SHORT-TIME NEAR-THE-MONEY SKEW IN ROUGH FRACTIONAL VOLATILITY MODELS

C. BAYER, P. K. FRIZ, A. GULISASHVILI, B. HORVATH, B. STEMPER

Abstract. We consider rough stochastic volatility models where the driving noise of volatility has fractional scaling, in the "rough" regime of Hurst parameter $H < 1/2$. This regime recently attracted a lot of attention both from the statistical and option pricing point of view. With focus on the latter, we sharpen the large deviation results of Forde-Zhang (2017) in a way that allows us to zoom-in around the money while maintaining full analytical tractability. More precisely, this amounts to proving higher order moderate deviation estimates, only recently introduced in the option pricing context. This in turn allows us to push the applicability range of known at-the-money skew approximation formulae from CLT type log-moneyness deviations of order $t^{1/2}$ (recent works of Alòs, León & Vives and Fukasawa) to the wider moderate deviations regime.

Contents

1. Introduction

Since the groundbreaking work of Gatheral, Jaisson and Rosenbaum [GJR14a], the past two years have brought about a gradual shift in volatility modeling, leading away from classical diffusive stochastic volatility models towards so-called rough

Date: March 16, 2017.

2010 *Mathematics Subject Classification.* 91G20, 60H30, 60F10, 60H07, 60G22, 60G18.

Key words and phrases. rough stochastic volatility model, European option pricing, small-time asymptotics, moderate deviations.

We gratefully acknowledge financial support through DFG research grants FR2943/2 and BA5484/1 (C. Bayer, P.K. Friz, B. Stemper), European Research Council Grant CoG-683166 (P.K. Friz), and SNF Early Postdoc Mobility Grant 165248 (B. Horvath) respectively.

volatility models. The term was coined in [GJR14a] and [BFG16], and it essentially describes a family of (continuous-path) stochastic volatility models where the driving noise of the volatility process has Hölder regularity lower than Brownian motion, typically achieved by modeling the fundamental noise innovations of the volatility process as a fractional Brownian motion with Hurst exponent (and hence Hölder regularity) $H < 1/2$. Here, we would also like to mention pioneering work on asymptotics for rough volatility models in [ALV07] and [Fuk11]. A major appeal of such rough volatility models lies in the fact that they effectively capture several stylized facts of financial markets both from a statistical [GJR14a, BLP16] and an option-pricing point of view [BFG16]. In particular, with regards to the latter point of view, a widely observed empirical phenomenon in equity markets is the "steepness of the smile on the short end" describing the fact that as time to maturity becomes small the empirical implied volatility skew follows a power law with negative exponent, and thus becomes arbitrarily large near zero. While standard stochastic volatility models with continuous paths struggle to capture this phenomenon, predicting instead a constant at-the-money implied volatility behaviour on the short end [Gat11], models in the fractional stochastic volatility family (and more specifically so-called rough volatility models) constitute a class, well-tailored to fit empirical implied volatilities for short dated options.

Typically, the popularity of asset pricing models hinges on the availability of efficient numerical pricing methods. In the case of diffusions, these include Monte Carlo estimators, PDE discretization schemes, asymptotic expansions and transform methods. With fractional Brownian motion being the prime example of a process beyond the semimartingale framework, most currently prevalent option pricing methods – particularly the ones assuming semimartingality or Markovianity – may not easily carry over to the rough setting. In fact, the memory property (aka non-Markovianity) of fractional Brownian motion rules out PDE methods, heat kernel methods and all related methods involving a Feynman-Kac-type Ansatz. Previous work has thus focused on finding efficient Monte Carlo simulation schemes [BFG16, BLP15, BFG$^+$17] or – in the special case of the Rough Heston model – on an explicit formula for the characteristic function of the log-price (see [ER16]), thus in this particular model making pricing amenable to Fourier based methods. In our work, we rely on small-maturity approximations of option prices. This is a well-studied topic. See, e.g., [ALV07, GVZ15] (the at-the-money (ATM) regime) or [DFJV14a, DFJV14b, GJR14b, GHJ16, GVZ15] (the out-of-the-money (OTM) regime, where large deviations results are used). We also refer the reader to the papers [Fuk11, Fuk17, FZ17] concerning large deviations, and to [MT16, Osa07, MS03, MS07] for related work. Based on the moderate deviations regime, Friz et al. [FGP17] have recently introduced another regime called Moderately-out-of-the-money (MOTM), which, in a sense, effectively navigates between the two regimes mentioned above, by rescaling the strike with respect to the time to maturity. This approach has various advantages. On the one hand, it reflects the market reality that as time to maturity approaches zero, strikes with acceptable bid-ask spreads tend to move closer to the money (see [FGP17] for more details). On the other hand, it allows us to zoom in on the term structure of implied volatility around the money at a high resolution scale. To be more specific, our paper adds to the existing literature in two ways. First, we obtain a generalization of the Osajima energy expansion [Osa15] to a non-Markovian case, and using the new expansion, we extend the analysis of [FGP17] to the case, where the volatility is driven by a rough ($H < 1/2$) fractional Brownian motion. Indeed, Laplace approximation methods on Wiener space in the spirit of Ben Arous [BA88] and Bismut [Bis84] remain valid in the present context, and our analysis builds upon this framework in a fractional setting. Second, we use an asymptotic expansion going back to Azencott [Aze85] to bypass the need for deriving an asymptotic expansion of the density of the underlying process to obtain asymptotics for option prices. We display the potential prowess of this approach by applying it to our specific model, and derive asymptotics for call prices directly, irrespectively of corresponding density asymptotics. Finally, using a version of the "*rough Bergomi model*" [BFG16], we demonstrate numerically that our implied volatility asymptotics capture very well the geometry of the term structure of implied volatility over a wide array of maturities, extending up to a year.

The paper is organized as follows: In Section 2 we set the scene, describing the class of models included in our framework ((2.1) and (2.2)) and recalling some known results ((2.4) and (2.8)), which are the starting point of our analysis. Most importantly, we argue that for small-time considerations it would suffice to restrict our attention to a class of stochastic volatility models of the form (2.3) with a volatility process driven by a Gaussian Volterra process such as in (2.2). We formulate general assumptions on the Volterra kernel (Assumptions 2.1 and 2.5) and on the function σ in (2.3) (Assumption 2.4) under which our results are valid. In Section 3 we gather our main results, concerning a higher order expansion of the energy (Theorem 3.1), and a general expansion formula for the corresponding call prices. We derive the classical Black-Scholes expansion for the call price, using the latter result mentioned above. In addition, in Section 3 we formulate moderate deviation expansions, which allow us to derive the corresponding asymptotic formulae for implied volatilities and implied volatility skews. Finally, Section 4 displays our simulation results. Sections 5, 6 and 7 are devoted to proofs of the energy expansion, the price expansion and the moderate deviations expansion, respectively. In the appendix, we have collected some auxiliary lemmas, which are used in different sections.

2. Exposition and assumptions

We consider a rough stochastic volatility model, normalized to $r = 0$ and $S_0 = 1$, of the form suggested by Forde-Zhang [FZ17]

$$\frac{dS_t}{S_t} = \sigma(\widehat{B}_t) d\left(\overline{\rho} W_t + \rho B_t\right). \tag{2.1}$$

Here (W, B) are two independent standard Brownian motions, $\rho \in (-1, 1)$ a correlation parameter, we use the by now standard notation that $\overline{\rho}^2 = 1 - \rho^2$. Then $\overline{\rho} W + \rho B$ is another standard Brownian motion which has constant correlation ρ with the factor B, which drives stochastic volatility

$$\sigma_{\text{stoch}}\,(t, \omega) := \sigma(\widehat{B}_t\,(\omega)) \equiv \sigma(\widehat{B}).$$

Here $\sigma : \mathbb{R} \to (0, \infty)$ is some real-valued function and we will denote by $\sigma_0 := \sigma(0)$ the spot volatility. Furthermore, $\widehat{B}$ is a Gaussian (Volterra) process of the form

$$\widehat{B}_t = \int_0^t K\,(t, s)\, dB_s, \quad t \geq 0, \tag{2.2}$$

for some kernel K, which shall be further specified in Assumptions 2.1 and 2.5 below. The log-price $\widetilde{X}_t = \log(S_t)$ satisfies

$$d\widetilde{X}_t = -\frac{1}{2}\sigma^2(\widehat{B}_t)dt + \sigma(\widehat{B}_t)d(\overline{\rho}W + \rho B),$$

but for the subsequent short-time considerations, it is enough (cf. [FZ17, Proof of Theorem 4.1] and [DZ09, Definition 4.2.10]) to study its driftless version

$$dX_t = \sigma(\widehat{B}_t)d(\overline{\rho}W_t + \rho B_t), \quad X_0 = 0. \tag{2.3}$$

Recall that by Brownian scaling, for fixed $t > 0$,

$$(B_{ts}, W_{ts})_{s\geq 0} \overset{law}{=} \varepsilon(B_s, W_s)_{s\geq 0}, \quad \text{where} \quad \varepsilon \equiv \varepsilon(t) \equiv t^{1/2}.$$

As a direct consequence, classical short-time SDE problems can be analyzed as small-noise problems on a unit time horizon. For our analysis, it will also be crucial to impose such a scaling property on the Gaussian process $\widehat{B}$ (more precisely on the kernel K in (2.2) driving the volatility process in our model:

Assumption 2.1 (Small time self-similarity)**.** There exists a number t_0 with $0 < t_0 \leq 1$ and a function $t \mapsto \widehat{\varepsilon} = \widehat{\varepsilon}(t)$, $0 \leq t \leq t_0$, such that

$$(\widehat{B}_{ts} : 0 \leq s \leq t_0) \overset{law}{=} (\widehat{\varepsilon}\widehat{B}_s : 0 \leq s \leq t_0).$$

In fact, we will always have

$$\widehat{\varepsilon} \equiv \widehat{\varepsilon}(t) \equiv t^H = \varepsilon^{2H},$$

which covers the examples of interest, in particular standard fractional Brownian motion $\widehat{B} = B^H$ or Riemann-Liouville fBM with explicit kernel $K(t,s) = \sqrt{2H}\,|t-s|^{H-1/2}$. (This is very natural, even from a general perspective of self-similar processes, see [Lam62].)

We insist that no (global) self-similarity of $\widehat{B}$ is required, as only $\widehat{B}|_{[0,t]}$ for arbitrarily small t, matters.

Remark 2.2. In all likelihood, it should be possible to replace the fractional Brownian motion by a certain fractional Ornstein-Uhlenbeck process in the results obtained in this paper. Intuitively, this replacement creates a negligible perturbation (for $t << 1$) of the fBm environment. A similar situation was in fact encountered in [CF10], where fractional scaling at times near zero was important. To quantify the perturbation, the authors of [CF10] introduced an easy to verify coupling condition (see Corollary 2 in [CF10]). In our opinion, a version of this condition can be employed in the present paper to justify the replacement mentioned above. We will however not pursue this point further here.

Remark 2.3. Throughout this article, one can consider a classical (Markovian, diffusion) stochastic volatility setting by taking $K \equiv 1$, or equivalently $H \equiv 1/2$, by simply ignoring all hats ($\widehat{\cdot}$) in the sequel. In particular then, $\frac{\widehat{\varepsilon}}{\varepsilon} \equiv 1$ in all subsequent formulae.

General facts on large deviations of Gaussian measures on Banach spaces [DS89] such as the path space $C([0,1],\mathbb{R}^3)$ imply that a large deviation principle holds for the triple $\{\widehat{\varepsilon}(W,B,\widehat{B}) : \widehat{\varepsilon} > 0\}$, with speed $\widehat{\varepsilon}^2$ and rate function

$$\begin{cases} \frac{1}{2}\|h\|^2_{H^1_0} + \frac{1}{2}\|f\|^2_{H^1_0}, & f,h \in H^1_0 \text{ and } \widehat{f} = K\dot{f}, \\ +\infty, & \text{else,} \end{cases} \tag{2.4}$$

where

$$K\dot{f}(t) := \int_0^t K(t,s)\,\dot{f}(s)ds$$

for $f \in H_0^1$, and

$$H_0^1 := \left\{ f : [0,1] \to \mathbb{R} \;\middle|\; \|f\|_{H_0^1}^2 := \int_0^1 \left|\dot{f}(s)\right|^2 ds < \infty,\ f(0) = 0 \right\}. \tag{2.5}$$

This enables us to derive a large deviations principle for the X in (2.3): The (local) small-time self similarity property of $\widehat{B}$ (Assumption 2.1) implies that $X_t \overset{law}{=} X_1^\varepsilon$ where

$$dX_t^\varepsilon = \sigma(\widehat{\varepsilon}\widehat{B}_t)\varepsilon d\left(\overline{\rho}W_t + \rho B_t\right), \quad X_0^\varepsilon = 0.$$

For what follows, it will be convenient to consider a rescaled version of (2.3)

$$d\widehat{X}_t^\varepsilon \equiv d\left(\frac{\widehat{\varepsilon}}{\varepsilon}X_t^\varepsilon\right) = \sigma(\widehat{\varepsilon}\widehat{B}_t)\widehat{\varepsilon}d\left(\overline{\rho}W_t + \rho B_t\right), \quad \widehat{X}_0^\varepsilon = 0.$$

since indeed $\widehat{X}_1^\varepsilon \equiv \Phi_1(\widehat{\varepsilon}W, \widehat{\varepsilon}B, \widehat{\varepsilon}\widehat{B})$ in terms of an "Itô-type map[1]"

$$\Phi_1(W, B, \widehat{B}) := \int_0^1 \sigma(\widehat{B})d\left(\overline{\rho}W + \rho B\right). \tag{2.6}$$

Thanks to the (extended) contraction principle, a large deviations principle also holds for $(\widehat{X}_1^\varepsilon)$, again with speed $\widehat{\varepsilon}^2$. With

$$\varphi_1(h, f) := \Phi_1(h, f, \widehat{f}) = \int_0^1 \sigma(\widehat{f})d\left(\overline{\rho}h + \rho f\right), \tag{2.7}$$

we so obtain the following rate function

$$\begin{aligned} I(x) &= \inf_{h,f \in H_0^1} \left\{ \frac{1}{2}\int_0^1 \dot{h}^2 dt + \frac{1}{2}\int_0^1 \dot{f}^2 dt : \varphi_1(h,f) = x \right\} \\ &= \inf_{f \in H_0^1} \left\{ \frac{1}{2} \frac{\left(x - \rho\left\langle \sigma(\widehat{f}), \dot{f} \right\rangle\right)^2}{\overline{\rho}^2 \left\langle \sigma^2(\widehat{f}), 1 \right\rangle} + \frac{1}{2}\int_0^1 \dot{f}^2 dt \right\}, \end{aligned} \tag{2.8}$$

the exact proof of which may be found in [FZ17].

