda-1}} \xi^{pr} (1 + \Upsilon^r) \Delta_N^{\lambda r} \\
&\quad =: \mathcal{C}_3 \Delta_N^{\lambda r}.
\end{aligned}$$

It remains to notice that, by (III.26) and (III.28), the condition (III.35) implies that $\mathbb{E}[\mathcal{C}_3] < \infty$. The second estimate can be obtained in a similar manner. ■

III.4 One-sided sandwich case

The drift-implicit Euler approximation scheme described in Section III.3 for the two-sided sandwich can also be adapted for the one-sided setting that corresponds to Assumption III.2.1 on the SDE (III.1). However, in the two-sided sandwich case the process Y was bounded (which was utilized, e.g., in Lemma III.3.6) and, moreover, the behaviour of Y was similar near both φ and ψ so that it was sufficient to analyze only one of the bounds. In the one-sided case, each $Y(t)$, for $t \in [0,T]$, is not a bounded random variable, therefore the approach from Section III.3 has to be adjusted. For this, we will be using the inequalities (III.10).

Let the noise Z satisfy **(Z1)**–**(Z2)**, $Y(0)$ and b satisfy Assumption III.2.1 and $Y = \{Y(t),\ t \in [0,T]\}$ be the unique solution of the SDE (III.8). In line with Section III.3, we consider a uniform partition $\{0 = t_0 < t_1 < ... < t_N = T\}$ of $[0,T]$, $t_k := \frac{Tk}{N}$, $k = 0, 1, ..., N$, with the mesh $\Delta_N := \frac{T}{N}$ such that

$$c_3 \Delta_N < 1, \tag{III.36}$$

where c_3 is an upper bound for $\frac{\partial b}{\partial y}$ from assumption **(A4)**. The backward Euler approximation $\widehat{Y}(t)$ is defined in a manner similar to (III.22), i.e.

$$\begin{aligned} \widehat{Y}(0) &= Y(0), \\ \widehat{Y}(t_{k+1}) &= \widehat{Y}(t_k) + b(t_{k+1}, \widehat{Y}(t_{k+1}))\Delta_N + (Z(t_{k+1}) - Z(t_k)), \\ \widehat{Y}(t) &= \widehat{Y}(t_k), \quad t \in [t_k, t_{k+1}), \end{aligned} \tag{III.37}$$

where the second expression is considered as an equation with respect to $\widehat{Y}(t_{k+1})$.

Remark III.4.1. Just as in the two-sided sandwich case, each $\widehat{Y}(t_k)$, $k = 1, ..., N$, is well defined since the equation

$$y - b(t,y)\Delta_N = z$$

has a unique solution w.r.t. y such that $y > \varphi(t)$ for any fixed $t \in [0,T]$ and any $z \in \mathbb{R}$. To understand this, note that assumption **(A4)** together with (III.36) imply that

$$(y - b(t,y)\Delta_N)'_y > 0. \tag{III.38}$$

Second, by **(A3)**,

$$y - b(t,y)\Delta_N \to -\infty, \quad y \to \varphi(t)+. \tag{III.39}$$

Next, by **(A2)**, for any (s, y_1), $(s, y_2) \in \overline{\mathcal{D}_1} := \{(u,y) \in [0,T] \times \mathbb{R}_+, y \in [\varphi(u)+1, \infty)\}$ we have that

$$|b(s,y_1) - b(s,y_2)| \le c_1 |y_1 - y_2|,$$

i.e.

$$\sup_{(s,y)\in\overline{\mathcal{D}_1}} \left| \frac{\partial b}{\partial y}(s,y) \right| < \infty.$$

Using this, **(A4)** and the mean value theorem, for any positive $y \geq \varphi(t) + 1$

$$\begin{aligned} b(t,y) &= b(t,\varphi(t)+1) + \frac{\partial b}{\partial y}(t,\theta_y)(y-1-\varphi(t)) \\ &\leq \max_{s\in[0,T]} b(t,\varphi(t)+1) + \max_{s\in[0,T]} |1+\varphi(s)| \sup_{(s,y)\in\overline{\mathcal{D}_1}} \left|\frac{\partial b}{\partial y}(s,y)\right| + c_3 y \\ &=: C + c_3 y, \end{aligned}$$

whence

$$y - b(t,y)\Delta_N \geq -C\Delta_N + (1 - c_3\Delta_N)y \to \infty, \quad y \to \infty. \tag{III.40}$$

Existence and uniqueness of the solution then follows from (III.38)–(III.40).

Remark III.4.2. Similarly to the two-sided sandwich case, the value of $\widehat{Y}(t)$ for $t \in [0,T] \setminus \{t_0, ..., t_N\}$ can also be defined via linear interpolation with no changes in formulations of the results and almost no variations in the proofs.

Our strategy for proving the convergence of $\widehat{Y}$ to Y will be similar to what we have done in section III.3. Therefore we will be omitting the details highlighting only the points which are different from the two-sided sandwich case. We start with some useful properties of $\widehat{Y}$ and Y.

Lemma III.4.3. *Let Z satisfy **(Z1)**–**(Z2)**, Assumption III.2.1 hold and the mesh of the partition Δ_N satisfy* (III.36). *Then there exist deterministic constants L_1, $L_2 > 0$ depending only on $Y(0)$, the shape of the drift b and λ, such that*

$$\widehat{Y}(t_n) \geq \varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \quad a.s.,$$

*where Λ is from assumption **(Z2)** and γ is from assumption **(A3)**. Moreover, there exist constants L_3, $L_4 > 0$ that also depend only on $Y(0)$, the shape of the drift b and λ such that*

$$\widehat{Y}(t_n) \leq L_3 + L_4\Lambda, \quad n = 0,1,...,N, \quad a.s.$$

*for all partitions with the mesh satisfying $\frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N < 1$ with c_1 and p being from **(A2)**.*

Proof. The proof of

$$\widehat{Y}(t_n) \geq \varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}$$

is identical to the corresponding one in Lemma III.3.4 and will be omitted. Let us prove that

$$\widehat{Y}(t_n) \leq L_3 + L_4\Lambda \quad a.s.$$

Fix $\omega \in \Omega$ for which $Z(\omega,t)$ is Hölder continuous, consider a partition with the mesh satisfying $\frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N < 1$ and fix an arbitrary $n = 0,1,...,N-1$.

