**Weierstraß-Institut
für Angewandte Analysis und Stochastik**

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Weak error rates for option pricing under linear rough volatility

Christian Bayer[1], Eric Hall[2], Raúl F. Tempone[3,4]

submitted: January 26, 2022

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: christian.bayer@wias-berlin.de

[2] University of Dundee
School of Science and Engineering
Mathematics Division
Dundee DD1 4HR, UK
E-Mail: ehall001@dundee.ac.uk

[3] RWTH Aachen
Departement for Mathematics
Kackertstraße 9
52072 Aachen
Germany
E-Mail: tempone@uq.rwth-aachen.de

[4] KAUST
Computer, Electrical and Mathematical
Sciences & Engineering Division (CEMSE)
Thuwal $23955-6900$
Saudi Arabia
E-Mail: raul.tempone@kaust.edu.sa

No. 2916
Berlin 2022

2020 *Mathematics Subject Classification.* 91G60, 91G20, 65C20.

Key words and phrases. Rough volatility, option pricing, weak error, Euler–Maruyama, non-Markovian dynamics, rough Stein–Stein model.

This work was supported by the KAUST Office of Sponsored Research (OSR) under Award No. OSR-2019-CRG8-4033 and the Alexander von Humboldt Foundation. R. Tempone is a member of the KAUST SRI Center for Uncertainty Quantification in Computational Science and Engineering. A portion of this work was carried out while E. Hall was a Postdoctoral Research Scientist in the Chair of Mathematics for Uncertainty Quantification at RWTH Aachen University. C. Bayer gratefully acknowledges support by the German Research Council (DFG) via the Research Unit FOR 2402. We are grateful to Andreas Neuenkirch for pointing out the simpler proof in the case of quadratic payoffs.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Weak error rates for option pricing under linear rough volatility

Christian Bayer, Eric Hall, Raúl F. Tempone

Abstract

In quantitative finance, modeling the volatility structure of underlying assets is vital to pricing options. Rough stochastic volatility models, such as the rough Bergomi model [Bayer, Friz, Gatheral, Quantitative Finance 16(6), 887-904, 2016], seek to fit observed market data based on the observation that the log-realized variance behaves like a fractional Brownian motion with small Hurst parameter, $H < 1/2$, over reasonable timescales. Both time series of asset prices and option-derived price data indicate that H often takes values close to 0.1 or less, i.e., rougher than Brownian motion. This change improves the fit to both option prices and time series of underlying asset prices while maintaining parsimoniousness. However, the non-Markovian nature of the driving fractional Brownian motion in rough volatility models poses severe challenges for theoretical and numerical analyses and for computational practice. While the explicit Euler method is known to converge to the solution of the rough Bergomi and similar models, its strong rate of convergence is only H. We prove rate $H + 1/2$ for the weak convergence of the Euler method for the rough Stein–Stein model, which treats the volatility as a linear function of the driving fractional Brownian motion, and, surprisingly, we prove rate one for the case of quadratic payoff functions. Indeed, the problem of weak convergence for rough volatility models is very subtle; we provide examples demonstrating the rate of convergence for payoff functions that are well approximated by second-order polynomials, as weighted by the law of the fractional Brownian motion, may be hard to distinguish from rate one empirically. Our proof uses Talay–Tubaro expansions and an affine Markovian representation of the underlying and is further supported by numerical experiments. These convergence results provide a first step toward deriving weak rates for the rough Bergomi model, which treats the volatility as a nonlinear function of the driving fractional Brownian motion.

1 Introduction

Rough stochastic volatility models form an increasingly popular paradigm in quantitative finance, as they simultaneously address two empirical challenges. Firstly, time series of realized variance indicate that variance is rough in the sense of having Hölder regularity $H \ll 1/2$, see [16, 8, 14]. Secondly, rough volatility models recover the power-law explosion of the at the money implied volatility skew of the form $\tau^{-\gamma}$ for $\gamma \sim 1/2$ as time to maturity $\tau \to 0$. In fact, these two constants are linked by $\gamma = 1/2 - H$, giving further evidence of regularity H being small, say around 0.1. We refer to [5] for the pricing perspective.

To fix notation, we consider a rough stochastic volatility model for an asset price process S_t of the form

$$\mathrm{d}S_t = \sqrt{v_t} S_t \mathrm{d}Z_t \,,$$

where Z is a Brownian motion (Bm). There are two classes of rough volatility models which differ in the specification of the instantaneous variance component v_t. The *rough Heston model* ([13]) is an example of one kind, with v_t given as a solution to a Volterra stochastic differential equation (SDE) with a power law kernel $K(r) \sim r^{H-1/2}$, $r > 0$. This paper will consider an alternative where the variance process is an explicit function of a fractional Brownian motion (fBm) W_t^H, which does not need to be the classical fBm. For instance, the *rough Bergomi model* ([5]) is specified by the choice

$$v_t := \xi(t) \exp\left(\eta W_t^H - \frac{1}{2}\eta^2 t^{2H}\right) , \tag{1.1}$$

DOI 10.20347/WIAS.PREPRINT.2916 Berlin 2022

where $\xi(t)$ denotes the *forward variance* and W_t^H denotes the *Riemann–Liouville* fBm given by

$$W_t^H \coloneqq \int_0^t K(t-s)\mathrm{d}W_s, \quad K(r) \coloneqq \sqrt{2H} r^{H-1/2}, \tag{1.2}$$

for a Bm W with correlation ρ with Z. A related model where the variance process is an explicit function of the fBm is the *fractional* or *rough Stein–Stein model* ([1]), given by

$$\begin{aligned} \mathrm{d}S_t &= v_t S_t \mathrm{d}Z_t\,, \\ v_t &= v_0(t) + \int_0^T K(t,s)\kappa v_s \mathrm{d}s + \int_0^T K(t,s)\eta \mathrm{d}W_s\,, \end{aligned} \tag{1.3}$$

for a Volterra kernel K and for arbitrary correlation ρ between Z and W. This extends the classic Stein–Stein model ([29]) and its generalization ([28]). For the particular choice $\kappa = 0$ and $K(t,s) = K(t-s)$ from (1.2), the volatility term (1.3) is a linear function of the fBm,

$$v_t = v_0(t) + \eta W_t^H\,.$$

For the later volatility, the rough Stein–Stein model can be viewed as a simplified fractional SABR model that enables explicit computations of certain quantities of interest ([15]).

The modelling advantages gained by capturing these two empirical challenges, i.e., low Hölder regularity ($H \ll 1/2$) and the power-law explosion, using a rough stochastic volatility model are paid for both on the theoretical and the numerical side. Indeed, rough stochastic volatility models are neither semi-martingales nor Markov processes. Despite the former, rough volatility models do not violate the no-arbitrage-condition, as the asset price process itself is a martingale. On the other hand, the difficulties caused by the lack of Markov property are more severe. In particular, there is no finite dimensional pricing PDE anymore (although we refer to [22, 7] for implementations of an infinite-dimensional pricing PDE based on machine learning). For some rough volatility models of affine Volterra type, for instance, the rough Heston model, there is still a semi-explicit formula for the asset price's characteristic function in terms of a deterministic fractional ODE. Otherwise, the rough stochastic volatility approach necessitates simulation-based methods.

On the numerical side, $W_{t_1}^H, \dots, W_{t_N}^H$ can be exactly sampled at discrete-time points as W^H is a Gaussian process with known covariance function. (The hybrid scheme of [9] is a popular alternative to exact simulation, sacrificing accuracy for speed.) However, simulation of S_t requires discretization of a stochastic integral, even in the case of the rough Bergomi and rough Stein–Stein models. As we shall see in further detail later, we essentially need to compute stochastic integrals of the form

$$\int_0^T \psi(t, W_t^H)\mathrm{d}W_t, \tag{1.4}$$

for some deterministic, "nice" function ψ. In particular, note that the integrand is adapted and square-integrable (under appropriate conditions). Hence, the stochastic integral exists in the classical Itō sense, and strong convergence of the numerical scheme

$$\sum_{i=0}^{n-1} \psi(t_i, W_{t_i}^H)(W_{t_{i+1}} - W_{t_i}) \tag{1.5}$$

is also classical. The speed of convergence is considerably less clear. Indeed, Neuenkirch and Shalaiko [25] proved strong convergence with rate H for a very similar problem, i.e., phrased in terms of classical fBm, and strong rate H is widely expected to hold also for the approximation scheme (1.5) to (1.4). Using techniques from regularity structures, in particular, renormalization by an exploding constant, [4] proved essentially the same strong rate for a Wong–Zakai type approximation of (1.4).

Combining our observations—that volatility is rough ($H \approx 0.1$) and typical schemes converge with strong rate H— we run into problems, as the rate of convergence is so small as to make it indistinguishable from lack of convergence in many cases of practical importance. Indeed, suppose that $H = 0.1$ and we need n time steps to reach an error tolerance ϵ. If we now decrease our tolerance by a factor ten, i.e., we require one additional significant digit, then the number of time-steps needed is increased by a factor 10^{10} in the asymptotic regime.

For most applications we really require weak as opposed to strong convergence of the numerical scheme. For instance, the price of a European option with payoff φ is $\mathbf{E}[\varphi(S_T)]$, and its computation relies on weak convergence of the scheme. Weak approximation of stochastic integrals is often much faster than strong approximation. Consider the Euler scheme for standard SDEs (the case $H = 1/2$). Generically, i.e., when the problem is sufficiently "nice", the weak rate of convergence is one, whereas the strong rate is $1/2$. This poses the interesting question about the relation between the Hölder regularity ($H = 1/2$), the weak rate of convergence (1) and the strong rate of convergence ($1/2$). Indeed, [25] showed us that the strong rate is equal to the Hölder regularity H, but there are several plausible candidates for the weak rate: $2H$, $H + 1/2$, and 1 (independent of H).[1] We stress that only the last two alternatives allow for feasible numerical simulations in the truly rough regime. Bluntly put, if the true weak error only decays proportionally to n^{-2H} in the number of time steps n, then simulation methods are not viable numerical methods for option pricing in rough volatility models.

Despite the importance of the problem of determining the weak rate, only little work has been done. Horvath, Jacquier and Muguruza [20] study a Donsker theorem for a rough volatility model, which translates into a week tree-type approximation. The rate of convergence of their method is H in the number of time-steps. At this stage, we should note that the trees are non-recombining, implying that the memory load increases exponentially in the number of time-steps. To the best of our knowledge, this work provides the only rigorous weak convergence result in the literature of rough volatility models. Indeed, it is worth pointing out that standard proof techniques for diffusions, see [30], strongly rely on the Markov property, and are, hence, not applicable in this setting.

At the same time, discretization-based simulation methods are often used in the literature, with great success. While convergence is rarely considered (not even empirically), we would expect to see difficulties emerge in the very rough cases $H \approx 0.1$ if the convergence rates were truly as bad as only H or $2H$. In fact, the few available empirical studies (for instance, [6]) indicate a much larger weak rate of convergence. In fact, the authors of [6] observe a weak rate of one which is stable enough to allow accelerated convergence by Richardson extrapolation.

In this paper we prove novel weak rates for the convergence of the left-hand rule (1.5) to (1.4):

Theorem 1.1. *The left-point approximation* (1.5) *to the rough stochastic integral* (1.4) *converges with weak rate* $H + 1/2$ *for* $\psi(t, W_t^H) = W_t^H$ *– i.e., in the rough Stein–Stein model. For the case that the payoff* φ *is a quadratic polynomial the convergence is with weak rate one.*

We refer to Theorems 2.1 and 4.1 for more precise statements. Some remarks are in order:

- The problem of weak convergence in this setting is very subtle; if we restrict ourselves to quadratic polynomials as payoff functions, then the weak rate of convergence is actually one, see Lemma 4.2. This implies the rate of convergence for payoff functions φ that can be well approximated by quadratic polynomials, as seen from the law of the solution, may be hard to distinguish from rate one empirically, due to prevalence of higher order terms (see Figure 3). Note that the result 4.2 and its proof were communicated to us by Andreas Neuenkirch [24] prior to starting this work; Lemma 4.2 indicates rate one for quardatic payoffs φ for a more

[1] Anecdotally, we asked several experts on stochastic numerics in early stages of working on this problem, and all three possibilities were put forward.

general class of ψ (i.e., including rough Bergomi) but it is unclear how to generalize this result to a broader class of payoffs.

- We do not have a lower bound establishing that the weak rate of convergence cannot be better than $H + 1/2$ in the generic case. We do offer numerical evidence for this assertion, though, see Figures 1 to 3.
- We do not doubt that the proof extends to the general case of non-linear ψ, which includes the rough Bergomi model. Indeed, the present paper is partly motivated to expose a possible proof strategy for the general case. Extending the method of proof using Faà di Bruno's formula poses some technical challenges, mainly due to the needed to control more complicated formulas.

Our proof for Theorem 1.1 relies on deriving Taylor expansions for the weak error using an affine Markovian representation of the underlying. The basic flavor of this approach, i.e., obtaining a Markovian extended variable system to facilitate analysis, is a strategy utilized in other non-Markovian stochastic dynamical systems such as the Generalized Langevin equation (see, e.g., [17, 12]) and open Hamiltonian systems ([26]). In the context of rough volatility models, Markovian approximations were also used in [2].

Outline of the paper

In Section 2 we provide the setting and the main result and discuss the general strategy of the proof. Section 3 introduces auxiliary, Markovian approximations to both (1.4) and (1.5) based on [10]. This high dimensional Markovian problem will serve as a surrogate problem for most of the convergence analysis. Section 4 considers the special case of quadratic payoff functions, for which the general proof strategy simplifies considerably. We contrast this with a specific proof only applicable to quadratic payoffs, which also works for general non-linear ψ. The proof of Theorem 1.1 (and Theorem 2.1) is then carried out in Section 5.

2 Problem setting: weak rate of convergence for Euler scheme is $H + 1/2$

We consider a smooth, bounded payoff function $\varphi(X_T)$ for an underlying

$$X_t := \int_0^t \psi(s, W_s^H) \mathrm{d}W_s \,, \tag{2.1}$$

where W_t^H is a Riemann–Liouville fBm given by (1.2) with Hurst parameter $H \in (0, 1/2)$. A simplified model of rough stochastic volatility, (2.1) retains keys features of the rough Bergomi model (1.1) and the rough Stein-Stein model (1.3). Namely, the X_t in (2.1) is non-Markovian as W_t^H, and hence $\psi(t, W_t^H)$, depends on the full history of $(W_s)_{s\in[0,t]}$ (cf. ψ to the instantaneous variance v_t in (1.1)). In fact, for the purposes of European option, the rough Bergomi model can be reduced to (2.1) in the following way (often attributed to [27]). First, Itō's formula implies that

$$S_T = S_0 \exp\left(-\frac{1}{2} \int_0^T v_s \mathrm{d}s + \int_0^T \sqrt{v_s} \mathrm{d}Z_s \right).$$

We can now replace the Bm Z by $\rho W + \sqrt{1-\rho^2} W^\perp$ for an independent Bm $W^\perp$. *Conditionally on* W, S_T has a log-normal distribution with parameters

$$\mu := \log S_0 - \frac{1}{2} \int_0^T v_s \mathrm{d}s + \rho \int_0^T \sqrt{v_s} \mathrm{d}W_s, \quad \sigma^2 := (1-\rho^2) \int_0^T v_s \mathrm{d}s.$$

If we denote the Black–Scholes price for the payoff function φ at maturity T by $C_{BS}(S_0, \sigma_{BS}^2 T, \varphi)$, for interest rate $r = 0$ and volatility σ_{BS}, then we get

$$\mathbf{E}[\varphi(S_T)] = \mathbf{E}\left[C_{BS}\left(S_0 \exp\left[-\frac{\rho^2}{2}\int_0^T v_s \mathrm{d}s + \rho\int_0^T \sqrt{v_s}\mathrm{d}W_s\right], (1-\rho^2)\int_0^T v_s \mathrm{d}s,\, \varphi\right)\right]. \tag{2.2}$$

Computation of the right hand side of (2.2) requires simulation of the Lebesgue integral $\int_0^T v_s \mathrm{d}s$ as well as simulation of

$$\int_0^T \sqrt{v_s}\mathrm{d}W_s = \int_0^T \sqrt{\xi(s)} \exp\left(\frac{\eta}{2}W_s^H - \frac{\eta^2}{4}s^{2H}\right)\mathrm{d}W_s, \tag{2.3}$$

which is of the form (2.1).

Presently, we derive weak rates of convergence,

$$\left|\mathbf{E}[\varphi(X_T) - \varphi(\overline{X}_T^{\Delta t})]\right| = O(\Delta t^\gamma), \tag{2.4}$$

for the left-hand scheme (1.5) with step-size Δt such that $n\Delta t = T$. Restricting to the the rough Stein–Stein model $\psi(s, W_s^H) = W_s^H$, the main finding of this work, in Theorem 2.1 (and implying the first statement in Theorem 1.1), is that the weak rate is $\gamma = H + 1/2$ for the Hurst parameter H.

Theorem 2.1 (Weak rate). *For general $\varphi \in C_b^\eta$, for integer $\eta = \lceil \frac{1}{H} \rceil$, and the rough Stein–Stein model $\psi(s, W_s^H) = W_s^H$, we have*

$$|\mathrm{Err}(T, \Delta t)| = \left|\mathbf{E}[\varphi(X_T) - \varphi(\overline{X}_T^{\Delta t})]\right| = O(\Delta t^{H+1/2}),$$

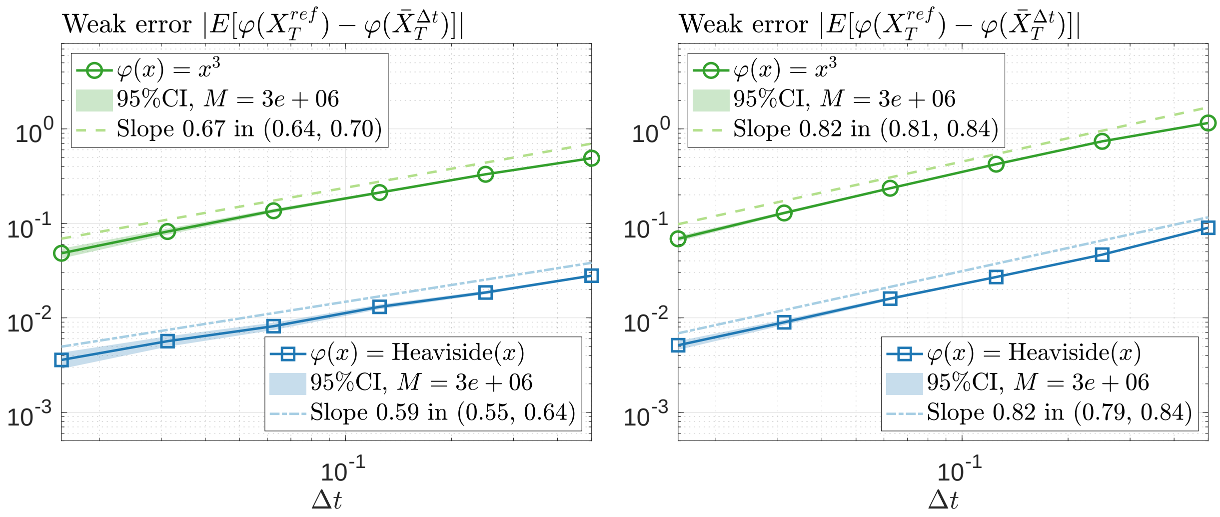
i.e. the Euler method is weak rate $H + 1/2$.

The proof of Theorem 2.1 is presented in Section 5. Before diving into the machinery needed for the proof, we first consider some numerical evidence that supports the rates in Theorem 1.1 and the accompanying remarks. Details of the implementation are outlined in Appendix A.