We have not yet said anything about the restrictions on the function σ.

Assumption 2.4. Assume $\sigma : \mathbb{R} \to (0, \infty)$ is smooth and such that

$$E\left[\int_0^1 \sigma^2(\widehat{B}_t)dt\right] < \infty.$$

This is satisfied (trivially) for σ bounded, no matter the precise nature of $\widehat{B}$. But also $\sigma(x) = \exp(\eta x)$ with fractional Brownian motion $\widehat{B}$ and $\eta \in \mathbb{R}$ is covered. In addition to Assumption 2.1, we impose from now on further conditions on the kernel K.

Assumption 2.5. The kernel K satisfies

(i) $\widehat{B}_t = \int_0^t K(t,s)dB_s$ has a continuous (in t) version on $[0,1]$.
(ii) $\forall t \in [0,1]: \int_0^t K(t,s)^2 ds < \infty$.

[1] Note that Φ_1 is measurable, but not necessarily continuous.

Note that the Riemann-Liouville kernel $K(t,s) = \sqrt{2H}(t-s)^{\gamma}$, $\gamma = H - 1/2$ satisfies Assumption 2.5.

Remark 2.6. Assumption 2.5 implies that the Cameron-Martin space $\mathcal{H}$ of $\widehat{B}$ is given by the image of H_0^1 under K, i.e.,

$$\mathcal{H} = \{K\dot{f} \mid f \in H_0^1\}.$$

See Lemma 5.3 and Remark 5.4 for more details. A reference and also a sufficient condition for Assumption 2.5 (i) can be can be found in [Dec05, Section 3].

3. Main results

The following result can be seen as a non-Markovian extension of work by Osajima [Osa15]. The statement here is a combination of Theorem 5.10 and Proposition (5.13) below. Recall that $\sigma_0 = \sigma(0)$ represents spot-volatility. We also set $\sigma_0' \equiv \sigma'(0)$.

Theorem 3.1 (Energy expansion)**.** *The rate function (or energy) I in* (2.8) *is smooth in a neighbourhood of $x = 0$ (at-the-money) and it is of the form*

$$I(x) = \frac{1}{\sigma_0^2}\frac{x^2}{2} - \left(6\rho\frac{\sigma_0'}{\sigma_0^4}\int_0^1\int_0^t K(t,s)dsdt\right)\frac{x^3}{3!} + \mathcal{O}(x^4).$$

The next result is an exact representation of call prices, valid in a non-Markovian generality, and amenable to moderate- and large-deviation analysis (Theorem 3.4 below) as well as to full asymptotic expansions, which will be explored in forthcoming work.

Theorem 3.2 (Pricing formula)**.** *For a fixed log-strike $x \geq 0$ and time to maturity $t > 0$, set $\widehat{x} := \frac{\varepsilon}{\widehat{\varepsilon}}x$, where $\varepsilon = t^{1/2}$ and $\widehat{\varepsilon} = t^H = \varepsilon^{2H}$, as before. Then we have*

$$\begin{aligned} c(\widehat{x},t) &= E\left[(\exp(X_t) - \exp\widehat{x})^+\right] \\ &= e^{-\frac{I(x)}{\widehat{\varepsilon}^2}} e^{\frac{\varepsilon}{\widehat{\varepsilon}}x} J(\varepsilon,x), \end{aligned}$$

where

$$J(\varepsilon,x) := E\left[e^{-\frac{I'(x)}{\widehat{\varepsilon}^2}\widehat{U}^{\varepsilon}}\left(\exp\left(\tfrac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^{\varepsilon}\right) - 1\right)e^{I'(x)R_2}\ \mathbf{1}_{\widehat{U}^{\varepsilon}\geq 0}\right]$$

and $\widehat{U}^{\varepsilon}$ is a random variable of the form

$$\widehat{U}^{\varepsilon} = \widehat{\varepsilon}g_1 + \widehat{\varepsilon}^2 R_2$$

with g_1 a centred Gaussian random variable, explicitly given in equation (6.3) below and R_2 is a (random) remainder term, in the sense of a stochastic Taylor expansion in $\widehat{\varepsilon}$, see Lemma (6.2) for more details.

Example 3.3. In the case of Black-Scholes one has $\sigma(\cdot) \equiv \sigma > 0$, $R_2 \equiv 0$ (recall from (2.3) that we consider the driftless version) and $\widehat{\varepsilon} = \varepsilon$. Moreover, here $g_1 \equiv \sigma W_1 \overset{law}{=} N(0,\sigma^2)$, the energy is $I(x) = \frac{x^2}{2\sigma^2}$ and

$$\begin{aligned} J(\varepsilon,x) &= E\left(e^{-\frac{I'(x)}{\varepsilon}g_1}\left(e^{\varepsilon g_1} - 1\right)\ \mathbf{1}_{U\geq 0}\right) \\ &= M\left(-\frac{I'(x)\sigma}{\varepsilon} + \varepsilon\sigma\right) - M\left(-\frac{I'(x)\sigma}{\varepsilon}\right) \end{aligned}$$

where $M(\alpha) := e^{\alpha^2/2}F(\alpha)$ and F the standard Gaussian distribution function. Using $J(\varepsilon,x) \sim M'\left(-\frac{I'(x)\sigma}{\varepsilon}\right)\varepsilon\sigma$ and $M'(\alpha) \sim (2\pi)^{-1/2}|\alpha|^{-2}$ as $\alpha \downarrow -\infty$ one deduces that, as long as $x/\varepsilon \to \infty$,

$$\begin{aligned} J(\varepsilon,x) &\sim \frac{1}{\sqrt{2\pi}}\left|\frac{I'(x)\sigma}{\varepsilon}\right|^{-2}\varepsilon\sigma \\ &= \frac{1}{\sqrt{2\pi}}\frac{\varepsilon^3\sigma^3}{x^2}. \end{aligned}$$

This analysis is valid in the large deviations regime with fixed $x > 0$. But we can also take $x = x_\varepsilon \sim c_1\varepsilon^{2\beta}$ for some $c_1 > 0$, and (recall the moderate regime, with $\varepsilon^2 = t$) as long as $\beta \in [0,1/2)$ the above analysis is justified. In particular, the term $J(\varepsilon,x_\varepsilon) \sim c_2\varepsilon^{c_3}$ in the pricing formula is of polynomial order in ε and so J is negligible on the moderate / large deviation scale, since, for any $\theta > 0$, we have $\varepsilon^\theta \log J(\varepsilon,x_\varepsilon) \to 0$ as $\varepsilon \to 0$. Consequently, with $k_t = kt^\beta$, for $t = \varepsilon^2$, $k > 0$, $\beta \in [0,1/2)$ we get the "moderate" Black-Scholes call price expansion,

$$-\log c_{BS}(k_t,t) = \frac{1}{t^{1-2\beta}}\frac{k^2}{2\sigma^2}\left(1+o\left(1\right)\right) \text{ as } t \downarrow 0.$$

While the above can be confirmed by elementary analysis of the Black–Scholes formula, the following theorem exhibits it as an instance of a general principle. See [FGP17] for a general diffusion statement.

In what follows, we assume that the kernel is such that "fractional scaling" applies, i.e., $\widehat{\varepsilon} = t^H$.

Theorem 3.4 (Moderate Deviations)**.** *In the rough volatility regime $H \in (0,1/2]$, consider log-strikes of the form*

$$k_t = kt^{\frac{1}{2}-H+\beta} \quad \textit{for a constant} \quad k \geq 0.$$

(i) For $\beta \in (0,H)$, we have

$$-\log c(k_t,t) = \frac{I''(0)}{t^{2H-2\beta}}\frac{k^2}{2} + \mathcal{O}\left(t^{3\beta-2H}\right) + \mathcal{O}(\log\frac{1}{t}) \textit{ as } t \downarrow 0.$$

(ii) For $\beta \in (0,\frac{2}{3}H)$ we have

$$-\log c(k_t,t) = \frac{I''(0)}{t^{2H-2\beta}}\frac{k^2}{2} + \frac{I'''(0)}{t^{2H-3\beta}}\frac{k^3}{6} + \mathcal{O}\left(t^{4\beta-2H}\right) + \mathcal{O}(\log\frac{1}{t}).$$

Moreover,

$$\begin{aligned} I''(0) &= \frac{1}{\sigma_0^2}, \\ I'''(0) &= -6\rho\frac{\sigma_0'}{\sigma_0^4}\int_0^1\int_0^t K(t,s)dsdt = -6\rho\frac{\sigma_0'}{\sigma_0^4}\langle K1,1\rangle, \end{aligned}$$

and $\langle\cdot,\cdot\rangle$ is the inner product in $L^2([0,1])$.

Proof. We apply Theorem 3.2 with $\widehat{x} = k_t = kt^{1/2-H+\beta}$, i.e., with $x = kt^\beta$. In Proposition 7.1, we will show that $\log J(\varepsilon,x) \sim \log(\varepsilon)$, uniformly for x in a neighborhood of 0. Furthermore, it is clear that $\frac{\varepsilon}{\widehat{\varepsilon}} = \mathcal{O}(\log\frac{1}{t})$, and hence we have

$$-\log c(k_t,t) = \frac{I(kt^\beta)}{t^{2H}} + \mathcal{O}(\log\frac{1}{t}).$$

The theorem now follows immediately from the Taylor expansion of $I(x)$ around $x = 0$ (see Theorem 3.1), plugging in $x = kt^{\beta}$. □

Fix real numbers $k > 0$, $0 < H < \frac{1}{2}$, $0 < \beta < H$, and an integer $n \geq 2$. For every $t > 0$, set

$$k_t = kt^{\frac{1}{2}-H+\beta},$$

and denote

$$\phi_{n,H,\beta}(t) = \max\left\{t^{2H-2\beta}\log\frac{1}{t}, t^{(n-1)\beta}\right\}.$$

It is clear that for all small t,

$$\phi_{n,H,\beta}(t) = t^{2H-2\beta}\log\frac{1}{t} \Leftrightarrow 2H-2\beta \leq (n-1)\beta \Leftrightarrow \frac{2H}{n+1} \leq \beta,$$

while

$$\phi_{n,H,\beta}(t) = t^{(n-1)\beta} \Leftrightarrow 2H-2\beta > (n-1)\beta \Leftrightarrow \beta < \frac{2H}{n+1}.$$

The following statement provides an asymptotic formula for the implied variance.

Theorem 3.5. *Suppose* $0 < \beta < \frac{2H}{n}$. *Then as* $t \to 0$,

$$\begin{aligned}\sigma_{\mathrm{impl}}(k_t,t)^2 = &\sum_{j=0}^{n-2}\frac{(-1)^j 2^j}{I''(0)^{j+1}}\left(\sum_{i=3}^{n}\frac{I^{(i)}(0)}{i!}k^{i-2}t^{(i-2)\beta}\right)^j \\ &+ O\left(\phi_{n,H,\beta}(t)\right).\end{aligned} \tag{3.1}$$

The O-estimate in (3.1) depends on n, H, β, *and* k. *It is uniform on compact subsets of* $[0,\infty)$ *with respect to the variable* k.

Remark 3.6. Using the multinomial formula, we can represent the expression on the left-hand side of (3.1) in terms of certain powers of t. However, the coefficients become rather complicated.

Remark 3.7. Let an integer $n \geq 2$ be fixed, and suppose we would like to use only the derivatives $I^{(i)}(0)$ for $2 \leq i \leq n$ in formula (3.1) to approximate $\sigma_{\mathrm{impl}}(k_t,t)^2$. Then, the optimal range for β is the following: $\frac{2H}{n+1} \leq \beta < \frac{2H}{n}$. On the other hand, if β is outside of the interval $[\frac{2H}{n+1}, \frac{2H}{n})$, more derivatives of the energy function at zero may be needed to get a good approximation of the implied variance in formula (3.1).

We will next derive from Theorem 3.5 several asymptotic formulas for the implied volatility. In the next corollary, we take $n = 2$.

Corollary 3.8. *As* $t \to 0$,

$$\sigma_{\mathrm{impl}}(k_t,t) = \sigma_0 + O(\phi_{2,H,\beta}(t)). \tag{3.2}$$

Corollary 3.8 follows from Theorem 3.5 with $n = 2$, the equality

$$I''(0) = \sigma_0^{-2} \tag{3.3}$$

given in Theorem 3.4, and the Taylor expansion $\sqrt{1+h} = 1 + O(h)$ as $h \to 0$.

In the next corollary, we consider the case where $n = 3$.

Corollary 3.9. *Suppose* $\beta < \frac{2H}{3}$. *Then, as* $t \to 0$,

$$\sigma_{\mathrm{impl}}(k_t,t) = \sigma_0 + \rho\frac{\sigma_0'}{\sigma_0}\langle K1,1\rangle kt^{\beta} + O(\phi_{3,H,\beta}(t)). \tag{3.4}$$

Corollary 3.9 follows from Theorem 3.5 with $n = 3$, formula (3.3), the equality

$$I'''(0) = -6\rho\frac{\sigma_0'}{\sigma_0^4}\langle K1, 1\rangle \tag{3.5}$$

(see Theorem 3.4), and the expansion $\sqrt{1+h} = 1 + \frac{1}{2}h + O(h^2)$ as $h \to 0$.

Using Corollary 3.9, we establish the following implied volatility skew formula in the moderate deviation regime.

Corollary 3.10. *Let $0 < H < \frac{1}{2}$, $0 < \beta < \frac{2}{3}H$, and fix $y, z > 0$ with $y \neq z$. Then as $t \to 0$,*

$$\frac{\sigma_{\mathrm{impl}}(yt^{\frac{1}{2}-H+\beta}, t) - \sigma_{\mathrm{impl}}(zt^{\frac{1}{2}-H+\beta}, t)}{(y-z)t^{\frac{1}{2}-H+\beta}} \sim \rho\frac{\sigma_0'}{\sigma_0}\langle K1, 1\rangle t^{H-\frac{1}{2}}. \tag{3.6}$$

Remark 3.11. Corollary 3.10 complements earlier works of Alòs et al. [ALV07] and Fukasawa [Fuk11, Fuk17]. For instance, the following formula can be found in [Fuk17, p. 6], see also [Fuk11, p. 14]:

$$\frac{\sigma_{\mathrm{impl}}(yt^{\frac{1}{2}}, t) - \sigma_{\mathrm{impl}}(zt^{\frac{1}{2}}, t)}{(y-z)t^{\frac{1}{2}}} \sim \rho C(H)\frac{\sigma_0'}{\sigma_0}t^{H-\frac{1}{2}}. \tag{3.7}$$

In formula (3.7), we employ the notation used in the present paper. Our analysis shows that the applicability range of skew approximation formulas is by no means restricted to the Central Limit Theorem type log-moneyness deviations of order $t^{1/2}$. It also includes the moderate deviations regime of order $t^{1/2-H+\beta}$. The previous rate is clearly $>> t^{1/2}$ as $t \to 0$.