Assume that $\widehat{Y}(t_{n+1}) > \varphi(t_{n+1}) + (Y(0) - \varphi(0))$ (otherwise the claim of the lemma holds automatically). Put

$$\kappa(n) := \max\{k = 0, 1, ..., n \mid \widehat{Y}(t_k) \le \varphi(t_k) + (Y(0) - \varphi(0))\}$$

and observe that $(t_k, \widehat{Y}(t_k)) \in \mathcal{D}_{Y(0)-\varphi(0)}$ for any $k = \kappa(n) + 1, ..., n + 1$, where $\mathcal{D}_{Y(0)-\varphi(0)}$ is defined via (III.7). Next, by **(A2)**, for any $y \in \mathcal{D}_{Y(0)-\varphi(0)}$

$$\begin{aligned}
|b(t, y)| - \Big|b\Big(t, \varphi(t) + (Y(0) - \varphi(0))\Big)\Big| & \\
\le \Big|b(t, y) - b\Big(t, \varphi(t) + (Y(0) - \varphi(0))\Big)\Big| & \\
\le \frac{c_1}{(Y(0) - \varphi(0))^p}|y - \varphi(t) - (\varphi(0) - Y(0))| & \\
\le \frac{c_1}{(Y(0) - \varphi(0))^p}|y| + \frac{c_1}{(Y(0) - \varphi(0))^p}|\varphi(t) + (\varphi(0) - Y(0))|, &
\end{aligned}$$

i.e. there exists a constant $C > 0$ that does not depend on the partition such that

$$|b(t, y)| \le C + \frac{c_1}{(Y(0) - \varphi(0))^p}|y|. \tag{III.41}$$

Next, observe that, for any $k = \kappa(n) + 1, ..., n + 1$, we have

$$\begin{aligned}
\widehat{Y}(t_k) &= \widehat{Y}(t_{\kappa(n)}) + \sum_{i=\kappa(n)+1}^{k} b(t_i, \widehat{Y}(t_i))\Delta_N + Z(t_k) - Z(t_{\kappa(n)}) \\
&\le \varphi(t_{\kappa(n)}) + (Y(0) - \varphi(0)) + \sum_{i=\kappa(n)+1}^{k} b(t_i, \widehat{Y}(t_i))\Delta_N + \Lambda(t_k - t_{\kappa(n)})^\lambda \\
&\le \left|\max_{s\in[0,T]} \varphi(s) + (Y(0) - \varphi(0))\right| + T^\lambda\Lambda + \sum_{i=\kappa(n)+1}^{k} b(t_i, \widehat{Y}(t_i))\Delta_N.
\end{aligned}$$

Therefore, using (III.41) and

$$\widehat{Y}(t_k) > \varphi(t_k) \ge \min_{s\in[0,T]} \varphi(s),$$

one can write

$$\begin{aligned}
|\widehat{Y}(t_k)| &\le \left|\min_{s\in[0,T]} \varphi(s)\right| + \left|\max_{s\in[0,T]} \varphi(s) + (Y(0) - \varphi(0))\right| \\
&\quad + T^\lambda\Lambda + \sum_{i=\kappa(n)+1}^{k} |b(t_i, \widehat{Y}(t_i))|\Delta_N \\
&\le \left|\min_{s\in[0,T]} \varphi(s)\right| + \left|\max_{s\in[0,T]} \varphi(s) + (Y(0) - \varphi(0))\right| \\
&\quad + T^\lambda\Lambda + C\sum_{i=\kappa(n)+1}^{k} \Delta_N + \frac{c_1}{(Y(0) - \varphi(0))^p}\sum_{i=\kappa(n)+1}^{k} |\widehat{Y}(t_i)|\Delta_N,
\end{aligned}$$

where $C > 0$ is some positive constant that does not depend on the partition.

Now we want to apply the discrete version of the Gronwall inequality from [20, Lemma A.3]. In order to do that, we observe that

$$|\widehat{Y}(t_{\kappa(n)+1})| \le C + T^\lambda \Lambda + \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N |\widehat{Y}(t_k)|,$$

and, for any $k = \kappa(n)+2, ..., n+1$,

$$|\widehat{Y}(t_k)| \le C + T^\lambda \Lambda + \frac{c_1}{(Y(0)-\varphi(0))^p} \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)|\Delta_N + \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N |\widehat{Y}(t_k)|.$$

Now, since $\frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N < 1$, we can write that

$$\left(1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N\right)|\widehat{Y}(t_{\kappa(n)+1})| \le C + T^\lambda \Lambda$$

and, for all $k = \kappa(n)+2, ..., n+1$,

$$\left(1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N\right)|\widehat{Y}(t_k)| \le C + T^\lambda \Lambda + \frac{c_1}{(Y(0)-\varphi(0))^p} \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)|\Delta_N.$$

Put

$$N_0 := \min\left\{N \ge 1 : \ \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N < 1\right\} = \left[\frac{Tc_1}{(Y(0)-\varphi(0))^p}\right] + 1$$

with $[x]$ being the greatest integer less than or equal to x and observe that, for all $N \ge N_0$,

$$1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N \ge 1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_{N_0}.$$