The first group of numerical experiments, in Figure 1, provide support for rate $H+1/2$ in Theorem 2.1. In Figure 1, the weak error rate is observed to depend on H for the general (i.e. non-quadratic) payoff functions $\varphi(x) = x^3$ and $\varphi(x) = \mathrm{Heaviside}(x)$. Indeed, the best fits (least squares) of the weak error to Δt, as well as the extremes suggested by the upper and lower 95% confidence interval for the mean based on $M = 3 \times 10^6$ samples, is consistent with the rate $H + 1/2$. Comparing Figure 1A to Figure 1B, the rate increases (and by approximately 0.1) as H increases from $H = 0.05$ to 0.15. Although the function $\varphi(x) = \mathrm{Heaviside}(x)$ is not continuous and therefore does not fit precisely into our theory, the consistency of the observed rates in Figure 1 hint at the generality of the findings in Theorem 2.1 to, e.g., digital call options.

In Figure 2B, we observe that for $H = 1/2$, i.e. standard Brownian motion, the best fit of the weak error rate is consistent with the known weak rate one for general payoff functions. However, in contrast to the rates observed in Figure 1, the behavior of quadratic payoffs looks decidedly different. We observe in Figure 2A that the weak rate for quadratic $\varphi(x) = x^2$ appears to be $\gamma = 1$ even for small $H = 0.05$ and $H = 0.15$. Weak rate one for quadratic payoff functions is recorded in Theorem 4.1 and Lemma 4.2 in Section 4; this surprising finding, that the rate depends on the payoff function, will be readily explained using the asymptotic expansions that are at the center of our approach.

Finally, in Figure 3, we observe that the best fit of weak rate to Δt for the shifted-cubic $\varphi(x) = (x + 1.5)^3$ is consistent with rate 1 even for small $H = 0.05$ and $H = 0.15$ (cf. compare the rates in Figure 3 to those forthe cubic payoff $\varphi(x) = x^3$ in Figures 1A and 1B). As seen from the law of the



(A) $H = 0.05$ (B) $H = 0.15$

Figure 1: For small Hurst parameters, (A) $H = 0.05$ and (B) $H = 0.15$, the best fit slope for the weak error for scheme (1.5), together with extremes suggested by the 95% CI based on M observations, are consistent with the rate $H + 1/2$ obtained in Theorem 2.1 for general payoff functions φ. In particular, the rate holds for the discontinuous $\varphi(x) = \text{Heaviside}(x)$ suggesting our findings are robust. Here $\Delta t \in [2^{-6}, 2^{-1}]$ and the reference mesh is $\Delta t^{ref} = 2^{-12}$.

solution, the shifted cubic is better approximated by quadratic polynomials and therefore its rate of convergence is much harder to distinguish from rate one. This numerical experiment not only drives home the subtlety of the problem of deriving weak rates for rough stochastic volatility models, but also leads us to be optimistic that efficient numerical methods can be obtained for a wide array of real-world problems where the *effective* rate of convergence is not as bad as the theoretical rate.

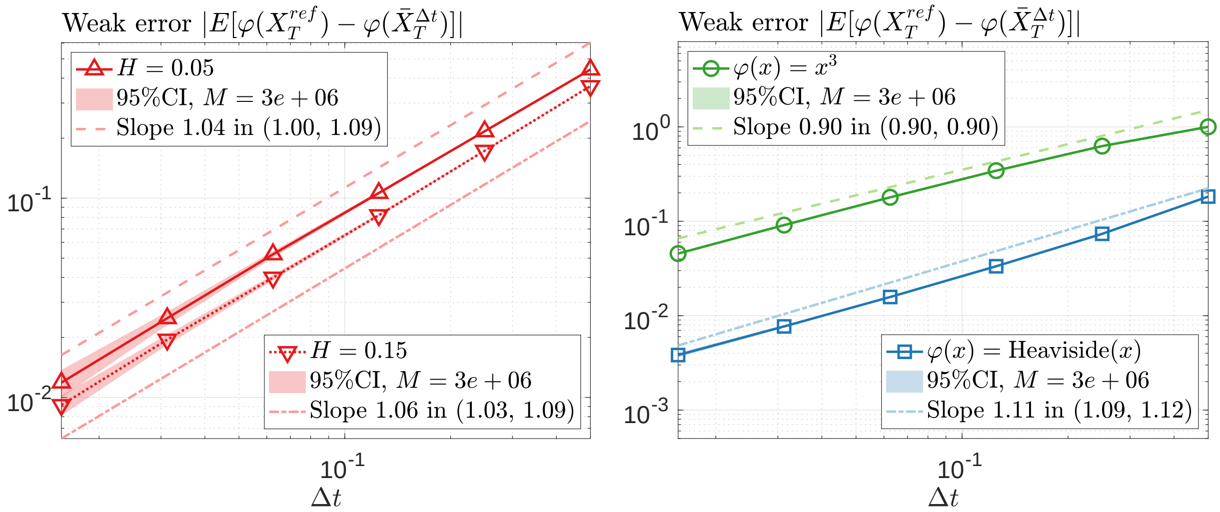
Remark 2.2 (Financial applications). Although the assumptions of Theorem 2.1 seem extremely strong, they do reflect meaningful financial situations. In particular, note that (2.2) allows us to replace the (generally non-smooth) payoff functions of European options by their smooth Black–Scholes prices. Additionally, put-call-parity may allow us to assume bounded payoffs. Linearity of ψ is, admittedly, a very strong assumption, which should be seen as the first stepping stone to the general result. We conjecture that Theorem 2.1 holds in the setting of the rough Bergomi model, i.e., for non-linear ψ as given in (2.3).

Remark 2.3 (Scheme). For the simple model problem (1.4) the numerical integration scheme (1.5) is the left-point approximation. If the problem were not trivialized to a stochastic integral, then in general $\overline{X}_T^{\Delta t}$ would correspond to the Euler–Maruyama approximation for the underlying SDE and we will refer to the scheme interchangeably as both.

In the next section, we introduce the notation and concepts that will be used to derive asymptotic expansions for the weak error in powers of Δt. In particular, we first use these expansions to derive weak rate one for quadratic payoffs, see Theorem 4.1, in Section 4. Finally in Section 5, a proof, following the approach used for Theorem 4.1 as a guide, is given for Theorem 2.1 obtaining weak rate $H + 1/2$ for general payoff functions. Taken together, the statements of Theorems 2.1 and 4.1 imply Theorem 1.1.

3 Markovian extended state space formulation

We first consider a well-known affine representation for the driving fBm. Discretizing this affine representation yields an extended state space for the dynamics of the underlying. A novelty of our



(A) Quadratic φ, small H (B) General (non-quadratic) φ, $H = 0.5$

Figure 2: (A) Surprisingly, the best fit line for the weak error for scheme (1.5) for the quadratic payoff $\varphi(x) = x^2$ is consistent with weak rate one even for small H, as found in Theorem 4.1. (B) For Hurst parameter $H = 1/2$, the weak rate in Theorem 2.1 for scheme (1.5) is consistent with the expected rate one (for standard Bm), as illustrated by the best fit slope for the weak error for $\varphi(x) = x^3$ and $\varphi(x) = \mathrm{Heaviside}(x)$ (cf. weak rate $H + 1/2$ observed in Figure 1 for small H). Here $\Delta t \in [2^{-6}, 2^{-1}]$ and the reference mesh is $\Delta t^{ref} = 2^{-12}$.

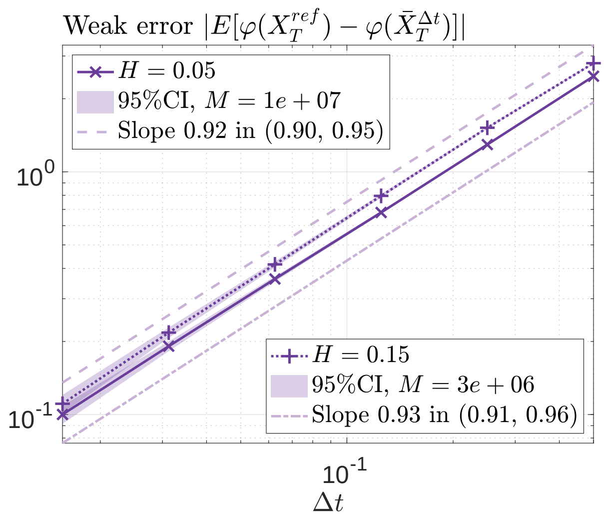


Figure 3: The weak error for scheme (1.5) for the shifted cubic payoff $\varphi(x) = (x + 1.5)^3$ achieves a higher rate than $\varphi(x) = x^3$ as the shifted cubic is better approximated by a quadratic in the support of the distribution for the underlying (cf. Figure 1). Here $\Delta t \in [2^{-6}, 2^{-1}]$ and the reference mesh is $\Delta t^{ref} = 2^{-12}$.

approach is to utilize this formulation to obtain asymptotic expansions for the weak error. In particular, we utilize the Markovian structure of the extended state space to show that (3.7) admits a Taylor expansion in Δt where the coefficients can be controlled independently of the choice of parameters used to obtain the extended state space formulation.

3.1 Affine representations for small Hurst index

Over the Hurst parameter regime of interest, the fBm (1.2) admits an affine representation as a linear functional of an infinite-dimensional family of Ornstein–Uhlenbeck (OU) processes ([10]).

Lemma 3.1 (Affine representation). *For* $0 < H < 1/2$*,*

$$W_t^H = \widetilde{c}_H \int_0^\infty \widetilde{Y}_t(\theta)\theta^{-(H+\frac{1}{2})}\mathrm{d}\theta\,, \tag{3.1}$$

where

$$\widetilde{Y}_t(\theta) = \int_0^t e^{-\theta(t-s)}\mathrm{d}W_s$$

and $\widetilde{c}_H$ *is a positive and finite constant depending on* H*.*

Although this statement is well-known we provide key details of the proof that will be referenced later for the convenience of the reader. The full proof can be found in, e.g., [10, 19] (see also [11, 23, 18] where [11] gives a Markovian representation for $H > 1/2$, [23] a time-homogeneous Markovian representation that is also defined for $t \in (-\infty, 0)$, and [18] gives bounds on tails and derivatives of the affine representation).

Proof. Writing the kernel appearing in (1.2) as a Laplace transform,

$$(t-s)^{H-\frac{1}{2}} = \frac{1}{\Gamma(\frac{1}{2}-H)}\int_0^\infty \theta^{-(H+\frac{1}{2})}e^{-\theta(t-s)}\mathrm{d}\theta\,,$$

and then using stochastic Fubini one obtains the desired result,

$$\begin{aligned} W_t^H &= \int_0^t \frac{\sqrt{2H}}{\Gamma(\frac{1}{2}-H)}\int_0^\infty \theta^{-(H+\frac{1}{2})}e^{-\theta(t-s)}\mathrm{d}\theta\mathrm{d}W_s \\ &= \int_0^\infty \widetilde{c}_H \int_0^t e^{-\theta(t-s)}\mathrm{d}W_s\theta^{-(H+\frac{1}{2})}\mathrm{d}\theta \\ &= \widetilde{c}_H \int_0^\infty \widetilde{Y}_t(\theta)\theta^{-(H+\frac{1}{2})}\mathrm{d}\theta\,, \end{aligned}$$

where $\widetilde{c}_H := \sqrt{2H}/\Gamma(\frac{1}{2}-H) < \infty$. □

A key tool in our proof of the weak rates will be to utilize the Markovian structure of a projection of the fBm obtained by discretizing the affine representation Lemma 3.1. We observe that the integral in (3.1) has a singularity at $\theta = 0$, but behaves essentially like $\theta^{-(H+\frac{1}{2})}$ before $\widetilde{Y}_t(\theta)$ vanishes in the limit of θ. To make (3.1) more amenable to quadrature we remove the singularity by introducing the change of variable,

$$\vartheta = \theta^{-(H+\frac{1}{2}-1)} = \theta^{\frac{1}{2}-H}\,,$$

thereby obtaining the representation

$$W_t^H = c_H \int_0^\infty \widetilde{Y}_t(\vartheta^{2/(1-2H)})\mathrm{d}\vartheta = c_H \int_0^\infty Y_t(\theta)\mathrm{d}\theta\,, \tag{3.2}$$

where the constant,

$$c_H := \frac{\widetilde{c}_H}{\frac{1}{2} - H} = \frac{\sqrt{2H}}{\Gamma(\frac{3}{2} - H)}\,,$$

is an increasing function of $H \in (0, \frac{1}{2})$ such that $0 < c_H < 1$. In (3.2),

$$Y_t(\theta) = \int_0^t e^{-(t-s)\theta^p} \mathrm{d}W_s \tag{3.3}$$

is an OU process with speed of mean-reversion given by θ^p with a positive power

$$p := 2/(1 - 2H) > 2\,. \tag{3.4}$$

One realization of $Y_t(\theta)$ is plotted in Figure 4 together with an envelope illustrating plus/minus two standard deviations of $Y_t(\theta)$, computed using the formula for the covariance, i.e.

$$\mathrm{Cov}(Y_t(\theta), Y_t(\eta)) = \frac{1}{\theta^p + \eta^p}(1 - e^{-(\theta^p + \eta^p)t})\,.$$

Replacing the integral in (3.2) with a quadrature rule in the parameter θ yields a projection of the fBm onto a finite state space.

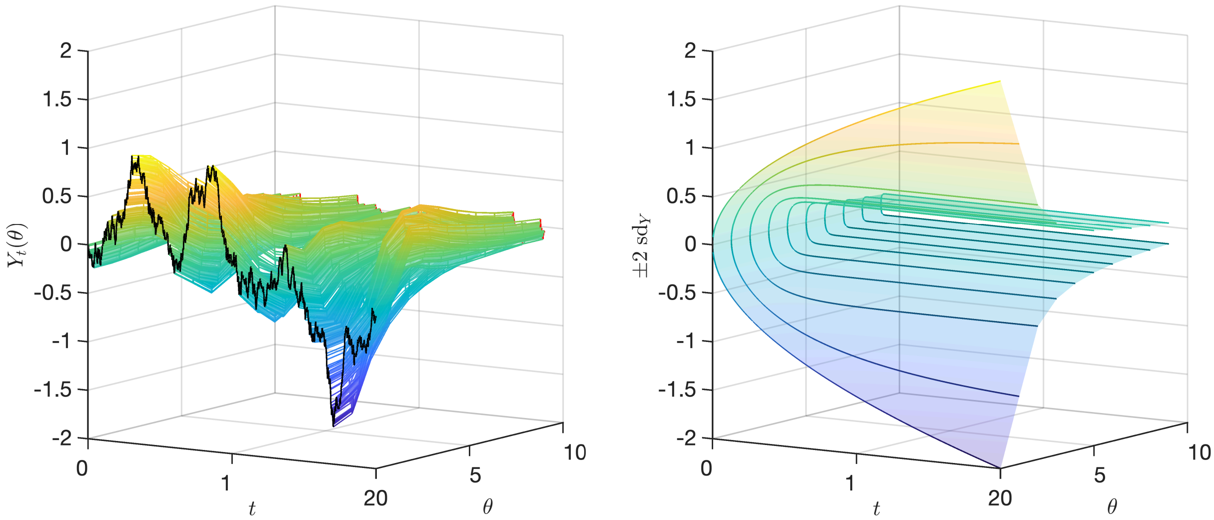


Figure 4: A sample of $Y_t(\theta)$ in (3.3), at left, with speed of mean reversion $\theta^{2/(1-2H)}$ for $H = 0.07$ plotted together with an envelope demonstrating plus/minus two standard deviations, at right, (cf. time series data of asset prices and option derived price data indicate that H often takes values close to 0.1 or even smaller [16]). $Y_t(\theta)$ is a smooth analytic function of θ and discretizing in θ yields an extended variable state space which we utilize in our analysis.

Lemma 3.2 (Approximate affine representation)**.** *For $0 < H < 1/2$, let*

$$\widehat{W}_t^H = c_H \sum_{l=1}^{N_L} Y_t^l \Delta\theta_l =: \mathcal{S}(\boldsymbol{Y}_t)\,, \tag{3.5}$$

depend on N_L degrees of freedom $\boldsymbol{Y}_t = (Y_t^1, \ldots, Y_t^{N_L})$ where $Y_t^l := Y_t(\theta_l)$ are OU process in (3.3) with speed of mean-reversion θ_l^p for p in (3.4). Then $\widehat{W}^H$ converges to W^H as $L, N_L \to \infty$ in $L^2(\Omega; C([0,T]))$.

Proof. We obtain an approximate affine representation of the fBm by discretizing (3.2) in two steps. First, we divide the integral in (3.2) into two parts,

$$W_t^H = c_H \int_0^L Y_t(\theta)\mathrm{d}\theta + \underbrace{c_H \int_L^\infty Y_t(\theta)\mathrm{d}\theta}_{=:R_L(Y_t)} ,$$

where R_L denotes the error in restricting the integral to a fixed computational domain $L > 1$. Second, we consider a quadrature rule

$$W_t^H = c_H \sum_{l=1}^{N_L} Y_t(\theta_l)\Delta\theta_l + R_{N_L}(Y_t) + R_L(Y_t) ,$$

with points $0 \le \theta_1 < \cdots < \theta_{N_L} \le L$ and weights $\Delta\theta_l = \theta_{l+1} - \theta_l$ where R_{N_L} denotes the quadrature truncation error.

That $\widehat{W}^H$ converges to W^H in the limit of L, N_L essentially follows from the "strong rates" in [18]. The R_{N_L} can be made arbitrarily small as $Y_t(\theta)$ is a smooth bounded (even analytic) function of θ (e.g. see Figure 4), i.e. the regularity in θ allows one to approximate efficiently using arbitrarily higher order quadrature rules if desired ([18]). The $R_L(Y_t)$ is a mean zero Gaussian process, since for all $\delta \in [0, H)$,

$$\sup_{L\in[1,\infty]} L^\delta \Big\| \sup_{t\in[0,T]} |R_L(Y_t)| \Big\|_{L^p(\Omega)} < \infty ,$$

guarantees integrability ([18, Lemma 1(b)]). Then R_L can be made arbitrarily small for sufficiently large L by observing that the variance,

$$\mathrm{Var}[R_L(Y_t(\theta))] = c_H^2 \int_L^\infty \int_L^\infty \mathrm{Cov}(Y_t(\theta), Y_t(\eta))\mathrm{d}\theta\mathrm{d}\eta \le c_H^2 \frac{2\pi}{4} \int_L^\infty \theta^{1-p}\mathrm{d}\theta = c_H^2 \frac{\pi}{2} \frac{L^{2-p}}{p-2} ,$$

decays in L since $p > 2$. □

We split the weak error (2.4) using the representations in Lemmas 3.1 and 3.2,

$$\begin{aligned} \mathbf{E}\left[\varphi(X_T(W^H)) - \varphi(X_T(\widehat{W}^H))\right] + \mathbf{E}\left[\varphi(X_T(\widehat{W}^H)) - \varphi(\overline{X}_T^{\Delta t}(\widehat{W}^H))\right] \\ + \mathbf{E}\left[\varphi(\overline{X}_T^{\Delta t}(\widehat{W}^H)) - \varphi(\overline{X}_T^{\Delta t}(W^H))\right] \end{aligned} \tag{3.6}$$

where we emphasize the dependence of the underlying on the driving process. The first and third terms both correspond to the error in approximating W^H with $\widehat{W}^H$ and therefore vanish by Lemma 3.2. Indeed, we have the following result.