Finally, we turn our attention to the case where $n = 4$. We will next provide a general asymptotic formula for the implied volatility that uses the fourth derivative $I^{(4)}(0)$.

Corollary 3.12. *Suppose $\beta < \frac{H}{2}$. Then as $t \to 0$,*

$$\begin{aligned}\sigma_{\mathrm{impl}}(k_t, t) &= \sigma_0 + \rho\frac{\sigma_0'}{\sigma_0}\langle K1, 1\rangle k t^{\beta} \\ &+ \left(\rho^2\frac{(\sigma_0')^2}{\sigma_0^3}\langle K1, 1\rangle^2 - \frac{I^{(4)}(0)\sigma_0^3}{24}\right)k^2 t^{2\beta} + O(\phi_{4,H,\beta}(t))).\end{aligned} \tag{3.8}$$

Remark 3.13 (Symmetry). Write $\Phi_1(W, B, \widehat{B}; \rho; \sigma)$ for the "Itô-type map" introduced in (2.6). It equals, in law, $\Phi_1(W, -B, -\widehat{B}; -\rho; \sigma(-\cdot))$, and indeed all our formulae are invariant under this transformation. In particular, the skew remains unchanged when the pair (ρ, σ_0') is replaced by $(-\rho, -\sigma_0')$.

4. Simulation results

We verify our theoretical results numerically with a variant of the *rough Bergomi model* [BFG16] which fits nicely into the general rough volatility framework considered in this paper. As before, the model has been normalized such that $S_0 = 1$ and $r = 0$. We let (W, B) be two independent Brownian motions and $\rho \in (-1, 1)$ with $\overline{\rho}^2 = 1 - \rho^2$ such that $Z = \overline{\rho}W + \rho B$ is another Brownian motion having constant correlation ρ with B. For some spot volatility σ_0 and volatility of volatility parameter η, we then assume the following dynamics for some asset S:

$$\frac{dS_t}{S_t} = \sigma(\widehat{B}_t)dZ_t \tag{4.1}$$

$$\sigma(x) = \sigma_0 \exp\left(\frac{1}{2}\eta\widehat{B}_t\right) \tag{4.2}$$

where $\widehat{B}$ is a Riemann-Liouville fBM given by

$$\widehat{B}_t = \sqrt{2H}\int_0^t |t-s|^{H-1/2} dB_s.$$

The approach taken for the Monte Carlo simulations of the quantities we are interested in is the one initially explored in the original *rough Bergomi* pricing paper [BFG16]. That is, exploiting their joint Gaussianity, where we use the well-known Cholesky method to simulate the joint paths of $(Z, \widehat{B})$ on some discretization grid $\mathcal{D}$. With (4.2) being an explicit function in terms of the rough driver, an Euler discretisation of the Ito SDE (4.1) on $\mathcal{D}$ then yields estimates for the price paths.

The Cholesky algorithm critically hinges on the availability and explicit computability of the joint covariance matrix of $(Z, \widehat{B})$ whose terms we readily compute below.[2]

Lemma 4.1. *For convenience, define constants* $\gamma = \frac{1}{2} - H \in [0, \frac{1}{2})$ *and* $D_H = \frac{\sqrt{2H}}{H+\frac{1}{2}}$ *and define an auxiliary function* $G : [1,\infty) \to \mathbb{R}$ *by*

$$G(x) = 2H\left(\frac{1}{1-\gamma}x^{-\gamma} + \frac{\gamma}{1-\gamma}x^{-(1+\gamma)}\frac{1}{2-\gamma}\,{}_2F_1(1, 1+\gamma, 3-\gamma, x^{-1})\right) \tag{4.3}$$

where ${}_2F_1$ *denotes the Gaussian hypergeometric function* [Olv10]*. Then the joint process* $(Z, \widehat{B})$ *has zero mean and covariance structure governed by*

$$\begin{cases} \mathrm{Var}[\widehat{B}_t^2] = t^{2H}, & \textit{for } t \geq 0, \\ \mathrm{Cov}[\widehat{B}_s\widehat{B}_t] = t^{2H}G\left(s/t\right), & \textit{for } s > t \geq 0, \\ \mathrm{Cov}[\widehat{B}_s Z_t] = \rho D_H\left(s^{H+\frac{1}{2}} - (s - \min(t,s))^{H+\frac{1}{2}}\right), & \textit{for } t, s \geq 0, \\ \mathrm{Cov}[Z_t Z_s] = \min(t,s), & \textit{for } t, s \geq 0. \end{cases}$$

Numerical simulations[3] confirm the theoretical results obtained in the last section. In particular - as can be seen in Figure 1 – the asymptotic formula for the implied volatility (3.4) captures very well the geometry of the term structure of implied volatility, with particularly good results for higher H and worsening results as $H \downarrow 0$. Quite surprisingly, despite being an asymptotic formula, it seems to be fairly accurate over a wide array of maturities extending up to a single year.

5. Proof of the energy expansion

Consider

$$\begin{aligned} dX &= \sigma(Y)\, d\left(\overline{\rho}dW + \rho dB\right), \quad X_0 = 0 \\ dY &= d\widehat{B}, \; Y_0 = 0 \end{aligned}$$

[2] Note that expressions for the exact same scenario have have been computed before in the original pricing paper [BFG16], yet in that version the expression for the autocorrelation of the fBM $\widehat{B}$ was incorrect. We compute and state here all the relevant terms for the sake of completeness.

[3] The Python 3 code used to run the simulations can be found at github.com/RoughStochVol.

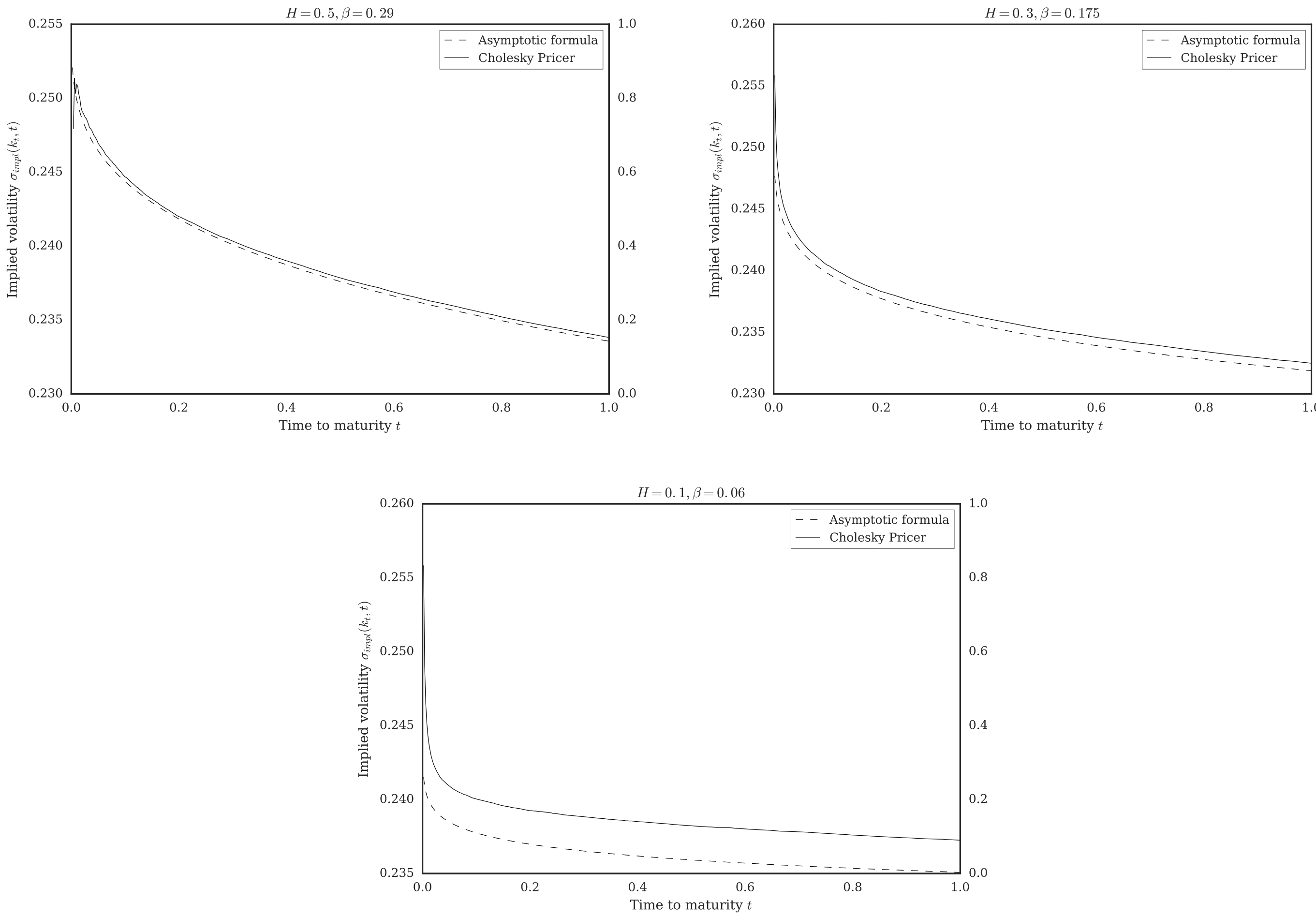

FIGURE 1. Illustration of the term structure of implied volatility of the Modified Rough Bergomi model in the Moderate deviations regime with time-varying log-strike $k_t = 0.4t^{\beta}$. Depicted are the asymptotic formula (Eq. (3.4), dashed line) and an estimate based on $N = 10^8$ samples of a MC Cholesky Option Pricer (solid line) with 500 time steps. Model parameters are given by spot vol $\sigma_0 \approx 0.2557$, vvol $\eta = 0.2928$ and correlation parameter $\rho = -0.7571$.

where $\widehat{B}_t = \int_0^t K(t,s)\, dB_s$ for a fixed Volterra kernel (recall (2.3) in the previous section) We study the small noise problem $(X^\varepsilon, Y^\varepsilon)$ where $\left(W, B, \widehat{B}\right)$ is replaced by $\left(\varepsilon W, \varepsilon B, \widehat{\varepsilon}\widehat{B}\right)$. The following proposition roughly says that

$$\mathbb{P}\left(X_1^\varepsilon \approx \frac{\varepsilon}{\widehat{\varepsilon}} x\right) \approx \exp\left(-\frac{I(x)}{\widehat{\varepsilon}^2}\right).$$

Proposition 5.1 (Forde-Zhang [FZ17])**.** *The rescaled process* $\left(\frac{\widehat{\varepsilon}}{\varepsilon} X_1^\varepsilon : \varepsilon \geq 0\right)$ *satisfies an LDP (with speed* $\widehat{\varepsilon}^2$*) and rate function*

$$I(x) = \inf_{f \in H_0^1} \left[\frac{(x - \rho G(f))^2}{2\overline{\rho}^2 F(f)} + \frac{1}{2} E(f)\right] \equiv \inf_{f \in H_0^1} \mathcal{I}_x(f) \tag{5.1}$$

where

$$\begin{aligned} G(f) &= \int_0^1 \sigma\left(\left(K\dot{f}\right)(s)\right) \dot{f}_s ds \equiv \left\langle \sigma\left(K\dot{f}\right), \dot{f}\right\rangle \equiv \left\langle \sigma(\widehat{f}), \dot{f}\right\rangle \\ F(f) &= \int_0^1 \sigma\left(\left(K\dot{f}\right)(s)\right)^2 ds \equiv \left\langle \sigma^2\left(K\dot{f}\right), 1\right\rangle \equiv \left\langle \sigma^2(\widehat{f}), 1\right\rangle \\ E(f) &= \int_0^1 \left|\dot{f}(s)\right|^2 ds \equiv \left\langle \dot{f}, \dot{f}\right\rangle \end{aligned}$$

Next we derive the first order optimality condition for the above minimization problem.

Proposition 5.2 (First order optimality condition)**.** *For any* $x \in \mathbb{R}$ *we have at any local minimizer* $f = f^x$ *of the functional* $\mathcal{I}_x$ *in* (5.1) *that*

$$\begin{aligned} f_t^x = & \frac{\rho(x - \rho G(f^x)) \left\{\left\langle \sigma\left(K\dot{f}^x\right), 1_{[0,t]}\right\rangle + \left\langle \sigma'\left(K\dot{f}^x\right)\dot{f}^x, K1_{[0,t]}\right\rangle\right\}}{\overline{\rho}^2 F(f^x)} \\ & + \frac{(x - \rho G(f^x))^2}{\overline{\rho}^2 F^2(f^x)} \left\langle (\sigma\sigma')\left(K\dot{f}^x\right), K1_{[0,t]}\right\rangle, \end{aligned} \tag{5.2}$$

for all $t \in [0,1]$.

