Quantitative Finance

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rquf20

Short-dated smile under rough volatility: asymptotics and numerics

Peter K. Friz, Paul Gassiat & Paolo Pigato

To cite this article: Peter K. Friz, Paul Gassiat & Paolo Pigato (2022) Short-dated smile under rough volatility: asymptotics and numerics, Quantitative Finance, 22:3, 463-480, DOI: 10.1080/14697688.2021.1999486

To link to this article: https://doi.org/10.1080/14697688.2021.1999486

Published online: 07 Dec 2021.

Quantitative Finance, 2022
Vol. 22, No. 3, 463–480, https://doi.org/10.1080/14697688.2021.1999486

Short-dated smile under rough volatility: asymptotics and numerics

PETER K. FRIZ*†, PAUL GASSIAT‡ and PAOLO PIGATO§

†Technische Universität Berlin and Weierstraß-Institut Berlin, Berlin, Germany
‡Université Paris-Dauphine, PSL University, Paris Cedex 16, France
§Department of Economics and Finance, University of Roma Tor Vergata, Roma, Italy

(*Received 27 October 2020; accepted 15 October 2021; published online 7 December 2021*)

In Friz *et al.* [Precise asymptotics for robust stochastic volatility models. *Ann. Appl. Probab*, 2021, **31**(2), 896–940], we introduce a new methodology to analyze large classes of (classical and rough) stochastic volatility models, with special regard to short-time and small-noise formulae for option prices, using the framework [Bayer *et al.*, A regularity structure for rough volatility. *Math. Finance*, 2020, **30**(3), 782–832]. We investigate here the fine structure of this expansion in large deviations and moderate deviations regimes, together with consequences for implied volatility. We discuss computational aspects relevant for the practical application of these formulas. We specialize such expansions to prototypical rough volatility examples and discuss numerical evidence.

Keywords: Rough volatility; European option pricing; Implied volatility; Small-time asymptotics; Rough paths; Regularity structures; Karhunen–Loeve

JEL Classifications: 91G20, 91G60, 60L30, 60L90, 60H30, 60F10, 60G22, 60G18

1. Introduction

In Friz *et al.* (2021), precise short-time asymptotics were established for call and put option prices under stochastic volatility, under a set of abstract conditions satisfied by most classical and rough volatility (RoughVol) models. These results are refinements of large deviation statements, providing the higher-order, algebraic term in an asymptotic expression, known as Laplace expansion. For RoughVol models, short-dated large deviation pricing is due to Forde and Zhang (2017), as is the induced implied volatility expansion (FZ expansion), which can be seen as a "rough" BBF (Berestycki–Busca–Florent Berestycki *et al.* 2004) formula. Our precise asymptotics provide a mechanism to compute refined implied volatility expansions, for log-strike $k_t = xt^{1/2-H}$, of the form

$$\sigma_{BS}^2(t,k_t) = \Sigma^2(x) + t^{2H}a(x) + o(t^{2H}) \text{ as } t \downarrow 0, \qquad (1)$$

where the zero-order $\Sigma(x)$ term corresponds to the rough BBF formula in Forde and Zhang (2017). The next-order term is seen of order t^{2H} and hence increasingly important for small Hurst parameter H, the basic premise of RoughVol modeling. Inclusion of this term hinges on an accurate evaluation of a. In this paper, we assume that the volatility process is of the form $\sigma(\widehat{W}_t, t^{2H})$, where $\widehat{W}$ is the Riemann–Liouville fractional Brownian motion (fBM) given by the self-similar Gaussian Volterra process in (A3). It has Hurst exponent $H \in (0, 1/2]$ and it is ρ-correlated with the Brownian driving the asset.

The functions $\Sigma(x)$ and $a(x)$ do not have explicit expressions and we discuss how to compute them numerically. Following Forde and Zhang (2017), $\Sigma(x)$ can be computed using the Ritz method. Moreover, we propose a method for computing $a(x)$ based on a Karhunen–Loeve (KL) decomposition of the Brownian motions. (This entails a numerical approximation to an infinite-dimensional Carleman–Fredholm determinant.)

We also derive near-the-money (meaning, as $x \to 0$) expansions of $\Sigma(x)$ and of the term structure $a(x)$ which can alternatively be used for numerics (and have the advantage of being explicit functions of model parameters). From these asymptotics, we derive consequences for at-the-money (ATM) implied skew and curvature. We also refine some moderate deviation asymptotics for call prices and implied volatilities, cf. (Friz *et al.* 2017, Bayer *et al.* 2019, Gulisashvili 2020, Jacquier and Pannier 2020).

Being able to evaluate $\Sigma(x)$ and $a(x)$ allows us to test the accuracy of the short-time asymptotics in practice. We

*Corresponding author. Email: friz@math.tu-berlin.de

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

do so with a numerical case study of the rough Bergomi (rBergomi) model. To exploit our general framework, we look at a volatility given by

$$\sigma(\widehat{W}_t, t^{2H}) = \sigma_0 \exp\left(\frac{\eta}{2}\widehat{W}_t - \frac{\theta\eta^2}{4}t^{2H}\right), \tag{2}$$

so that for $\theta = 1$ we get the rBergomi model considered in Bayer *et al.* (2016) and Bennedsen *et al.* (2017) with constant forward variance, for $\theta = 0$ the rBergomi version in Bayer *et al.* (2019) and Forde and Zhang (2017). Note, however, that (2) is a genuine rBergomi model for any value of θ, as discussed in Remark 4.1. We compare our approximation to the FZ expansion from Forde and Zhang (2017) and to the Edgeworth asymptotics in El Euch *et al.* (2019). We consider how smiles vary as θ varies in (2) and as expiry t increases. We discuss and test the volatility term structure and its slope ATM, and observe how the term $a(x)t^{2H}$ improves the asymptotics as H decreases. We observe the same feature when we implement the moderate deviation asymptotics for implied volatility, where for H small the inclusion of the term structure correction $a(0)t^{2H}$ significantly improves on the numerical results presented in Bayer *et al.* (2019).