Therefore,

$$\begin{aligned} |\widehat{Y}(t_{\kappa(n)+1})| &\le \frac{C}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N} + \frac{T^\lambda}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N}\Lambda \\ &\le \frac{C}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_{N_0}} + \frac{T^\lambda}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_{N_0}}\Lambda \\ &=: C_1 + C_2\Lambda \end{aligned}$$

and, for all $k = \kappa(n)+2, ..., n+1$,

$$|\widehat{Y}(t_k)| \le \frac{C}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N} + \frac{T^\lambda}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N}\Lambda$$

$$
\begin{aligned}
&\quad + \frac{c_1}{(Y(0)-\varphi(0))^p} \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)| \frac{\Delta_N}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_N} \\
&\le \frac{C}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_{N_0}} + \frac{T^\lambda}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_{N_0}}\Lambda \\
&\quad + \frac{c_1}{(Y(0)-\varphi(0))^p} \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)| \frac{\Delta_N}{1 - \frac{c_1}{(Y(0)-\varphi(0))^p}\Delta_{N_0}} \\
&=: C_1 + C_2\Lambda + C_3 \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)|\Delta_N.
\end{aligned}
$$

Using a discrete version of the Gronwall inequality, we now obtain that for all $k = \kappa(n)+1, ..., n+1$

$$
\begin{aligned}
|\widehat{Y}(t_k)| &\le (C_1 + C_2\Lambda) \exp\left\{ C_3 \sum_{i=\kappa(n)+1}^{k-1} \Delta_N \right\} \le (C_1 + C_2\Lambda)\exp\{TC_3\} \\
&=: L_3 + L_4\Lambda.
\end{aligned}
$$

which ends the proof. ■

Remark III.4.4. It is clear that constants L_1, L_2, L_3 and L_4 can be chosen jointly for Y and $\widehat{Y}$, so that the inequalities

$$
\varphi(t) + \frac{L_1}{(L_2+\Lambda)^\alpha} \le Y(t) \le L_3 + L_4\Lambda, \quad t \in [0,T],
$$

and

$$
\varphi(t_n) + \frac{L_1}{(L_2+\Lambda)^\alpha} \le \widehat{Y}(t_n) \le L_3 + L_4\Lambda, \quad n = 0, 1, ..., N,
$$

hold simultaneously with probability 1.

Next, corresponding to Lemma III.3.6 in the two-sided case, Y enjoys Hölder continuity with the Hölder constant being integrable provided that Λ has moments of sufficiently high order. This is summarized in the lemma below.

Lemma III.4.5. *Let Z satisfy **(Z1)**–**(Z2)** and assumptions **(A1)**–**(A3)** hold.*

1) *There exists a positive random variable Υ such that with probability 1*

$$
|Y(t) - Y(s)| \le \Upsilon |t-s|^\lambda, \quad t, s \in [0,T].
$$

2) *If, for some $r \ge 1$,*

$$
\mathbb{E}\left[\Lambda^{\frac{r(p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}}\right] < \infty, \tag{III.42}
$$

*where λ and Λ are from **(Z2)**, p is from **(A2)** and γ is from **(A3)**, then one can choose Υ such that*

$$
\mathbb{E}[\Upsilon^r] < \infty.
$$

Proof. By (III.10),

$$Y(t) \geq \varphi(t) + \frac{L_1}{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \quad a.s.,$$

i.e. with probability 1 $(t, Y(t)) \in \mathcal{D}_{\frac{1}{\xi}}$, $t \in [0,T]$, where

$$\xi := \frac{(L_2+\Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}{L_1} \tag{III.43}$$

and $\mathcal{D}_{\frac{1}{\xi}}$ is defined in (III.7). Denote $\phi(t) := \varphi(t)+1$ and notice that $(t, \phi(t)) \in \mathcal{D}_{\frac{1}{\xi}}$, $t \in [0,T]$, since $\frac{1}{\xi} \leq Y(0) - \varphi(0)$. Thus, using the same arguments as applied in (III.27), we can write that, with probability 1, for any $0 \leq s < t \leq T$:

$$|Y(t)-Y(s)| \leq c_1\xi^p \int_s^t |Y(u)-\phi(u)|du + \max_{u\in[0,T]} |b(u,\phi(u))|(t-s) + \Lambda(t-s)^\lambda,$$

where c_1 is from **(A2)**. Now, again by (III.10),

$$Y(t) \leq L_3 + L_4\Lambda \quad a.s.,$$

hence with probability 1

$$\begin{aligned}
|Y(t)-Y(s)| &\leq c_1\xi^p \int_s^t |Y(u)-\phi(u)|du + \max_{u\in[0,T]} |b(u,\phi(u))|(t-s) + \Lambda(t-s)^\lambda \\
&\leq c_1\xi^p(L_3+L_4\Lambda)(t-s) + c_1\xi^p \max_{u\in[0,T]} |\phi(u)|(t-s) \\
&\quad + \max_{u\in[0,T]} |b(u,\phi(u))|(t-s) + \Lambda(t-s)^\lambda \\
&\leq C(1+\xi^p\Lambda+\xi^p+\Lambda)(t-s)^\lambda,
\end{aligned}$$

where C is a positive constant. Now one can put

$$\Upsilon := C(1+\xi^p\Lambda+\xi^p+\Lambda) \tag{III.44}$$

and observe that

$$\mathbb{E}[\Upsilon^r] < \infty$$

whenever (III.42) holds. ∎

Corollary III.4.6. *Using Lemma III.4.3 and following the proof of Lemma III.4.5, it is easy to obtain that, for any partition with the mesh satisfying*

$$\max\left\{c_3, \frac{c_1}{(Y(0)-\varphi(0))^p}\right\}\Delta_N < 1 \tag{III.45}$$

there is a random variable Υ independent of the partition such that with probability 1

$$|\widehat{Y}(t_k) - \widehat{Y}(t_n)| \leq \Upsilon|t_k-t_n|^\lambda, \quad k,n = 0,...,N. \tag{III.46}$$

Furthermore, just like in Lemma III.3.6, for $r > 0$

$$\mathbb{E}[\Upsilon^r] < \infty$$

provided that

$$\mathbb{E}\left[\Lambda^{\frac{r(p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}}\right] < \infty.$$

Finally, such Υ *can be chosen jointly for* Y *and* $\widehat{Y}$*, so that*

$$|Y(t) - Y(s)| \le \Upsilon|t-s|^\lambda, \quad t, s \in [0,T],$$

holds simultaneously with (III.46) *with probability 1.*

Lemma III.4.7. *Let* Z *satisfy* ***(Z1)****–****(Z2)****, Assumption III.2.1 hold and the mesh of the partition* Δ_N *satisfy* (III.36)*.*