Proof. We denote $a \approx b$ whenever $a = b + o(\delta)$, for a small parameter δ, we expand

$$\begin{aligned} E(f + \delta g) &\approx E(f) + 2\delta \left\langle \dot{f}, \dot{g}\right\rangle \\ F(f + \delta g) &\approx F(f) + \delta \left\langle (\sigma^2)'\left(K\dot{f}\right), K\dot{g}\right\rangle \\ G(f + \delta g) &\approx G(f) + \delta \left\{\left\langle \sigma\left(K\dot{f}\right), \dot{g}\right\rangle + \left\langle \sigma'\left(K\dot{f}\right)\dot{f}, K\dot{g}\right\rangle\right\} \end{aligned}$$

If $f = f^x$ is a minimizer then $\delta \mapsto \mathcal{I}_x(f+\delta g)$ has a minimum at $\delta = 0$ for all g. We expand

$$\begin{aligned}
\mathcal{I}_x(f+\delta g) &= \frac{(x-\rho G(f+\delta g))^2}{2\overline{\rho}^2 F(f+\delta g)} + \frac{1}{2}E(f+\delta g) \\
&\approx \frac{\left(x-\rho G(f) - \delta\rho\left\{\left\langle \sigma\left(K\dot{f}\right), \dot{g}\right\rangle + \left\langle \sigma'\left(K\dot{f}\right)\dot{f}, K\dot{g}\right\rangle\right\}\right)^2}{2\overline{\rho}^2\left[F(f) + \delta\left\langle (\sigma^2)'\left(K\dot{f}\right), K\dot{g}\right\rangle\right]} \\
&\quad + \frac{1}{2}E(f) + \delta\left\langle \dot{f}, \dot{g}\right\rangle \\
&\approx \frac{(x-\rho G(f))^2 - \delta 2\rho(x-\rho G(f))\left\{\left\langle \sigma\left(K\dot{f}\right), \dot{g}\right\rangle + \left\langle \sigma'\left(K\dot{f}\right)\dot{f}, K\dot{g}\right\rangle\right\}}{2\overline{\rho}^2 F(f)\left[1 + \frac{\delta}{F(f)}\left\langle (\sigma^2)'\left(K\dot{f}\right), K\dot{g}\right\rangle\right]} \\
&\quad + \frac{1}{2}E(f) + \delta\left\langle \dot{f}, \dot{g}\right\rangle \\
&\approx \frac{(x-\rho G(f))^2 - \delta 2\rho(x-\rho G(f))\left\{\left\langle \sigma\left(K\dot{f}\right), \dot{g}\right\rangle + \left\langle \sigma'\left(K\dot{f}\right)\dot{f}, K\dot{g}\right\rangle\right\}}{2\overline{\rho}^2 F(f)} \\
&\quad - \frac{(x-\rho G(f))^2}{2\overline{\rho}^2 F(f)}\frac{\delta}{F(f)}\left\langle (\sigma^2)'\left(K\dot{f}\right), K\dot{g}\right\rangle + \frac{1}{2}E(f) + \delta\left\langle \dot{f}, \dot{g}\right\rangle .
\end{aligned}$$

As a consequence, we must have, for $f = f^x$ and every $\dot{g} \in L^2[0,1]$

$$\begin{aligned}
0 &= \frac{d}{d\delta}\left\{\mathcal{I}_x(f+\delta g)\right\}_{\delta=0} = -\frac{\rho(x-\rho G(f))\left\{\left\langle \sigma\left(K\dot{f}\right), \dot{g}\right\rangle + \left\langle \sigma'\left(K\dot{f}\right)\dot{f}, K\dot{g}\right\rangle\right\}}{\overline{\rho}^2 F(f)} \\
&\quad - \frac{(x-\rho G(f))^2}{\overline{\rho}^2 F^2(f)}\left\langle (\sigma\sigma')\left(K\dot{f}\right), K\dot{g}\right\rangle + \left\langle \dot{f}, \dot{g}\right\rangle .
\end{aligned}$$

Recall $f_0^x = 0$, any x. We now test with $\dot{g} = 1_{[0,t]}$ for a fixed $t \in [0,1]$ and obtain

$$\begin{aligned}
f_t^x &= \frac{\rho(x-\rho G(f^x))\left\{\left\langle \sigma\left(K\dot{f}^x\right), 1_{[0,t]}\right\rangle + \left\langle \sigma'\left(K\dot{f}^x\right)\dot{f}^x, K\mathbf{1}_{[0,t]}\right\rangle\right\}}{\overline{\rho}^2 F(f^x)} \\
&\quad + \frac{(x-\rho G(f^x))^2}{\overline{\rho}^2 F^2(f^x)}\left\langle (\sigma\sigma')\left(K\dot{f}^x\right), K\mathbf{1}_{[0,t]}\right\rangle . \qquad \square
\end{aligned}$$

5.1. **Smoothness of the energy.** Having formally identified the first order condition for minimality in (5.1), we will now show that the energy $x \mapsto I(x)$ is a smooth function. More precisely, we will use the implicit function theorem to show that the minimizing configuration f^x is a smooth function in x (locally at $x = 0$). As $\mathcal{I}_x$ is a smooth function, too, this will imply smoothness of $x \mapsto \mathcal{I}_x(f^x) = I(x)$, at least in a neighborhood of 0.

As the Cameron-Martin space $\mathcal{H}$ of the process $\widehat{B}$ continuously embeds into $C([0,1])$, K maps H_0^1 continuously into $C([0,1])$, i.e., there is a constant $C > 0$ such that for any $f \in H_0^1$ we have

$$\left\|K\dot{f}\right\|_\infty \le C\,\|f\|_{H_0^1} . \tag{5.3}$$

This result will follow from

Lemma 5.3. *Let $(V_t : 0 \le t \le 1)$ be a continuous, centred Gaussian process and $\mathcal{H}$ its Cameron-Martin space. Then we have the continuous embedding $\mathcal{H} \hookrightarrow C[0,1]$. That is, for some constant C,*

$$\|h\|_\infty \le C \|h\|_{\mathcal{H}}.$$

Proof. By a fundamental result of Fernique, applied to the law of V as Gaussian measure on the Banach space $(C[0,1], \|\cdot\|_\infty)$, the random variable $\|V\|_\infty$ has Gaussian integrability. In particular,

$$\sigma^2 := \mathbb{E}(\|V\|_\infty^2) < \infty.$$

On the other hand, a generic element $h \in \mathcal{H}$ can be written as $h_t = E[V_t Z]$ where Z is a centred Gaussian random variable with variance $\|h\|_{\mathcal{H}}^2$. By Cauchy–Schwarz,

$$|h_t| \le E[|V_t|]^{1/2} \|h\|_{\mathcal{H}} \le \sigma \|h\|_{\mathcal{H}}$$

and conclude by taking the sup over on the l.h.s. over $t \in [0,1]$. □

Remark 5.4. Assume V is of Volterra form, i.e. $V_t = \int_0^t K(t,s)\, dB_s$. Then it can be shown (see [Dec05, Section 3]) that $\mathcal{H}$ is the image of L^2 under the map

$$K : \dot{f} \mapsto \widehat{f} := \left(t \mapsto \int_0^t K(t,s) \dot{f}_s ds \right)$$

and $\left\| K\dot{f} \right\|_{\mathcal{H}} = \left\| \dot{f} \right\|_{L^2}$. In particular then, applying the above with $h = K\dot{f} \in \mathcal{H}$, gives

$$\left\| K\dot{f} \right\|_\infty \le C \left\| K\dot{f} \right\|_{\mathcal{H}} = C \left\| \dot{f} \right\|_{L^2} = C \|f\|_{H_0^1}.$$

5.1.1. *The uncorrelated case.* We start with the case $\rho = 0$ as the formulas are much simpler in this case.

By Proposition 5.2, any local optimizer $f = f^x$ of the functional $\mathcal{I}_x : H_0^1 \to \mathbb{R}$ in the uncorrelated case $\rho = 0$ satisfies for any $t \in [0,1]$

$$f_t = \frac{x^2}{F^2(f)} \left\langle (\sigma\sigma') \left(K\dot{f} \right), K\mathbf{1}_{[0,t]} \right\rangle.$$

We define a map $H : H_0^1 \times \mathbb{R} \to H_0^1$ by

$$H(f,x)(t) := f_t - \frac{x^2}{F^2(f)} \left\langle (\sigma\sigma') \left(K\dot{f} \right), K\mathbf{1}_{[0,t]} \right\rangle. \tag{5.4}$$

Hence, for given $x \in \mathbb{R}$, any local optimizer f must solve $H(f,x) = 0$. As one particular solution is given by the pair $(0,0)$, we are in the realm of the implicit function theorem. We need to prove that

- $(f,x) \mapsto H(f,x)$ is locally smooth (in the sense of Fréchet);
- $DH(f,x) := \frac{\partial}{\partial f} H(f,x)$ is invertible in $(0,0)$.

Note that invertibility should hold for x small enough, as $DH(f,x) = \mathrm{id}_{H_0^1} - x^2 R$ for some R, which is invertible as long as R has a bounded norm for sufficiently small x.

Remark 5.5. The method of proof in this section is purely local in H_0^1. Hence, we do not really need C^∞-boundedness of σ, smoothness locally around 0 is enough. Note, however, that stochastic Taylor expansions used in Section 6 will actually require global smoothness of σ.

Lemma 5.6. *The functions $F : H_0^1 \to \mathbb{R}$ and $R_1 : H_0^1 \to C([0,1])$ defined by*

$$R_1(f)(t) \coloneqq \left\langle (\sigma\sigma')\left(K\dot{f}\right), K\mathbf{1}_{[0,t]}\right\rangle, \quad t \in [0,1],$$

are smooth in the sense of Fréchet.

Proof. For $N \geq 1$ we note that the Gateaux derivative of F satisfies

$$D^N F(f)\cdot(g_1,\dots,g_N) = \int_0^1 \frac{d^N}{dx^N}\sigma^2(K\dot{f})K\dot{g}_1\cdots K\dot{g}_N ds.$$

By Lemma 5.3, we can bound

$$\begin{aligned} \left|D^N F(f)\cdot(g_1,\dots,g_N)\right| &\leq \text{const} \int_0^1 |K\dot{g}_1(s)|\cdots|K\dot{g_N}(s)|\,ds \\ &\leq \text{const } \|K\dot{g}_1\|_\infty \cdots \|K\dot{g}_N\|_\infty \\ &\leq \text{const } C^N \|g_1\|_{H_0^1}\cdots\|g_N\|_{H_0^1}, \end{aligned}$$

for const $= \left\|\frac{d^n}{dx^n}\sigma^2\right\|_\infty$. Thus, $D^N F(f)$ is a multi-linear form on H_0^1 with operator norm $\left\|D^N F(f)\right\| \leq \left\|\frac{d^n}{dx^n}\sigma^2\right\|_\infty C^N$ independent of f. As $f \mapsto D^N F(f)$ is continuous, we conclude that $D^N F(f)$ as given above is, in fact, a Fréchet derivative.

Let us next consider the functional R_1. Note that

$$\left(D^N R_1(f)\cdot(g_1,\dots,g_N)\right)(t) = \left\langle \mathfrak{s}_N(K\dot{f})K\dot{g}_1\cdots K\dot{g}_N, K\mathbf{1}_{[0,t]}\right\rangle$$

for $\mathfrak{s}_N(x) \coloneqq \frac{d^N}{dx^N}\sigma(x)\sigma'(x)$. Hence, Assumption 2.5 implies that

$$\begin{aligned} \left\|D^N R_1(f)\cdot(g_1,\dots,g_N)\right\|_{H_0^1}^2 &= \int_0^1 \left(\int_t^1 \mathfrak{s}_N\left((K\dot{f})(s)\right)\prod_{i=1}^N (K\dot{g}_i)(s)K(s,t)ds\right)^2 dt \\ &\leq \|\mathfrak{s}_N\|_\infty^2 \prod_{i=1}^N \|K\dot{g}_i\|_\infty^2 \int_0^1\int_t^1 K(s,t)^2 dsdt \\ &\leq \|\mathfrak{s}_N\|_\infty^2\ C^{2N}\prod_{i=1}^N \|g_i\|_{H_0^1}^2 \int_0^1\int_0^s K(s,t)^2 dtds \\ &\leq \|\mathfrak{s}_N\|_\infty^2\ C^{2N}\int_0^1\int_0^s K(s,t)^2 dtds \prod_{i=1}^N \|g_i\|_{H_0^1}^2. \end{aligned}$$

We see that the multi-linear map $D^N R_1(f)$ has operator norm bounded by

$$\left\|D^N R_1(f)\right\| \leq \|\mathfrak{s}_N\|_\infty\ C^N \sqrt{\int_0^1\int_0^s K(s,t)^2 dtds},$$

independent of f. From continuity of $f \mapsto D^N R_1(f)$, it follows that $D^N R_1(f)$ is the N'th Fréchet derivative. □

Theorem 5.7 (Zero correlation)**.** *Assuming $\rho = 0$, the energy $I(x)$ is smooth in a neighborhood of $x = 0$.*

Proof. By construction, we have

$$DH(f,x) = \mathrm{id}_{H_0^1} - x^2 A(f)$$

for $A : H_0^1 \to \mathcal{L}(H_0^1, H_0^1)$ defined by

$$A(f) \coloneqq R(f) \otimes DF^{-2}(f) + F^{-2}(f) DR_1(f).$$

Here,

$$\big(R(f) \otimes DF^{-2}(f)\big) \cdot g = \underbrace{(DF^{-2}(f) \cdot g)}_{\in \mathbb{R}} \underbrace{R_1(f)}_{\in H_0^1}.$$

As verified above, H is smooth in the sense of Fréchet. Trivially, $DH(0,0) = \mathrm{id}_{H_0^1}$ is invertible and $H(0,0) = 0$. Therefore, the implicit function theorem implies that there are open neighborhoods U and V of $0 \in H_0^1$ and $0 \in \mathbb{R}$, respectively, and a smooth map $x \mapsto f^x$ from V to U such that $H(f^x, x) \equiv 0$ and f^x is unique in U with this property.

For the energy, we prove that $I(x) = \mathcal{I}_x(f^x)$ in a neighborhood of $x = 0$. First of all, we show that a minimizer exists. If not, there is a function $g \in H_0^1$ with $\mathcal{I}_x(g) < \mathcal{I}_x(f^x)$. For small enough x such a g must be inside a ball with radius ϵ around $0 \in H_0^1$, as $\mathcal{I}_x(g) \geq \frac{1}{2} \|g\|_{H_0^1}^2$ and $\lim_{x \to 0} \mathcal{I}_x(f^x) = 0$. Then note that for any $g \in H_0^1$

$$D^2\mathcal{I}_0(0) \cdot (g, g) = \|g\|_{H_0^1}^2 > 0,$$

where $D^2\mathcal{I}_x(f)$ denotes the second derivative of $f \mapsto \mathcal{I}_x(f)$. By continuity, $D^2\mathcal{I}_x(f)$ stays positive definite for (x, f) in a neighborhood of $(0,0)$. As noted, for x small enough, both g and f^x (and the line connecting them) lie in this neighborhood. For $h \coloneqq g - f^x$, this implies

$$\mathcal{I}_x(g) - \mathcal{I}_x(f^x) = D\mathcal{I}_x(f_x) \cdot h + \int_0^1 D^2\mathcal{I}_x(f^x + th) \cdot (h, h)\, dt > 0,$$

since $D\mathcal{I}_x(f_x) \cdot h = 0$ and $D^2\mathcal{I}_x(f^x + tsh) \cdot (h, h) > 0$. This contradicts the assumption that $\mathcal{I}_x(g) < \mathcal{I}_x(f^x)$, and we conclude that f^x is, indeed, a minimizer of $\mathcal{I}_x$, implying that $I(x) = \mathcal{I}_x(f^x)$ locally.

Finally, as $x \mapsto f^x$ is smooth and $(f, x) \mapsto \mathcal{I}_x(f) = \frac{x^2}{2F(f)} + \frac{1}{2} \|f\|_{H_0^1}^2$ is smooth, we see that $x \mapsto I(x) = \mathcal{I}_x(f^x)$ is smooth in a neighborhood of 0. $\square$

Remark 5.8. Classical counter-examples in the context of the *direct method* of calculus of variations show that the step of verifying the existence of a minimizer should not be taken too lightly. For instance, the functional

$$J(u) \coloneqq \int_0^1 \left[(u'(s)^2 - 1)^2 + u(s)^2\right] ds$$

does not have a minimizer in H_0^1, but J can be made arbitrarily close to 0 by choosing piecewise-linear functions u with slope $|u'| = 1$ oscillating around 0. We refer to any text book on calculus of variations. In the situation above, local "convexity" in the sense of a positive definite second derivative prevents this phenomenon. An alternative method of proof for the existence of a minimizer is to show that J is (lower semi-) continuous in the weak sense.