Proofs rely on stochastic Taylor expansions, rate function representations in Forde and Zhang (2017) and Bayer *et al.* (2019) and on the local analysis on the Wiener space introduced in Friz *et al.* (2021) and Bayer *et al.* (2020). The classical Gao–Lee results (Gao and Lee 2014) are used to go from option prices to implied volatility asymptotics both in large and moderate deviation regimes.

Rough Volatility. It has been shown in recent years that RoughVol models provide great fits to observed volatility surfaces (Bayer *et al.* 2016) capturing fundamental stylized facts of implied volatility in a parsimonious way. Specifically, this class of models can reproduce the steep short end of the smile, displaying exploding implied skew (Alòs *et al.* 2007, Fukasawa 2011, 2017), and they are the only models consistent with the power law of the skew (Bayer *et al.* 2016, Lee 2005) not admitting arbitrage (Fukasawa 2021). RoughVol is also supported by statistical and time series analysis (Gatheral *et al.* 2018, Fukasawa *et al.* 2019, Bennedsen *et al.* 2021) and by market microstructure considerations (El Euch *et al.* 2018). Many authors have even argued that $H \approx 0$, such as to be consistent with a skew explosion close to $t^{-1/2}$ (Bayer *et al.* 2016, 2021). One main aspect of RoughVol is non-Markovianity. This is a serious complication when it comes to pricing, as Monte Carlo methods become more expensive and PDE methods are not available. For this reason, efficient simulation schemes have been proposed (Bayer *et al.* 2020, Bennedsen *et al.* 2017, McCrickerd and Pakkanen 2018). Fourier-based methods are available for the rough Heston model (El Euch and Rosenbaum 2019). Deep and machine learning approaches have also recently been discussed in Bayer *et al.* (2019) and Goudenège *et al.* (2020). Small maturity approximations are used in this context to obtain starting points for calibration procedures, which are then based on numerical evaluations.

Asymptotic option pricing. Classical motivation for (semi-closed form) asymptotic pricing includes fast calibration and a quantitative understanding of the impact of model parameters on relevant quantities such as implied skew and curvature/convexity along the moneyness dimension or slope along the term-structure dimension. Explicit expressions for such quantities (that follow in this setting from our expansion) and their shape characteristics are also used to choose the most appropriate model to be fitted to data (Ait-Sahalia *et al.* 2020), leave alone being the origin of some widely used parametrisations of the volatility surface. An interesting, if recent, addition to this list comes from a machine learning perspective: the form of an expansion such as (1) may be viewed as *expert knowledge*, which significantly narrows the learning task to finer information such as the error in that expansions; it is equally conceivable to learn $a = a(x)$ and other components in the expansion.

Under Markovian stochastic volatility, expansion (2) is analogous, e.g. to the result derived in Forde *et al.* (2012) for the Heston model. There, the term structure is $a(x)t$ (due to the diffusive scaling of the volatility), whereas here the correction term is $a(x)t^{2H}$ (due to the rough scaling of the volatility). Similar expansions are derived also in Osajima (2015), for more general Markovian models, and (formally) in Medvedev and Scaillet (2003) and Medvedev and Scaillet (2007) for Markov stochastic volatility models with jumps.

In recent years several authors have studied the short-time behavior of RoughVol models. Theoretical results on short-time skew and curvature are given in Fukasawa (2017) and Alòs and León (2017). A second-order short-time expansion is given in El Euch *et al.* (2019) for general (rough) stochastic volatility models. In Jacquier *et al.* (2018), the pathwise large deviation behavior under rBergomi dynamics is studied. Pathwise large and moderate deviation principles for (possibly rough) Gaussian stochastic volatility models are established in Gulisashvili (2020) and Gulisashvili (2020), together with asymptotic results at the central limit (Edgeworth) regime. For the rough Heston model, the recent work (Forde *et al.* 2020) provides call expansions of the same type as ours, involving the energy function and the first-order algebraic term, at the same large deviations regime $k_t = xt^{1/2-H}$. (The rigid infinite-dimensional affine structure which underlies (Forde *et al.* 2020) is not available for rBergomi type models as considered in this work.) As already mentioned, our work builds on the large deviations principle proved in Forde and Zhang (2017) for models with volatility $\sigma(\widehat{W}_t)$, and on Bayer *et al.* (2019), where the at-the-money behavior of the Forde–Zhang rate function is used to prove moderate deviation principles and implied volatility asymptotics for the same type of models. The theoretical foundations of the present paper are given in Friz *et al.* (2021).

In Section 2, we explain our RoughVol setting. In Section 3, we state and comment our results. In Section 4 we discuss and implement our results in the case of the rBergomi model. In Section 5, we show how Σ and a can be computed using Ritz method and KL decomposition. We collect all the proofs in Section 6.