1) *For any* $r \ge 1$*, there exists a positive random variable* $\mathcal{C}_4$ *that does not depend on the partition such that*

$$\sup_{k=0,1,\dots,N} |Y(t_k) - \widehat{Y}(t_k)|^r \le \mathcal{C}_4\Delta_N^{\lambda r} \quad a.s.$$

2) *If, additionally,*

$$\mathbb{E}\left[\Lambda^{\frac{r(2p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}}\right] < \infty, \tag{III.47}$$

where λ *and* Λ *are from* ***(Z2)****,* p *is from* ***(A2)*** *and* γ *is from* ***(A3)****, then one can choose* $\mathcal{C}_4$ *such that* $\mathbb{E}[\mathcal{C}_4] < \infty$*, i.e. there exists a deterministic constant* C *that does not depend on the partition such that*

$$\mathbb{E}\left[\sup_{k=0,1,\dots,N} |Y(t_k) - \widehat{Y}(t_k)|^r\right] \le C\Delta_N^{\lambda r}.$$

Proof. Following the proof of Lemma III.3.8, one can easily obtain that for any $n = 0, 1, ..., N$

$$|Y(t_n) - \widehat{Y}(t_n)| \le C\left(\sum_{i=1}^n \int_{t_{i-1}}^{t_i} |b(s, Y(s)) - b(t_i, Y(t_i))|\, ds\right)^r.$$

Next, note that $(t, Y(t)) \in \mathcal{D}_{\frac{1}{\xi}}$, where ξ is defined by (III.43), so, by **(A2)** and

Lemma III.4.5,

$$
\begin{aligned}
&\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |b(s,Y(s)) - b(t_i, Y(t_i))|\, ds\right)^r \\
&\le C\xi^{pr}\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |s-t_i|^{\lambda}\, ds\right)^r + C\xi^{pr}\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |Y(s)-Y(t_i)|\, ds\right)^r \\
&\le C\xi^{pr}\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |s-t_i|^{\lambda}\, ds\right)^r + C\xi^{pr}\Upsilon^r\left(\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |s-t_i|^{\lambda}\, ds\right)^r \\
&= C\xi^{pr}(1+\Upsilon^r)\left(\sum_{i=1}^{n}\frac{1}{(1+\lambda)}\Delta_N^{1+\lambda}\right)^r \\
&\le C\xi^{pr}(1+\Upsilon^r)\Delta_N^{\lambda r},
\end{aligned}
$$

i.e.

$$
\sup_{n=0,\dots,N} |Y(t_n) - \widehat{Y}(t_n)|^r \le C\xi^{pr}(1+\Upsilon^r)\Delta_N^{\lambda r}. \tag{III.48}
$$

In order to conclude the proof, it remains to notice that (III.43), (III.44) and (III.47) imply that

$$
\mathbb{E}\left[\xi^{pr}(1+\Upsilon^r)\right] < \infty.
$$

■

Now we are ready to formulate the two main results of this section.

Theorem III.4.8. *Let Z satisfy **(Z1)**–**(Z2)**, Assumption III.2.1 hold and the mesh of the partition Δ_N satisfy* (III.45)*.*

1) *For any $r \ge 1$, there exists a random variable $\mathcal{C}_5$ that does not depend on the partition such that*

$$
\sup_{t\in[0,T]} |Y(t) - \widehat{Y}(t)|^r \le \mathcal{C}_5 \Delta_N^{\lambda r} \quad a.s.
$$

2) *If, additionally,*

$$
\mathbb{E}\left[\Lambda^{\frac{r(2p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}}\right] < \infty,
$$

*where λ and Λ are from **(Z2)**, p is from **(A2)** and γ is from **(A3)**, then one can choose $\mathcal{C}_5$ such that $\mathbb{E}[\mathcal{C}_5] < \infty$, i.e. there exists a deterministic constant C that does not depend on the partition such that*

$$
\mathbb{E}\left[\sup_{t\in[0,T]} |Y(t) - \widehat{Y}(t)|^r\right] \le C\Delta_N^{\lambda r}.
$$

Proof. The proof is similar to the one of Theorem III.3.9 but instead of Lemmas III.3.6, III.3.8 and bound (III.34) one should apply Lemmas III.4.5, III.4.7 and bound (III.48). ■

Theorem III.4.9. *Let Z satisfy **(Z1)**–**(Z2)**, Assumption III.2.1 hold and the mesh of the partition Δ_N satisfy* (III.45)*.*

1) *For any $r \geq 1$, there exists a random variable $\mathcal{C}_6$ that does not depend on the partition such that*

$$\sup_{n=0,1,\ldots,N} \left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r \leq \mathcal{C}_6 \Delta_N^{\lambda r} \quad a.s.$$

2) *If, additionally,*

$$\mathbb{E}\left[\Lambda^{\frac{r(2+2p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}}\right] < \infty, \tag{III.49}$$

*where λ and Λ are from **(Z2)**, p is from **(A2)** and γ is from **(A3)**, then one can choose $\mathcal{C}_6$ such that $\mathbb{E}[\mathcal{C}_6] < \infty$, i.e. there exists a deterministic constant C that does not depend on the partition such that*

$$\mathbb{E}\left[\sup_{n=0,1,\ldots,N} \left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r\right] \leq C \Delta_N^{\lambda r}.$$

Proof. The proof is similar to Theorem III.3.10 and is omitted. ■

III.5 Examples and simulations

The algorithms presented in (III.22) and (III.37) imply that, in order to generate $\widehat{Y}(t_{n+1})$, one has to solve an equation that potentially can be challenging from the computational point of view. However, in some cases that are relevant for applications this equation has a simple explicit solution.