5.1.2. *The general case.* In the general case (cf. Proposition 5.2), we define the function $H : H_0^1 \times \mathbb{R} \to H_0^1$ by

$$
\begin{aligned}
H(f,x)(t) &\coloneqq f_t - \frac{\rho\,(x-\rho G\,(f))\left\{\left\langle \sigma\left(K\dot{f}\right), \mathbf{1}_{[0,t]}\right\rangle + \left\langle \sigma'\left(K\dot{f}\right)\dot{f}, K\mathbf{1}_{[0,t]}\right\rangle\right\}}{\overline{\rho}^2 F\,(f)} \\
&\quad + \frac{(x-\rho G\,(f))^2}{\overline{\rho}^2 F^2\,(f)}\left\langle (\sigma\sigma')\left(K\dot{f}\right), K\mathbf{1}_{[0,t]}\right\rangle \\
&= f_t - \frac{\rho\,(x-\rho G(f))}{\overline{\rho}^2 F(f)}\left(R_2(f)(t)+R_3(f)(t)\right) + \frac{(x-\rho G(f))^2}{\overline{\rho}^2 F(f)^2} R_1(f)(t),
\end{aligned}
\tag{5.5}
$$

where $R_2, R_3 : H_0^1 \to H_0^1$ are defined by

$$
R_2(f)(t) \coloneqq \left\langle \sigma(K\dot{f}), \mathbf{1}_{[0,t]}\right\rangle, \tag{5.6}
$$

$$
R_3(f)(t) \coloneqq \left\langle \sigma'(K\dot{f})\dot{f}, K\mathbf{1}_{[0,t]}\right\rangle, \tag{5.7}
$$

$t \in [0,1]$.

One easily checks that G, R_2, R_3 are smooth in the Fréchet sense.

Lemma 5.9. *The functions $G : H_0^1 \to \mathbb{R}$, $R_2 : H_0^1 \to H_0^1$ and $R_3 : H_0^1 \to H_0^1$ are smooth in Fréchet sense.*

Proof. The proof of smoothness is clear. We report the actual derivatives. For G we get

$$
\begin{aligned}
D^N G(f)\cdot(g_1,\dots,g_N) &= \left\langle \sigma^{(N)}\left(K\dot{f}\right)\dot{f}, \prod_{i=1}^N K\dot{g}_i\right\rangle + \\
&\quad + \sum_{k=1}^N \left\langle \sigma^{(N-1)}\left(K\dot{f}\right), \dot{g}_k \prod_{i\neq k} K\dot{g}_i\right\rangle .
\end{aligned}
$$

For R_2 and, respectively, R_3, we obtain

$$
\left(D^N R_2(f)\cdot(g_1,\dots,g_N)\right)(t) = \int_0^t \sigma^{(N)}\left((K\dot{f})(s)\right)\prod_{i=1}^N (K\dot{g}_i)(s)ds,
$$

and

$$
\begin{aligned}
\left(D^N R_3(f)\cdot(g_1,\dots,g_N)\right)(t) &= \left\langle \sigma^{(N+1)}\left(K\dot{f}\right)\dot{f}K\mathbf{1}_{[0,t]}, \prod_{i=1}^N K\dot{g}_i\right\rangle + \\
&\quad + \sum_{k=1}^N \left\langle \sigma^{(N)}\left(K\dot{f}\right)K\mathbf{1}_{[0,t]}, \dot{g}_k \prod_{i\neq k} K\dot{g}_i\right\rangle . \qquad \square
\end{aligned}
$$

Theorem 5.10. *Assume σ smooth and (for simplicity only) bounded with bounded derivatives of all order. Then the energy $I(x)$ is smooth in a neighborhood of $x = 0$.*

Proof. The proof is similar to the proof of Theorem 5.7. In fact, the only difference is in establishing invertibility of $DH(0,0)$ and the existence of a minimizer.

Note that (5.5) contains three terms. The derivative of the first term ($f \mapsto f$) is always equal to $\mathrm{id}_{H_0^1}$. For the second term, we note that

$$
(x-\rho G(f))|_{x=0,\, f=0} = 0.
$$

Hence, the only non-vanishing contribution to the derivative of the second term evaluated in direction $g \in H_0^1$ at $x = 0$, $f = 0$ and $t \in [0,1]$ is

$$\frac{\rho^2 DG(0) \cdot g}{\overline{\rho}^2 F(0)} (R_2(0) + R_3(0)) = \frac{\rho^2 \sigma_0 g(1)}{\overline{\rho}^2 \sigma_0^2} (\sigma_0 t + 0) = \frac{\rho^2}{\overline{\rho}^2} g(1) t.$$

For the same reason, the derivative of the third term at $(f, x) = (0, 0)$ vanishes entirely. Hence,

$$(DH(0,0) \cdot g)(t) = g(t) + \frac{\rho^2}{\overline{\rho}^2} g(1) t.$$

It is easy to see that $g \mapsto DH(0,0) \cdot g$ is invertible. Indeed, let us construct the pre-image $g = DH(0,0)^{-1} \cdot h$ of some $h \in H_0^1$. At $t = 1$ we have

$$\frac{\overline{\rho}^2 + \rho^2}{\overline{\rho}^2} g(1) = h(1),$$

implying $g(1) = \overline{\rho}^2 h(1)$. For $0 \le t < 1$, we then get

$$g(t) + \frac{\rho^2}{\overline{\rho}^2} g(1) t = g(t) + \frac{\rho^2}{\overline{\rho}^2} \overline{\rho}^2 h(1) t = g(t) + \rho^2 h(1) t = h(t),$$

or $g(t) = h(t) - \rho^2 h(1) t$.

For existence of the minimizer, note that

$$D^2 J_0(0) \cdot (g, g) = \frac{\rho^2}{\overline{\rho}^2} g(1)^2 + \|g\|_{H_0^1}^2,$$

which is again positive definite. □

5.2. **Energy expansion.** Having established smoothness of the energy I as well as of the minimizing configuration $x \mapsto f^x$ locally around $x = 0$, we can proceed with computing the Taylor expansion of f^x around $x = 0$. We will once more rely on the first order optimality condition given in Proposition 5.2. Plugging the Taylor expansion of f^x into $\mathcal{I}_x$ will then give us the local Taylor expansion of $I(x)$.

5.2.1. *Expansion of the minimizing configuration.*

Theorem 5.11. *We have*

$$\begin{aligned} f_t^x &= \alpha_t x + \beta_t \frac{x^2}{2} + O\left(x^3\right), \\ \alpha_t &= \frac{\rho}{\sigma_0} t, \\ \beta_t &= 2 \frac{\sigma_0'}{\sigma_0^3} \left[\rho^2 \left\langle K1 \, , \mathbf{1}_{[0,t]} \right\rangle + \left\langle K\mathbf{1}_{[0,t]} \, , 1 \right\rangle - 3\rho^2 t \left\langle K1 \, , 1 \right\rangle \right]. \end{aligned}$$

Remark 5.12 (Non-Markovian transversality). In the RL-fBM case, $K(t,s) = \sqrt{2H} \, |t-s|^\gamma$ with $\gamma = H - 1/2$ one computes

$$\left\langle 1, K1_{[0,t]} \right\rangle = \frac{1}{(1+\gamma)(2+\gamma)} \left\{ 1 - (1-t)^{2+\gamma} \right\} \in C^1 [0,1].$$

Interestingly, the transversality condition known from the Markovian setting ($q_1 = 0$, which readily translates to $\dot{f}_1^x = 0$ there) remains valid here (for $\rho = 0$), at least to order x^2, in the sense that

$$\dot{f}_t^x \approx \beta_t \frac{x^2}{2} = (\text{const}) \, (1-t)^{1+\gamma} \, |_{t=1} = 0$$

Proof of Theorem 5.11. **First order expansion:**
Up to the order needed in order to get the first order term, we have

$$\begin{aligned}
f_t^x &= \alpha_t x + O(x^2),\\
\dot{f}_t^{\,x} &= \dot\alpha_t x + O(x^2),\\
\sigma(K\dot f^x) &= \sigma_0 + \sigma_0' K\dot\alpha\ x + O(x^2),\\
\sigma'(K\dot f^x) &= \sigma_0' + \sigma_0'' K\dot\alpha\ x + O(x^2),\\
F(f^x) &= \langle \sigma^2(K\dot f^x), 1\rangle\\
&= \sigma_0^2 + O(x),\\
G(f^x) &= \langle \sigma(K\dot f^x), \dot f^x\rangle\\
&= \langle \sigma_0\,, \dot\alpha\rangle\, x + O(x^2).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\langle \sigma(K\dot f^x), \mathbf{1}_{[0,t]}\rangle &= \sigma_0 t + O(x),\\
\langle \sigma'(K\dot f^x)\dot f^x, K\mathbf{1}_{[0,t]}\rangle &= O(x),\\
\langle \sigma\sigma'(K\dot f^x), K\mathbf{1}_{[0,t]}\rangle &= O(1),\\
x - \rho G(f^x) &= (1-\rho\sigma_0\alpha_1)x + O(x^2),\\
(x-\rho G(f^x))^2 &= O(x^2).
\end{aligned}$$

This yields for the first order term in (5.2)

$$\alpha_t = \frac{\rho(1-\rho\sigma_0\alpha_1)}{\overline{\rho}^2\sigma_0} t.$$

Setting $t=1$, we get

$$\alpha_1 = \frac{\rho}{\overline{\rho}^2\sigma_0} - \frac{\rho^2}{\overline{\rho}^2}\alpha_1,$$

which is solved by $\alpha_1 = \frac{\rho}{\sigma_0}$. Inserting this term back into the equation for α_t, we get

$$\alpha_t = \frac{\rho}{\sigma_0} t. \tag{5.8}$$

Second order expansion:

Using (5.8) and the ansatz $f_t^x = \alpha_t x + \frac{1}{2}\beta_t x^2 + O(x^3)$, we re-compute the relevant terms appearing in the (5.2). We have

$$\sigma(K\dot f^x(s)) = \sigma_0 + \sigma_0' \frac{\rho}{\sigma_0}(K1)(s)x + O(x^2)$$

and analogously for σ replaced by σ', $\sigma\sigma'$. This implies

$$\begin{aligned}
\left\langle \sigma(K\dot f^x)\,, \mathbf{1}_{[0,t]}\right\rangle &= \sigma_0 t + \sigma_0'\frac{\rho}{\sigma_0}\left\langle K1\,, \mathbf{1}_{[0,t]}\right\rangle x + O(x^2),\\
\left\langle \sigma'(K\dot f^x)\dot f^x\,, K\mathbf{1}_{[0,t]}\right\rangle &= \rho\frac{\sigma'}{\sigma_0}\left\langle K\mathbf{1}_{[0,t]}\,, 1\right\rangle x + O(x^2),\\
\left\langle \sigma\sigma'(K\dot f^x)\,, K\mathbf{1}_{[0,t]}\right\rangle &= \sigma_0\sigma_0'\left\langle K\mathbf{1}_{[0,t]}\,, 1\right\rangle + O(x).
\end{aligned}$$

Using the notation introduced earlier, we have

$$F(f^x) = \sigma_0^2 + 2\sigma_0'\rho \langle K1\,, 1\rangle\, x + O(x^2),$$
$$G(f^x) = \rho x + \left(\frac{1}{2}\sigma_0\beta_1 + \rho^2\frac{\sigma_0'}{\sigma_0^2}\langle K1\,, 1\rangle\right) x^2 + O(x^3).$$

This directly implies

$$x - \rho G(f^x) = \overline{\rho}^2 x - \rho\left(\frac{1}{2}\sigma_0\beta_1 + \rho^2\frac{\sigma_0'}{\sigma_0^2}\langle K1\,, 1\rangle\right) x^2 + O(x^3),$$
$$(x-\rho G(f^x))^2 = \overline{\rho}^4 x^2 - 2\overline{\rho}^2\rho\left(\frac{1}{2}\sigma_0\beta_1 + \rho^2\frac{\sigma_0'}{\sigma_0^2}\langle K1\,, 1\rangle\right) x^3 + O(x^4).$$

We next compute some auxiliary terms appearing in (5.2).

$$\begin{aligned} N_1 &\coloneqq \rho(x-\rho G(f^x))\left(\left\langle \sigma(K\dot{f}^x)\,, \mathbf{1}_{[0,t]}\right\rangle + \left\langle \sigma'(K\dot{f}^x)\dot{f}^x\,, K\mathbf{1}_{[0,t]}\right\rangle\right) \\ &= \rho\overline{\rho}^2\sigma_0 t x + \left[\rho^2\overline{\rho}^2\frac{\sigma_0'}{\sigma_0}\left(\left\langle K1\,, \mathbf{1}_{[0,t]}\right\rangle + \left\langle K\mathbf{1}_{[0,t]}\,, 1\right\rangle\right)\right. \\ &\quad \left. - \rho^4\frac{\sigma_0'}{\sigma_0} t\langle K1\,, 1\rangle - \frac{1}{2}\rho^2\sigma_0^2 t\beta_1\right] x^2 + O(x^3) \end{aligned}$$

The corresponding denominator is $\overline{\rho}^2 F(f^x)$. Using the formula

$$\frac{a_1 x + a_2 x^2 + O(x^3)}{b_0 + b_1 x + O(x^2)} = \frac{a_1}{b_0}x + \frac{a_2 b_0 - a_1 b_1}{b_0^2}x^2 + O(x^3),$$

we obtain

$$\begin{aligned} \frac{N_1}{\overline{\rho}^2 F(f^x)} &= \frac{\rho}{\sigma_0} t x + \left[\rho^2\frac{\sigma_0'}{\sigma_0^3}\left(\left\langle K1\,, \mathbf{1}_{[0,t]}\right\rangle + \left\langle K\mathbf{1}_{[0,t]}\,, 1\right\rangle\right)\right. \\ &\quad \left. - \left(\frac{\rho^4}{\overline{\rho}^2} + 2\rho^2\right)\frac{\sigma_0'}{\sigma_0^3} t\langle K1\,, 1\rangle - \frac{1}{2}\frac{\rho^2}{\overline{\rho}^2}\beta_1 t\right] x^2 + O(x^3) \end{aligned} \tag{5.9}$$