2. Preliminaries on rough volatility

We consider the following RoughVol model, with $H \in (0, 1/2]$, normalized to rate $r = 0$ and $S_0 = 1$

$$\frac{\mathrm{d}S_t}{S_t} = \sigma(\widehat{W}_t, t^{2H}) d(\rho W_t + \overline{\rho}\overline{W}_t), \tag{3}$$

where $W, \overline{W}$ are independent Brownian motions (BM) and $\rho \in (-1, 1)$, $\rho^2 + \overline{\rho}^2 = 1$. We also write $\widetilde{W} = \rho W + \overline{\rho}\overline{W}$. Moreover, $\widehat{W} = (\widehat{W}_t)_{t\geq 0}$ is a Gaussian Volterra process of the form

$$\widehat{W}_t = (K * \dot{W})_t = \int_0^t K(t, s)\, \mathrm{d}W_s, \tag{4}$$

for a kernel $K(t, s)$ such that $\widehat{W}$ is self-similar with exponent $H \in (0, 1/2]$, meaning

$$\mathrm{Law}(\widehat{W}_{\varepsilon^2 t} : t \leq \bar{t}) = \varepsilon^{2H} \mathrm{Law}(\widehat{W}_t : t \leq \bar{t}), \quad \text{for some } \bar{t} > 0. \tag{5}$$

The BM W drives the stochastic "rough" volatility, meaning (with abusive notation) that $\sigma(t, \omega) = \sigma(\widehat{W}_t, t^{2H})$, where $\sigma(x, y)$ is a smooth deterministic real-valued function. We denote $\sigma'(x, y) = \partial_x \sigma(x, y)$, $\sigma''(x, y) = \partial_{xx}\sigma(x, y)$, $\dot{\sigma}(x, y) = \partial_y \sigma(x, y)$. We also denote $\sigma_0 = \sigma(0, 0) > 0$ the spot volatility and

$$\sigma_0' = \sigma'(0, 0), \quad \sigma_0'' = \sigma''(0, 0), \quad \dot{\sigma}_0 = \dot{\sigma}(0, 0), \tag{6}$$

the derivatives of the volatility function at the initial condition. We consider a dependence in t^{2H} in $\sigma(\cdot)$, because this is the scaling of the variance of the fBm at time t. For this reason, this is the scaling of the time-dependent term in the rBergomi model, and also the scaling such that we observe a dependence in $\dot{\sigma}_0$ in our precise asymptotics. We apply the abstract results proved in Friz *et al.* (2021) for $K(t, s) = const \times (t - s)^{H-1/2}$. However, we expect these approximations to hold in greater generality: the same type of expansions should hold for other kernels such that $\widehat{W}$ in (4) satisfies (5). Self-similarity is equivalent to the fact that K can be written in the following form

$$K(t, s) = (t - s)^{H-1/2} f_K(s/t), \tag{7}$$

for a suitable function f_K (see Jost 2007, Lemma 2.4), so that all such kernels can be seen as a perturbation of $(t - s)^{H-1/2}$. Two classical processes of this form are the Mandelbrot–Van Ness and the Riemann–Liouville fBMs (see Appendix). Without loss of generality, we also assume $K(t, s) = 0$ for $t < s$.

A similar setting has been considered in Forde and Zhang (2017) and Bayer *et al.* (2019). The main difference in the structure of the model is that here we allow for a direct dependence on time in $\sigma(t, \omega) = \sigma(\widehat{W}_t, t^{2H})$, whereas in Forde and Zhang (2017) and Bayer *et al.* (2019) the volatility function depends only on the fBM, so $\sigma(t, \omega) = \sigma(\widehat{W}_t)$. As mentioned in the introduction, assuming that the volatility is a deterministic function only of the fBM rules out the rBergomi model $\sigma(\widehat{W}_t, t^{2H}) = \sigma_0 \exp(\eta \widehat{W}_t/2 - \eta^2 t^{2H}/4)$, see Bayer *et al.* (2016) and Bennedsen *et al.* (2017), from the analysis, so a modified version of rBergomi is considered in Bayer *et al.* (2019). We discuss in detail both versions of this model in Section 4. With a volatility function $\sigma(\widehat{W}_t, t^{2H})$, one can write the dynamics of the log-price $X = \log S$ as

$$X_t = \int_0^t \sigma\left(\widehat{W}_s, s^{2H}\right) d(\overline{\rho}\overline{W} + \rho W)_s - \frac{1}{2}\int_0^t \sigma^2(\widehat{W}_s, s^{2H})\mathrm{d}s. \tag{8}$$

In this case, a LDP holds, writing $\widehat{\varepsilon} = \varepsilon^{2H}$, for

$$\overline{X}_1^{\varepsilon} = \int_0^1 \sigma\left(\widehat{\varepsilon}\widehat{W}_t, \widehat{\varepsilon}^2 t^{2H}\right) \widehat{\varepsilon} d(\overline{\rho}\overline{W} + \rho W)_t - \frac{1}{2}\varepsilon\widehat{\varepsilon}\int_0^1 \sigma^2\left(\widehat{\varepsilon}\widehat{W}_t, \widehat{\varepsilon}^2 t^{2H}\right) \mathrm{d}t, \tag{9}$$

with speed $\widehat{\varepsilon}^2$ and rate function

$$\Lambda(x) := \inf\left\{\tfrac{1}{2}\|h, \overline{h}\|_{H^1}^2 : \int_0^1 \sigma\left(\widehat{h}, 0\right) d\left(\overline{\rho}\overline{h} + \rho h\right) = x\right\} \equiv \tfrac{1}{2}\|h^x, \overline{h}^x\|_{H^1}^2, \tag{10}$$

where $\widehat{h}_t = (K * \dot{h})_t$ and $\|\cdot\|_{H^1}$ is the Cameron–Martin norm. The existence of a minimizer above is obtained from a standard compactness argument. Through the space-time scaling $t = \varepsilon^2$ and the fact that, in law, $\overline{X}_1^{\varepsilon} = \frac{\widehat{\varepsilon}}{\varepsilon} X_{\varepsilon^2}$, this small-noise LDP translates to a short-time LDP. This result was proved for $\sigma(\widehat{W}_t, t^{2H}) = \sigma(\widehat{W}_t)$ in Forde and Zhang (2017) and then extended to possible dependence in t^{2H} in Friz *et al.* (2021, Section 7.3). In general, when looking only at large (or moderate) deviations, the t^{2H}-dependence in $\sigma(\cdot)$ does not affect the analysis, and the large (or moderate) deviations behavior is the same one would get with volatility $\sigma(\widehat{W}_t, 0)$. In Friz *et al.* (2021), we consider a general asymptotic setting, obtaining for generic stochastic volatility models (including RoughVol ones) precise asymptotics that refine such large deviations asymptotics. For such refinement, this t^{2H}-dependence actually affects the asymptotics. In the present paper, we provide computationally relevant results that allow for the practical usage of such refined pricing asymptotics and discuss their consequences on the Black–Scholes implied volatility.