Lemma 3.3. *Assume that φ and ψ are Lipschitz, the latter uniformly in time, with Lipschitz constants K_φ and K_ψ, respectively. Then the first and the third term of* (3.6) *converge to zero as $N_L, L \to \infty$. Regarding the third term, the convergence is uniform with respect to Δt.*

Proof. We consider the third term first. By basic probabilistic estimates using the Lipschitz property of φ and ψ, we have

$$\begin{aligned}\left|\mathbf{E}\left[\varphi(\overline{X}_T^{\Delta t}(\widehat{W}^H)) - \varphi(\overline{X}_T^{\Delta t}(W^H))\right]\right| &\leq K_\varphi\, \mathbf{E}\left[\left(\overline{X}_T^{\Delta t}(\widehat{W}^H) - \overline{X}_T^{\Delta t}(W^H)\right)^2\right]^{1/2} \\ &= K_\varphi\, \mathbf{E}\left[\left(\sum_{i=0}^{n-1}\left(\psi(s_i, \widehat{W}_{s_i}^H) - \psi(s_i, W_{s_i}^H)\right) W_{s_i, s_{i+1}}\right)^2\right]^{1/2} \\ &= K_\varphi \left(\sum_{i=0}^{n-1}\mathbf{E}\left[\left(\psi(s_i, \widehat{W}_{s_i}^H) - \psi(s_i, W_{s_i}^H)\right)^2\right](s_{i+1}-s_i)\right)^{1/2} \\ &\leq K_\varphi K_\psi \left(\sum_{i=0}^{n-1}\mathbf{E}\left[\left(\widehat{W}_{s_i}^H - W_{s_i}^H\right)^2\right](s_{i+1}-s_i)\right)^{1/2} \\ &\leq K_\varphi K_\psi \left\|\widehat{W}^H - W^H\right\|_{L^2(\Omega;C([0,T]))} T^{1/2} \to 0\end{aligned}$$

as $N_L, L \to \infty$ by Lemma 3.2. The result for the first term follows in the same manner. □

Remark 3.4 (Convergence rates in N_L, L). Following [18, Theorem 1], convergence rates in N_L and L for the first and third terms of (3.6) could undoubtedly be established. Keep in mind, however, that we only use the scheme $\overline{X}_T^{\Delta t}(\widehat{W}^H)$ as a tool for the analysis of the scheme $\overline{X}_T^{\Delta t}(W^H)$, i.e., with exact simulation of W^H. Consequently, rates of the convergence in N_L and L are not required to get rates of convergence of $\overline{X}_T^{\Delta t}(W^H)$ in terms of Δt. Indeed, with respect to the actual scheme $\overline{X}_T^{\Delta t}(W^H)$ analyzed in this paper, the error contributions from the first and third terms vanish.

The sole remaining term in (3.6),

$$\mathrm{Err}(T, \Delta t) := \mathbf{E}[\varphi(X_T(\widehat{W}^H))] - \mathbf{E}[\varphi(\overline{X}_T^{\Delta t}(\widehat{W}^H))]\,, \tag{3.7}$$

that gives the weak error in the Euler scheme, depends on the approximate affine representation in Lemma 3.2. Indeed, suppose that we are given an error tolerance ε. By Lemma 3.3, we can find $L = L(\varepsilon, H, T, \varphi, \psi)$ and $N_L = N_L(\varepsilon, H, T, \varphi, \psi)$ such that the first and third terms of (3.6) are bounded by $\varepsilon/3$ each, the third one irrespectively of Δt. Our task is now to choose time steps Δt such that also the second term is bounded by $\varepsilon/3$ for the given L, N_L. In the next section, we will obtain an extended variable system for the dynamics of the underlying that we will use to obtain an asymptotic expansions for (3.7).

Remark 3.5 (Quadrature). In the interest of keeping our arguments constructive, we first fixed a computational domain L and then introduced a quadrature based on N_L points without specifying the precise rule. One could also choose, e.g., a Gauss–Laguerre quadrature suitable for the half-line thereby reducing the number of parameters to one (see also [18]). The splitting (3.7) still holds.

3.2 Forward Euler scheme for extended variable system

Substituting (3.5) into the underlying (2.1), yields

$$\widehat{X}_t := \int_0^t \psi(s, \widehat{W}_s^H)\mathrm{d}W_s = \int_0^t \psi(s, \mathcal{S}(\boldsymbol{Y}_s))\mathrm{d}W_s\,, \tag{3.8}$$

a finite dimensional Markovian approximation $\widehat{X}_t = X_t(\widehat{W}^H)$ of the underlying X_t that appears in the weak error (3.7). The dynamics of (3.8) are described by

$$\boldsymbol{Z} = (\widehat{X}, Y^1, \ldots, Y^{N_L})\,,$$

a d-dimensional extended variable state space ($d = N_L + 1$), solving the system

$$\mathrm{d}\boldsymbol{Z}_t = -b(\boldsymbol{Z}_t, t)\mathrm{d}t + \sigma(\boldsymbol{Z}_t, t)\mathrm{d}W_t\,, \qquad \boldsymbol{Z}_0 = 0\,, \tag{3.9}$$

with d-vectors b and σ given by,

$$\begin{aligned} b(\boldsymbol{Z}_t, t) &:= (0, Y_t^1\theta_1^p, \ldots, Y_t^{N_L}\theta_{N_L}^p)\,, \\ \sigma(\boldsymbol{Z}_t, t) &:= (\psi(t, \mathcal{S}(\boldsymbol{Y}_t)), 1, \ldots, 1)\;, \end{aligned}$$

where there is a single driving Brownian motion, i.e. (3.9) is a degenerate system.

For the interval $[0, T]$, we define the uniform time grid discretization $t_i := i\Delta t$ for $i = 0, \ldots, n-1$, where $n := T/\Delta t$, and consider the Euler–Maruyama scheme for the underlying,

$$\overline{X}_{t_{i+1}} = \overline{X}_{t_i} + \psi(t_i, \mathcal{S}(\boldsymbol{Y}_{t_i}))\Delta W_{t_i}\,, \quad \overline{X}_{t_0} = 0\,, \tag{3.10}$$

where

$$\Delta W_{t_i} := W_{t_i, t_{i+1}} = W_{t_{i+1}} - W_{t_i}$$

are the increments of the driving Wiener process and where at each time step the vector of extended variables $\boldsymbol{Y}_\cdot = (Y_\cdot^l)_{l=1,\ldots,N_L}$ is sampled exactly. That is, for the Euler update at t_{i+1}, one can form the joint distribution

$$(\boldsymbol{Y}_\tau, \Delta W_\tau)_{\tau=t_0,\ldots,t_i}\,, \tag{3.11}$$

an $(N_L+1)\times(i+1)$-dimensional Gaussian. The variance-covariance matrix for (3.11) can be obtained using the known covariances $\mathrm{Cov}(Y_{t_i}^k, Y_{t_j}^l)$, $\mathrm{Cov}(Y_{t_i}^k, \Delta W_{t_j})$, and $\mathrm{Cov}(\Delta W_{t_i}, \Delta W_{t_j})$, and then the target variables $\boldsymbol{Y}_{t_i}$ required in (3.10) can be sampled exactly from the joint distribution, e.g., using the Cholesky decomposition. We extend $\overline{X}$ in (3.10) to all $t \in [0, T]$ by the interpolation,

$$\overline{X}(t) = \int_0^t \psi(\kappa_s, \widehat{W}_{\kappa_s}^H)\mathrm{d}W_s = \int_0^t \psi\Big(\kappa_s, c_H \sum_{l=1}^{N_L} \Delta\theta_l Y_{\kappa_s}^l\Big)\mathrm{d}W_s\,, \tag{3.12}$$

where $\kappa_s = t_i$ if $s \in [t_i, t_{i+1})$ for each $i = 0, \ldots, n-1$.

Coupling the interpolant for the Euler scheme with the exact dynamics of the OU variables leads us to define the "discretized" extended variable system $\overline{\boldsymbol{Z}} = (\overline{X}, Y^1, \ldots, Y^{N_L})$ embedded in the SDE

$$\mathrm{d}\overline{\boldsymbol{Z}}_s = -\bar{b}(\overline{\boldsymbol{Z}}_s)\mathrm{d}s + \bar{\sigma}(\overline{\boldsymbol{Z}}_s)\mathrm{d}W_s\,, \quad s \in [0, T]\,, \qquad \overline{\boldsymbol{Z}}_0 = 0\,, \tag{3.13}$$

with coefficients

$$\begin{aligned} \bar{b}(\overline{\boldsymbol{Z}}_s) &= (0, Y_s^1\theta_1^p, \ldots, Y_s^{N_L}\theta_{N_L}^p) = b(\boldsymbol{Z}_s, s)\,, \\ \bar{\sigma}(\overline{\boldsymbol{Z}}_s) &= (\psi(\kappa_s, \mathcal{S}(\boldsymbol{Y}_{\kappa_s})), 1, \ldots, 1) = \sigma(\boldsymbol{Z}_{\kappa_s}, \kappa_s)\,. \end{aligned}$$

For (3.13) and Section 3.2, we are able to formulate a corresponding Kolmogorov backward equation. For a smooth and bounded payoff $\varphi(\boldsymbol{Z}_T) = \varphi(Z_T^1) = \varphi(\widehat{X}_T)$, we consider the value function,

$$u(\boldsymbol{z}, t) := \mathbf{E}[\varphi(\boldsymbol{Z}_T) \mid \boldsymbol{Z}_t = \boldsymbol{z}] = \mathbf{E}[\varphi(\widehat{X}_T) \mid \boldsymbol{Z}_t = \boldsymbol{z}]\,, \tag{3.14}$$

that is the conditional expected value of the payoff at time $t < T$ given the starting value $\boldsymbol{Z}_t = \boldsymbol{z}$ for $\boldsymbol{z} = (z_1, \ldots, z_d) = (x, y_1, \ldots, y_{N_L})$. The associated Kolmogorov backward equation is given by

$$\begin{aligned} &\partial_t u(\boldsymbol{z}, t) - b^j(\boldsymbol{z}, t)D_j u(\boldsymbol{z}, t) + \frac{1}{2}A^{jk}(\boldsymbol{z}, t)D_{jk}u(\boldsymbol{z}, t) = 0\,, \quad t < T\,, \quad \boldsymbol{z} \in \mathbf{R}^d\,, \\ &u(\boldsymbol{z}, T) = \varphi(z_1)\,, \end{aligned} \tag{3.15}$$

where repeated indices indicate summation (over $1,\dots,d$), $b^j = b^j(\boldsymbol{z},t)$ is the jth component of the d-vector,

$$b(\boldsymbol{z},t) := (0, z_2\theta_1^p, \dots, z_{N_L+1}\theta_{N_L}^p)\,,$$

and $A^{jk} = A^{jk}(\boldsymbol{z},t)$ are elements of the $d\times d$-matrix $A = (\sigma\sigma^*)$,

$$A^{11} = \psi(t,\mathcal{S}(\boldsymbol{z}))^2\,, \qquad A^{1j} = A^{j1} = \psi(t,\mathcal{S}(\boldsymbol{z}))\,,\ j>1\,, \qquad A^{jk} = 1\,,\ j,k>1\,, \tag{3.16}$$

i.e. that contains ones except along the first row and column, where we slightly abuse notation,

$$\mathcal{S}(\boldsymbol{z}) := 0\cdot z_1 + c_H\sum_{j=2}^{d} z_j\Delta\theta_{j+1} = c_H\sum_{l=1}^{N_L} y_l\Delta\theta_l\,.$$

Remark 3.6 (Kolmogorov backward equation). In the presentation of the Kolmogorov backward equation, we assume necessary regularity conditions on φ, i.e. smoothness and boundedness, as a matter of convenience. From the context of the problem this is not a strong assumption, see Remark 2.2.

3.3 Local weak error representation

Throughout the remainder of this work we consider the case when $\psi(s,W_s^H) = W_s^H$ in (2.1). Returning to (3.7), we obtain a representation for the weak error in terms of local errors. First, we write the discretization error as a telescoping sequence in the value function (3.14),

$$\begin{aligned}\mathrm{Err}(T,\Delta t) &= \mathbf{E}[\varphi(\widehat{X}_T) - \varphi(\overline{X}_{t_n})]\\ &= -\Big(\mathbf{E}\,u(\overline{\boldsymbol{Z}}_{t_n},T) - \mathbf{E}\,u(\overline{\boldsymbol{Z}}_{t_0},0)\Big)\\ &= -\sum_{i=0}^{n-1}\mathbf{E}\left[u(\overline{\boldsymbol{Z}}_{t_{i+1}},t_{i+1}) - u(\overline{\boldsymbol{Z}}_{t_i},t_i)\right]\,,\end{aligned} \tag{3.17}$$

using that

$$\mathbf{E}\,\varphi(\overline{X}_{t_n}) = \mathbf{E}\left[\,\mathbf{E}[\varphi(Z_T^1)\mid \boldsymbol{Z}_T = \overline{\boldsymbol{Z}}_{t_n}]\right] = \mathbf{E}\,u(\overline{\boldsymbol{Z}}_{t_n},T)$$

and

$$\mathbf{E}\,\varphi(\widehat{X}_T) = \mathbf{E}\left[\,\mathbf{E}[\varphi(Z_T^1)\mid \boldsymbol{Z}_0 = \overline{\boldsymbol{Z}}_{t_0}]\right] = \mathbf{E}\,u(\overline{\boldsymbol{Z}}_{t_0},0)\,.$$

We then represent each difference appearing in (3.17) as a stochastic differential over a small time increment. Over the interval $[t_i,t_{i+1})$, we have that

$$\begin{aligned}\mathbf{E}[u(\overline{\boldsymbol{Z}}_{t_{i+1}},t_{i+1}) - u(\overline{\boldsymbol{Z}}_{t_i},t_i)] &= \mathbf{E}\int_{t_i}^{t_{i+1}}\mathrm{d}u(\overline{\boldsymbol{Z}}_s,s)\\ &= \mathbf{E}\int_{t_i}^{t_{i+1}}\Big(\partial_t u(\overline{\boldsymbol{Z}}_s,s) - b^j(\overline{\boldsymbol{Z}}_s,s)D_j u(\overline{\boldsymbol{Z}}_s,s) + \frac{1}{2}A^{jk}(\overline{\boldsymbol{Z}}_{t_i},t_i)D_{jk}u(\overline{\boldsymbol{Z}}_s,s)\Big)\mathrm{d}s\,,\end{aligned}$$

using Itō's formula applied to (3.13) where repeated indices indicate summation over $1,\dots,d$. Subtracting off the Kolmogorov backward equation (3.15) evaluated at $(\overline{\boldsymbol{Z}}_s,s)$ then yields

$$\mathbf{E}[u(\overline{\boldsymbol{Z}}_{t_{i+1}},t_{i+1}) - u(\overline{\boldsymbol{Z}}_{t_i},t_i)] = \frac{1}{2}\,\mathbf{E}\int_{t_i}^{t_{i+1}}\Big(A^{jk}(\overline{\boldsymbol{Z}}_{t_i},t_i) - A^{jk}(\overline{\boldsymbol{Z}}_s,s)\Big)\,D_{jk}u(\overline{\boldsymbol{Z}}_s,s)\mathrm{d}s\,. \tag{3.18}$$

We note that the non-zero terms correspond to differences along the first row and column of A in (3.16), and thus (3.18) simplifies to the following expression for the local weak error in the value function,

$$\begin{aligned}\mathbf{E}[u(\overline{\boldsymbol{Z}}_{t_{i+1}},t_{i+1})-u(\overline{\boldsymbol{Z}}_{t_i},t_i)] &= \frac{1}{2}\,\mathbf{E}\int_{t_i}^{t_{i+1}}\Big(A^{11}(\overline{\boldsymbol{Z}}_{t_i},t_i)-A^{11}(\overline{\boldsymbol{Z}}_s,s)\Big)\,D_{11}u(\overline{\boldsymbol{Z}}_s,s)\mathrm{d}s\\ &\quad+\mathbf{E}\int_{t_i}^{t_{i+1}}\sum_{j=2}^{d}\Big(A^{j1}(\overline{\boldsymbol{Z}}_{t_i},t_i)-A^{j1}(\overline{\boldsymbol{Z}}_s,s)\Big)\,D_{j1}u(\overline{\boldsymbol{Z}}_s,s)\mathrm{d}s\\ &= -\frac{1}{2}\,\mathbf{E}\int_{t_i}^{t_{i+1}}\mathcal{S}(\boldsymbol{Y}.)^2_{t_i,s}D_{11}u(\overline{\boldsymbol{Z}}_s,s)\mathrm{d}s\\ &\quad-\mathbf{E}\int_{t_i}^{t_{i+1}}\mathcal{S}(\boldsymbol{Y}.)_{t_i,s}\sum_{j=2}^{d}D_{j1}u(\overline{\boldsymbol{Z}}_s,s)\mathrm{d}s\,,\end{aligned} \tag{3.19}$$

where we express the differences in components of A in terms of the increments of the approximate fBm

$$\mathcal{S}(\boldsymbol{Y}.)^k_{t_i,s}:=\mathcal{S}(\boldsymbol{Y}_s)^k-\mathcal{S}(\boldsymbol{Y}_{t_i})^k\,,\qquad k=1,2\,.$$

Observe that the derivatives of the value function appearing in (3.19) can be further resolved by directly computing fluxes. Here we consider the simplification $\psi(s,\widehat{W}^H_s)=\widehat{W}^H_s$ in (2.1) (i.e. "linear ψ").

Lemma 3.7 (Fluxes)**.** *Let $\psi(s,W^H_s)=W^H_s$. For the value function $u(\boldsymbol{z},t)$ defined in* (3.14)*,*

$$D^\beta u(\boldsymbol{z},t)=c_H^{|\beta|-\beta_1}\,\mathbf{E}\Big[\varphi^{(|\beta|)}(\widehat{X}_T)\prod_{j=1}^{N_L}(\Delta\theta_j M^j_{t,T})^{\beta_{j+1}}\mid \boldsymbol{Z}_t=\boldsymbol{z}\Big]\,, \tag{3.20}$$

for a multi-index $\beta=(\beta_1,\dots,\beta_d)$ where

$$M^j_{t,T}:=\int_t^T e^{-(r-t)\theta^p_j}\mathrm{d}W_r\,,\quad j=1,\dots,N_L\,.$$

Proof. Let $\boldsymbol{Z}^{t,\boldsymbol{z}}_s$ be the Markov process started at $\boldsymbol{Z}_t=\boldsymbol{z}$ with components

$$\boldsymbol{Z}^{t,\boldsymbol{z}}_s=(\widehat{X}^{t,x}_s,\boldsymbol{Y}^{t,\boldsymbol{y}}_s)=(\widehat{X}^{t,x}_s,Y^{t,y_1}_s,\dots,Y^{t,y_{N_L}}_s)\,;$$

here we drop the index $Y^{t,y_l}=Y^{l;t,y_l}$ when the index is clear from the initial condition. We recall that Y^{t,y_l}_s started at the value y_l at time t is given by,

$$Y^{t,y_l}_s=e^{-(s-t)\theta^p_l}y_l+\int_t^s e^{-(s-r)\theta^p_l}\mathrm{d}W_r\,,\quad t<s\,, \tag{3.21}$$

and, similarly, that $\widehat{X}^{t,x}_s$ is given by,

$$\widehat{X}^{t,x}_s=x+\int_t^s\mathcal{S}(\boldsymbol{Y}^{t,\boldsymbol{y}}_r)\mathrm{d}W_r=x+\int_t^s c_H\sum_{l=1}^{N_L}Y^{t,y_l}_r\Delta\theta_l\mathrm{d}W_r\,,\quad t<s\,.$$

Working directly with (3.21) and Section 3.3, derivatives of the underlying with respect to the initial conditions are given by

$$\frac{\partial\widehat{X}^{t,x}_s}{\partial x}=1\,,$$

and, for $l=1,\dots,N_L$,

$$\frac{\partial\widehat{X}^{t,x}_s}{\partial y_l}=c_H\Delta\theta_l\int_t^s e^{-(r-t)\theta^p_l}\mathrm{d}W_r=:c_H\Delta\theta_l M^l_{t,s}\,.$$

The formula (3.20) follows as all higher derivatives of $\widehat{X}^{t,x}_s$ vanish. □

Returning to our expression for the local weak error in the value function (3.19) we apply (3.20) thereby obtaining,

$$\begin{aligned}\mathbf{E}[u(\overline{\boldsymbol{Z}}_{t_{i+1}},t_{i+1})-u(\overline{\boldsymbol{Z}}_{t_i},t_i)] &= -\frac{1}{2}\,\mathbf{E}\int_{t_i}^{t_{i+1}}\mathcal{S}(\boldsymbol{Y}_{\cdot})^2_{t_i,s}\,\mathbf{E}[\varphi''(\widehat{X}_T)\mid \boldsymbol{Z}_s=\overline{\boldsymbol{Z}}_s]\mathrm{d}s\\ &\quad-\sum_{l=1}^{N_L}\mathbf{E}\int_{t_i}^{t_{i+1}}\mathcal{S}(\boldsymbol{Y}_{\cdot})_{t_i,s}\,\mathbf{E}[\varphi''(\widehat{X}_T)c_H\Delta\theta_l M^l_{s,T}\mid \boldsymbol{Z}_s=\overline{\boldsymbol{Z}}_s]\mathrm{d}s\,.\end{aligned}\tag{3.22}$$

We introduce deterministic functions of $\boldsymbol{z}$,

$$\nu(\boldsymbol{z},s):=\mathbf{E}[\varphi''(\widehat{X}_T)\mid \boldsymbol{Z}_s=\boldsymbol{z}]\,,$$

and

$$\widetilde{\nu}(\boldsymbol{z},s):=\mathbf{E}\Big[\varphi''(\widehat{X}_T)(c_H\textstyle\sum_l M^l_{s,T}\Delta\theta_l)\mid \boldsymbol{Z}_s=\boldsymbol{z}\Big]\,.$$

Rewriting (3.22) with this new notation leads to an expression for the local weak error in the value function,

$$\begin{aligned}\mathbf{E}[u(\overline{\boldsymbol{Z}}_{t_{i+1}},t_{i+1})-u(\overline{\boldsymbol{Z}}_{t_i},t_i)] &= -\frac{1}{2}\,\mathbf{E}\int_{t_i}^{t_{i+1}}\mathcal{S}(\boldsymbol{Y}_{\cdot})^2_{t_i,s}\nu(\overline{\boldsymbol{Z}}_s,s)\mathrm{d}s-\mathbf{E}\int_{t_i}^{t_{i+1}}\mathcal{S}(\boldsymbol{Y}_{\cdot})_{t_i,s}\widetilde{\nu}(\overline{\boldsymbol{Z}}_s,s)\mathrm{d}s\\ &=: J+\widetilde{J}\,,\end{aligned}\tag{3.23}$$

that will serve as our starting point for the convergence rates. Next we derive an expansion for (3.23) in powers of Δt from which we obtain convergence rates.