For the second term in (5.2), let

$$N_2 \coloneqq (x-\rho G(f^x))^2\left\langle (\sigma\sigma')(K\dot{f}^x)\,, K\mathbf{1}_{[0,t]}\right\rangle = \overline{\rho}^4\sigma_0\sigma_0'\left\langle K\mathbf{1}_{[0,t]}\,, 1\right\rangle x^2 + O(x^3).$$

The corresponding denominator is $\overline{\rho}^2 F(f^x)^2 = \overline{\rho}^2\sigma_0^4 + O(x)$. Hence,

$$\frac{N_2}{\overline{\rho}^2 F(f^x)^2} = \overline{\rho}^2\frac{\sigma_0'}{\sigma_0^3}\left\langle K\mathbf{1}_{[0,t]}\,, 1\right\rangle x^2 + O(x^3). \tag{5.10}$$

Combining (5.9) and (5.10), we get

$$\begin{aligned} f_t^x = \frac{\rho}{\sigma_0} t x &+ \left[\rho^2\frac{\sigma_0'}{\sigma_0^3}\left(\left\langle K1\,, \mathbf{1}_{[0,t]}\right\rangle + \left\langle K\mathbf{1}_{[0,t]}\,, 1\right\rangle\right) - \frac{\rho^4}{\overline{\rho}^2}\frac{\sigma_0'}{\sigma_0^3} t\langle K1\,, 1\rangle\right. \\ &\left. - \frac{1}{2}\frac{\rho^2}{\overline{\rho}^2}\beta_1 t - 2\rho^2\frac{\sigma_0'}{\sigma_0^3} t\langle K1\,, 1\rangle + \overline{\rho}^2\frac{\sigma_0'}{\sigma_0^3}\left\langle K\mathbf{1}_{[0,t]}\,, 1\right\rangle\right] x^2 + O(x^3) \end{aligned}$$

We shall next compute β_1. Taking the second order terms on both sides and letting $t = 1$, we obtain

$$\frac{1}{2}\beta_1 = \rho^2 \frac{\sigma_0'}{\sigma_0^3} 2 \langle K1\,, 1\rangle - \frac{\rho^4}{\overline{\rho}^2}\frac{\sigma_0'}{\sigma_0^3} \langle K1\,, 1\rangle \\ - \frac{1}{2}\frac{\rho^2}{\overline{\rho}^2}\beta_1 - 2\rho^2 \frac{\sigma_0'}{\sigma_0^3} \langle K1\,, 1\rangle + \overline{\rho}^2 \frac{\sigma_0'}{\sigma_0^3} \langle K1\,, 1\rangle .$$

Moving β_1 to the other side with $1 + \frac{\rho^2}{\overline{\rho}^2} = \frac{1}{\overline{\rho}^2}$ and collecting terms on the right hand side, we arrive at

$$\frac{1}{2}\frac{1}{\overline{\rho}^2}\beta_1 = \frac{\sigma_0'}{\sigma_0^3} \langle K1\,, 1\rangle \left(2\rho^2 - \frac{\rho^4}{\overline{\rho}^2} - 2\rho^2 + \overline{\rho}^2\right) = \frac{1-2\rho^2}{\overline{\rho}^2}\frac{\sigma_0'}{\sigma_0^3} \langle K1\,, 1\rangle$$

We conclude that

$$\beta_1 = 2(1-2\rho^2)\frac{\sigma_0'}{\sigma_0^3} \langle K1\,, 1\rangle$$

Hence, we obtain

$$\beta_t = 2\frac{\sigma_0'}{\sigma_0^3}\left[\rho^2 \left\langle K1\,, \mathbf{1}_{[0,t]}\right\rangle + \left\langle K\mathbf{1}_{[0,t]}\,, 1\right\rangle - 3\rho^2 t \langle K1\,, 1\rangle\right] . \qquad \square$$

5.2.2. *Energy expansion in the general case.* Now we compute the Taylor expansion of $I(x)$ as defined in Proposition 5.1. We start with the second term. Plugging in the optimal path $f_t^x = \alpha_t x + \frac{1}{2}\beta_t x^2 + O(x^3)$ (and using $\left\langle \dot\beta\,, 1\right\rangle = \beta_1$ as $\beta_0 = 0$) we obtain

$$\frac{1}{2}\left\langle \dot f^x\,, \dot f^x\right\rangle = \frac{1}{2}\frac{\rho^2}{\sigma_0^2}x^2 + \frac{1}{2}\frac{\rho}{\sigma_0}\beta_1 x^3 + O(x^4).$$

Inserting $\beta_1 = 2(1-2\rho^2)\frac{\sigma_0'}{\sigma_0^3}\langle K1\,, 1\rangle$ into the above formula for $(x - \rho G(f^x))^2$, we get

$$(x - \rho G(f^x))^2 = \overline{\rho}^4 x^2 - 2\overline{\rho}^4 \rho \frac{\sigma_0'}{\sigma_0^2}\langle K1\,, 1\rangle x^3 + O(x^4).$$

Recall the denominator

$$2\overline{\rho}^2 F(f^x) = 2\overline{\rho}^2\sigma_0^2 + 4\overline{\rho}^2\sigma_0'\rho \langle K1\,, 1\rangle x + O(x^2).$$

Using the expansion of a fraction

$$\frac{a_2x^2 + a_3x^3 + O(x^4)}{b_0 + b_1 x + O(x^2)} = \frac{a_2}{b_0}x^2 + \frac{a_3b_0 - a_2b_1}{b_0^2}x^3 + O(x^4),$$

we obtain from

$$\begin{aligned}\frac{(x - \rho G(f^x))^2}{2\overline{\rho}^2 F(f^x)} &= \frac{\overline{\rho}^4}{2\overline{\rho}^2\sigma_0^2}x^2 + \\ &\quad + \frac{\left(-2\overline{\rho}^4\rho\frac{\sigma_0'}{\sigma_0^2}\langle K1\,, 1\rangle\right)2\overline{\rho}^2\sigma_0^2 - \overline{\rho}^4\left(4\overline{\rho}^2\sigma_0'\rho\langle K1\,, 1\rangle\right)}{4\overline{\rho}^4\sigma_0^4}x^3 + O(x^4) \\ &= \frac{\overline{\rho}^2}{2\sigma_0^2}x^2 - 2\overline{\rho}^2\rho\frac{\sigma_0'}{\sigma_0^4}\langle K1\,, 1\rangle x^3 + O(x^4).\end{aligned}$$

We note that

$$\frac{1}{2}\frac{\rho}{\sigma_0}\beta_1 - 2\overline{\rho}^2\rho\frac{\sigma_0'}{\sigma_0^4}\langle K1\,, 1\rangle = \left((1-2\rho^2) - 2(1-\rho^2)\right)\rho\frac{\sigma_0'}{\sigma_0^4}\langle K1\,, 1\rangle = -\rho\frac{\sigma_0'}{\sigma_0^4}\langle K1\,, 1\rangle .$$

Adding both terms, we arrive at the

Proposition 5.13. *The energy expansion to third order gives*

$$I(x) = \frac{1}{2\sigma_0^2}x^2 - \rho\frac{\sigma_0'}{\sigma_0^4}\langle K1\,,1\rangle\,x^3 + O(x^4).$$

5.2.3. *Energy expansion for the Riemann-Liouville kernel.* Let us specialize the energy expansion given in Proposition 5.13 for the Riemann-Liouville fBm. Choose $\gamma = H - \frac{1}{2}$ and recall that the kernel K takes the form $K(t,s) = (t-s)^\gamma$. We get

$$(K1)(t) = \int_0^t K(t,s)ds = \int_0^t (t-s)^\gamma ds = \frac{t^{1+\gamma}}{1+\gamma}.$$

The key term $\langle K1\,,1\rangle$ appearing in the energy expansion now gives

$$\langle K1\,,1\rangle = \int_0^1 (K1)(t)dt = \int_0^1 \frac{t^{1+\gamma}}{1+\gamma}dt = \frac{1}{(1+\gamma)(2+\gamma)} = \frac{1}{(H+1/2)(H+3/2)}.$$

Plugging formula (5.2.3) into the energy expansion, we obtain the energy expansion for the Riemann-Liouville fractional Browian motion

$$I(x) = \frac{1}{2\sigma_0^2}x^2 - \frac{\rho}{(H+1/2)(H+3/2)}\frac{\sigma_0'}{\sigma_0^4}x^3 + O(x^4).$$

For completeness, let us also fully describe the time-dependence of the second order term β_t in the expansion of the optimal trajectory f_t^x. Unlike the first order time, here we do not have a linear movement any more. Indeed

$$\left\langle K1\,,\mathbf{1}_{[0,t]}\right\rangle = \int_0^t (K1)(s)ds = \int_0^t \frac{s^{1+\gamma}}{1+\gamma}ds = \frac{t^{2+\gamma}}{(1+\gamma)(2+\gamma)}, \tag{5.11}$$

$$\left\langle K\mathbf{1}_{[0,t]}\,,1\right\rangle = \frac{1}{(1+\gamma)(2+\gamma)}\left(1-(1-t)^{2+\gamma}\right). \tag{5.12}$$

6. Proof of the pricing formula

Fix $x \geq 0$ and $\widehat{x} = \frac{\varepsilon}{\widehat{\varepsilon}}x$ where $\varepsilon = t^{1/2}$ and $\widehat{\varepsilon} = t^H = \varepsilon^{2H}$. We have

$$\begin{aligned} c(\widehat{x},t) &= E\left(\exp\left(X_t\right) - \exp\widehat{x}\right)^+ \\ &= E\left(\exp\left(X_1^\varepsilon\right) - \exp\widehat{x}\right)^+ \\ &= E\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{X}_1^\varepsilon\right) - \exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}x\right)\right)^+ \end{aligned}$$

where we recall

$$\widehat{X}_1^\varepsilon \equiv \frac{\widehat{\varepsilon}}{\varepsilon}X_1^\varepsilon = \int_0^1 \sigma(\widehat{\varepsilon}\widehat{B})\widehat{\varepsilon}d\left(\overline{\rho}W + \rho B\right),\ \ \widehat{X}_0^\varepsilon = 0.$$

Consider a Cameron-Martin perturbation of $\widehat{X}_1^\varepsilon$. That is, for a Cameron-Martin path $\mathrm{h} = (h,f) \in H_0^1 \times H_0^1$ consider a measure change corresponding to a transformation $\widehat{\varepsilon}(W,B) \rightsquigarrow \widehat{\varepsilon}(W,B) + (h,f)$ (transforming the Brownian motions to Brownian motions with drift), we obtain the Girsanov density

$$G_\varepsilon = \exp\left(-\frac{1}{\widehat{\varepsilon}}\int \dot{h}dW - \frac{1}{\widehat{\varepsilon}}\int \dot{f}dB - \frac{1}{2\widehat{\varepsilon}^2}\int\left(\dot{h}^2 + \dot{f}^2\right)dt\right). \tag{6.1}$$

For further details we refer to [BO11, Theorem 2.4]. Under the new measure, $\widehat{X}_1^\varepsilon$ can become $\widehat{Z}_1^\varepsilon$, where

$$\widehat{Z}_1^\varepsilon = \int_0^1 \sigma(\widehat{\varepsilon}\widehat{B} + \widehat{f}) \left[\widehat{\varepsilon} d\left(\overline{\rho} W + \rho B\right) + d\left(\overline{\rho} h + \rho f\right)\right].$$

Definition 6.1. For fixed $x \geq 0$, write $(h, f) \in \mathcal{K}^x$ if $\Phi_1\left(h, f, \widehat{f}\right) = x$. Call such (h, f) admissible for arrival at log-strike x. Call (h^x, f^x) the cheapest admissible control, which attains

$$I(x) = \inf_{h, f \in H_0^1} \left\{ \frac{1}{2} \int_0^1 \dot{h}^2 dt + \frac{1}{2} \int_0^1 \dot{f}^2 dt : \Phi_1\left(h, f, \widehat{f}\right) = x \right\},$$

where we recall that $\widehat{f} = K \dot{f}$ and

$$\Phi_1(h, f, \widehat{f}) = \int_0^1 \sigma(\widehat{f}) d\left(\overline{\rho} h + \rho f\right).$$

A look at (6) reveals that for any Cameron-Martin path (h, f), the perturbed random variable $\widehat{Z}_1^\varepsilon$ admits a stochastic Taylor expansion with respect to $\widehat{\varepsilon}$.

Lemma 6.2. *Fix $(h, f) \in \mathcal{K}^x$ and define $\widehat{Z}_1^\varepsilon$ accordingly. Then*

$$\widehat{Z}_1^\varepsilon = x + \widehat{\varepsilon} g_1 + \widehat{\varepsilon}^2 R_2, \tag{6.2}$$

where g_1 is a Gaussian random variable, given explicitly by

$$g_1 = \int_0^1 \{\sigma(\widehat{f}_t) d\left(\overline{\rho} W_t + \rho B_t\right) + \sigma'(\widehat{f}_t) \widehat{B}_t d\left(\overline{\rho} h_t + \rho f_t\right)\}, \tag{6.3}$$

and

$$\begin{aligned} R_2 = {} & \int_0^1 \sigma'\left(\widehat{f}_t\right) \widehat{B}_t d\left(\overline{\rho} W_t + \rho B_t\right) + \\ & + \frac{1}{2\widehat{\varepsilon}^2} \int_0^{\widehat{\varepsilon}} \int_0^1 \sigma''\left(\zeta \widehat{B}_t + \widehat{f}_t\right) \widehat{B}_t^2 \left[\widehat{\varepsilon} d\left(\overline{\rho} W_t + \rho B_t\right) + d\left(\overline{\rho} h_t + \rho f_t\right)\right] \left(\widehat{\varepsilon} - \zeta\right) d\zeta. \end{aligned} \tag{6.4}$$

Proof. By a stochastic Taylor expansion for the controlled process $\widehat{Z}_t^\varepsilon$ with control $(h, f) \in \mathcal{K}^x$ as in Definition 6.1 and thanks to $\sigma \in C^2$, we have at $t = 1$

$$\begin{aligned} \widehat{Z}_1^\varepsilon &= \int_0^1 \sigma(\widehat{\varepsilon}\widehat{B} + \widehat{f}) \left[\widehat{\varepsilon} d\left(\overline{\rho} W + \rho B\right) + d\left(\overline{\rho} h + \rho f\right)\right] \\ &= \int_0^1 \sigma(\widehat{f}) d\left(\overline{\rho} h + \rho f\right) + \widehat{\varepsilon} \int_0^1 \{\sigma(\widehat{f}) d\left(\overline{\rho} W + \rho B\right) + \sigma'(\widehat{f}) \widehat{B} d\left(\overline{\rho} h + \rho f\right)\} + O\left(\widehat{\varepsilon}^2\right). \end{aligned}$$

Collecting terms in powers of $\widehat{\varepsilon}$ and with $g_1(\omega)$ as in (6.3), we have

$$\widehat{Z}_1^\varepsilon = \int_0^1 \sigma(\widehat{f}) d\left(\overline{\rho} h + \rho f\right) + \widehat{\varepsilon} g_1(\omega) + O(\varepsilon^2),$$

furthermore, since $(h, f) \in \mathcal{K}^x$, by the definition of Φ_1, it holds that

$$\int_0^1 \sigma(\widehat{f}) d\left(\overline{\rho} h + \rho f\right) = x.$$

This proves the statement (6.2) and the statement that g_1 is Gaussian is immediate from the form (6.3). $\square$

Lemma 6.3. *We have*

$$\int \dot{h}^x dW + \int \dot{f}^x dB = I'(x)\, g_1(\omega).$$

Proof. Appendix □

We are now ready to prove the pricing formula from Section 3.