3. Results

We consider call and put prices under model (8), i.e.

$$c(t, k) = E[(\exp X_t - \exp k)^+], \quad p(t, k) = E[(\exp k - \exp X_t)^+],$$

where k is the log-strike (or log-moneyness). In Friz *et al.* (2021, Theorem 1.1) we obtain precise small-noise price expansions for generic (classical and rough) volatility dynamics. As in the classical Brownian case, such small-noise results can be translated into short-time results writing $t = \varepsilon^2$. In this paper, we focus on the short-time setting. We write $\sim$ for asymptotic equivalence, $f_t \sim g_t$ if $f_t/g_t \to 1$ as $t \to 0$, and "$\approx$" for "is close to" in informal terms. We also write $\sigma_x^2 = 2\Lambda(x)/\Lambda'(x)^2$.

ASSUMPTION 3.1 Throughout the paper, we assume K in (4) is of the form

$$K(t,s) = const \times (t-s)^{H-1/2}.$$

In short-time, Friz *et al.* (2021, Theorem 1.1) reads as follows:

THEOREM 3.2 *Let $H \in (0,1/2]$ and $k_t = xt^{1/2-H}$. Assume that a LDP holds for c, p above, and the existence of 1^+ moments for $\exp X_t$. Then , for $x>0$ small enough, the rate function $\Lambda = \Lambda(x)$ is continuously differentiable at x and*

$$c(t,k_t) \sim \exp\left(-\frac{\Lambda(x)}{t^{2H}}\right) t^{1/2+2H} \frac{A(x)}{(\Lambda'(x))^2\sigma_x\sqrt{2\pi}} \text{ as } t \downarrow 0,$$

for some function $A(x)$ with $A(x) \to 1$ as $x \downarrow 0$. Similarly, for $x<0$, close enough to 0, we have

$$p(t,k_t) \sim \exp\left(-\frac{\Lambda(x)}{t^{2H}}\right) t^{1/2+2H} \frac{A(x)}{(\Lambda'(x))^2\sigma_x\sqrt{2\pi}} \text{ as } t \downarrow 0,$$

for some function $A(x)$ with $A(x) \to 1$ as $x \uparrow 0$. Moreover, such A can be expressed as

$$A(x) = \begin{cases} E\big[\exp(\Lambda'(x)\Delta_2^x)\big], & \text{if } H < 1/2, \\ e^x E\big[\exp(\Lambda'(x)\Delta_2^x)\big], & \text{if } H = 1/2, \end{cases} \tag{11}$$

where Δ_2^x is a certain quadratic Wiener functional (specified in Friz et al. (2021, Equation (7.4)), *see also* (36) *below).*

REMARK 3.3 The fact that $x>0$ above has to be taken small enough is in order for the minimizer $(h^x, \overline{h}^x)$ in (10) to be unique and non-degenerate. The latter means, in a nutshell, that the Hessian of $I(h,\overline{h}) := \frac{1}{2}\|h,\overline{h}\|^2_{H^1}$ is strictly positive when restricted to those $(h,\overline{h})$ such that $\int_0^1 \sigma(\widehat{h},0)d(\overline{\rho}\overline{h} + \rho h) = x$, and is equivalent to the finiteness of $A(x)$ defined above.

We write $Kf(t) = \int_0^t K(t,s)f(s)\mathrm{d}s$, $K^2 f(t) = \int_0^t K^2(t,s)f(s)\,\mathrm{d}s$ and $\langle\cdot,\cdot\rangle$ for the inner product in $L^2[0,1]$. We also denote $\overline{K}$ the adjoint of K in $L^2[0,1]$ so that $\overline{K}1(u) = \int_u^1 K(t,u)\mathrm{d}t$. Fully explicit expressions are computable in the case of the Riemann–Liouville fBM (Appendix) and in particular in the case of standard BM (this is the classical case of Markovian stochastic volatility). We denote

$$\begin{aligned} C_{K,\rho} &= \frac{\langle K^2 1,1\rangle}{2} - \frac{3}{2}\langle(\overline{K}1)^2,1\rangle \\ &\quad + \rho^2\left(\frac{17}{2}\langle K1,1\rangle^2 - \frac{3}{2}\langle(K1)^2,1\rangle - 3\langle K1,\overline{K}1\rangle\right), \\ \overline{C}_{K,\rho} &= \frac{\langle K^2 1,1\rangle}{2} - \frac{3}{2}\rho^2\langle(K1)^2,1\rangle. \end{aligned}$$

LEMMA 3.4 Fine structure of A *For $H \in (0,1/2]$, the following expansion holds for $A(x)$ as $x \to 0$:*

$$\begin{aligned} A(x) &= 1 - x\frac{\rho\sigma_0'\langle K1,1\rangle}{\sigma_0^2} \\ &\quad + x^2\left(\frac{(\sigma_0')^2}{\sigma_0^4}C_{K,\rho} + \frac{\sigma_0''}{\sigma_0^3}\overline{C}_{K,\rho} + \frac{\dot{\sigma}_0}{(2H+1)\sigma_0^3}\right) \\ &\quad + \left(\frac{x}{2} + \frac{x^2}{8}\right)\mathbf{1}_{\{H=1/2\}} + O(x^3). \end{aligned} \tag{12}$$

As a consequence of Theorem 3.2 the following expansion holds for the Black–Scholes implied volatility (by a standard application of Gao and Lee (2014), detailed in Friz *et al.* (2021, Appendix D)).