3.4 Taylor expansions and conditional independence

Starting with the local weak error (3.23), we derive asymptotic expansions for J (and $\widetilde{J}$) in in powers of Δt by Taylor expanding ν (and $\widetilde{\nu}$) at $\overline{\boldsymbol{Z}}_{t_i}$ and applying a conditioning argument.

We observe that $\overline{\boldsymbol{Z}}=(\overline{X},\boldsymbol{Y})$ in (3.13), i.e. the interpolation (3.12) together with the exact dynamics of the OU extended variables, is linear with respect to the increment over $[t_i,s]$,

$$\overline{\boldsymbol{Z}}_s-\overline{\boldsymbol{Z}}_{t_i}=(\mathcal{S}(\boldsymbol{Y}_{t_i})W_{t_i,s}\,,Y^1_{t_i,s}\,,\dots\,,Y^{N_L}_{t_i,s})\tag{3.24}$$

where

$$W_{t_i,s}:=W_s-W_{t_i}\quad\text{and}\quad Y^l_{t_i,s}:=Y^l_s-Y^l_{t_i}\,,\qquad\text{for } s\in[t_i,t_{i+1})\,,$$

are increments of the driving Brownian motion and extended variable OU processes, respectively. Using the linearization (3.24), the Taylor expansion of ν at $\overline{\boldsymbol{Z}}_{t_i}$ is,

$$\nu(\overline{\boldsymbol{Z}}_s,s)=\nu(\overline{\boldsymbol{Z}}_{t_i},s)+\sum_{\beta\in\mathcal{I}_\kappa}D^\beta\nu(\overline{\boldsymbol{Z}}_{t_i},s)\cdot(\mathcal{S}(\boldsymbol{Y}_{t_i})W_{t_i,s})^{\beta_1}(\boldsymbol{Y}_{t_i,s})^{\hat{\beta}}+\mathcal{R}_\kappa(\nu)\,,\tag{3.25}$$

for sums over multiindices in the set

$$\mathcal{I}_\kappa=\{\beta=(\beta_1,\hat{\beta})=(\beta_1,\beta_2,\dots,\beta_d)\;:\;1\le|\beta|\le\kappa-1\}$$

where

$$(\boldsymbol{Y}_{t_i,s})^{\hat{\beta}}=\prod_{l=1}^{N_L}(Y^l_{t_i,s})^{\beta_{l+1}}=(Y^1_{t_i,s})^{\beta_2}\cdots(Y^{N_L}_{t_i,s})^{\beta_d}$$

and the remainder is given in integral form by,

$$\mathcal{R}_\kappa(\nu) = \frac{1}{\kappa!}\sum_{|\beta|=\kappa}(\mathcal{S}(\boldsymbol{Y}_{t_i})W_{t_i,s})^{\beta_1}(\boldsymbol{Y}_{t_i,s})^{\hat{\beta}}\int_0^1 D^\beta\nu(\xi_\tau,s)\mathrm{d}\tau\,, \qquad \xi_\tau := \overline{\boldsymbol{Z}}_{t_i} + \tau(\mathcal{S}(\boldsymbol{Y}_{t_i})W_{t_i,s}\,,Y^1_{t_i,s}\,,\dots\,,Y^{N_L}_{t_i,s})\,. \tag{3.26}$$

In (3.25), the terms ν and $D^\beta\nu$ are deterministic functions of the $\mathcal{F}_{t_i}$-measurable random variable $\boldsymbol{Z}_{t_i}$ and are therefore $\mathcal{F}_{t_i}$-measurable. Analogous expressions hold for (3.25) and (3.26) with $\widetilde{\nu}$ in place of ν.

Plugging the ν-expansion (3.25) into (3.23), we obtain an asymptotic expansion for J,

$$\begin{aligned} 2J = &-\mathbf{E}\int_{t_i}^{t_{i+1}}\nu(\overline{\boldsymbol{Z}}_{t_i},s)\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}\mid\mathcal{F}_{t_i}]\mathrm{d}s \\ &-\sum_{\beta\in\mathcal{I}_\kappa}\mathbf{E}\int_{t_i}^{t_{i+1}}D^\beta\nu(\overline{\boldsymbol{Z}}_{t_i},s)\mathcal{S}(\boldsymbol{Y}_{t_i})^{\beta_1}\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}(W_{t_i,s})^{\beta_1}(\boldsymbol{Y}_{t_i,s})^{\hat{\beta}}\mid\mathcal{F}_{t_i}]\mathrm{d}s \\ &-\mathbf{E}\int_{t_i}^{t_{i+1}}\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}\mathcal{R}_\kappa(\nu)\mid\mathcal{F}_{t_i}]\mathrm{d}s\,, \end{aligned} \tag{3.27}$$

by conditional independence. An expansion analogous to (3.27) holds for $\widetilde{J}$ with $\widetilde{\nu}$ in place of ν. The only terms that depend on Δt in (3.27) (and in the analogously expansion for $\widetilde{J}$) are the conditional expectations involving products of the increments $\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}$ or $\mathcal{S}(\boldsymbol{Y}_\cdot)_{t_i,s}$ together with powers of $W_{t_i,s}$ and $\boldsymbol{Y}_{t_i,s}$. The key obtain weak error rates will be to show after isolating the order in Δt using these expansions that the expansion coefficients, which depend on the extended state space variables $\boldsymbol{Y}$, are controlled with respect to summation in the parameter(s) θ.

In the next section, we observe that for quadratic payoffs, the expansions in ν and $\widetilde{\nu}$ truncate after the first term since ν and $\widetilde{\nu}$ already depend on two derivatives of φ. In this special case, we derive weak rate one in Theorem 4.1. In Section 5, we prove that in general the weak rate is $H + 1/2$, as reported in Theorem 2.1, also using the asymptotic expansions approach.

4 Weak rate one for quadratic payoffs

Using the preceding machinery, we will now derive rates of convergence for the weak error via Taylor expansions in powers of Δt such that all terms stay integrable in θ. For quadratic payoff functions, we obtain that the weak error is $O(\Delta t)$, i.e. rate one in Theorem 4.1 below, which is supported by numerical evidence, recall Figure 2A. The mechanism by which rate one is achieved can be observed in the expansions; the expansion coefficients depend on derivatives of the payoff function and higher-order terms that reduce the rate vanish when φ is quadratic.

4.1 Asymptotic expansion approach to weak rate one

Returning to the increment of the value functional (3.23), if $\varphi \in \mathcal{P}^2$, that is, is a quadratic polynomial, then the derivatives of ν and $\widetilde{\nu}$ (as defined in Section 3.3) vanish and only the first terms in the expansion (3.25) remain. Then,

$$\begin{aligned} J + \widetilde{J} &= -\frac{1}{2}\mathbf{E}\int_{t_i}^{t_{i+1}}\nu(\overline{\boldsymbol{Z}}_{t_i},s)\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}\mid\mathcal{F}_{t_i}]\mathrm{d}s - \mathbf{E}\int_{t_i}^{t_{i+1}}\widetilde{\nu}(\boldsymbol{Z}_{t_i},s)\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)_{t_i,s}\mid\mathcal{F}_{t_i}]\mathrm{d}s \\ &=: \tfrac{1}{2}J_0 + \widetilde{J}_0\,, \end{aligned} \tag{4.1}$$

and estimating J_0 and $\widetilde{J}_0$ yields the weak rate corresponding to quadratic payoff functions.

Theorem 4.1 (Weak rate quadratic payoff)**.** *Let $\varphi \in \mathcal{P}_2$ and let $\psi(s, \widehat{W}^H_s) = \widehat{W}^H_s$, then*

$$\mathrm{Err}(T, \Delta t, \varphi) := \mathbf{E}\big[\varphi(\widehat{X}_T) - \varphi(\overline{X}_{t_n})\big] \lesssim O(\Delta t)\,,$$

i.e. the Euler method is weak rate one.

Proof. We estimate the terms J_0 and $\widetilde{J}_0$ in (4.1) beginning with $\widetilde{J}_0$. Working directly with the increments of the extended variables,

$$\mathbf{E}\big[\mathcal{S}(\boldsymbol{Y}_\cdot)_{t_i,s} \mid \mathcal{F}_{t_i}\big] = -c_H \sum_{l=1}^{N_L} \mathbf{E}\big[Y^l_{t_i,s} \mid \mathcal{F}_{t_i}\big]\Delta\theta_l = 0\,,$$

and thus we conclude

$$\widetilde{J}_0 = -\,\mathbf{E}\int_{t_i}^{t_{i+1}} \widetilde{\nu}(\overline{\boldsymbol{Z}}_{t_i}, s)\,\mathbf{E}\big[\mathcal{S}(\boldsymbol{Y}_\cdot)_{t_i,s} \mid \mathcal{F}_{t_i}\big] = 0\,.$$

In the case of quadratic φ, we observe that $\|\nu\|_\infty = O(1)$ and deterministic. From the definition of Y_s (e.g. see (3.21)), working again directly with the increments of the extended variables we have that

$$\begin{aligned}
\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} &= -c_H^2 \sum_{k,l=1}^{N_L} (Y^k_{t_i}Y^l_{t_i} - Y^k_s Y^l_s)\Delta\theta_k\Delta\theta_l \\
&= -c_H^2 \sum_{k,l=1}^{N_L} \Big\{(1 - e^{-(s-t_i)(\theta_k^p+\theta_l^p)})Y^k_{t_i}Y^l_{t_i} - e^{-(s-t_i)\theta_k^p}Y^k_{t_i}\int_{t_i}^s e^{-(s-r)\theta_l^p}\mathrm{d}W_r \\
&\quad - e^{-(s-t_i)\theta_l^p}Y^l_{t_i}\int_{t_i}^s e^{-(s-r)\theta_k^p}\mathrm{d}W_r - \int_{t_i}^s e^{-(s-r)\theta_k^p}\mathrm{d}W_r \int_{t_i}^s e^{-(s-r)\theta_l^p}\mathrm{d}W_r\Big\}\Delta\theta_k\Delta\theta_l\,.
\end{aligned}$$

From this the key conditional expectation term reduces to

$$\mathbf{E}\big[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} \mid \mathcal{F}_{t_i}\big] = -c_H^2 \sum_{l,k=1}^{N_L} \Delta\theta_l\Delta\theta_k(1 - e^{-(s-t_i)(\theta_l^p+\theta_k^p)})\big(Y^l_{t_i}Y^k_{t_i} - \tfrac{1}{\theta_l^p+\theta_k^p}\big)\,. \tag{4.2}$$

Only the component $1 - e^{-\Delta t(\theta_l^p+\theta_k^p)}$ will contribute to the estimate for the weak rate, provided that the sums over θ in (4.2) converge.

Using (4.2) we find

$$\begin{aligned}
J_0 &= -2c_H^2 \int_{t_i}^{t_{i+1}} \sum_{l,k=1}^{N_L} \Delta\theta_l\Delta\theta_k\,\mathbf{E}[Y^l_{t_i}Y^k_{t_i} - \tfrac{1}{\theta_l^p+\theta_k^p}](1 - e^{-(s-t_i)(\theta_l^p+\theta_k^p)})\mathrm{d}s \\
&= -2c_H^2 \sum_{l,k=1}^{N_L} \Delta\theta_l\Delta\theta_k\,\mathbf{E}\big[Y^l_{t_i}Y^k_{t_i} - \tfrac{1}{\theta_l^p+\theta_k^p}\big]\int_0^{\Delta t} g(s)\mathrm{d}s\,.
\end{aligned} \tag{4.3}$$

The integrand,

$$g(s) := 1 - e^{-s(\theta_l^p+\theta_k^p)}\,, \tag{4.4}$$

is a function of $s \in [0, \Delta t)$ such that $g(0) = 0$ and the associated Lipschitz constant K is given by,

$$K = \frac{\partial}{\partial s}g(s)\Big|_{s=0} = \theta_l^p + \theta_k^p\,.$$

Then, since $|g(s) - g(0)| \le Ks$, we have

$$|J_0| \le 2c_H^2 \sum_{l,k=1}^{N_L} \Delta\theta_l\Delta\theta_k\,\Big|\mathbf{E}\big[Y^l_{t_i}Y^k_{t_i}(\theta_l^p + \theta_k^p) - 1\big]\Big|\int_0^{\Delta t} s\mathrm{d}s\,.$$

Computing the covariance appearing above,

$$\mathbf{E}[Y^l_{t_i} Y^k_{t_i}] = \int_0^{t_i} e^{-(t_i-r)(\theta_l^p+\theta_k^p)}\mathrm{d}r = \frac{1-e^{-t_i(\theta_l^p+\theta_k^p)}}{\theta_l^p+\theta_k^p}\,,$$

we see that

$$|J_0| \le 2c_H^2\Big(\sum_{l=1}^{N_L}\Delta\theta_l e^{-t_i\theta_l^p}\Big)^2\Delta t^2 \le \frac{c_H^2}{p^2}\Gamma\big(\tfrac{1}{p}\big)^2 t_i^{-2/p}\Delta t^2 \lesssim O(\Delta t^2)\,, \tag{4.5}$$

since

$$\sum_{l=1}^{N_L}\Delta\theta_l e^{-t_i\theta_l^p} \le \int_0^L e^{-t_i\theta_l^p}\mathrm{d}\theta_l \le \int_0^\infty e^{-t_i\theta_l^p}\mathrm{d}\theta_l = \frac{1}{p}\Gamma\big(\tfrac{1}{p}\big)t_i^{-1/p}\,.$$

Importantly, $t^{-2/p}$ is integrable on $[0,T]$ when $p>2$ and therefore the $t_i^{-2/p}$ appearing in (4.5) will remain uniformly bounded when summing over t_i in (3.17).

Turning now to the telescoping representation of the weak error (3.17) and using (4.5), we obtain the desired rate

$$\begin{aligned}
|\mathrm{Err}(T,\Delta t,\varphi)| &= \Big|\sum_{i=0}^{n-1}\mathbf{E}\big[u(\overline{\boldsymbol{Z}}_{t_{i+1}},t_{i+1}) - u(\overline{\boldsymbol{Z}}_{t_i},t_i)\big]\Big| \\
&= \Big|\sum_{i=0}^{n-1}\mathbf{E}\int_{t_i}^{t_{i+1}}\nu(\overline{\boldsymbol{Z}}_{t_i},s)\,\mathbf{E}\big[\mathcal{S}(\boldsymbol{Y}.)^2_{t_i,s}\mid\mathcal{F}_{t_i}\big]\mathrm{d}s\Big| \\
&\le \frac{c_H^2}{p^2}\Gamma\big(\tfrac{1}{p}\big)^2\sum_{i=0}^{n-1}\Delta t^2 t_i^{-2/p} \\
&\le \frac{c_H^2}{p(p-2)}\Gamma\big(\tfrac{1}{p}\big)^2 T^{(p-2)/p}\Delta t \\
&\le T^{2H}\Delta t\,.
\end{aligned}$$

□

Although initially surprising that the rate depends on the payoff, the Taylor expansion for the weak error provides insight into this behavior. The expansion coefficients in (3.25) depend on increasingly higher order derivatives of the payoff function φ through derivatives of ν and $\widetilde{\nu}$ (in Section 3.3). Unlike for quadratic φ where terms in (3.25) vanish, the higher order derivative terms persist for general payoff functions. Oddly, it is only the next higher term (compared to those involved in the expansion for quadratic φ) that reduces the overall rate for general φ. In this context, one might hope to obtain an *effective rate* that is independent of H for payoff functions well approximated by quadratic polynomials.

4.2 A simpler proof for weak rate one

Before moving on to the proof of the main result in Section 5, we first present a simpler proof of weak rate one for quadratic payoff functions that is also applicable to nonlinear ψ. This proof as well as the weak rate itself was communicated to us by A. Neuenkirch [24].

Lemma 4.2. *Suppose that $\psi \in C^1_{\mathrm{pol}}$, i.e. $\psi \in C^1$ and $\psi, \partial_t\psi, \partial_x\psi$ have polynomial growth, and $\varphi(x) = x^2$. Then*

$$\Big|\mathbf{E}\big[\varphi(X_T) - \varphi(\widetilde{X}_T)\big]\Big| = O(\Delta t)\,,$$

where

$$\widetilde{X}_T \coloneqq \int_0^T \psi(s, W^H_{\kappa_s})\mathrm{d}W_s\,,$$

for $\kappa_s = t_i$ if $s \in [t_i, t_{i+1})$ for each $i = 0,\dots,n-1$.

Proof. By the Itō isometry, we have

$$\mathbf{E}\left[\varphi\left(X_T\right)\right] = \int_0^T \mathbf{E}\left[\psi^2(s, s^H W_1^H)\right] \mathrm{d}s = \int_0^T g(s)\mathrm{d}s\,,$$
$$\mathbf{E}[\varphi(\widetilde{X}_T)] = \int_0^T \mathbf{E}\left[\psi^2(\kappa_s, \kappa_s^H W_1^H)\right] \mathrm{d}s = \int_0^T g(\kappa_s)\mathrm{d}s\,,$$

where

$$g(s) := \mathbf{E}\left[\psi^2(s, s^H V)\right], \quad V \sim \mathsf{N}(0,1)\,.$$

Note that g is differentiable with integrable derivative and we assume the time derivative of ψ is bounded. Indeed,

$$g'(t) = 2\,\mathbf{E}\left[\partial_t \psi(t, t^H V)\right] + 2\,\mathbf{E}\left[\partial_x \psi(t, t^H V) V\right] t^{H-1}\,,$$

which is of order t^{H-1} and, hence, integrable.