Regarding the numerical examples that follow, we remark that:

1) all the simulations are performed in the R programming language on the system with Intel Core i9-9900K CPU and 64 Gb RAM;

2) in order to simulate paths of fractional Brownian motion, R package somebm is used;

3) in Example III.5.3, discrete samples of the multifractional Brownian motion (mBm) values are simulated using the Cholesky decomposition of the corresponding covariance matrix (for covariance structure of the mBm, see e.g. [5, Proposition 4]) and the R package nleqslv is used for solving (III.22) numerically.

Example III.5.1. (*Generalized CIR processes*) Let $\varphi \equiv 0$, Z satisfy **(Z1)**–**(Z2)** with $\lambda \in \left(\frac{1}{2}, 1\right)$, $Y(0)$, κ_1, $\kappa_2 > 0$, $\gamma > \frac{1}{\lambda} - 1$ be given and $\{Y(t),\ t \in [0,T]\}$ satisfy the SDE of the form

$$Y(t) = Y(0) + \int_0^t \left(\frac{\kappa_1}{Y(s)} - \kappa_2 Y(s) \right) ds + Z(t), \quad t \in [0,T]. \tag{III.50}$$

This process fits into the framework of Section III.4 and the equation for $\hat{Y}(t_{k+1})$ from (III.37) reads as follows:

$$\widehat{Y}(t_{k+1}) = \widehat{Y}(t_k) + \left(\frac{\kappa_1}{\widehat{Y}(t_{k+1})} - \kappa_2 \widehat{Y}(t_{k+1}) \right) \Delta_N + Z(t_{k+1}) - Z(t_k).$$

It is easy to see that it has a unique positive solution

$$\widehat{Y}(t_{k+1}) = \frac{\widehat{Y}(t_k) + (Z(t_{k+1}) - Z(t_k))}{2(1+\kappa_2\Delta_N)} + \frac{\sqrt{\left(\widehat{Y}(t_k) + (Z(t_{k+1}) - Z(t_k))\right)^2 + 4\kappa_1\Delta_N(1+\kappa_2\Delta_N)}}{2(1+\kappa_2\Delta_N)}.$$

Fig. III.1 contains 10 sample paths of the process (III.50) driven by a fractional Brownian motion with $H = 0.7$. In all simulation we take $N = 10000$, $T = 1$ and $Y(0) = 1 = \kappa_1 = \kappa_2 = 1$. In order to illustrate the convergence,

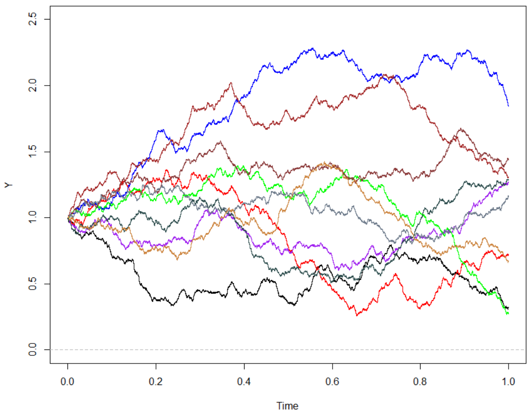
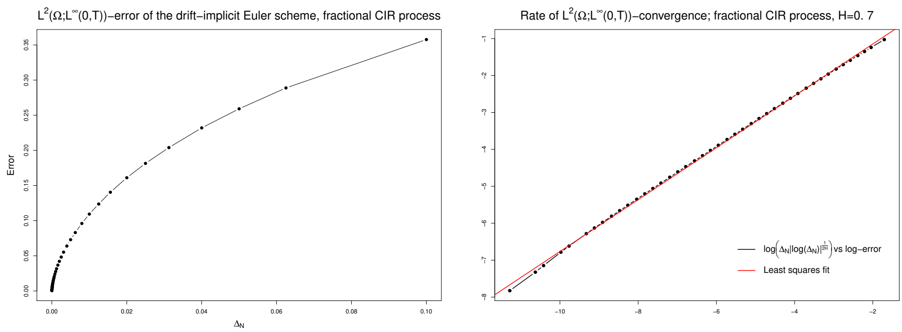


Figure III.1: Ten sample paths of (III.50) generated using the drift-implicit Euler approximation scheme; $N = 10000$, $T = 1$, $Y(0) = \kappa_1 = \kappa_2 = 1$, Z is a fractional Brownian motion with $H = 0.7$.

we also simulate the drift-implicit approximation $\widehat{Y}$ with a small step size 10^{-6} (it will serve as the "exact" solution). Then, using the same path of Z, we generate the drift-implicit Euler approximations with step sizes of the form $1/N$, where N runs over all divisors of 10^6. Afterwards, we compute the $L^\infty([0,T])$-distances between the "exact" solution and its approximations with larger step sizes. This procedure is performed 10000 times and the mean square error of each $L^\infty([0,T])$-distance is computed. The resulting values serve as consistent estimators of the corresponding $L^2(\Omega; L^\infty([0,T]))$-errors and are depicted on Figure III.2(a).



(a) (b)

Figure III.2: Convergence analysis of the drift-implicit Euler approximation scheme for (III.50); $T = 1$, $Y(0) = \kappa_1 = \kappa_2 = 1$, Z is a fractional Brownian motion with $H = 0.7$. On panel (a), $L^2(\Omega; L^\infty([0,T]))$-errors are depicted. Panel (b) contains the values of $\log\left(\Delta_N \left|\log(\Delta_N)\right|^{\frac{1}{2H}}\right)$ plotted against the logarithms of the corresponding $L^2(\Omega; L^\infty([0,T]))$-errors (black) as well as the line fitted with the least squares method (red). The slope of the red line is $0.7022687 \approx 0.7 = H$.