COROLLARY 3.5 Asymptotic smile and term structure at the large deviations regime

Writing $k_t = xt^{1/2-H}$, we have the following expansion, for $x \in \mathbb{R}\setminus\{0\}$ such that Theorem 3.2 *holds*:

$$\sigma_{BS}^2(t,k_t) = \Sigma^2(x) + t^{2H}a(x) + o(t^{2H}) \text{ as } t \downarrow 0, \tag{13}$$

where

$$\Sigma(x) = \frac{|x|}{\sqrt{2\Lambda(x)}} \tag{14}$$

and

$$a(x) = \begin{cases} \dfrac{x^2}{2\Lambda(x)^2}\log\left(\dfrac{2A(x)\Lambda(x)}{\Lambda'(x)x}\right) & \text{if } H < 1/2, \\ \dfrac{x^2}{2\Lambda(x)^2}\log\left(\dfrac{2A(x)\Lambda(x)}{\Lambda'(x)x\exp(x/2)}\right) & \text{if } H = 1/2. \end{cases} \tag{15}$$

REMARK 3.6 In general, from a LDP for call prices follows the celebrated BBF formula for implied volatility (Berestycki–Busca–Florent Berestycki *et al.* 2004, see also Pham Pham 2010 for a derivation). Under RoughVol pricing with $\sigma(\omega,t) = \sigma(\widehat{W}_t)$, this has been extended in Forde and Zhang (2017) to

$$\sigma_{BS}^2(t,k_t) \sim \frac{x^2}{2\Lambda(x)}, \tag{16}$$

holding for fixed x, in short-time, with $k_t = xt^{1/2-H}$. Thanks to the A-term in (12), we can extend this approximation, adding the term structure $t^{2H}a(x)$. Note that the expansions hold for $H \in (0,1/2]$, but for $H = 1/2$ their functional form is different, as some additional terms appear in $A(x)$ and in the term structure of the Black–Scholes implied volatility $a(x)$.

We denote now

$$\begin{aligned} D_{K,\rho} &= \langle K^2 1,1\rangle - \langle(\overline{K}1)^2,1\rangle \\ &\quad + \rho^2\big(3\langle K1,1\rangle^2 - \langle(K1)^2,1\rangle - 2\langle K1,\overline{K}1\rangle\big), \\ \overline{D}_{K,\rho} &= \langle K^2 1,1\rangle - \rho^2\langle(K1)^2,1\rangle. \end{aligned} \tag{17}$$

The short-time implied volatility coefficients in the previous statement can be expanded as follows near-the-money.

THEOREM 3.7 At-the-money expansion of the coefficients

For $x \to 0$, the Σ coefficient has the following expansion:

$$\Sigma(x) = \sigma_0 + x\Sigma'(0) + x^2\frac{\Sigma''(0)}{2} + O(x^3), \tag{18}$$

where

$$\Sigma'(0) = \frac{\rho\sigma_0'\langle K1,1\rangle}{\sigma_0},$$

$$\frac{\Sigma''(0)}{2} = \frac{(\sigma_0')^2}{\sigma_0^3}\left\{-3\rho^2\langle K1,1\rangle^2 + \frac{\rho^2}{2}\langle (K1)^2,1\rangle + \frac{1}{2}\langle(\overline{K}1)^2,1\rangle + \rho^2\langle K1,\overline{K}1\rangle\right\} + \frac{\sigma_0''}{\sigma_0^2}\frac{\rho^2}{2}\langle (K1)^2,1\rangle.$$

The term structure coefficient, at the first order in x at 0, is

$$a(x) = a_0 + O(x), \tag{19}$$

with

$$a_0 = (\sigma_0')^2 D_{K,\rho} + \sigma_0\sigma_0'' \overline{D}_{K,\rho} + \frac{\sigma_0\dot{\sigma}_0}{H+1/2} + \rho\sigma_0'\sigma_0^2\langle K1,1\rangle \mathbf{1}_{\{H=1/2\}}.$$

Remark 3.8 From definition (14)–(15) and from the fact that Λ is quadratic in x we see that (19) implies a relation between A and Λ for $x \to 0$.

Remark 3.9 Implied variance expansion (13) reads as follows on implied volatility

$$\sigma_{BS}(t,k_t) \approx \frac{|x|}{\sqrt{2\Lambda(x)}} + t^{2H}\frac{a(x)}{|x|}\sqrt{\frac{\Lambda(x)}{2}}. \tag{20}$$

In order to implement these expansions, one can use the methods discussed in Section 5, computing numerically the rate function $\Lambda(x)$ and $\Sigma(x)$ using FZ expansion, and then computing $a(x)$ using KL. However, this last step can be computationally expensive, since a large number of basis functions are needed for the KL decomposition to be accurate, for H close to 0. As an alternative, one can use approximation

$$\sigma_{BS}(t,k_t) \approx \Sigma(x) + t^{2H}\frac{a_0}{2\sigma_0} = \frac{|x|}{\sqrt{2\Lambda(x)}} + t^{2H}\frac{a_0}{2\sigma_0}, \tag{21}$$

for implied volatility, which follows from implied variance expansion (13) and (19). If the rate function cannot be computed, we can use (18) to expand the implied volatility as

$$\sigma_{BS}(t,k_t) \approx \Sigma(0) + \Sigma'(0)x + \frac{\Sigma''(0)}{2}x^2 + t^{2H}\frac{a_0}{2\sigma_0}. \tag{22}$$

In particular, we get the following explicit expansion for the ATM term structure:

$$\sigma_{BS}(t,0) = \sigma_0 + t^{2H}\frac{a_0}{2\sigma_0} + o(t^{2H}). \tag{23}$$

Remark 3.10 The term structure of implied volatility From the expansion of the ATM term structure (23) we also see, in the short end, that $\sigma_{BS}^2(t,0)$ is increasing in t if $a_0 > 0$ and decreasing if $a_0 < 0$. This may be compared with a large body of literature concerning monotonicity properties of the term structure of implied volatility, see e.g. (Camara *et al.* 2011, Guo *et al.* 2014, Krylova *et al.* 2009, Vasquez 2017).