Proof of Theorem 3.2. With a Girsanov factor (all integrals on $[0,1]$), evaluated at the minimizer,

$$\begin{aligned} G_\varepsilon &= e^{-\frac{1}{\varepsilon}\int \dot{h} dW - \frac{1}{\varepsilon}\int \dot{f} dB - \frac{1}{2\widehat{\varepsilon}^2}\int \left(\dot{h}^2 + \dot{f}^2\right) dt} \\ G_\varepsilon|_* &= e^{-\frac{I(x)}{\widehat{\varepsilon}^2}} e^{-\frac{I'(x) g_1(\omega)}{\widehat{\varepsilon}}} \end{aligned}$$

we have, setting $\widehat{U}^\varepsilon := \widehat{Z}_1^\varepsilon - x = \widehat{\varepsilon} g_1 + \widehat{\varepsilon}^2 R_2$

$$\begin{aligned} c(\widehat{x}, t) &= E\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{Z}_1^\varepsilon\right) - \exp\left(\frac{\varepsilon}{\widehat{\varepsilon}} x\right)\right)^+ G_\varepsilon|_*\right] \\ &= e^{\frac{\varepsilon}{\widehat{\varepsilon}} x} E\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^\varepsilon\right) - 1\right)^+ G_\varepsilon|_*\right] \\ &= e^{-\frac{I(x)}{\widehat{\varepsilon}^2}} e^{\frac{\varepsilon}{\widehat{\varepsilon}} x} E\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^\varepsilon\right) - 1\right)^+ e^{-\frac{I'(x) g_1(\omega)}{\widehat{\varepsilon}}}\right] \\ &= e^{-\frac{I(x)}{\widehat{\varepsilon}^2}} e^{\frac{\varepsilon}{\widehat{\varepsilon}} x} E\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^\varepsilon\right) - 1\right) e^{-\frac{I'(x)}{\widehat{\varepsilon}^2}\widehat{U}^\varepsilon} e^{I'(x) R_2} \mathbf{1}_{\widehat{U}^\varepsilon \geq 0}\right]. \\ &= e^{-\frac{I(x)}{\widehat{\varepsilon}^2}} e^{\frac{\varepsilon}{\widehat{\varepsilon}} x} J(\varepsilon, x). \end{aligned}$$

□

7. Proof of the moderate deviation expansions

Higher-order moderate deviations expansions in Theorem (3.4) follow from the pricing formula, provided we can show that the remainder term $J(\varepsilon, x)$ is bounded from above and below by a power in ε. By a large deviation estimate, it is enough to do so for

$$\begin{aligned} J_\delta(\varepsilon, x) &= E_\delta\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^\varepsilon\right) - 1\right) e^{-\frac{I'(x)}{\widehat{\varepsilon}^2}\widehat{U}^\varepsilon} e^{I'(x) R_2}\right] \\ \text{with } E_\delta[\,\cdot\,] &= E[(\cdot)\mathbf{1}_{\widehat{U}^\varepsilon \geq 0}\mathbf{1}_{\widehat{Z}_1^\varepsilon \in B(\mathfrak{h}^0, \delta)}] \end{aligned}$$

where $\mathfrak{h}^0 = (h^x, f^x) \in \mathcal{K}^x$ is an optimal control, and $B(\mathfrak{h}^0, \delta) \subset C^\lambda([0,1], \mathbb{R}^2)$ denotes a δ ball for a $0 < \lambda < H$ around the optimal control $\mathfrak{h}^0$ in the λ-Hölder topology

$$||f||_\lambda := ||f||_\infty + \sup_{0 \leq s \leq t \leq 1} \frac{|f(t) - f(s)|}{(t-s)^\lambda}, \quad \text{for} \quad f \in C^\lambda([0,1], \mathbb{R}^2).$$

By a large deviations estimate

$$|J(\varepsilon, x) - J_\delta(\varepsilon, x)| \leq e^{-d/\epsilon^2}$$

for some $d > 0$. We refer to [BA88, Lemma 1.32]. Note that $J(\varepsilon)$ as defined in [BA88] contains the factor $exp(-a/\varepsilon^2)$ with $a = I(x)$. See also [BO15, Section 4, Step 1] for a straight-forward adaptation of this to a fractional setting. Note that

R_2 depends on both x and $\widehat{\varepsilon}$. Nonetheless we know from [BO11, Section A.1] (see also [BA88]) that there is $c_2 > 0$, uniformly for x, $\widehat{\varepsilon}$ small enough

$$P_\delta\left[|R_2| > r\right] \lesssim \exp\left(-c_2 r\right)$$

so that for x, $\widehat{\varepsilon}$ small enough (so that $I'(x)$ arbitrarily small) $M^{x,\varepsilon} = e^{\left|I'(x)R_2\right|}$ has finite expectation.

Upper bound. Since $I'(x) \geq 0$ for x small enough,

$$\begin{aligned} J_\delta(\varepsilon, x) &\leq E_\delta\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^\varepsilon\right) - 1\right) e^{I'(x)R_2}\right] \\ &\leq \left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\delta\right) - 1\right) E_\delta\left[e^{I'(x)R_2}\right] \\ &\leq C\frac{\varepsilon}{\widehat{\varepsilon}}\delta \end{aligned}$$

uniformly for x and ε in a neighbourhood of $0+$.

Lower bound. For $p > 1$ with Hölder conjugate p', remember $1/p + 1/p' = 1$, we have

$$\begin{aligned} & E_\delta\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^\varepsilon\right) - 1\right)^{\frac{1}{p}} e^{-\frac{I'(x)}{\widehat{\varepsilon}^2 p}\widehat{U}^\varepsilon}\right] \\ \leq\ & E_\delta\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^\varepsilon\right) - 1\right)^{\frac{1}{p}} e^{-\frac{I'(x)}{\widehat{\varepsilon}^2 p}\widehat{U}^\varepsilon} e^{-\frac{1}{p}I'(x)R_2} e^{\frac{1}{p}I'(x)R_2}\right] \\ \lesssim\ & \underbrace{E_\delta\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^\varepsilon\right) - 1\right) e^{-\frac{I'(x)}{\widehat{\varepsilon}^2}\widehat{U}^\varepsilon} e^{I'(x)R_2}\right]}_{J_\delta(\varepsilon,x)}^{\frac{1}{p}} E_\delta\left[e^{-\frac{p'}{p}I'(x)R_2}\right]^{\frac{1}{p'}}. \end{aligned}$$

For fixed $\delta, p \in (1, \infty)$, uniformly in x small enough

$$M^x_{\delta,p} := E_\delta\left[-e^{\frac{p'}{p}I'(x)R_2}\right]^{\frac{1}{p'}} < \infty.$$

On the other hand, by elementary analysis, for suitable non-zero $\gamma = \gamma(\varepsilon), \overline{\delta}(\varepsilon)$ polynomial in ε,

$$\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}u\right) - 1\right)^{\frac{1}{p}} e^{-\frac{I'(x)}{\widehat{\varepsilon}^2 p}u} \geq \gamma u^{\frac{1}{p}} \text{ for } u \in [0, \overline{\delta}],$$

so that

$$E_\delta\left[\left(\exp\left(\frac{\varepsilon}{\widehat{\varepsilon}}\widehat{U}^\varepsilon\right) - 1\right)^{\frac{1}{p}} e^{-\frac{I'(x)}{\widehat{\varepsilon}^2 p}\widehat{U}^\varepsilon}\right] \geq \gamma E_{\delta\wedge\overline{\delta}}\left[\left|\widehat{U}^\varepsilon\right|^{\frac{1}{p}}\right] =: N^x_{\varepsilon,\delta,p}.$$

With $\widehat{U}^\varepsilon \sim \widehat{\varepsilon}g_1(\omega)$, one sees that $N^x_{\varepsilon,\delta,p}$ scales as power of ε. All in all,

$$N^x_{\varepsilon,\delta,p} \lesssim J_\delta(\varepsilon, x)^{\frac{1}{p}} M^x_{\delta,p}$$

which implies the lower bound. Summarizing, we obtain

Proposition 7.1. *There are exponents* $p_1, p_2 > 0$ *and constants* $C_1, C_2 > 0$ *such that the following inequality holds uniformly for* x *in a neighborhood of* 0*:*

$$C_1\varepsilon^{p_1} \leq J(\varepsilon, x) \leq C_2\varepsilon^{p_2}.$$

We next turn to the implied volatility expansion.

Proof of Theorem 3.5. We will use an asymptotic formula for the dimensionless implied variance

$$V_t^2 = t\sigma_{\mathrm{impl}}(k_t,t)^2, \quad t>0,$$

obtained in [GL14]. It follows from the first formula in Remark 7.3 in [GL14] that

$$V_t^2 - \frac{k_t^2}{2L_t} = O\left(\frac{k_t^2}{L_t^2}(k_t + |\log k_t| + \log L_t)\right), \quad t\to 0, \tag{7.1}$$

where $L_t = -\log c(k_t,t)$, $t>0$.

We will need the following formula that was established in the proof of Theorem 3.4:

$$L_t = \frac{I(kt^\beta)}{t^{2H}} + O(\log\frac{1}{t}) \tag{7.2}$$

as $t\to 0$, for all $x\geq 0$ and $\beta\in[0,H)$. Let us first assume $\frac{2H}{n+1}\leq\beta<\frac{2H}{n}$. Using the energy expansion, we obtain from (7.2) that

$$\begin{aligned} L_t &= \sum_{i=2}^{n}\frac{I^{(i)}(0)}{i!}k^i t^{i\beta-2H} + O\left(\log\frac{1}{t}\right) = \frac{I''(0)}{2}k^2t^{2\beta-2H} \\ &\quad\times\left[1+\sum_{i=3}^{n}\frac{2I^{(i)}(0)}{i!I''(0)}k^{i-2}t^{(i-2)\beta} + O\left(t^{2H-2\beta}\log\frac{1}{t}\right)\right] \end{aligned} \tag{7.3}$$

as $t\to 0$. The second term in the brackets on the right-hand side of (7.3) disappears if $n=2$.

Remark 7.2. Suppose $n\geq 2$ and $\frac{2H}{n+1}\leq\beta<\frac{2H}{n}$. Then formula (7.3) is optimal. Next, suppose $n\geq 2$ and $0<\beta<\frac{2H}{n+1}$. In this case, there exists $m\geq n+1$ such that $\frac{2H}{m+1}\leq\beta<\frac{2H}{m}$, and hence (7.3) holds with m instead of n. However, we can replace m by n, by making the error term worse. It is not hard to see that the following formula holds for all $n\geq 2$ and $0<\beta<\frac{2H}{n+1}$:

$$\begin{aligned} L_t &= \sum_{i=2}^{n}\frac{I^{(i)}(0)}{i!}k^i t^{i\beta-2H} + O\left(t^{(n+1)\beta-2H}\right) = \frac{I''(0)}{2}k^2t^{2\beta-2H} \\ &\quad\times\left[1+\sum_{i=3}^{n}\frac{2I^{(i)}(0)}{i!I''(0)}k^{i-2}t^{(i-2)\beta} + O\left(t^{(n-1)\beta}\right)\right] \end{aligned} \tag{7.4}$$

as $t\to 0$

Let us continue the proof of Theorem 3.5. Since $k_t\approx t^{\frac{1}{2}-H+\beta}$ and $L_t\approx t^{2\beta-2H}$ as $t\to 0$, (7.1) implies that

$$V_t^2 = \frac{k^2t^{1-2H+2\beta}}{2L_t} + O\left(t^{1+2H-2\beta}\log\frac{1}{t}\right), \quad t\to 0. \tag{7.5}$$

Next, using the Taylor formula for the function $u\mapsto\frac{1}{1+u}$, and setting

$$u = \sum_{i=3}^{n}\frac{2I^{(i)}(0)}{i!I''(0)}k^{i-2}t^{(i-2)\beta} + O(t^{2H-2\beta}\log\frac{1}{t}),$$

we obtain from (7.3) that

$$(2L_t)^{-1} = \frac{t^{2H-2\beta}}{k^2 I''(0)} \left[\sum_{j=0}^{n-2} (-1)^j u^j + O(u^{n-1}) \right]$$

as $t \to 0$. It follows from $\frac{2H}{n+1} \leq \beta < \frac{2H}{n}$ that $(n-1)\beta \geq 2H - 2\beta$, and hence

$$\begin{aligned}
(2L_t)^{-1} &= \frac{t^{2H-2\beta}}{k^2 I''(0)} \left[\sum_{j=0}^{n-2} (-1)^j u^j \right] + O(t^{4H-4\beta} \log \frac{1}{t}) \\
&= \frac{t^{2H-2\beta}}{k^2 I''(0)} \left[\sum_{j=0}^{n-2} (-1)^j \left(\sum_{i=3}^{n} \frac{2I^{(i)}(0)}{i! I''(0)} k^{i-2} t^{(i-2)\beta} \right)^j \right] + O(t^{4H-4\beta} \log \frac{1}{t})
\end{aligned}$$

as $t \to 0$. Now, (7.5) gives

$$\begin{aligned}
V_t^2 = \frac{t}{I''(0)} &\left[\sum_{j=0}^{n-2} (-1)^j \left(\sum_{i=3}^{n} \frac{2I^{(i)}(0)}{i! I''(0)} k^{i-2} t^{(i-2)\beta} \right)^j \right] \\
&+ O\left(t^{1+2H-2\beta} \log \frac{1}{t} \right)
\end{aligned}$$

as $t \to 0$. Finally, by cancelling a factor of t in the previous formula, we obtain formula (3.1) for $\frac{2H}{n+1} \leq \beta < \frac{2H}{n}$. The proof in the case where $\beta \leq \frac{2H}{n+1}$ is similar. Here we take into account Remark 7.2. This completes the proof of Theorem 3.5. $\square$

We will now provide a proof for the general asymptotic formula for the implied volatility that uses the fourth derivative $I^{(4)}(0)$.