Note that the drift-implicit Euler scheme for (III.50) driven by the fractional Brownian motion was the main subject of [18] and [22], but in both cases the convergence of $\widehat{Y}$ to Y is established only on $[0,T]$ with T being small (see e.g. [18, Eq. (8) and Remark 3.1]). Our results fill this gap and prove that convergence holds on arbitrary $[0,T]$ for any model parameters. It should be noted though that the convergence rate in Theorem III.4.8 is not optimal and can be improved for the fractional Brownian driver. It is well-known that paths of a fractional Brownian motion are Hölder continuous up to (but not including) its Hurst index H and whence Theorem III.4.8 indicates that the exact convergence speed of the drift-implicit Euler scheme is better than $O\left(\Delta_N^\lambda\right)$ for any $\lambda \in (0, H)$. In turn, [18] uses the results on the modulus of the continuity of the fractional Brownian motion and establishes that the exact speed of convergence is $O\left(\Delta_N^H \sqrt{|\log(\Delta_N)|}\right)$ (provided that T is small enough). On Fig. III.2(b), we plot the values of $\log\left(\Delta_N \left|\log(\Delta_N)\right|^{\frac{1}{2H}}\right)$ against the logarithms of the corresponding $L^2(\Omega; L^\infty([0,T]))$-errors from Figure III.2(a). The resulting points (depicted in black) turn out to be located along the line with the slope $0.7022687 \approx 0.7 = H$ (depicted in red; least squares method was used to estimate the slope). This gives an empirical evidence to the conjecture that additional conditions on T in [18] can be lifted and the speed $O\left(\Delta_N^H \sqrt{|\log(\Delta_N)|}\right)$ is still preserved.

Example III.5.2. *(Sandwiched process of the TSB type)* Consider a sandwiched

SDE of the form

$$Y(t) = Y(0) + \int_0^t \left(\frac{\kappa_1}{Y(s) - \varphi(s)} - \frac{\kappa_2}{\psi(s) - Y(s)} - \kappa_3 Y(s) \right) ds + Z(t), \quad \text{(III.51)}$$

$t \in [0, T]$, where Z satisfies **(Z1)–(Z2)** with $\lambda \in \left(\frac{1}{2}, 1\right)$. This equation fits into the framework of Section III.4 and the scheme (III.22) leads to N cubic equations of the form

$$\widehat{Y}^3(t_{n+1}) + B_{2,n}\widehat{Y}^2(t_{n+1}) + B_{1,n}\widehat{Y}(t_{n+1}) + B_{0,n} = 0, \quad n = 0, ..., N-1, \quad \text{(III.52)}$$

where

$$B_{0,n} := \frac{-\varphi(t_{n+1})\psi(t_{n+1})\left(\widehat{Y}(t_n) + \Delta Z_n\right) + \Delta_N\left(\kappa_1\psi(t_{n+1}) + \kappa_2\varphi(t_{n+1})\right)}{1 + \Delta_N\kappa_3},$$

$$B_{1,n} := \varphi(t_{n+1})\psi(t_{n+1}) + \frac{(\varphi(t_{n+1}) + \psi(t_{n+1}))(\widehat{Y}(t_n) + \Delta Z_n) - \Delta_N(\kappa_1 + \kappa_2)}{1 + \Delta_N\kappa_3},$$

$$B_{2,n} := -\varphi(t_{n+1}) - \psi(t_{n+1}) - \frac{\widehat{Y}(t_n) + \Delta Z_n}{1 + \Delta_N\kappa_3},$$

Note this equation can be solved explicitly using, e.g., the celebrated Cardano method. Namely, define

$$Q_{1,n} := B_{1,n} - \frac{B_{2,n}^2}{3}, \quad Q_{2,n} := \frac{2B_{2,n}^3}{27} - \frac{B_{2,n}B_{1,n}}{3} + B_{0,n}$$

and put

$$Q_n := \left(\frac{Q_{1,n}}{3}\right)^3 + \left(\frac{Q_{2,n}}{2}\right)^2,$$

$$\alpha_n := \sqrt[3]{-\frac{Q_{2,n}}{2} + \sqrt{Q_n}}, \quad \beta_n := \sqrt[3]{-\frac{Q_{2,n}}{2} - \sqrt{Q_n}},$$

where among possible complex values of α_n and β_n one should take those for which $\alpha_n\beta_n = -\frac{Q_{1,n}}{3}$. Then the three roots of the cubic equation (III.52) are

$$y_{1,n} = \alpha_n + \beta_n, \qquad y_{2,n} = -\frac{\alpha_n + \beta_n}{2} + i\frac{\alpha_n - \beta_n}{2}\sqrt{3},$$
$$y_{3,n} = -\frac{\alpha_n + \beta_n}{2} - i\frac{\alpha_n - \beta_n}{2}\sqrt{3},$$

and $\widehat{Y}(t_{n+1})$ is equal to the root which belongs to $(\varphi(t_{n+1}), \psi(t_{n+1}))$ (note that there is exactly one root in that interval).

Fig. III.3 contains 10 sample paths of the process (III.51) driven by a fractional Brownian motion with $H = 0.7$. In all simulations, we take $\varphi \equiv -1$, $\psi \equiv 1$, $N = 10000$, $T = 1$ and $Y(0) = 0$, $\kappa_1 = \kappa_2 = \frac{1}{2}$, $\kappa_3 = 0$ (this case corresponds to the TSB equation described in Example III.2.10). Simulation is performed by direct implementation of the Cardano's method in R. On

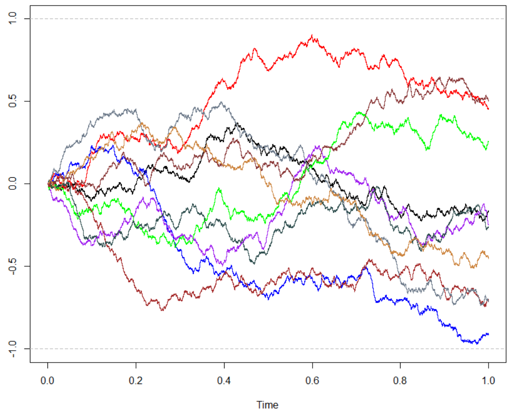


Figure III.3: Ten sample paths of (III.51) generated using the drift-implicit Euler approximation scheme; $\varphi \equiv -1$, $\psi \equiv 1$, $N = 10000$, $T = 1$, $Y(0) = 0$, $\kappa_1 = \kappa_2 = \frac{1}{2}$, $\kappa_3 = 0$, Z is a fractional Brownian motion with $H = 0.7$.