Corollary 3.11 Skew and curvature at the large deviation regime

Let $k_t = xt^{1/2-H}$, for $x \in \mathbb{R}\setminus\{0\}$. Then, if $H < 1/2$, for $t \downarrow 0$

$$\frac{\sigma_{BS}(t,k_t) - \sigma_{BS}(t,-k_t)}{2k_t} \sim \frac{\Sigma(x) - \Sigma(-x)}{2x}t^{H-1/2}. \tag{24}$$

$$\frac{\sigma_{BS}(t,k_t) + \sigma_{BS}(t,-k_t) - 2\sigma_{BS}(t,0)}{k_t^2} \sim \frac{\Sigma(x) + \Sigma(-x) - 2\Sigma(0)}{x^2}t^{2H-1}. \tag{25}$$

Remark 3.12 The quantities in the rhs of the equivalences converge as $x \downarrow 0$ to $\Sigma'(0)$, $\Sigma''(0)$ given in Theorem 3.7. The quantities in the lhs of the equivalences are finite difference approximations of ATM implied volatility skew $\partial_k\sigma_{BS}(t,0)$ and curvature $\partial_{kk}\sigma_{BS}(t,0)$. Such finite differences are relevant because only a finite number of prices are observable on real markets. They give skew and curvature at the large deviation regime, a result that complements (Fukasawa 2017, El Euch *et al.* 2019) (skew and curvature at central limit regime), Bayer *et al.* (2019) (skew at moderate deviation regime), Forde *et al.* (2020) (skew and curvature at large deviations regime for rough Heston), Alòs and León (2017) (true skew and curvature).

From these formulas, we also infer the sign of implied skew and of implied curvature (convexity). Indeed, if $\sigma_0, \sigma_0' \neq 0$, it is clear that $sgn(\Sigma'(0)) = sgn(\rho)$ and that

$$\begin{cases} \Sigma''(0) = 0 \text{ iff} \\ \quad \rho^2 = \dfrac{\langle(\overline{K}1)^2,1\rangle}{6\langle K1,1\rangle^2 - \langle (K1)^2,1\rangle - 2\langle K1,\overline{K}1\rangle - \frac{\sigma_0''\sigma_0}{(\sigma_0')^2}\langle (K1)^2,1\rangle}, \\ \Sigma''(0) < 0 \text{ iff} \\ \quad \rho^2 > \dfrac{\langle(\overline{K}1)^2,1\rangle}{6\langle K1,1\rangle^2 - \langle (K1)^2,1\rangle - 2\langle K1,\overline{K}1\rangle - \frac{\sigma_0''\sigma_0}{(\sigma_0')^2}\langle (K1)^2,1\rangle} > 0, \\ \Sigma''(0) > 0 \text{ otherwise.} \end{cases} \tag{26}$$

Theorem 3.13 Moderate deviations *Assume that Λ is $i \in \mathbb{N}$ times continuously differentiable. Let $H \in (0,1/2)$, $\beta > 0$ and $n \in \mathbb{N}$ such that $\beta \in (\frac{2H}{n+1}, \frac{2H}{n}]$. Set $k_t = xt^{1/2-H+\beta}$. Then*

$$c(t,k_t) \sim \exp\left(-\sum_{i=2}^{n}\frac{\Lambda^{(i)}(0)}{i!}x^i t^{i\beta-2H}\right)t^{1/2+2H-2\beta}\frac{\sigma_0^3}{x^2\sqrt{2\pi}}.$$

Moreover

$$\sigma_{BS}^2(t,k_t) = \sum_{j=0}^{n-2}(-1)^j 2^j\sigma_0^{2(j+1)}\left(\sum_{i=3}^{n}\frac{\Lambda^{(i)}(0)}{i!}x^{i-2}t^{(i-2)\beta}\right)^j + o(t^{2H-2\beta}). \tag{27}$$

Remark 3.14 An implied volatility expansion similar to (27) was proved in Bayer *et al.* (2019), in the case $\sigma(t,\omega) = \sigma(\widehat{W}_t)$, for $\beta \in [\frac{2H}{n+1}, \frac{2H}{n})$, with remainder of order max $(t^{2H-2\beta-\varepsilon}, t^{(n-1)\beta})$. The derivatives of the rate function were computed until $\Lambda'''(0)$, here we also computed $\Lambda^{(4)}(0)$ (cf.

Lemma 6.1). This allows us to use the second-order moderate deviation (instead of first order as in Bayer *et al.* 2019)

$$\sigma(t,k_t) = \Sigma(0) + \Sigma'(0)xt^{\beta} + \frac{\Sigma''(0)}{2}x^2t^{2\beta} + o(t^{2H-2\beta}).$$

Moreover, even if it does not show up in the asymptotics, the term structure can be incorporated as follows

$$\sigma(t,k_t) \approx \Sigma(0) + \Sigma'(0)xt^{\beta} + \frac{\Sigma''(0)}{2}x^2t^{2\beta} + \frac{a_0}{2\sigma_0}t^{2H},$$

and this provides a sensible improvement in the implementation of such short-time result (cf. figure 5).