Setting

$$\zeta_t := \min\{t_i \mid t_i \geq t\}\,,$$

we conclude with

$$\begin{aligned}
\left|\int_0^T g(s)\mathrm{d}s - \int_0^T g(\kappa_s)\mathrm{d}s\right| &\leq \int_0^T \left|g(\kappa_s) + \int_{\kappa_s}^s g'(t)\mathrm{d}t - g(\kappa_s)\right| \mathrm{d}s \\
&\leq \int_0^T \int_{\kappa_s}^s \left|g'(t)\right| \mathrm{d}t\mathrm{d}s \\
&= \int_0^T \int_t^{\zeta_t} \mathrm{d}s \left|g'(t)\right| \mathrm{d}t \\
&\leq \max_{i=0,\dots,n-1} |t_{i+1} - t_i| \left\|g'\right\|_{L^1([0,T])}\,.
\end{aligned}$$

□

For simplicity, we assumed that $\psi(s, \widehat{W}_s^H) = \widehat{W}_s^H$ in Theorem 4.1. In contrast Lemma 4.2 is applicable to any ψ including nonlinear functions. However, it is not clear how to extend the approach of the simple proof for Lemma 4.2 to more general payoff functions φ. In the next section we use the asymptotic expansions to prove the main result, Theorem 2.1, obtaining the weak rate $H + 1/2$ for general payoff functions for $\psi(s, \widehat{W}_s^H) = \widehat{W}_s^H$.

5 Proof of Theorem 2.1

Our proof of Theorem 2.1 follows the expansion approach used to obtain rate one for quadratic payoffs in Section 4.1. We derive asymptotic expansions in powers of Δt for increments of the value function in (3.23), i.e., for J and $\widetilde{J}$. For the case of general payoff functions, this requires two rounds of Taylor expansions. The first round expands ν and $\widetilde{\nu}$ at $\overline{\boldsymbol{Z}}_{t_i}$ using (3.25), as was done in Section 4.1. Then after applying a conditioning argument, we explicitly deal with correlations by expressing our expansions in terms of the extended variables and making a second round of Taylor expansions with respect to select components of $\boldsymbol{Y}_{t_i}$. Again, a key point in the proof is that all the terms in the expansions are controlled with respect to θ.

Returning to the local weak error (3.19), we use the ν-expansion (3.25) to find that J, the term corresponding to the increment $\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}$, up to order $\kappa = 3$ is given by,

$$
\begin{aligned}
2J = &- \mathbf{E}\int_{t_i}^{t_{i+1}} \nu(\overline{\boldsymbol{Z}}_{t_i}, s)\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} \mid \mathcal{F}_{t_i}]\mathrm{d}s \\
&- \mathbf{E}\int_{t_i}^{t_{i+1}} D_1\nu(\overline{\boldsymbol{Z}}_{t_i}, s)\mathcal{S}(\boldsymbol{Y}_{t_i})\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} W_{t_i,s} \mid \mathcal{F}_{t_i}]\mathrm{d}s \\
&- \sum_{j=2}^{d}\mathbf{E}\int_{t_i}^{t_{i+1}} D_j\nu(\overline{\boldsymbol{Z}}_{t_i}, s)\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} Y^{j-1}_{t_i,s} \mid \mathcal{F}_{t_i}]\mathrm{d}s \\
&- \mathbf{E}\int_{t_i}^{t_{i+1}} D_{11}\nu(\overline{\boldsymbol{Z}}_{t_i}, s)\mathcal{S}(\boldsymbol{Y}_{t_i})^2\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} (W_{t_i,s})^2 \mid \mathcal{F}_{t_i}]\mathrm{d}s \\
&- 2\sum_{j=2}^{d}\mathbf{E}\int_{t_i}^{t_{i+1}} D_{j1}\nu(\overline{\boldsymbol{Z}}_{t_i}, s)\mathcal{S}(\boldsymbol{Y}_{t_i})\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} W_{t_i,s} Y^{j-1}_{t_i,s} \mid \mathcal{F}_{t_i}]\mathrm{d}s \\
&- \sum_{j,k=2}^{d}\mathbf{E}\int_{t_i}^{t_{i+1}} D_{jk}\nu(\overline{\boldsymbol{Z}}_{t_i}, s)[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} Y^{j-1}_{t_i,s} Y^{k-1}_{t_i,s} \mid \mathcal{F}_{t_i}]\mathrm{d}s \\
&- \mathbf{E}\int_{t_i}^{t_{i+1}} \mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}\mathcal{R}_3(\nu) \mid \mathcal{F}_{t_i}]\mathrm{d}s \\
=: &\; J_0 + J_{1,0} + J_{1,1} + J_{2,0} + J_{2,1} + J_{2,2} - \mathbf{E}\int_{t_i}^{t_{i+1}} \mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}\mathcal{R}_3(\nu) \mid \mathcal{F}_{t_i}]\mathrm{d}s\,.
\end{aligned} \tag{5.1}
$$

Here J_0 is as before and the $J_{k,i}$ involve kth order derivatives of ν that do not necessarily vanish for general payoff functions φ (here the second index $i \le k$ is the number of the derivatives that correspond to extended variable directions and hence the number of sums over extended variable indices). Analogously, for $\widetilde{J}$, the term corresponding to the increment $\mathcal{S}(\boldsymbol{Y}_\cdot)_{t_i,s}$, we have,

$$
\widetilde{J} = \widetilde{J}_0 + \widetilde{J}_{1,0} + \widetilde{J}_{1,1} + \widetilde{J}_{2,0} + \widetilde{J}_{2,1} + \widetilde{J}_{2,2} - \mathbf{E}\int_{t_i}^{t_{i+1}} \mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)_{t_i,s}\mathcal{R}_3(\widetilde{\nu}) \mid \mathcal{F}_{t_i}]\mathrm{d}s\,, \tag{5.2}
$$

with $\widetilde{\nu}$ in place of ν and the increment $\mathcal{S}(\boldsymbol{Y}_\cdot)_{t_i,s}$ in place of $\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}$ compared to (5.1). In the sequel, we will simply write

$$
J_k = \sum_{i=0}^{k} J_{k,i} \qquad \text{and} \qquad \widetilde{J}_k = \sum_{i=0}^{k} \widetilde{J}_{k,i}\,, \qquad k > 0\,, \tag{5.3}
$$

for the sum of all terms involving kth order derivatives of ν and $\widetilde{\nu}$. In what follows, we first take the fBm view and assume deterministic $\|D\nu\|_\infty = O(1)$ and similarly for $\widetilde{\nu}$. Since the terms corresponding to ν and $\widetilde{\nu}$ contribute only to the constant and not to the rate, this assumption allows us to easily deduce the order in Δt, namely, that terms J_k are at least order $O(\Delta t^{H+3/2})$. Expressing the J_k in extended variables, as in (4.3), it is then possible to carrying out a second round of Taylor expansions to demonstrates that the constants are controlled.

5.1 Estimate for general payoffs: J_0 is $O(\Delta t^{H+3/2})$

In Section 4.1, ν is deterministic and $O(1)$ since it depends on two derivatives of the quadratic payoff φ. Continuing as in (4.2), our starting point for the full estimate for J_0 is

$$
J_0 = -c_H^2\,\mathbf{E}\int_{t_i}^{t_{i+1}} \nu(\overline{X}_{t_i}, \boldsymbol{Y}_{t_i}\,; s) \sum_{k,l=1}^{N_L} \Delta\theta_k\Delta\theta_l[Y^k_{t_i}Y^l_{t_i} - \tfrac{1}{\theta_k^p+\theta_l^p}]g(s)\mathrm{d}s\,, \tag{5.4}
$$

where we emphasize the dependence of ν on $\boldsymbol{Y}_{t_i}$. We define an auxiliary function,

$$
f_s^{kl}(Y^k_{t_i}, Y^l_{t_i}) := \mathbf{E}[\nu(\overline{X}_{t_i}, \boldsymbol{Y}_{t_i}\,; s) \mid Y^k_{t_i}, Y^l_{t_i}]\,, \tag{5.5}
$$

and expand f_s^{kl} in a Taylor series at zero,

$$f_s^{kl}(Y_{t_i}^k, Y_{t_i}^l) = f_s^{kl}(\mathbf{0}) + \sum_{\alpha \in \mathcal{J}_{\alpha_{\max}}} \frac{1}{|\alpha|!} \partial_{kl}^{\alpha} f_s^{kl}(\mathbf{0}) (Y_{t_i}^k Y_{t_i}^l)^{\alpha} + R_{\alpha_{\max}}(Y_{t_i}^k, Y_{t_i}^l) , \tag{5.6}$$

for a set of multiindices $\mathcal{J}_{\alpha_{\max}} = \{\alpha = (\alpha_1, \alpha_2) \ : \ 1 \le |\alpha| < \alpha_{\max}\}$ where we use the notation,

$$\partial_{l_1 \ldots l_j}^{\alpha} = \partial_{l_1}^{\alpha_1} \ldots \partial_{l_j}^{\alpha_j} , \qquad \alpha = (\alpha_1, \ldots, \alpha_j) ,$$

(as opposed to D) to emphasize that the derivatives are taken with respect to y_l directions only. The remainder is given by,

$$R_{\alpha_{\max}}(Y_{t_i}^k, Y_{t_i}^l) := \sum_{|\alpha| = \alpha_{\max}} \frac{1}{\alpha_{\max}!} \partial_{kl}^{\alpha} f_s^{kl}(\xi_k, \xi_l) (Y_{t_i}^k Y_{t_i}^l)^{\alpha} , \tag{5.7}$$

for an intermediate point (ξ_k, ξ_l).

Plugging this second round of Taylor expansions (5.6) into (5.4), yields

$$\begin{aligned} J_0 = -c_H^2 \Bigg(& \sum_{k,l=1}^{N_L} \mathbf{E}[Y_{t_i}^k Y_{t_i}^l - \tfrac{1}{\theta_k^p + \theta_l^p}] \int_{t_i}^{t_{i+1}} f_s^{kl}(\mathbf{0}) g(s - t_i) \mathrm{d}s \Delta\theta_k \Delta\theta_l \\ & + \sum_{\alpha \in \mathcal{J}} \frac{1}{|\alpha|!} \sum_{k,l=1}^{N_L} \mathbf{E}[(Y_{t_i}^k)^{\alpha_1} (Y_{t_i}^l)^{\alpha_2} (Y_{t_i}^k Y_{t_i}^l - \tfrac{1}{\theta_k^p + \theta_l^p})] \int_{t_i}^{t_{i+1}} \partial_{kl}^{\alpha} f_s^{kl}(\mathbf{0}) g(s - t_i) \mathrm{d}s \Delta\theta_k \Delta\theta_l \\ & + \sum_{k,l=1}^{N_L} \mathbf{E}[R_{\alpha_{\max}}(Y_{t_i}^k, Y_{t_i}^l)(Y_{t_i}^k Y_{t_i}^l - \tfrac{1}{\theta_k^p + \theta_l^p})] \Delta\theta_k \Delta\theta_l \int_{t_i}^{t_{i+1}} g(s - t_i) \mathrm{d}s \Bigg) , \end{aligned} \tag{5.8}$$

where $g(s)$ as in (4.4). For the remainder term in (5.8), we use the Hölder regularity of the fBm to estimate the derivative of the auxiliary function evaluated at an intermediate point,

$$\sum_{k,l=1}^{N_L} \mathbf{E}[|R_{\alpha_{\max}}(Y_{t_i}^k, Y_{t_i}^l)| |Y_{t_i}^k| |Y_{t_i}^l|] \sim \mathbf{E}[|W_{t_i, t_{i+1}}^H|^{\alpha_{\max}}] \lesssim \Delta t^{H \alpha_{\max}} .$$

From this last expression, the number of terms in the auxiliary expansion, $\alpha_{\max}$, is finite and the contribution from the remainder can be made to be order one by choosing

$$\alpha_{\max} := \lceil \tfrac{1}{H} \rceil .$$

For the remaining terms in (5.8), the f^{kl} can be estimated by the payoff φ (see Lemma B.1 in Appendix B) and therefore we write for convenience that f^{kl} and all derivatives are bounded by a constant Q,

$$f^{kl} \in C_b^{\alpha_{\max}} \qquad \text{and} \qquad \|D^{\alpha} f^{kl}\|_{\infty} \le Q \quad \text{a.s.} \tag{5.9}$$

To estimate the first term in (5.8), we use the Lipschitz argument from the proof of weak rate one for quadratic payoffs in (4.4) together with (5.9), and find that,

$$c_H^2 Q \Delta t^2 \sum_{k,l=1}^{N_L} |\,\mathbf{E}[Y_{t_i}^k Y_{t_i}^l (\theta_k^p + \theta_l^p) - 1]| \Delta\theta_k \Delta\theta_l \le \frac{c_H^2}{2p^2} Q \Gamma(\tfrac{1}{p})^2 t_i^{-2/p} \Delta t^2 , \tag{5.10}$$

for a constant proportional to $t_i^{-1/p}$ as in (4.5). The key to obtaining the rate in Δt for terms of higher order in α again depends on demonstrating summability in k and l, as in (5.10). The higher order terms in (5.8) are of the form,

$$\left\{ \mathbf{E}\left[(Y_{t_i}^k)^{\alpha_1 + 1} (Y_{t_i}^l)^{\alpha_2 + 1} \right] - \mathbf{E}\left[(Y_{t_i}^k)^{\alpha_1} (Y_{t_i}^l)^{\alpha_2} \right] \tfrac{1}{\theta_k^p + \theta_l^p} \right\} \int_0^{\Delta t} \partial_{kl}^{\alpha} f_{s - t_i}(\mathbf{0}) g(s) \mathrm{d}s ,$$

and we use Isserlis' theorem to expand the expectations into products of covariances of the extended variables. For $|\alpha|$ odd, e.g. when $|\alpha| = 1$ as $\alpha_1 = 1, \alpha_2 = 0$ or $\alpha_1 = 0, \alpha_2 = 1$, then the term is zero by Isserlis'. For $|\alpha|$ even, we use the exact expression for the covariance and then check the summability in k and l. In contrast to (5.10), for these higher order terms we obtain the rate $O(\Delta t^{H+3/2})$.

For example when $|\alpha| = 2$, we apply Isserlis' theorem and obtain a term containing,

$$\mathbf{E}[Y^k_{t_i}Y^k_{t_i}]\,\mathbf{E}[Y^l_{t_i}Y^l_{t_i}] + 2\,\mathbf{E}[Y^k_{t_i}Y^l_{t_i}]^2 - \mathbf{E}[Y^k_{t_i}Y^l_{t_i}]\tfrac{1}{\theta^p_k+\theta^p_l} = \frac{1-e^{-2\theta^p_k t_i}}{2\theta^p_k}\frac{1-e^{-2\theta^p_l t_i}}{2\theta^p_l} + 2\left(\frac{1-e^{-(\theta^p_k+\theta^p_l)t_i}}{\theta^p_k+\theta^p_l}\right)^2 - \frac{1-e^{-(\theta^p_k+\theta^p_l)t_i}}{(\theta^p_k+\theta^p_l)^2}\,. \tag{5.11}$$

Using the Lipschitz argument as in (5.10), the final two terms in (5.11), containing powers of $(\theta^p_k+\theta^p_l)$ in the denominator, are summable in k and l yielding the estimate $O(\Delta t^2)$. Neglecting $\partial^\alpha f_s$, we focus on the contribution to J_0 from first term in (5.11),

$$F(\theta^p_k)F(\theta^p_l)\int_0^{\Delta t}(1-e^{-s(\theta^p_k+\theta^p_l)})\mathrm{d}s\Delta\theta_k\Delta\theta_l \le \underbrace{F(\theta^p_k)F(\theta^p_l)\left(1-e^{-(\theta^p_k+\theta^p_l)\Delta t}\right)}_{=:G(\theta^p_k,\theta^p_l,\Delta t)}\Delta t\Delta\theta_k\Delta\theta_l$$

where we introduce the notation

$$F(u) := \frac{1-\exp(-2ut_i)}{2u}\,.$$

We thus consider the sum

$$\sum_{(k,l)\in\mathbf{N}^2} G(\theta^p_k,\theta^p_l,\Delta t)\Delta t\Delta\theta_k\Delta\theta_l \tag{5.12}$$

and obtain an upper bound $\tilde{C}\Delta t^{H+3/2}$, for a constant independent of Δt, by partitioning according to the four cases below; we let $C, C', C'' > 0$ and $\alpha \in (0,1)$ be constants that are independent of Δt.

For the first case, we consider

$$\mathcal{N}_1 := \{(k,l)\in\mathbf{N}^2 : \theta^p_k+\theta^p_l \le C\}\,.$$

Since $1-\exp(-u) \le \min(1,u)$, for $u>0$, the summand in (5.12) is bounded by

$$G(\theta^p_k,\theta^p_l,\Delta t) \le t_i^2 C\Delta t\,,$$

for $(k,l)\in\mathcal{N}_1$, and thus

$$\sum_{(k,l)\in\mathcal{N}_1} G(\theta^p_k,\theta^p_l,\Delta t)\Delta t\Delta\theta_k\Delta\theta_l = O(\Delta t^2)\,.$$

In the second case, we consider

$$\mathcal{N}_2 := \{(k,l)\in\mathbf{N}^2 : C \le \theta^p_k+\theta^p_k \le C'\Delta t^{-\alpha}\}\,.$$

For $(k,l)\in\mathcal{N}_2$, we estimate

$$1-e^{-(\theta^p_k+\theta^p_l)\Delta t} \le 1-e^{-C'\Delta t^{1-\alpha}} \le C'\Delta t^{1-\alpha}$$

and also

$$F(\theta^p) = \frac{1-e^{-2\theta^p t_i}}{2\theta^p} \le \min(t_i, \tfrac{1}{2\theta^p}) =: m(\theta^p)\,,$$

so that the summand in (5.12) is bounded by

$$G(\theta_k^p, \theta_l^p, \Delta t) \leq C' \Delta t^{1-\alpha} m(\theta_k^p) m(\theta_l^p)\,.$$

Although $\mathcal{N}_2$ grows as $\Delta t \to 0$, the order one contribution from m is overtaken by the decay in θ, and thus,

$$\sum_{(k,l)\in\mathcal{N}_2} G(\theta_k^p, \theta_l^p, \Delta t)\Delta t \Delta\theta_k \Delta\theta_l = O(\Delta t^{2-\alpha})\,.$$

In the third case, we let

$$\mathcal{N}_3 := \{(k,l) \in \mathbf{N}^2 : C'\Delta t^{-\alpha} \leq \theta_k^p + \theta_l^p \leq C''\Delta t^{-1}\}\,;$$

this is the most critical case, where our estimate must be sharpest. In paricular, first observe that the function $1 - \exp(-(\theta_k^p + \theta_l^p)\Delta t)$ varies from $O(\Delta t^{1-\alpha})$ for values $\theta_k^p + \theta_l^p = O(\Delta t^{-\alpha})$ up to $O(1)$ for values $\theta_k^p + \theta_l^p = O(\Delta t^{-1})$ and we exploit this variation to achieve our estimate. Recalling that $1 - \exp(-u) \leq u$, for $0 \leq u \leq 1$, then for $(k,l) \in \mathcal{N}_3$ (taking $C' = 1$) we bound

$$1 - e^{-(\theta_k^p+\theta_l^p)\Delta t} \leq (\theta_k^p + \theta_l^p)\Delta t\,.$$