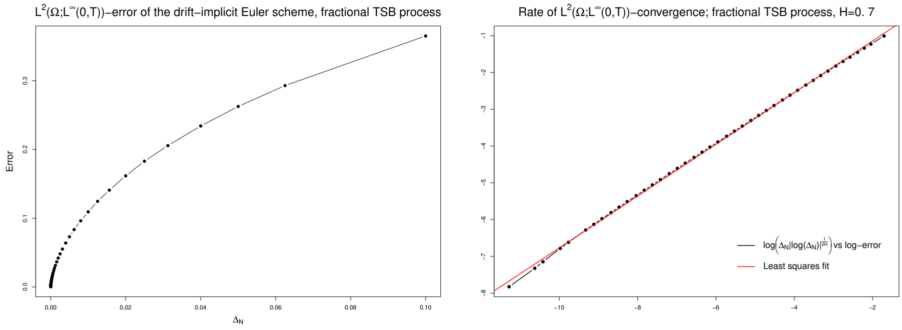
Fig. III.4(a), the $L^2(\Omega; L^\infty([0,T]))$-errors are depicted. Just as in Example III.5.1, behaviour of the modulus of continuity of the fractional Brownian motion allows to suggest that the exact convergence speed of the numerical scheme is $O\left(\Delta_N^H \sqrt{|\log(\Delta_N)|}\right)$. Fig. III.4(b) gives an empirical evidence to this conjecture: the values of $\log\left(\Delta_N \, |\log(\Delta_N)|^{\frac{1}{2H}}\right)$ plotted against the logarithms of the corresponding $L^2(\Omega; L^\infty([0,T]))$-errors (black) are located along the line (red) with the slope $0.7033434 \approx 0.7 = H$ (least squares fit was used).

In both Examples III.5.1 and III.5.2, equations for computing $\widehat{Y}$ could be explicitly solved but the Hölder continuity of the noise could not be less then $1/2$. The next example shows that the drift-implicit Euler scheme can be applied in the rough case as well.

Example III.5.3. *(Sandwiched process driven by multifractional Brownian motion)* Consider the sandwiched SDE of the form

$$Y(t) = Y(0) + \int_0^t \left(\frac{\kappa_1}{(Y(s) - \varphi(s))^4} - \frac{\kappa_2}{(\psi(s) - Y(s))^4} \right) ds + Z(t), \qquad \text{(III.53)}$$

$t \in [0,T]$. In this case, Theorem III.2.3 guarantees existence and uniqueness of the solution for λ-Hölder Z with $\lambda > \frac{1}{5}$ (note that this equation fits the framework of Example III.2.11 from Section III.2). On Fig. III.5, one can see paths of the process (III.53) with $\kappa_1 = \kappa_2 = 1$, $\varphi(t) = \sin(10t)$, $\psi(t) = \sin(10t) + 2$ driven by multifractional Brownian motion (mBm) with functional Hurst parameter $H(t) = \frac{1}{5}\sin(2\pi t) + \frac{1}{2}$ (note that the lowest value of the functional Hurst parameter is $H\left(\frac{3}{4}\right) = 0.3$; for more details on mBm, see [5] as well as [17, Lemma 3.1] for results on Hölder continuity of its paths). Fig III.6 contains



(a) (b)

Figure III.4: Convergence analysis of the drift-implicit Euler approximation scheme for (III.51); $\varphi \equiv -1$, $\psi \equiv 1$, $T = 1$, $Y(0) = 0$, $\kappa_1 = \kappa_2 = \frac{1}{2}$, $\kappa_3 = 0$, Z is a fractional Brownian motion with $H = 0.7$. On panel (a), $L^2(\Omega; L^\infty([0,T]))$-errors are depicted. Panel (b) contains the values of $\log\left(\Delta_N \left|\log(\Delta_N)\right|^{\frac{1}{2H}}\right)$ plotted against the logarithms of the corresponding $L^2(\Omega; L^\infty([0,T]))$-errors (black) as well as the line fitted with the least squares method (red). The slope of the red line is $0.7033434 \approx 0.7 = H$.

the $L^2(\Omega; L^\infty([0,T]))$-errors of approximation. Note a much slower rate of convergence: the multifractional Brownian motion Z under consideration is Hölder continuous up to the order 0.3, therefore Theorem III.3.9 guarantees convergence speed of only $O(\Delta_N^\lambda)$ with $\lambda \in (0, 0.3)$.

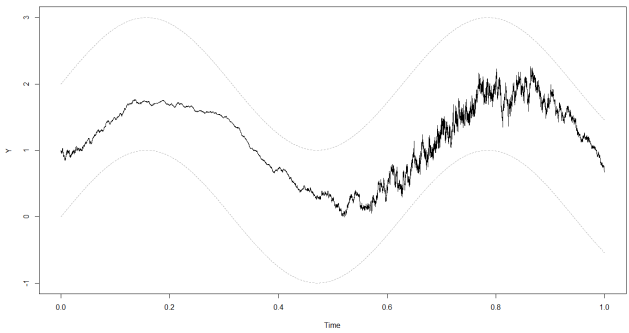


Figure III.5: A sample path of (III.53) generated using the backward Euler approximation scheme; $N = 10000$, $T = 1$, $Y(0) = 1$, $\kappa_1 = \kappa_2 = 1$, $\varphi(t) = \sin(10t)$, $\psi(t) = \sin(10t) + 2$, Z is a multifractional Brownian motion with functional Hurst parameter $H(t) = \frac{1}{5}\sin(2\pi t) + \frac{1}{2}$.