4. A case study: the rough Bergomi model

4.1. The rough Bergomi model

Introduced in Bayer *et al.* (2016), as a modification of the classical Bergomi model where the exponential (Ornstein–Uhlenbeck) kernel is replaced by a power-law kernel, the rBergomi model provides great fits of empirical implied volatility surfaces with a very small number of parametres. In such model, the volatility is given by the "Wick" exponential of a Riemann–Liouville fBM

$$\sigma(t,\omega) = \sigma_0 \exp\left(\frac{\eta}{2}\widehat{W}_t - \frac{\eta^2}{4}t^{2H}\right). \tag{28}$$

In the most general framework (Bayer *et al.* 2016), the constant σ_0^2 is replaced by the forward variance curve, which is a function of time observable on the market (so it plays the role of an initial condition, cf. also Remark 4.1). The specific volatility in (28) did not fit in the framework of Forde and Zhang (2017) and Bayer *et al.* (2019), as in these papers the volatility is assumed to be $\sigma(\widehat{W}_t)$. For this reason, in Bayer *et al.* (2019), the following version of the rBergomi model is considered

$$\sigma(t,\omega) = \sigma_0 \exp\left(\frac{\eta}{2}\widehat{W}_t\right). \tag{29}$$

In this work we consider (2), a version of the rBergomi model with one additional parameter $\theta \in \mathbb{R}$, that includes both the previous ones (for $\theta = 0, 1$). The volatility function in (3) is

$$\sigma(x,y) = \sigma_0 \exp\left(\frac{\eta}{2}x - \frac{\theta\eta^2}{4}y\right). \tag{30}$$

The interpretation of the parameters is the following: σ_0 is the spot volatility and η represents the volatility of volatility. The parameters of the driving noise are the Hurst exponent H of $\widehat{W}$ and the correlation parameter ρ between the BM $\widetilde{W}$ driving the asset and W in (4). We can interpret the newly introduced θ parameter as a damping coefficient of the volatility.

Remark 4.1 Note that the forward variance curve model $\xi_t(u) = E[\sigma^2(\widehat{W}_u, u^{2H})|\mathcal{F}_t]$ (and $\mathcal{F}_t$ filtration generated by W) induced by (30) and (2), is a genuine rBergomi model for any value of θ, with different values of θ corresponding to different specifications of the initial variance curve. More precisely, for fixed θ,

$$\xi_0(u) = E[\sigma^2(\widehat{W}_u, u^{2H})] = \sigma_0^2 \exp\left(\frac{(1-\theta)\eta^2}{2}u^{2H}\right).$$

Coming now to short-time pricing, Lemma 6.1 holds for the general model in (2), so that we are able to compare our asymptotics with large or moderate deviations results for the different versions of rBergomi in Bayer *et al.* (2019), Forde and Zhang (2017) and Jacquier *et al.* (2018). However, in Corollary 3.5, $\Sigma^2(x)$ is not affected by the value of θ, but the term structure $a(x)$ is.

From the volatility function (30) we get

$$\sigma_0 = \sigma_0, \quad \sigma_0' = \frac{\sigma_0\eta}{2}, \ \sigma_0'' = \frac{\sigma_0\eta^2}{4}, \ \dot{\sigma}_0 = -\frac{\theta\eta^2}{4}\sigma_0,$$

so all constants can be simplified. In particular condition (26) for the convexity of the short-time smile (with $\sigma_0, \eta \neq 0$) simplifies to

$$\begin{cases} \Sigma''(0) = 0 \text{ iff} \\ \quad \rho^2 = \dfrac{\langle(\overline{K}1)^2, 1\rangle}{6\langle K1, 1\rangle^2 - 2\langle(K1)^2, 1\rangle - 2\langle K1, \overline{K}1\rangle}, \\ \Sigma''(0) < 0 \text{ iff} \\ \quad \rho^2 > \dfrac{\langle(\overline{K}1)^2, 1\rangle}{6\langle K1, 1\rangle^2 - 2\langle(K1)^2, 1\rangle - 2\langle K1, \overline{K}1\rangle} > 0, \\ \Sigma''(0) > 0 \text{ otherwise.} \end{cases}$$

(note the dependence only on H, through K, and ρ). On calibrated parameters (for example in Bayer *et al.* 2016) we have that the condition for vanishing second derivative is almost satisfied. This means that the short-time ATM curvature is very close to 0, and indeed observed smiles are almost linear ATM.

All the constants in previous expansions depend on the kernel K. For the Riemann–Liouville kernel (A4) the K-functionals involved are explicit, given in (A5).

4.2. Implementation of rough Bergomi

Our goal in this section is to compare expansion (20) with other known implied volatility expansions under RoughVol. We consider:

- Implied volatility from Monte Carlo pricing, using the hybrid scheme for rBergomi in Bennedsen *et al.* (2017) with $\kappa = 2$ (note that a slight modification of the implementation is necessary for $\theta \neq 1$).
- Our implied volatility expansion, where the term structure coefficient $a(x)$ is computed using KL, so that we have (20), or where $a(x)$ is expanded at 0, so that we have (21).
- The FZ expansion (16). In Forde and Zhang (2017), Forde and Zhang show that this asymptotics holds for volatilities of type $\sigma(t,\omega) = \sigma(\widehat{W}_t)$, with no direct dependence on t, so this applies to (30) for $\theta = 0$. However, as we have shown in Friz