Proof of Corollary 3.12. It follows from Theorem 3.5 with $n = 4$, (3.3), and (3.5) that as $t \to 0$,

$$\begin{aligned}
\sigma_{\text{impl}}(k_t, t)^2 &= \frac{1}{I''(0)} - \frac{2}{I''(0)^2} \left(\frac{I'''(0)}{6} k t^\beta + \frac{I^{(4)}(0)}{24} k^2 t^{2\beta} \right) \\
&\quad + \frac{4}{I''(0)^3} \left(\frac{I'''(0)^2}{36} k^2 t^{2\beta} + \frac{I'''(0) I^{(4)}(0)}{72} k^3 t^{3\beta} + \frac{I^{(4)}(0)^2}{576} k^4 t^{4\beta} \right) \\
&\quad + O(\phi_{4,H,\beta}(t)) \\
&= \frac{1}{I''(0)} - \frac{I'''(0)}{3 I''(0)^2} k t^\beta + \left(\frac{I'''(0)^2}{9 I''(0)^3} - \frac{I^{(4)}(0)}{12 I''(0)^2} \right) k^2 t^{2\beta} \\
&\quad + O(\phi_{4,H,\beta}(t)) \\
&= \sigma_0^2 + 2\rho \sigma_0' \langle K1, 1 \rangle k t^\beta + \left(\frac{4\rho^2 (\sigma_0')^2}{\sigma_0^2} \langle K1, 1 \rangle^2 - \frac{I^{(4)}(0) \sigma_0^4}{12} \right) k^2 t^{2\beta} \\
&\quad + O(\phi_{4,H,\beta}(t)) \\
&= \sigma_0^2 \left[1 + 2\rho \frac{\sigma_0'}{\sigma_0^2} \langle K1, 1 \rangle k t^\beta + \left(\frac{4\rho^2 (\sigma_0')^2}{\sigma_0^4} \langle K1, 1 \rangle^2 - \frac{I^{(4)}(0) \sigma_0^2}{12} \right) k^2 t^{2\beta} \right] \\
&\quad + O(\phi_{4,H,\beta}(t)).
\end{aligned}$$

Now, it is not hard to see that Corollary 3.12 can be derived from the previous formula and the expansion $\sqrt{1+h} = 1 + \frac{1}{2}h - \frac{1}{4}h^2 + O(h^3)$ as $h \to 0$. $\square$

Appendix A. Auxiliary lemmas

Lemma A.1. *Assume $\sigma(.) \geq \underline{\sigma} > 0$ and $|\rho| < 1$. Then $\mathcal{K}^x$ is a Hilbert manifold near any $\mathfrak{h} := (h, f) \in \mathcal{K}^x \subset \mathfrak{H} := H_0^1 \times H_0^1$.*

Proof. Similar to Bismut [Bis84, p. 25] we need to show that $D\varphi_1(\mathfrak{h})$ is surjective where $\varphi_1(\mathfrak{h}) : \mathfrak{H}^1 \to \mathbb{R}$ with

$$\varphi_1(\mathfrak{h}) = \varphi_1(h, f) = \int_0^1 \sigma(\widehat{f}) d(\overline{\rho} h + \rho f).$$

From

$$\begin{aligned}
\varphi_1(\mathfrak{h} + \delta\mathfrak{h}') &= \int_0^1 \sigma(\widehat{f} + \delta\widehat{f}') d(\overline{\rho} h + \rho f + \delta(\overline{\rho} h' + \rho f')) \\
&= \varphi_1(\mathfrak{h}) + \delta \int_0^1 \sigma(\widehat{f}) d(\overline{\rho} h' + \rho f') \\
&\quad + \delta \int_0^1 \sigma'(\widehat{f}) \widehat{f}' d(\overline{\rho} h + \rho f) + o(\delta).
\end{aligned}$$

the functional derivative $D\varphi_1(\mathfrak{h})$ can be computed explicitly. In fact, even the computation

$$(D\varphi_1(\mathfrak{h}), (h', 0)) = \overline{\rho} \int_0^1 \sigma(\widehat{f}) dh'$$

is sufficient to guarantee surjectivity of $D\varphi_1(\mathfrak{h})$. □

Lemma A.2. *(i) Any optimal control $\mathfrak{h}^0 = (h^x, f^x) \in \mathcal{K}^x$ is a critical point of*

$$\mathfrak{h} = (h, f) \mapsto -I\left(\varphi_1^{\mathfrak{h}}\right) + \frac{1}{2} \|\mathfrak{h}\|_{\mathfrak{H}}^2;$$

(ii) it holds that

$$\int_0^1 \dot{h}^x dW + \int_0^1 \dot{f}^x dB = I'(x) g_1.$$

Proof. (Step 1) Write $\mathfrak{h} = (h, f)$ and

$$\varphi_1(\mathfrak{h}) = \varphi_1(h, f) = \int_0^1 \sigma(\widehat{f}) d(\overline{\rho} h + \rho f).$$

Let $\mathfrak{h}^0 = (h^x, f^x) \in \mathcal{K}^x$ an optimal control. Then

$$\mathfrak{Ker} D\varphi_1\left(\mathfrak{h}^0\right) = T_{\mathfrak{h}^0}\mathcal{K}^x = \left\{\mathfrak{h} \in \mathfrak{H}^1 : D\varphi_1(\mathfrak{h}) = 0\right\}.$$

(This requires $\mathcal{K}^x$ to be a Hilbert manifold near $\mathfrak{h}^0$, as was seen in the last lemma.)
(Step 2) For fixed $\mathfrak{h} \in \mathfrak{H}$, define

$$u(t) := -I\left(\varphi_1^{\mathfrak{h}^0 + t\mathfrak{h}}\right) + \frac{1}{2} \left\|\mathfrak{h}^0 + t\mathfrak{h}\right\|_{\mathfrak{H}}^2 \geq 0$$

with equality at $t = 0$ (since $x = \varphi_1^{\mathfrak{h}^0}$ and $I(x) = \frac{1}{2}\left\|\mathfrak{h}^0\right\|_{\mathfrak{H}}^2$) and non-negativity for all t because $\mathfrak{h}^0 + t\mathfrak{h}$ is an admissible control for reaching $\widetilde{x} = \varphi_1^{\mathfrak{h}^0 + t\mathfrak{h}}$ (so that $I(\widetilde{x}) = \inf\{...\} \leq \frac{1}{2}\left\|\mathfrak{h}^0 + t\mathfrak{h}\right\|_{\mathfrak{H}}^2$.)
(Step 3) We note that $\dot{u}(0) = 0$ is a consequence of $u \in C^1$ near 0, $u(0) = 0$ and $u \geq 0$. In other words, $\mathfrak{h}^0$ is a critical point for

$$\mathfrak{H}^1 \ni \mathfrak{h} \mapsto -I\left(\varphi_1^{\mathfrak{h}}\right) + \frac{1}{2} \|\mathfrak{h}\|_{\mathfrak{H}}^2.$$

(Step 4) The functional derivative of this map at $\mathfrak{h}^0$ must hence be zero. In particular, for all $\mathfrak{h} \in \mathfrak{H}$,

$$\begin{aligned} 0 &\equiv -I'\left(\varphi_1^{\mathfrak{h}^0}\right)\left\langle D\varphi_1\left(\mathfrak{h}^0\right),\mathfrak{h}\right\rangle + \left\langle \mathfrak{h}^0,\mathfrak{h}\right\rangle \\ &= -I'\left(x\right)\left\langle D\varphi_1\left(\mathfrak{h}^0\right),\mathfrak{h}\right\rangle + \left\langle \mathfrak{h}^0,\mathfrak{h}\right\rangle. \end{aligned}$$

(Step 5) With $\mathfrak{h}^0 = (h^x, f^x)$ and $\mathfrak{h} = (h, f)$

$$\begin{aligned} \left\langle D\varphi_1\left(\mathfrak{h}^0\right),\mathfrak{h}\right\rangle &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\int_0^1 \sigma(\widehat{f}^x + \varepsilon\widehat{f})d\left(\overline{\rho}h^x + \rho f^x + \varepsilon\left(\overline{\rho}h + \rho f\right)\right) \\ &= \int_0^1 \sigma(\widehat{f}^x)d\left(\overline{\rho}h + \rho f\right) + \int_0^1 \sigma'(\widehat{f}^x)\widehat{f}d\left(\overline{\rho}h^x + \rho f^x\right) \end{aligned}$$

By continuous extension, replace $\mathfrak{h} = (h, f)$ by (W, B) above and note that

$$\left\langle D\varphi_1\left(\mathfrak{h}^0\right),(W,B)\right\rangle = g_1$$

since indeed $g_1 = \int_0^1 \{\sigma(\widehat{f}_t)d\left(\overline{\rho}W_t + \rho B_t\right) + \sigma'(\widehat{f}_t)\widehat{B}_t d\left(\overline{\rho}h_t + \rho f_t\right)$.

$$\int_0^1 \dot{h}^x dW + \int_0^1 \dot{f}^x dB = I'\left(x\right)g_1. \qquad \square$$

References

[ALV07] Elisa Alòs, Jorge A León, and Josep Vives. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, 11(4):571–589, 2007.

[Aze85] Robert Azencott. Petites perturbations aléatoires des systemes dynamiques: développements asymptotiques. *Bulletin des sciences mathématiques*, 109(3):253–308, 1985.

[BA88] Gérard Ben Arous. Methods de Laplace et de la phase stationnaire sur l'espace de Wiener. *Stochastics*, 25(3):125–153, 1988.

[BFG16] Christian Bayer, Peter K. Friz, and Jim Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904, 2016.

[BFG+17] Christian Bayer, Peter K. Friz, Paul Gassiat, Joerg Martin, and Benjamin Stemper. A regularity structure for rough volatility. In preparation, 2017.

[Bis84] Jean-Michel Bismut. *Large deviations and the Malliavin calculus*, volume 45 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1984.

[BLP15] Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. Hybrid scheme for Brownian semistationary processes. Preprint, 2015. arXiv:1507.03004.

[BLP16] Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. Decoupling the short- and long-term behavior of stochastic volatility. Preprint, 2016. arXiv:1610.00332.

[BO11] Fabrice Baudoin and Cheng Ouyang. Small-time kernel expansion for solutions of stochastic differential equations driven by fractional Brownian motions. *Stochastic Processes and their Applications*, 121(4):759–792, 2011.

[BO15] Fabrice Baudoin and Cheng Ouyang. *On Small Time Asymptotics for Rough Differential Equations Driven by Fractional Brownian Motions*, pages 413–438. Springer International Publishing, 2015.

[CF10] Thomas Cass and Peter Friz. Densities for rough differential equations under Hörmander's condition. *Annals of Mathematics*, pages 2115–2141, 2010.

[Dec05] Laurent Decreusefond. Stochastic integration with respect to Volterra processes. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 41(2):123 – 149, 2005.

[DFJV14a] Jean-Dominique Deuschel, Peter K. Friz, Antoine Jacquier, and Sean Violante. Marginal density expansions for diffusions and stochastic volatility I: Theoretical foundations. *Comm. Pure Appl. Math.*, 67(1):40–82, 2014.

[DFJV14b] Jean-Dominique Deuschel, Peter K. Friz, Antoine Jacquier, and Sean Violante. Marginal density expansions for diffusions and stochastic volatility II: Applications. *Comm. Pure Appl. Math.*, 67(2):321–350, 2014.

[DS89] Jean-Dominique Deuschel and Daniel W. Stroock. *Large deviations*, volume 137. Academic press Boston, 1989.

[DZ09] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*, volume 38. Springer Science & Business Media, 2009.

[ER16] Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. Preprint, 2016. arXiv:1609.02108.

[FGP17] Peter K. Friz, Stefan Gerhold, and Arpad Pinter. Option Pricing in the Moderate Deviations Regime. *Mathematical Finance*, 2017. To appear.

[Fuk11] Masaaki Fukasawa. Asymptotic analysis for stochastic volatility: martingale expansion. *Finance and Stochastics*, 15(4):635–654, 2011.

[Fuk17] Masaaki Fukasawa. Short-time at-the-money skew and rough fractional volatility. *Quantitative Finance*, 17(2):189–198, 2017.

[FZ17] Martin Forde and Hongzhong Zhang. Asymptotics for rough stochastic volatility models. *SIAM Journal on Financial Mathematics*, 8(1):114–145, 2017.

[Gat11] Jim Gatheral. *The volatility surface: a practitioner's guide*. John Wiley & Sons, 2011.

[GHJ16] Archil Gulisashvili, Blanka Horvath, and Antoine Jacquier. Mass at zero in the uncorrelated SABR model and implied volatility asymptotics. Preprint, 2016. arXiv:1502.03254v2.

[GJR14a] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. Preprint, 2014. arXiv:1410.3394.

[GJR14b] Hamza Guennoun, Antoine Jacquier, and Patrick Roome. Asymptotic behaviour of the fractional Heston model. Preprint, 2014. arXiv:1411.7653.

[GL14] Kun Gao and Roger Lee. Asymptotics of implied volatility to arbitrary order. *Finance and Stochastics*, 18(2):349–392, 2014.

[GVZ15] Archil Gulisashvili, Frederi Viens, and Xin Zhang. Small-time asymptotics for gaussian self-similar stochastic volatility models. Preprint, 2015. arXiv:1505.05256.

[Lam62] John Lamperti. Semi-stable stochastic processes. *Transactions of the American mathematical Society*, 104(1):62–78, 1962.

[MS03] Alexey Medvedev and Olivier Scaillet. A simple calibration procedure of stochastic volatility models with jumps by short term asymptotics. Preprint, 2003. Available at SSRN 477441.

[MS07] Alexey Medvedev and Olivier Scaillet. Approximation and calibration of short-term implied volatilities under jump-diffusion stochastic volatility. *Review of Financial Studies*, 20(2):427–459, 2007.

[MT16] Aleksandar Mijatović and Peter Tankov. A new look at short-term implied volatility in asset price models with jumps. *Mathematical Finance*, 26(1):149–183, 2016.

[Olv10] Frank WJ Olver. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.

[Osa07] Yasufumi Osajima. The asymptotic expansion formula of implied volatility for dynamic SABR model and FX hybrid model. Preprint, 2007. Available at SSRN 965265.

[Osa15] Yasufumi Osajima. *General Asymptotics of Wiener Functionals and Application to Implied Volatilities*, pages 137–173. Springer International Publishing, Cham, 2015.

WIAS Berlin, TU and WIAS Berlin, Ohio University, Imperial College London, TU and WIAS Berlin

E-mail address: christian.bayer@wias-berlin.de, friz@math.tu-berlin.de, gulisash@ohio.edu, b.horvath@imperial.ac.uk, stemper@math.tu-berlin.de