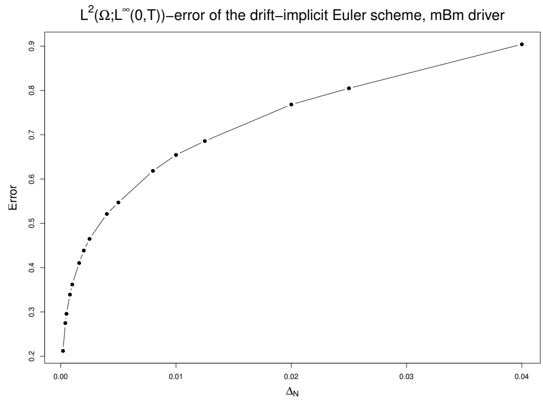


Figure III.6: $L^2(\Omega; L^\infty([0,T]))$-errors of the drift-implicit Euler approximation scheme for (III.53); $T = 1$, $Y(0) = 1$, $\kappa_1 = \kappa_2 = 1$, $\varphi(t) = \sin(10t)$, $\psi(t) = \sin(10t) + 2$, Z is a multifractional Brownian motion with functional Hurst parameter $H(t) = \frac{1}{5}\sin(2\pi t) + \frac{1}{2}$.

References

[1] Alfi, V. et al. “Roughness and finite size effect in the NYSE stock-price fluctuations”. In: *The European physical journal. B* vol. 55, no. 2 (2007), pp. 135–142.

[2] Alfonsi, A. “On the discretization schemes for the CIR (and Bessel squared) processes”. In: *Monte Carlo Methods Appl.* vol. 11, no. 4 (2005), pp. 355–384.

[3] Alfonsi, A. “Strong order one convergence of a drift implicit Euler scheme: application to the CIR process”. In: *Statist. Probab. Lett.* vol. 83, no. 2 (2013), pp. 602–607.

[4] Andersen, L. B. G. and Piterbarg, V. V. “Moment explosions in stochastic volatility models”. In: *Finance and Stochastics* vol. 11, no. 1 (Sept. 2006), pp. 29–50.

[5] Ayache, A., Cohen, S., and Vehel, J. L. “The covariance structure of multifractional Brownian motion, with application to long range dependence”. In: *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100)*. IEEE, 2002.

[6] Azmoodeh, E. et al. “Necessary and sufficient conditions for Hölder continuity of Gaussian processes”. In: *Statist. Probab. Lett.* vol. 94 (2014), pp. 230–235.

[7] Bayer, C., Friz, P., and Gatheral, J. “Pricing under rough volatility”. In: *Quant. Finance* vol. 16, no. 6 (2016), pp. 887–904.

[8] Chan, K. C. et al. "An empirical comparison of alternative models of the short-term interest rate". In: *The journal of finance* vol. 47, no. 3 (1992), pp. 1209–1227.

[9] Cox, J. C. "The constant elasticity of variance option pricing model". In: *The Journal of Portfolio Management* vol. 23, no. 5 (1996), pp. 15–17.

[10] Cox, J. C., Ingersoll, J. E., and Ross, S. A. "A Re-examination of Traditional Hypotheses about the Term Structure of Interest Rates". In: *The Journal of Finance* vol. 36, no. 4 (Sept. 1981), pp. 769–799.

[11] Cox, J. C., Ingersoll, J. E., and Ross, S. A. "A Theory of the Term Structure of Interest Rates". In: *Econometrica* vol. 53, no. 2 (Mar. 1985), p. 385.

[12] Cox, J. C., Ingersoll, J. E., and Ross, S. A. "An Intertemporal General Equilibrium Model of Asset Prices". In: *Econometrica* vol. 53, no. 2 (Mar. 1985), p. 363.

[13] d'Onofrio, A., ed. *Bounded Noises in Physics, Biology, and Engineering*. Springer New York, 2013.

[14] Dereich, S., Neuenkirch, A., and Szpruch, L. "An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process". In: *Proceedings of the Royal Society A. Mathematical, physical, and engineering sciences* vol. 468, no. 2140 (2012), pp. 1105–1115.

[15] Di Nunno, G., Mishura, Y., and Yurchenko-Tytarenko, A. "Sandwiched SDEs with unbounded drift driven by Hölder noises". In: *Advances in Applied Probability (to appear)* vol. 55, no. 3 (2023).

[16] Domingo, D., d'Onofrio, A., and Flandoli, F. "Properties of bounded stochastic processes employed in biophysics". In: *Stochastic Analysis and Applications* vol. 38, no. 2 (Dec. 2019), pp. 277–306.

[17] Dozzi, M. et al. "Asymptotic growth of trajectories of multifractional Brownian motion, with statistical applications to drift parameter estimation". In: *Statistical inference for stochastic processes* vol. 21, no. 1 (2018), pp. 21–52.

[18] Hong, J. et al. "Optimal strong convergence rate of a backward Euler type scheme for the Cox–Ingersoll–Ross model driven by fractional Brownian motion". In: *Stochastic Processes and their Applications* vol. 130, no. 5 (2020), pp. 2675–2692.

[19] Hu, Y., Nualart, D., and Song, X. "A singular stochastic differential equation driven by fractional Brownian motion". In: *Statistics & Probability Letters* vol. 78, no. 14 (Oct. 2008), pp. 2075–2085.

[20] Kruse, R. *Strong and weak approximation of semilinear stochastic evolution equations*. Springer International Publishing, 2014.

[21] Kubilius, K. and Medžiūnas, A. "Positive solutions of the fractional SDEs with non-Lipschitz diffusion coefficient". In: *Mathematics* vol. 9, no. 1 (2020).

[22] Zhang, S.-Q. and Yuan, C. “Stochastic differential equations driven by fractional Brownian motion with locally Lipschitz drift and their implicit Euler approximation”. In: *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* (2020), pp. 1–27.

Authors’ addresses

Giulia Gi Nunno Department of Mathematics, University of Oslo, Moltke Moes vei 35, 0851 Oslo, Norway, giulian@math.uio.no

Yuliya Mishura Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, 64/13, Volodymyrs’ka St., Kyiv 01601, Ukraine, myus@univ.kiev.ua

Anton Yurchenko-Tytarenko Department of Mathematics, University of Oslo, Moltke Moes vei 35, 0851 Oslo, Norway, antony@math.uio.no