Thus,

$$\sum_{(k,l)\in\mathcal{N}_3} G(\theta_k^p, \theta_l^p, \Delta t)\Delta t\Delta\theta_k\Delta\theta_l \leq \sum_{(k,l)\in\mathcal{N}_3} m(\theta_k^p)m(\theta_l^p)(\theta_k^p+\theta_l^p)\Delta t^2\Delta\theta_k\Delta\theta_l$$

so that we are left to estimate the quantity on the right-hand side which blows up with a certain rate on Δt as $\Delta t \to 0$. We bound the sum by the corresponding integral, i.e.,

$$\sum_{(k,l)\in\mathcal{N}_3} m(\theta_k^p)m(\theta_l^p)(\theta_k^p+\theta_l^p)\Delta\theta_k\Delta\theta_l \leq \iint\limits_{\substack{\Delta t^{-\alpha}\leq\theta_k^p+\theta_l^p\leq\Delta t^{-1}\\ \theta_j\geq 0}} \frac{\theta_k^p+\theta_l^p}{\max(t_i^{-1},\theta_k^p)\max(t_i^{-1},\theta_l^p)}\mathrm{d}\theta_k\mathrm{d}\theta_l$$

and further divide the right-hand side above into three integrals, I_1, I_2, and I_3, according to the regions

$$\begin{aligned}
R_1 &:= \{\Delta t^{-\alpha} \leq \theta_j^p \leq \Delta t^{-1} \quad \text{for } j = k,l\}\,,\\
R_2 &:= \{\Delta t^{-\alpha} \leq \theta_k^p \leq \Delta t^{-1} \quad \text{and} \quad 0 \leq \theta_l^p \leq \Delta t^{-\alpha}\}\,,\\
R_3 &:= \{\Delta t^{-\alpha} \leq \theta_k^p + \theta_l^p \leq \Delta t^{-1} \quad \text{and} \quad 0 \leq \theta_j^p \leq \Delta t^{-\alpha} \quad \text{for } j = k,l\}\,,
\end{aligned}$$

respectively. Then,

$$\begin{aligned}
I_1 &\leq \iint\limits_{R_1} \left(\theta_l^{-p} + \theta_k^{-p}\right)\mathrm{d}\theta_k\mathrm{d}\theta_l\\
&\leq 2\int\limits_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} \frac{\mathrm{d}\theta_k}{\theta_k^p} = -\frac{\theta_k^{-(p-1)}}{p-1}\Bigg|_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} \lesssim \Delta t^{\alpha(1-1/p)}\,,
\end{aligned}$$

and we note that $\Delta t^{\alpha(1-1/p)} \to 0$ as $\Delta t \to 0$, since $0 < \alpha < 1$ and $p > 2$, so this is not the dominant term. For R_2,

$$\begin{aligned}
I_2 &\lesssim \int\limits_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} \int\limits_{0}^{\Delta t^{-\alpha/p}} \frac{(\theta_k^p+\theta_l^p)}{\max(t_i^{-1},\theta_k^p)\max(t_i^{-1},\theta_l^p)}\mathrm{d}\theta_l\mathrm{d}\theta_k\\
&\lesssim \int\limits_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} \int\limits_{t_i^{-1/p}}^{\Delta t^{-\alpha/p}} \frac{\theta_k^p+\theta_l^p}{\theta_k^p\theta_l^p}\mathrm{d}\theta_l\mathrm{d}\theta_k + \int\limits_{0}^{t_i^{-1/p}} \int\limits_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} \frac{\theta_k^p+\theta_l^p}{\theta_k^p t_i^{-1}}\mathrm{d}\theta_k\mathrm{d}\theta_l\,.
\end{aligned}$$

The first integral is equal to

$$\int_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} \int_{t_i^{-1/p}}^{\Delta t^{-\alpha/p}} (\theta_l^{-p} + \theta_k^{-p}) \mathrm{d}\theta_k \mathrm{d}\theta_l = 2\Delta t^{-\alpha/p} \int_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} \theta_k^{-p} \mathrm{d}\theta_k = \Delta t^{-\alpha/p} \left. \frac{\theta^{-p+1}}{-p+1} \right|_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} = \Delta t^{\alpha(1-2/p)} ,$$

which we note converges to 0 as $\Delta t \to 0$ since $p > 2$, $0 < \alpha < 1$ and therefore this is not the dominant term. On the other hand, the second integral is equal to

$$\int_0^{t_i^{-1/p}} \int_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} \left(\frac{1}{t_i^{-1}} + \frac{\theta_l^p}{\theta_k^p t_i^{-1}} \right) \mathrm{d}\theta_k \mathrm{d}\theta_l = O(\Delta t^{-1/p}) + O(1) \left(\int_{\Delta t^{-\alpha/p}}^{\Delta t^{-1/p}} \theta_k^{-p} \mathrm{d}\theta_k \right) ,$$

which diverges as $\Delta t \to 0$ and therefore this is the dominant term. For R_3,

$$I_3 \lesssim \int_0^{\Delta t^{-\alpha/p}} \int_0^{\Delta t^{-\alpha/p}} \frac{(\theta_k^p + \theta_l^p)}{\max(t_i^{-1}, \theta_k^p) \max(t_i^{-1}, \theta_l^p)} \mathrm{d}\theta_l \mathrm{d}\theta_k ,$$

can be seen to converge to zero as $\Delta t \to 0$, following similar reasoning to I_2, and is therefore not the dominant term. In conclusion, we get that the contribution from $(k,l) \in \mathcal{N}_3$ (after summing over t_i) is

$$\Delta t^{1-1/p} = \Delta t^{H+1/2} ,$$

and it seems that we can take α as close to zero as we want.

In the fourth and final case, we consider

$$\mathcal{N}_4 = \{(k,l) \in N^2 : C'' \Delta t^{-1} \le \theta_k^p + \theta_l^p\} ,$$

where we bound the summand in (5.12),

$$G(\theta_k^p, \theta_l^p, \Delta t) \le m(\theta_k^p) m(\theta_l^p) ,$$

for $(k,l) \in \mathcal{N}_4$. Now estimating the sum by the corresponding integral yields,

$$\sum_{(k,l) \in \mathcal{N}_4} \frac{\Delta\theta_k \Delta\theta_l}{\max(t_i^{-1}, \theta_k^p) \max(t_i^{-1}, \theta_l^p)} \lesssim \iint_{\substack{\theta_k^p + \theta_l^p \ge C'' \Delta t^{-1} \\ \theta_j \ge 0}} \frac{\mathrm{d}\theta_k \mathrm{d}\theta_l}{\max(t_i^{-1}, \theta_k^p) \max(t_i^{-1}, \theta_l^p)} ,$$

and we again consider three integrals, I_1, I_2 and I_3, now corresponding to the regions

$$\begin{aligned} R_1' &:= \{\theta_j^p \ge \Delta t^{-1} \quad \text{for } j = k, l\} , \\ R_2' &:= \{\theta_k^p \ge \Delta t^{-1} \quad \text{and} \quad \theta_l^p \le \Delta t^{-1}\} , \\ R_3' &:= \{\theta_k^p + \theta_l^p \ge \Delta t^{-1} \quad \text{and} \quad \theta_j^p \le \Delta t^{-1} \quad \text{for } j = k, l\} , \end{aligned}$$

respectively. Then,

$$I_1 = \iint_{R_1'} \theta_k^{-p} \theta_l^{-p} \mathrm{d}\theta_k \mathrm{d}\theta_l = \left(\int_{\Delta t^{-1/p}}^{\infty} \theta_k^{-p} \mathrm{d}\theta_k \right)^2 = \left(- \left. \frac{\theta^{-p+1}}{-p+1} \right|_{\Delta t^{-1/p}}^{\infty} \right)^2 = O(\Delta t^{2(1-1/p)}) ,$$

$$I_2 = 2 \iint_{R_2'} \frac{\mathrm{d}\theta_k \mathrm{d}\theta_l}{\theta_k^p \max(t_i^{-1}, \theta_l^p)} = O(\Delta t^{1-1/p}) \underbrace{\left(t_i t_i^{-1/p} + \cdots \right)}_{O(1)} = O(\Delta t^{1-1/p}) ,$$

and

$$I_3 \simeq \iint_{R_3'} \frac{\mathrm{d}\theta_k \mathrm{d}\theta_l}{\max(\theta_k^p, t_i^{-1})\max(\theta_l^p, t_i^{-1})} \lesssim \underbrace{\int_0^{\Delta t^{-1/p}} \frac{\mathrm{d}\theta_l}{\max(\theta_l^p, t_i^{-1})}}_{O(1)} \underbrace{\int_{\frac{1}{2}\Delta t^{-1/p}}^{\Delta t^{-1/p}} \frac{\mathrm{d}\theta_k}{\max(\theta_k^p, t_i^{-1})}}_{O(\Delta t^{(p-1)/p})} = O(\Delta t^{1-1/p}).$$

Then taking the three integrals into account,

$$\sum_{(k,l)\in\mathcal{N}_4} G(\theta_k^p, \theta_l^p, \Delta t)\Delta t \Delta\theta_k \Delta\theta_l = O(\Delta t^{1-1/p})\Delta t,$$

and we observe that the estimate in this region does not depend on α and therefore, recalling $p = 2/(1-2H)$, we obtain $\Delta t^{H+1/2}$ as the rate (after summing over t_i).

Altogether, our full estimate for J_0 is then

$$|J_0| \lesssim C(H, Q, \alpha_{\max}, t_i)\Delta t^{H+3/2} \lesssim O(\Delta t^{H+3/2}), \tag{5.13}$$

for a constant C independent of Δt. In the next section, we consider estimates for the terms J_1 and $\widetilde{J}_1$ in (5.1) and (5.2). Inspired by the observation that we obtain the same rate in (5.10) as in (4.5) with only the constant changing, we first work directly with the fBm view to easily ascertain the rate. However, to obtain the full estimate, we proceed as in this section by utilizing the structure of the affine approximation to make a second round of Taylor expansions and subsequently observing that the coefficients are controlled.

5.2 Estimate for general payoffs: J_1 and $\widetilde{J}_1$ are also $O(\Delta t^{H+3/2})$

For the term $J_{1,0}$ in (5.1),

$$J_{1,0} = -\mathbf{E}\int_{t_i}^{t_{i+1}} D_1\nu(\overline{X}_{t_i}, \boldsymbol{Y}_{t_i}; s)\mathcal{S}(\boldsymbol{Y}_{t_i})\,\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} W_{t_i,s} \mid \mathcal{F}_{t_i}],$$

we first determine the order in Δt which arises from the conditional expectation. Working directly with the increment,

$$\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s} = \widehat{W}^H_{t_i,s}(\widehat{W}^H_s + \widehat{W}^H_{t_i}) = \widehat{W}^H_{t_i,s}(\widehat{W}^H_{t_i,s} + 2\widehat{W}^H_{t_i}),$$

we take the fBm view and express

$$\widehat{W}^H_{t_i,s} \sim W^H_{t_i,s} = \underbrace{\int_0^{t_i} K(s-r) - K(t_i - r)\mathrm{d}W_r}_{:=V^H_{t_i}(s)} + \int_{t_i}^s K(s-r)\mathrm{d}W_r,$$

using the power law kernel, i.e. $K(r) = \sqrt{2H}r^{H-1/2}$, (inverse discrete Laplace transform). We observe

$$\mathbf{E}[(W^H_{t_i,s})^2 W_{t_i,s} \mid \mathcal{F}_{t_i}] = 2V^H_{t_i}(s)\int_{t_i}^s K(s-r)\mathrm{d}r,$$

$$\mathbf{E}[W^H_{t_i,s} W_{t_i,s} \mid \mathcal{F}_{t_i}] = \int_{t_i}^s K(s-r)\mathrm{d}r,$$

and

$$|\mathbf{E}[W^H_{t_i} V^H_{t_i}(s)]| = \left|\int_0^{t_i} K(t_i - r)K(s-r) - K^2(t_i - r)\mathrm{d}r\right| \le \left|\int_0^{t_i} K^2(t_i - r)\mathrm{d}r\right| \sim t_i^{2H} \le 1.$$

Then,

$$\begin{aligned} J_{1,0} &= -\mathbf{E}\int_{t_i}^{t_{i+1}} D_1\nu(\overline{X}_{t_i},\boldsymbol{Y}_{t_i};s)\widehat{W}^H_{t_i}\,\mathbf{E}[\widehat{W}^H_{t_i,s}(\widehat{W}^H_{t_i,s}+2\widehat{W}^H_{t_i})W_{t_i,s}\mid\mathcal{F}_{t_i}]\mathrm{d}s \\ &\cong -\mathbf{E}\int_{t_i}^{t_{i+1}} D_1\nu(\overline{X}_{t_i},\boldsymbol{Y}_{t_i};s)W^H_{t_i}(\mathbf{E}[(W^H_{t_i,s})^2W_{t_i,s}\mid\mathcal{F}_{t_i}]+2W^H_{t_i}\,\mathbf{E}[W^H_{t_i,s}W_{t_i,s}\mid\mathcal{F}_{t_i}])\mathrm{d}s \\ &\cong -2\,\mathbf{E}\int_{t_i}^{t_{i+1}} D_1\nu(\overline{X}_{t_i},\boldsymbol{Y}_{t_i};s)W^H_{t_i}(V^H_{t_i}(s)+W^H_{t_i})\underbrace{\int_{t_i}^{s}K(s-r)\mathrm{d}r}_{\propto(s-t_i)^{H+1/2}}\mathrm{d}s\,, \end{aligned} \tag{5.14}$$

which suggests $J_{1,0}=O(\Delta t^{H+3/2})$ provided the coefficient is controlled. The order of $J_{1,0}$ in Δt is asymptotically exact in (5.14), i.e. there is only estimation of the constant.

In general, $D_1\nu$ is random and depends on $\boldsymbol{Y}_{t_i}$. This changes the coefficient in the estimate but not the order in Δt, as was observed by comparing (5.13) to (4.5). Following Section 5.1, we proceed to estimate the coefficient by making a second round of Taylor expansions by introducing an auxiliary function f. In this spirit, we start directly from (5.14), noting that

$$V^H_{t_i}(s)\sim\widehat{V}^H_{t_i}(s):=c_H\sum_{l=1}^{N_L}Y^l_{t_i}(e^{-\theta^p_l(s-t_i)}-1)\Delta\theta_l\,,$$

since

$$\begin{aligned} \int_0^{t_i}K(s-r)\mathrm{d}W_r &= \int_0^{t_i}\sqrt{2H}(s-r)^{H-1/2}\mathrm{d}W_r \\ &= \int_0^{t_i}\frac{\sqrt{2H}}{\Gamma(\frac{1}{2}-H)}\int_0^\infty\theta^{-(H+1/2)}e^{-\theta(s-r)}\mathrm{d}\theta\mathrm{d}W_r \\ &= \widetilde{c}_H\int_0^\infty\theta^{-(H+1/2)}e^{-\theta(s-t_i)}\int_0^{t_i}e^{-\theta(t_i-r)}\mathrm{d}W_r\mathrm{d}\theta \\ &= \widetilde{c}_H\int_0^\infty\theta^{-(H+1/2)}e^{-\theta(s-t_i)}\widetilde{Y}_{t_i}\mathrm{d}\theta \\ &\approx c_H\sum_{l=1}^{N_L}Y^l_{t_i}e^{-\theta^p_l(s-t_i)}\Delta\theta_l\,. \end{aligned}$$

Rewriting $J_{1,0}$ directly in terms of the extended variables, we obtain

$$\begin{aligned} J_{1,0} &= -2\,\mathbf{E}\int_{t_i}^{t_{i+1}} D_1\nu(\overline{X}_{t_i},\boldsymbol{Y}_{t_i};s)\widehat{W}^H_{t_i}(\widehat{V}^H_{t_i}+\widehat{W}^H_{t_i})g(s)\mathrm{d}s \\ &= -2c_H^2\sum_{k,l=1}^{N_L}\int_{t_i}^{t_{i+1}}\mathbf{E}\big[D_1\nu(\overline{X}_{t_i},\boldsymbol{Y}_{t_i};s)Y^k_{t_i}Y^l_{t_i}\big]e^{-\theta^p_l(s-t_i)}g(s)\mathrm{d}s\Delta\theta_k\Delta\theta_l\,, \end{aligned}$$

where

$$g(s):=\int_{t_i}^{s}K(s-r)\mathrm{d}r\propto(s-t_i)^{H+1/2}\,.$$

Recalling that $\nu(\overline{\boldsymbol{Z}}_{t_i},s)=\nu(\overline{X}_{t_i},\boldsymbol{Y}_{t_i};s)$ is a deterministic function of $\overline{\boldsymbol{Z}}_{t_i}$, we further write

$$\mathbf{E}[D_1\nu(\overline{X}_{t_i},\boldsymbol{Y}_{t_i};s)Y^k_{t_i}Y^l_{t_i}]=\mathbf{E}\big[\mathbf{E}[D_1\nu(\overline{X}_{t_i},\boldsymbol{Y}_{t_i};s)\mid Y^k_{t_i},Y^l_{t_i}]Y^k_{t_i},Y^l_{t_i}\big]\,,$$

where we define a new auxiliary function (here φ is already higher order compared to (5.5))

$$f^{kl}_s(Y^k_{t_i},Y^l_{t_i}):=\mathbf{E}\big[D_1\nu(\overline{X}_{t_i},\boldsymbol{Y}_{t_i};s)\mid Y^k_{t_i},Y^l_{t_i}\big]\,.$$

We Taylor expand f^{kl} at zero, yielding

$$J_{1,0} = -2c_H^2\Bigg(\sum_{k,l=1}^{N_L} \mathbf{E}[Y^k_{t_i} Y^l_{t_i}] \int_{t_i}^{t_{i+1}} f^{kl}_s(\mathbf{0}) e^{-\theta_l^p(s-t_i)} g(s)\mathrm{d}s \Delta\theta_k\Delta\theta_l$$
$$+ \sum_{\alpha\in\mathcal{J}_{\alpha_{\max}}} \frac{1}{|\alpha|!} \sum_{k,l=1}^{N_L} \mathbf{E}\big[(Y^k_{t_i})^{\alpha_1+1}(Y^l_{t_i})^{\alpha_2+1}\big] \int_{t_i}^{t_{i+1}} \partial^\alpha_{kl} f^{kl}_s(\mathbf{0}) e^{-\theta_l^p(s-t_i)} g(s)\mathrm{d}s\Delta\theta_k\Delta\theta_l$$
$$+ \sum_{k,l=1}^{N_L} \mathbf{E}\big[R_{\alpha_{\max}}(Y^k_{t_i}, Y^l_{t_i}) Y^k_{t_i} Y^l_{t_i}\big] \int_{t_i}^{t_{i+1}} e^{-\theta_l^p(s-t_i)} g(s)\mathrm{d}s\Delta\theta_k\Delta\theta_l\Bigg)$$

where $\mathcal{J}_{\alpha_{\max}} = \{\alpha = (\alpha_1,\alpha_2) : 1 \le |\alpha| < \alpha_{\max}\}$ and the remainder $R_{\alpha_{\max}}$ is given in the Lagrange form as in (5.7). Since f^{kl} is bounded by φ (see Lemma B.1 in Appendix B), we write the bound in terms of Q, as in (5.9), to obtain

$$|J_{1,0}| \lesssim C(H,Q,\alpha_{\max}) t_i^{1-2/p} \Delta t^{H+3/2} \lesssim O(\Delta t^{H+3/2}),$$

where we use Isserlis' theorem and the explicit form of the covariances, e.g.,

$$\begin{aligned}\sum_{k,l=1}^{N_L} \mathbf{E}[Y^k_{t_i} Y^l_{t_i}]\Delta\theta_k\Delta\theta_l &= \sum_{k,l=1}^{N_L} \frac{\Delta\theta_k\Delta\theta_l}{\theta_k^p+\theta_l^p}(1-e^{-(\theta_k^p+\theta_l^p)t_i}) \\ &\le \frac{2\pi}{4}\int_0^\infty \theta^{1-p}(1-e^{-\theta^p t_i})\mathrm{d}\theta \\ &\le \frac{2\pi}{4}\frac{1}{p-1}\Gamma(\tfrac{2}{p}) t_i^{1-2/p},\end{aligned}$$

to observe that the sums over θ_k and θ_l converge independently of L, N_L.

The term $J_{1,1}$ includes increments of OU processes arising from the extended variables. Taking the fBm view using the power law kernel,

$$\begin{aligned} J_{1,1} &= -\sum_{j=1}^{N_L} \mathbf{E}\int_{t_i}^{t_{i+1}} D_{j+1}\nu(\overline{X}_{t_i}, \mathbf{Y}_{t_i}; s)\, \mathbf{E}[\widehat{W}^H_{t_i,s}(\widehat{W}^H_{t_i,s} + 2\widehat{W}^H_{t_i}) Y^j_{t_i,s} \mid \mathcal{F}_{t_i}]\mathrm{d}s \\ &\cong -\sum_{j=1}^{N_L} \mathbf{E}\int_{t_i}^{t_{i+1}} D_{j+1}\nu(\overline{X}_{t_i}, \mathbf{Y}_{t_i}; s)\big((V^H_{t_i}(s))^2 + 2W^H_{t_i} V^H_{t_i}(s)\big) Y^j_{t_i}(e^{-\theta_j^p(s-t_i)}-1)\mathrm{d}s \qquad (5.15a) \\ &\quad -\sum_{j=1}^{N_L} \mathbf{E}\int_{t_i}^{t_{i+1}} D_{j+1}\nu(\overline{X}_{t_i}, \mathbf{Y}_{t_i}; s)(2V^H_{t_i}(s) + 2W^H_{t_i}) \int_{t_i}^s e^{-\theta_j^p(s-r)} K(s-r)\mathrm{d}r\mathrm{d}s \qquad (5.15b) \\ &\quad -\sum_{j=1}^{N_L} \mathbf{E}\int_{t_i}^{t_{i+1}} D_{j+1}\nu(\overline{X}_{t_i}, \mathbf{Y}_{t_i}; s) Y^j_{t_i}(e^{-\theta_j^p(s-t_i)}-1) \int_{t_i}^s K^2(s-r)\mathrm{d}r\mathrm{d}s, \qquad (5.15c)\end{aligned}$$

since

$$\begin{aligned}\mathbf{E}[(W^H_{t_i,s})^2 Y^j_{t_i,s} \mid \mathcal{F}_{t_i}] = Y^j_{t_i}(e^{-\theta_j^p(s-t_i)}-1)\int_{t_i}^s K^2(s-r)\mathrm{d}r + 2V^H_{t_i}(s)\int_{t_i}^s K(s-r)e^{-\theta_j^p(s-r)}\mathrm{d}r \\ +(V^H_{t_i}(s))^2 Y^j_{t_i}(e^{-\theta_j^p(s-t_i)}-1)\end{aligned}$$

and

$$\mathbf{E}[W^H_{t_i,s} Y^j_{t_i,s} \mid \mathcal{F}_{t_i}] = \int_{t_i}^s e^{-\theta_j^p(s-r)} K(s-r)\mathrm{d}r + V^H_{t_i}(s) Y^j_{t_i}(e^{-\theta_j^p(s-t_i)}-1).$$

We examine the contributions to the rate in Δt from each of the terms (5.15a) to (5.15c); after integrating over s we find that (5.15a) yields $O(\Delta t^2)$, (5.15b) yields $O(\Delta t^{H+3/2})$, and (5.15c) yields

$O(\Delta t^{2H+2})$. Since (5.15a) and (5.15c) are higher order in Δt we examine only (5.15b), where we note

$$\begin{aligned}\int_{t_i}^{s} e^{\theta_j^p(s-t_i)} K(s-r)\mathrm{d}r &= e^{-\theta_j^p s}\sqrt{2H}\int_{t_i}^{s} e^{\theta_j^p r}(s-r)^{H-1/2}\mathrm{d}r \\ &= \sqrt{2H}\theta_j^{-(H+1/2)p}\Big[\Gamma\big(H+\tfrac{1}{2},(s-r)\theta_j^p\big)\Big]_{r=t_i}^{r=s} \\ &= \sqrt{2H}\theta_j^{-(H+1/2)p}\Big[e^{s-r}(s-r)^{H+1/2}\Gamma(H+1/2,\theta_j^p)\Big]_{r=t_i}^{r=s} \\ &= -\sqrt{2H}\theta_j^{-(H+1/2)p}\Gamma(H+\tfrac{1}{2},\theta_j^p)e^{s-t_i}(s-t_i)^{H+1/2}\end{aligned} \tag{5.16}$$

by a change of variable in the argument of the incomplete gamma function where

$$q := (H+1/2)p = \frac{2H+1}{1-2H} > 1\,, \quad H \in (0,1/2)\,.$$

This suggests $J_{1,1} = O(\Delta t^{H+3/2})$ since the sum over θ_j converges (also helpful to note $\Gamma(H+\frac{1}{2},\theta_j^p) \to 0$ as $\theta_j \to \infty$).

For the full estimate of $J_{1,1}$, we note

$$D_{j+1}\nu(\overline{X}_{t_i}\boldsymbol{Y}_{t_i}\,;s) = c_H\Delta\theta_j\,\mathbf{E}\big[\varphi^{(3)}(\widehat{X}_T)M_{s,T}^j \mid (\widehat{X}_s,\boldsymbol{Y}_s) = (\overline{X}_{t_i},\boldsymbol{Y}_{t_i})\big]\,,$$

where we assume that $M_{s,T}^j < \infty$ (see (3.22) for definition of M). Omitting the higher order terms in Δt in (5.15a) and (5.15c), we return directly to (5.15b),

$$J_{1,1} = -2c_H\sum_{j=1}^{N_L}\mathbf{E}\int_{t_i}^{t_{i+1}}\mathbf{E}\big[\varphi^{(3)}(\widehat{X}_T)M_{s,T}^j \mid (\widehat{X}_s,\boldsymbol{Y}_s) = (\overline{X}_{t_i},\boldsymbol{Y}_{t_i})\big](\widehat{V}_{t_i}^H(s)+\widehat{W}_{t_i}^H)g(s)h(\theta_j)\mathrm{d}s\Delta\theta_j\,,$$

where we let

$$g(s) := -\sqrt{2H}e^{s-t_i}(s-t_i)^{H+1/2}$$

and

$$h(\theta_j) := \theta_j^{-q}\Gamma(H+\tfrac{1}{2},\theta_j^p)\,.$$

We define a new auxiliary function

$$f_s^k(Y_{t_i}^k) := \mathbf{E}\big[\mathbf{E}[\varphi^{(3)}(\widehat{X}_T)M_{s,T}^j \mid (\widehat{X}_s,\boldsymbol{Y}_s) = (\overline{X}_{t_i},\boldsymbol{Y}_{t_i})] \mid Y_{t_i}^k\big]\,,$$

where, although the inner expectation depends implicitly on j, here the index k refers to the general component $Y_{t_i}^k$ that we are conditioning against. Expanding f^k at zero, we find

$$\begin{aligned}J_{1,1} &= -2c_H^2\sum_{j,k=1}^{N_L}h(\theta_j)\int_{t_i}^{t_{i+1}}\mathbf{E}\big[f_s^k(Y_{t_i}^k)Y_{t_i}^k\big]e^{-\theta_k^p(s-t_i)}g(s)\mathrm{d}s\Delta\theta_j\Delta\theta_k \\ &= -2c_H^2\sum_{\alpha_1=1}^{\alpha_{\max}-1}\frac{1}{\alpha_1!}\sum_{j,k=1}^{N_L}h(\theta_j)\,\mathbf{E}\big[(Y_{t_i}^k)^{\alpha_1+1}\big]e^{-\theta_k^p(s-t_i)}\int_{t_i}^{t_{i+1}}\partial_k^{\alpha_1}f_s^k(\mathbf{0})g(s)\mathrm{d}s\Delta\theta_j\Delta\theta_k \\ &\quad - 2c_H^2\sum_{j,k=1}^{N_L}h(\theta_j)\,\mathbf{E}\big[R_{\alpha_{\max}}(Y_{t_i}^k)Y_{t_i}^k\big]e^{-\theta_k^p(s-t_i)}\int_{t_i}^{t_{i+1}}g(s)\mathrm{d}s\Delta\theta_j\Delta\theta_k\,.\end{aligned}$$

Following from the boundedness of f and its derivatives (as in (5.9)), we obtain the full estimate for $J_{1,1}$,

$$|J_{1,1}| \lesssim C(H,Q,\alpha_{\max})t_i^{1-2/p}\Delta t^{H+3/2} \lesssim O(\Delta t^{H+3/2})\,,$$

where we use Isserlis' theorem and the representation of the covariance for the extended variables to determine the coefficient depending on t_i (which we note is also summable over i).

The estimation of $\widetilde{J}_{1,0}$ and $\widetilde{J}_{1,1}$ follows the program above. For $\widetilde{J}_{1,0}$ we begin by taking the fBm view with the kernel K,

$$\begin{aligned}|\widetilde{J}_{1,0}| &= \Big|\mathbf{E}\int_{t_i}^{t_{i+1}} D_1\widetilde{\nu}(\overline{X}_{t_i}, \boldsymbol{Y}_{t_i}\,; s) W^H_{t_i}\, \mathbf{E}[W^H_{t_i,s} W_{t_i,s} \mid \mathcal{F}_{t_i}]\mathrm{d}s\Big| \\ &= c_H\Big|\sum_{k=1}^{N_L}\int_{t_i}^{t_{i+1}} D_1\widetilde{\nu}(\overline{X}_{t_i}, \boldsymbol{Y}_{t_i}\,; s) W^H_{t_i} \underbrace{\int_{t_i}^{s} K(s-r)\mathrm{d}r}_{(s-t_i)^{H+1/2}}\,\mathrm{d}s\Big|\end{aligned}$$

which suggests $\widetilde{J}_{1,0} = O(\Delta t^{H+3/2})$. We recall that

$$D_1\widetilde{\nu}(\overline{\boldsymbol{Z}}_{t_i}, s) = \mathbf{E}\big[\varphi^{(3)}(\widehat{X}_T)\big(c_H \textstyle\sum_l M^l_{s,T}\Delta\theta_l\big) \mid \boldsymbol{Z}_s = \overline{\boldsymbol{Z}}_{t_i}\big]\,,$$

where we assume that

$$c_H \sum_{l=1}^{N_L} M^l_{s,T}\Delta\theta_l < \infty\,,$$

and then define an auxiliary function

$$f^k_s(Y^k_{t_i}) := \mathbf{E}\Big[\mathbf{E}\big[\varphi^{(3)}(\widehat{X}_T)\big(c_H \textstyle\sum_l M^l_{s,T}\Delta\theta_l\big) \mid (\widehat{X}_s, \boldsymbol{Y}_s) = (\overline{X}_{t_i}, \boldsymbol{Y}_{t_i})\big] \mid Y^k_{t_i}\Big]\,.$$

Finally, Taylor expanding f^k_s at zero we encounter terms similar to those estimated previously yielding the full estimate,

$$|\widetilde{J}_{1,0}| \lesssim O(\Delta t^{H+3/2})\,.$$

Likewise for $\widetilde{J}_{1,1}$, taking the fBm view with the conditional expectation term $\mathbf{E}[\widehat{W}^H_{t_i,s} Y^j_{t_i,s} \mid \mathcal{F}_{t_i}]$ yields

$$|\widetilde{J}_{1,1}| = \Big|\sum_{j=1}^{N_L}\mathbf{E}\int_{t_i}^{t_{i+1}} D_{j+1}\widetilde{\nu}(\overline{X}_{t_i}, \boldsymbol{Y}_{t_i}; s)\int_{t_i}^{s} e^{-\theta^p_j(s-r)}K(s-r)\mathrm{d}r\mathrm{d}s\Big|\,,$$

where the estimate (5.16) (encountered in $J_{1,1}$) suggest the rate $J_{1,1} = O(\Delta t^{H+3/2})$. For the full estimate, we recall that

$$D_{j+1}\widetilde{\nu}(\overline{\boldsymbol{Z}}_{t_i}, s) = \mathbf{E}\big[\varphi^{(3)}(\widehat{X}_T)\big(c_H \textstyle\sum_l M^l_{s,T}\Delta\theta_l\big)^2 \mid \boldsymbol{Z}_s = \overline{\boldsymbol{Z}}_{t_i}\big]$$

and expand in a second round of Taylor expansions for an appropriate auxiliary function thereby obtaining

$$|\widetilde{J}_{1,1}| \lesssim O(\Delta t^{H+3/2})\,.$$

Taken together, the estimates Section 5.2 imply that the terms corresponding to first order derivatives of ν and $\widetilde{\nu}$ in (5.1) and (5.2), respectively, yields

$$|J_1| + |\widetilde{J}_1| = O(\Delta t^{H+3/2})\,,$$

using the notation in (5.3).

5.3 Estimate for general payoffs: remaining terms are higher order

Additional terms (5.1) and (5.2) appearing in the expansion are higher order than $H+3/2$. For example, We find $J_{2,0} = O(\Delta t^{2+2H})$ which can be seen by once again taking the fractional view,

$$\begin{aligned}\mathbf{E}[\mathcal{S}(\boldsymbol{Y}.)^2_{t_i,s}(W_{t_i,s})^2 \mid \mathcal{F}_{t_i}] &= \mathbf{E}[\widehat{W}^H_{t_i,s}(\widehat{W}^H_{t_i,s} + 2\widehat{W}^H_{t_i})(W_{t_i,s})^2] \\ &= \mathbf{E}[(\widehat{W}^H_{t_i,s})^2(W_{t_i,s})^2] + 2\widehat{W}^H_{t_i}\underbrace{\mathbf{E}[\widehat{W}^H_{t_i,s}(W_{t_i,s})^2]}_{\text{Isserlis'} \implies 0} \\ &= \mathbf{E}[(\widehat{W}^H_{t_i,s})^2]\,\mathbf{E}[(W_{t_i,s})^2] + 2(\mathbf{E}[\widehat{W}^H_{t_i,s}W_{t_i,s}])^2 \\ &= \Delta t^{2H}\Delta t - 2(\Delta t^{H+1/2})^2 =: g(\Delta t)\,,\end{aligned}$$

where $g(s) \propto (s-t_i)^{2H+1}$ and then expanding in a second round of Taylor expansions for a suitable auxiliary function and checking the control of the coefficient with respect to θ,

$$|J_{2,0}| \lesssim \frac{1}{2}\Big|\mathbf{E}\Big[(W^H_{t_i})^2\int_0^{\Delta t} D_{11}\nu(\overline{\boldsymbol{Z}}_{t_i}, s+t_i)s^{2H+1}\mathrm{d}s\Big]\Big| = O(\Delta t^{2+2H})\,.$$

The term $\widetilde{J}_{2,0}$ vanishes,

$$|\widetilde{J}_{2,0}| = \Big|\frac{1}{2}\,\mathbf{E}\Big[\int_{t_i}^{t_{i+1}} D_{11}\widetilde{\nu}(\overline{\boldsymbol{Z}}_{t_i}, s)(\widehat{W}^H_{t_i})^2\underbrace{\mathbf{E}[\widehat{W}^H_{t_i,s}(W_{t_i,s})^2]}_{\text{Isserlis'} \implies 0}\,\mathrm{d}s\Big]\Big| = 0\,,$$

by Isserlis' theorem (although one could argue that the integrand is formally contributing rate $H+2$ here for $\widetilde{J}_{2,0}$). The terms involving increments of the OU processes yield similar results (constants obtained would be better behaved owing to the additional decay).

5.4 Closure argument to finish proof of Theorem 2.1

Thus far we have estimated the first few terms (5.1) and (5.2) arising from the Taylor expansion in powers of Δt. Returning to the telescoping sum (3.17), we summarize our estimates up to order κ,

$$\begin{aligned}\mathrm{Err}(T,\Delta t) &= \mathbf{E}[\varphi(\widehat{X}_T) - \varphi(\overline{X}_T)] \\ &= \sum_{i=0}^{n-1}\left(\sum_{k=0}^{\kappa-1}\Big(J_k + \widetilde{J}_k\Big) + \mathcal{R}_\kappa(\nu) + \mathcal{R}_\kappa(\widetilde{\nu})\right) \\ &\le \sum_{i=0}^{n-1}\Big(C_1(t_i)\Delta t^{H+3/2} + C_0(t_i)\Delta t^2 + O(\Delta t^{H+2}) + \mathcal{R}_\kappa(\nu) + \mathcal{R}_\kappa(\widetilde{\nu})\Big)\,,\end{aligned} \tag{5.17}$$

using the notation in (5.3). Here the constants C_k, for $k = 0,\ldots,\kappa-1$, depending on t_i, H, and Q, are the coefficients that appear in the estimates (5.13) and Section 5.2, etc., for $J.$ and likewise for $\widetilde{J}.$. Importantly, each of these constants was obtained independently of N_L, L, that is, independently of the choice of parameter θ arising in the Markovian extended variable state space formulation. Moreover, each constant was summable in t_i. Interchanging the order of summation in (5.17) we obtain,

$$\begin{aligned}\mathrm{Err}(T,\Delta t,\varphi) &\lesssim C'_1\Delta t^{H+1/2} + C'_0\Delta t + O(\Delta t^{H+1}) \\ &\quad - \sum_{i=1}^{n-1}\mathbf{E}\int_{t_i}^{t_{i+1}}\Big(\mathbf{E}[\mathcal{S}(\boldsymbol{Y}.)^2_{t_i,s}\mathcal{R}_\kappa(\nu) \mid \mathcal{F}_{t_i}] + \mathbf{E}[\mathcal{S}(\boldsymbol{Y}.)_{t_i,s}\mathcal{R}_\kappa(\widetilde{\nu}) \mid \mathcal{F}_{t_i}]\Big)\mathrm{d}s\,,\end{aligned} \tag{5.18}$$

with new constants

$$C'_j = \sum_{i=1}^{n-1} C_j(t_i)\Delta t\,.$$

The remainders have a integral form (3.26), and thus the conditional expectations in (5.18) are,

$$\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}\mathcal{R}_\kappa(\nu) \mid \mathcal{F}_{t_i}] = \frac{1}{\kappa!}\sum_{|\beta|=\kappa}(\widehat{W}^H_{t_i})^{\beta_1}\,\mathbf{E}\left[(\widehat{W}^H_{t_i,s})^2(W_{t_i,s})^{\beta_1}(\boldsymbol{Y}_{t_i,s})^{\hat{\beta}}\int_0^1 D^\beta\nu(\boldsymbol{\xi}_\tau,s)\mathrm{d}\tau\right]$$

and

$$\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)_{t_i,s}\mathcal{R}_\kappa(\widetilde{\nu}) \mid \mathcal{F}_{t_i}] = \frac{1}{\kappa!}\sum_{|\beta|=\kappa}(\widehat{W}^H_{t_i})^{\beta_1}\,\mathbf{E}\left[\widehat{W}^H_{t_i,s}(W_{t_i,s})^{\beta_1}(\boldsymbol{Y}_{t_i,s})^{\hat{\beta}}\int_0^1 D^\beta\widetilde{\nu}(\boldsymbol{\xi}_\tau,s)\mathrm{d}\tau\right],$$

where

$$\mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)^2_{t_i,s}\mathcal{R}_\kappa(\nu) \mid \mathcal{F}_{t_i}] \sim (\widehat{W}^H_{t_i,s})^\kappa \qquad \text{and} \qquad \mathbf{E}[\mathcal{S}(\boldsymbol{Y}_\cdot)_{t_i,s}\mathcal{R}_\kappa(\widetilde{\nu}) \mid \mathcal{F}_{t_i}] \sim (\widehat{W}^H_{t_i,s})^\kappa .$$

We recall that

$$\mathbf{E}\,|W^H_{t_i,t_{i+1}}|^\gamma \lesssim \Delta t^{\gamma H},$$

by the Hölder continuity of the sample paths. Then by applying Cauchy–Schwarz, we obtain in (5.18) that the remainder terms yield $O(\Delta t^{\kappa H})$. Thus, κ can be chosen such that $\kappa > \frac{1}{H}$ yields a large but finite expansion for general payoff functions φ. Then the error is is given by

$$\mathrm{Err}(T,\Delta t,\varphi) \lesssim C_1'\Delta t^{H+1/2} + C_0'\Delta t + O(\Delta t^{H+1}),$$

with weak error rate $H + 1/2$ where all the coefficients are controlled.

Remark 5.1 (Kernel). In the proof of the main result Theorem 2.1 and of Theorem 4.1, the specific form of K in Section 3.1 and (1.2) is not relevant and the spirit of the proof follows with any relevant L^2 kernel where the integrability conditions need to be checked.

6 Conclusions and outlook

Rough stochastic volatility modles are increasingly popular for option pricing in quantitative finance. On the one hand, the rough stochastic volatility overcomes empirical challenges to deliver predictions consistent with observed market data. On the other hand, the non-Markovian nature of the fractional Brownian motion (fBm) driver is an impediment to both theory and numerics. Despite the widespread use of discretization-based simulation methods for option pricing under the rough Bergomi model and the rough Stein–Stein model, few works have studied the weak convergence rates that underpin this practice.

For the rough Stein–Stein model, which treats the volatility as a linear function of the driving fractional Brownian motion, we prove that the weak convergence of the Euler scheme depends on the Hurst parameter H of the fBm driver and is weak rate $H + 1/2$ for general payoff functions (see Theorems 1.1 and 2.1). Strong numerical evidence is provided to support our theory. Our proof relies on Taylor expansions for an extended variable system that is derived from an affine Markovian approximation of the fBm drive. Our novel approach also yields insights into unexpected behavior (that we suspect has contributed to the consternation among experts regarding the weak rate, as remarked in the footnote in Section 1). In particular, the expansions (see Theorem 4.1) easily explain the better weak rate 1 obtained for quadratic payoffs (see Lemma 4.2). This last point leads us to conjecture that the rate of convergence for payoff functions well approximated by quadratic polynomials, as seen from the law of the solution, may be hard to distinguish from rate 1 as illustrated in Figure 3. As stated in Remark 2.2, we do not doubt that the proof of Theorem 2.1 can be extended to nonlinear rough volatility models such as the rough Bergomi model and this is the subject of ongoing work.

References

[1] E. Abi Jaber. *The characteristic function of Gaussian stochastic volatility models: an analytic expression*. HAL Id hal-02946146v2. 2021. URL: https://hal.archives-ouvertes.fr/hal-02946146v2.

[2] E. Abi Jaber and O. El Euch. "Multifactor approximation of rough volatility models". In: *SIAM Journal on Financial Mathematics* 10.2 (2019), pp. 309–349. DOI: 10.1137/18M1170236.

[3] C. Bayer, P. K. Friz, A. Gulisashvili, B. Horvath, and B. Stemper. "Short-time near-the-money skew in rough fractional volatility models". In: *Quant. Financ.* 19.5 (2019), pp. 779–798. DOI: 10.1080/14697688.2018.1529420.

[4] C. Bayer, P. K. Friz, P. Gassiat, J. Martin, and B. Stemper. "A regularity structure for rough volatility". In: *Mathematical Finance* 30.3 (2020), pp. 782–832. DOI: 10.1111/mafi.12233.

[5] C. Bayer, P. K. Friz, and J. Gatheral. "Pricing under rough volatility". In: *Quant. Financ.* 16.6 (2016), pp. 887–904. DOI: 10.1080/14697688.2015.1099717.

[6] C. Bayer, C. B. Hammouda, and R. Tempone. "Hierarchical adaptive sparse grids and quasi-Monte Carlo for option pricing under the rough Bergomi model". In: *Quant. Financ.* 0.0 (2020), pp. 1–17. DOI: 10.1080/14697688.2020.1744700.

[7] C. Bayer, J. Qiu, and Y. Yao. "Pricing options under rough volatility with backward SPDEs". In: *arXiv preprint arXiv:2008.01241* (2020).

[8] M. Bennedsen, A. Lunde, and M. S. Pakkanen. "Decoupling the short-and long-term behavior of stochastic volatility". July 2017. arXiv: 1610.00332.

[9] M. Bennedsen, A. Lunde, and M. S. Pakkanen. "Hybrid scheme for Brownian semistationary processes". In: *Finance Stoch.* 21.4 (2017), pp. 931–965. DOI: 10.1007/s00780-017-0335-5.

[10] P. Carmona and L. Coutin. "Fractional Brownian Motion and the Markov Property". In: *Electron. Commun. Probab.* 3 (1998), pp. 95–107. DOI: 10.1214/ECP.v3-998.

[11] P. Carmona, L. Coutin, and G. Montseny. "Approximation of Some Gaussian Processes". In: *Stat. Inference Stoch. Process.* 3.1 (2000), pp. 161–171. DOI: 10.1023/A:1009999518898.

[12] G. Didier, S. A. McKinley, D. B. Hill, and J. Fricks. "Statistical challenges in microrheology". In: *J. Time Series Anal.* 33.5 (2012), pp. 724–743. DOI: 10.1111/j.1467-9892.2012.00792.x.

[13] O. El Euch and M. Rosenbaum. "The characteristic function of rough Heston models". In: *Mathematical Finance* 29.1 (2019), pp. 3–38. DOI: 10.1111/mafi.12173.

[14] M. Fukasawa, T. Takabatake, and R. Westphal. "Is Volatility Rough?" May 2019. arXiv: 1905.04852.

[15] J. Gatheral. *Fractional volatility models*. https://mfe.baruch.cuny.edu/wp-content/uploads/2012/09/FractionalVolatility2014BBQ.pdf. Presentation in the Bloomberg BBQ seminar. 2014.

[16] J. Gatheral, T. Jaisson, and M. Rosenbaum. "Volatility is Rough". In: *Quant. Financ.* 18.6 (2018), pp. 933–949. DOI: 10.1080/14697688.2017.1393551.

[17] E. J. Hall, M. A. Katsoulakis, and L. Rey-Bellet. "Uncertainty quantification for generalized Langevin dynamics". In: *J. Chem. Phys.* 145.22 (2016), p. 224108. DOI: 10.1063/1.4971433. URL: http://doi.org/bwdv.

[18] P. Harms. "Strong convergence rates for Markovian representations of fractional Brownian motion". Feb. 2019. arXiv: 1902.01471.

[19] P. Harms and D. Stefanovits. “Affine representations of fractional processes with applications in mathematical finance”. In: *Stochastic Process. Appl.* 129.4 (2019), pp. 1185 –1228. DOI: 10.1016/j.spa.2018.04.010. URL: http://www.sciencedirect.com/science/article/pii/S0304414918301418.

[20] B. Horvath, A. J. Jacquier, and A. Muguruza. “Functional central limit theorems for rough volatility”. May 2019. arXiv: 1711.03078.

[21] J. Huntley. *Generation of Random Variates*. MATLAB Central File Exchange. 2020. URL: https://www.mathworks.com/matlabcentral/fileexchange/35008-generation-of-random-variates (visited on 01/01/2020).

[22] A. J. Jacquier and M. Oumgari. “Deep PPDEs for rough local stochastic volatility”. 2019. arXiv: 1906.02551.

[23] A. A. Muravlev. “Representation of a fractional Brownian motion in terms of an infinite-dimensional Ornstein-Uhlenbeck process”. In: *Russian Math. Surveys* 66.2 (Apr. 2011), pp. 439–441. DOI: 10.1070/rm2011v066n02abeh004746.

[24] A. Neuenkirch. Personal communication. 2017.

[25] A. Neuenkirch and T. Shalaiko. “The order barrier for strong approximation of rough volatility models”. June 2016. arXiv: 1606.03854.

[26] L. Rey-Bellet. “Open classical systems”. In: *Open Quantum Systems II*. Ed. by S. Attal, A. Joye, and C. Pillet. Vol. 1881. Lecture Notes in Mathematics. Berlin: Springer, 2006, pp. 41–78. DOI: 10.1007/3-540-33966-3_2.

[27] M. Romano and N. Touzi. “Contingent claims and market completeness in a stochastic volatility model”. In: *Mathematical Finance* 7.4 (1997), pp. 399–412. DOI: 10.1111/1467-9965.00038.

[28] R. Schöbel and J. Zhu. “Stochastic Volatility With an Ornstein-Uhlenbeck Process: An Extension.” In: *European Finance Review* 3.1 (1999), pp. 23 –46. URL: http://search.ebscohost.com.libezproxy.dundee.ac.uk/login.aspx?direct=true&db=bth&AN=18657649&authtype=shib&site=ehost-live&scope=site.

[29] E. M. Stein and J. C. Stein. “Stock Price Distributions with Stochastic Volatility: An Analytic Approach”. In: *The Review of Financial Studies* 4.4 (1991), pp. 727–752. URL: http://www.jstor.org/stable/2962156.

[30] D. Talay and L. Tubaro. “Expansion of the global error for numerical schemes solving stochastic differential equations”. In: *Stochastic Anal. Appl.* 8.4 (1990), pp. 483–509. DOI: 10.1080/07362999008809220.

A Details of numerical implementation

We compute the weak error (here we assume $T = 1$)

$$\left|\mathbf{E}[\varphi(X_T^{ref}) - \varphi(X_T^{\Delta t})]\right|,$$

for various payoff functions, φ, by the left-point scheme

$$\overline{X}_T(n) = \sum_{i=0}^{n} W_{t_i}^H (W_{t_{i+1}} - W_{t_i}), \tag{A.1}$$

using a reference solution $\overline{X}_T^{ref} := \overline{X}_T(2^{12})$ and computations $X_T^{\Delta t} = \overline{X}_T(n)$ for $\log_2 n = \{6, \dots, 1\}$. We sample paths $(W_{t_i}^H, W_{t_i})_{i\in[0:n]}$ required for the Monte Carlo approximation of (A.1) at points of the reference mesh using the Cholesky decomposition method. For each $\{H,T,n\}$ we first form the eigenvalue decomposition (L,D) of the full covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} ,$$

with blocks $(\Sigma_{11})_{ij} = \mathrm{Cov}(W^H_{t_i}, W^H_{t_j})$, $(\Sigma_{12})_{ij} = \mathrm{Cov}(W^H_{t_i}, W_{t_j})$, $(\Sigma_{22})_{ij} = \mathrm{Cov}(W_{t_i}, W_{t_j})$. We refer to, e.g., Lemma 4.1 of [3], for the covariance function of the Riemann-Liouville fBm and to `pfq.m` from [21] to compute hypergeometric functions:

```
t = T/n:T/n:T;
% require pfq.m from MATLAB File Exchange
G = @(x) 2.0*H*( x.^(-gam)/(1.0-gam) + (gam*x.^(-1.0-gam)./(1.0-gam)) .* ...
    pfq([1.0, 1.0+gam], 3.0-gam, x.^(-1.0))/(2.0-gam) );
disp('Computing blocks S11 S12 S22')
[X,Y] = meshgrid(t,t);
Gmat = G((tril(Y./X)+tril(Y./X)') - eye(size(Y,1)).*diag(Y./X));
Gmat = Gmat.*~eye(size(Gmat)) + eye(size(Gmat)); % diag is 1 b/c G(1)=1
S11 = ((tril(X)+tril(X)') - eye(size(X,1)).*diag(X)).^(2*H) .* Gmat;
S12 = sqrt(2*H)*(Y.^(H+0.5)-(Y-min(Y,X)).^(H+0.5))./(H+0.5);
S22 = min(X,Y);
disp('Finished blocks S11 S12 S22')
% LDL decomposition; defaults to Cholesky method
disp('Computing L D via eig')
[L,D] = eig([S11 S12; S12.' S22]);
```

We then generate M samples using the LDL eigenvalue decomposition:

```
z = randn(2*n,M);
X = real(L*(D^0.5)*z);
WH = [zeros(1,M); X(1:n, :)];
W = [zeros(1,M); X(n+1:end, :)];
```

and note that for $H = 0.5$ the sample paths of W^H generated by this method is identical to W. One sample of X_T^{ref} and $X_T^{\Delta t}$ at the final time can then be computed by summing the appropriate terms using the reference paths:

```
XTref(1,:) = sum(WH(1:end-1,:) .* diff(W(1:end,:)));
XTdt = nan(numDt,M);
for j=1:numDt
    XTdt(j,:) =  sum(WH(1:2^(j+gap-1):end-1,:) .* diff(W(1:2^(j+gap-1):end,:)));
end
```

The code referenced above can be found at the git repository:
`https://bitbucket.org/datainformeduq/rbwc_code`.

B Bound auxiliary functions f by φ

Lemma B.1. *Let $\varphi(x)$ be the payoff function, and define*

$$f_s(Y^{l_1}_{t_i}, \dots, Y^{l_k}_{t_i}) := \mathbf{E}[\nu(\overline{\boldsymbol{Z}}_{t_i}, s) \mid Y^{l_1}_{t_i}, \dots, Y^{l_k}_{t_i}] = \mathbf{E}[\mathbf{E}[\varphi^{(m)}(\widehat{X}_T) \mid \boldsymbol{Z}_s = (\overline{X}_{t_i}, \mathbf{Y}_{t_i})] \mid Y^{l_1}_{t_i}, \dots, Y^{l_k}_{t_i}] ,$$

for components $(Y_{t_i}^{l_1}, \dots, Y_{t_i}^{l_k}) \subset \boldsymbol{Y}_{t_i}$. *Then for a multiindex* $\alpha = (\alpha_1, \dots, \alpha_k)$,

$$|\partial^\alpha f_s^{kl}(\mathbf{0})| \lesssim \sum_{j=0}^{|\alpha|} \|\varphi^{(m+j)}\|_\infty .$$

Proof. (Lemma B.1) Recall that $\nu(\overline{\boldsymbol{Z}}_{t_i}, s)$ is a deterministic function of the jointly Gaussian η-dimensional vector (with $\eta := (N_L + 1) \times (i+1)$),

$$\boldsymbol{\xi} = (\boldsymbol{Y}_\tau, \Delta W_\tau)_{\tau = t_0, \dots, t_i} ,$$

(cf. (3.11)). Denoting the density as $\varrho(\boldsymbol{\xi})$, we note $\boldsymbol{\xi}$ is mean zero and has variance-covariance $\boldsymbol{\Sigma}$ given by the known quantities $\mathrm{Cov}(Y_{t_i}^l, Y_{t_j}^k)$, $\mathrm{Cov}(Y_{t_i}^l, \Delta W_{t_j})$, and $\mathrm{Cov}(\Delta W_{t_i}, \Delta W_{t_j})$, which have closed form expressions. For each s, the variable $\nu(\overline{\boldsymbol{Z}}_{t_i}, s)$ has a density proportional to $\varrho(\boldsymbol{\xi})$,

$$\mathrm{d}P_\nu \propto \varrho(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} .$$

Writing $\boldsymbol{y} = (Y_{t_i}^{l_1}, \dots, Y_{t_i}^{l_k})$, a subset of components of $\boldsymbol{Y}_{t_i}$ of size $k = |\boldsymbol{y}|$ such that $k \le N_L$, the function

$$f_s(\boldsymbol{y}) = \mathbf{E}[\nu(\overline{\boldsymbol{Z}}_{t_i}, s) \mid \boldsymbol{y}] ,$$

a deterministic function of $\boldsymbol{y}$, can be expressed in terms of a conditional Gaussian density. Partitioning $\boldsymbol{\xi} = (\widetilde{\boldsymbol{\xi}}, \boldsymbol{y})$ and, likewise, the covariance matrix $\boldsymbol{\Sigma}$ into components $\boldsymbol{\Sigma}_{11}$ corresponding to $\widetilde{\boldsymbol{\xi}}$, $\boldsymbol{\Sigma}_{22}$ to $\boldsymbol{y}$, and $\boldsymbol{\Sigma}_{12}$ to the mixed terms, the the conditional Gaussian density $\varrho(\widetilde{\boldsymbol{\xi}} \mid \boldsymbol{y})$ has conditional mean

$$\widetilde{\boldsymbol{\mu}} = \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{y} ,$$

which is linear in $\boldsymbol{y}$, and conditional variance-covariance matrix

$$\widetilde{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{12}^\top \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} ,$$

that does not depend on $\boldsymbol{y}$.

Rewritten in terms of this conditional density, f is given by

$$f_s(\boldsymbol{y}) = \int_{\mathbf{R}^{\eta-k}} \nu(\overline{X}_{t_i}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{y}; s) \varrho(\widetilde{\boldsymbol{\xi}} \mid \boldsymbol{y}) \mathrm{d}\widetilde{\boldsymbol{\xi}} ,$$

where

$$\nu(\overline{X}_{t_i}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{y}; s) = \mathbf{E}[\varphi^{(m)}(\widehat{X}_T) \mid \boldsymbol{Z}_s = (\overline{X}_{t_i}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{y})] .$$

For a multiindex $\alpha = (\alpha_1, \dots, \alpha_k)$ corresponding to the components of $\boldsymbol{y}$,

$$\partial^\alpha = \frac{\partial^\alpha}{\partial \boldsymbol{y}^\alpha} = \frac{\partial^{\alpha_1}}{\partial y_{l_1}} \cdots \frac{\partial^{\alpha_k}}{\partial y_{l_k}} ,$$

taking the derivative inside the integral we obtain

$$\begin{aligned} \partial^\alpha f_s(\boldsymbol{y}) &= \int_{\mathbf{R}^{\eta-k}} \left\{ (\partial^\alpha \nu(\overline{X}_{t_i}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{y})) \varrho(\widetilde{\boldsymbol{\xi}} \mid \boldsymbol{y}) + \nu(\overline{X}_{t_i}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{y}) \partial^\alpha \varrho(\widetilde{\boldsymbol{\xi}} \mid \boldsymbol{y}) \right\} \mathrm{d}\widetilde{\boldsymbol{\xi}} \\ &= \int_{\mathbf{R}^{\eta-k}} \left\{ \partial^\alpha \nu(\overline{X}_{t_i}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{y}) + \nu(\overline{X}_{t_i}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{y}) P_k(\boldsymbol{y}) \right\} \varrho(\widetilde{\boldsymbol{\xi}} \mid \boldsymbol{y}) \mathrm{d}\widetilde{\boldsymbol{\xi}} , \end{aligned} \tag{B.1}$$

where

$$\partial^\alpha \varrho(\widetilde{\boldsymbol{\xi}} \mid \boldsymbol{y}) = P_k(\boldsymbol{y}) \varrho(\widetilde{\boldsymbol{\xi}} \mid \boldsymbol{y}) ,$$

for P_k a polynomial of degree k since

$$\varrho(\widetilde{\boldsymbol{\xi}} \mid \boldsymbol{y}) \propto \exp[-\tfrac{1}{2}(\widetilde{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\mu}})^\top \widetilde{\boldsymbol{\Sigma}} (\widetilde{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\mu}})] = \exp[-\tfrac{1}{2}(\widetilde{\boldsymbol{\xi}} - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{y})^\top \widetilde{\boldsymbol{\Sigma}} (\widetilde{\boldsymbol{\xi}} - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{y})] .$$

The remaining derivative in (B.1) follows similarly to the computation of the fluxes in Lemma 3.7,

$$\partial^\alpha \nu(\overline{X}_{t_i}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{y}; s) = c_H^{|\alpha|} \, \mathbf{E}\Big[\varphi^{(|\alpha|+m)}(\widehat{X}_T) \prod_{j=1}^{|\alpha|} (\Delta \theta_{l_j} M_{s,T}^{l_j})^{\alpha_j} \mid \boldsymbol{Z}_s = (\overline{X}_{t_i}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{y}) \Big] ,$$

and the estimate follows. □