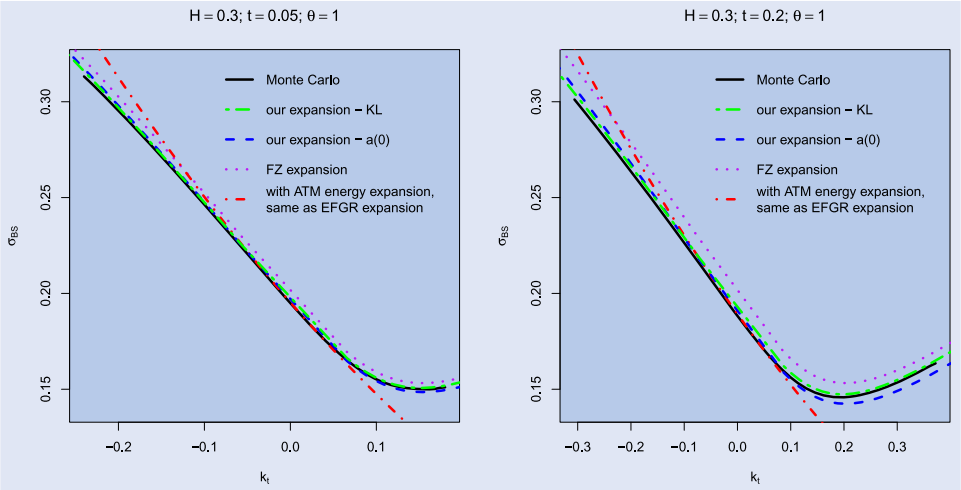

Figure 1. Implied volatility smile approximations for the rBergomi model with parameters $\theta = 1, \sigma_0 = 0.2, \eta = 1.5, \rho = -0.7, H = 0.3$, for expiry $t = 0.05, 0.2$. The Monte Carlo price is computed via the hybrid scheme for rBergomi in Bennedsen *et al.* (2017) with $\kappa = 2$, with 10^9 simulations and 500 time steps of length $t/500$. The rate function is computed using the Ritz method in Section 5.1 with $N = 8$ Haar basis functions, the coefficient $a(x)$ is computed using the Karhunen–Loeve decomposition with $N = 300$ Haar basis functions (KL). We also compare with $a(x)$ expanded at 0 ($a(x) \approx a(0)$).

et al. (2021, Section 7.3), the same large deviation behavior holds when $\theta \neq 0$. Therefore, the FZ expansion gives the same asymptotic smile, independently of the choice of θ.

- Expansion (21), with ATM expansion of Σ as in (22) (so, rate function is expanded as well). In case $\theta = 1$, one can check that this approximation is consistent with the expansion in El Euch *et al.* (2019, Section 5), that we refer to as "EFGR expansion". These two mathematical results ar different, since log-strikes are in our case (large deviation regime) $k_t = xt^{1/2-H}$ and in El Euch *et al.* (2019) (central limit regime) $k_t = xt^{1/2}$. However, when plotting for finite k and t the approximate implied volatility, the two curves are the same.

We first use the numerical methods detailed in next Section 5 to compute $\Sigma(x)$ and $a(x)$. In figure 1, we display implied volatility smiles in the rBergomi model with $\theta = 1$, for varying t, where the rate function is computed using the Ritz method in Section 5.1 and the coefficient $a(x)$ is computed using the KL decomposition from Section 5.3. For comparison, we also use approximation $a(x) \approx a_0$, and show (21). We notice that both implementations perform well, and the use of KL decomposition gives a better approximation of the right wing. On several simulations, this improvement of KL over expansion $a(x) \approx a_0$ is more evident when taking $\theta = 1$, less when $\theta = 0$.

Practically, implementation of the KL formula requires to approximate the infinite product (38), and we observed that for smaller values of H the convergence of this product was much slower, requiring a prohibitively large number of basis functions, which is why we present these results for $H = 0.3$. We leave the numerically efficient implementation of the KL decomposition method for small values of H as a topic for future research. In what follows we will consider the approximation $a(x) \approx a(0)$, which is faster while still producing accurate smiles.

First, in figure 2, we show implied volatilities under model (2), with realistic parameters (close to the calibrated parameter to the SPX volatility on February 4, 2010, see Bayer *et al.* 2016), varying θ from 0 to 1. We note how our approximation is general enough to be applicable for any θ, improving previous asymptotics in all cases. We also note a slight deterioration of the quality of the approximation in the right wing as $\theta \to 1$, that could be improved using KL to compute $a(x)$.

Then, instead of varying θ, we fix $\theta = 0$ and show in figure 3 the comparison with the same approximations as before, when the expiry t increases. We see how our expansion lifts the FZ expansion, improving the approximation of the Monte Carlo price. The difference between the two approximations is due to the term structure correction $a_0 t^{2H}$. Clearly, the effect of this correction becomes more evident as t increases. On a number of numerical experiments, it is also clear that this correction becomes more and more important as $H \to 0$, not surprisingly since t^{2H} is larger, for small t, when H vanishes.

Now we check how our approximations behave as time increases. To do so, in figure 4, we show the ATM term structure of implied volatility, comparing ATM implied volatilities computed using Monte Carlo simulations and expansion (23), for rBergomi with $\theta = 0$ and $\theta = 1$. We do so for parameters as in figures 2 and 3, with $H = 0.3$, and for a different choice of parameters with $H = 0.1$ and a smaller volatility of volatility η, as in Bayer *et al.* (2019, Section 4). The value of η and H affect the quality of the approximation, which is less accurate for H very close to 0 and $\eta > 1$. On the other hand, as we show in figure 4, for $H = 0.3$ and $\eta > 1$ or H very close to 0 and $\eta < 1$ the short-time approximation is very good. This is consistent with the considerations on the interplay of H and η in El Euch *et al.* (2019, Page 505). We also see how the term structure is increasing in case $\theta = 0$ and decreasing in

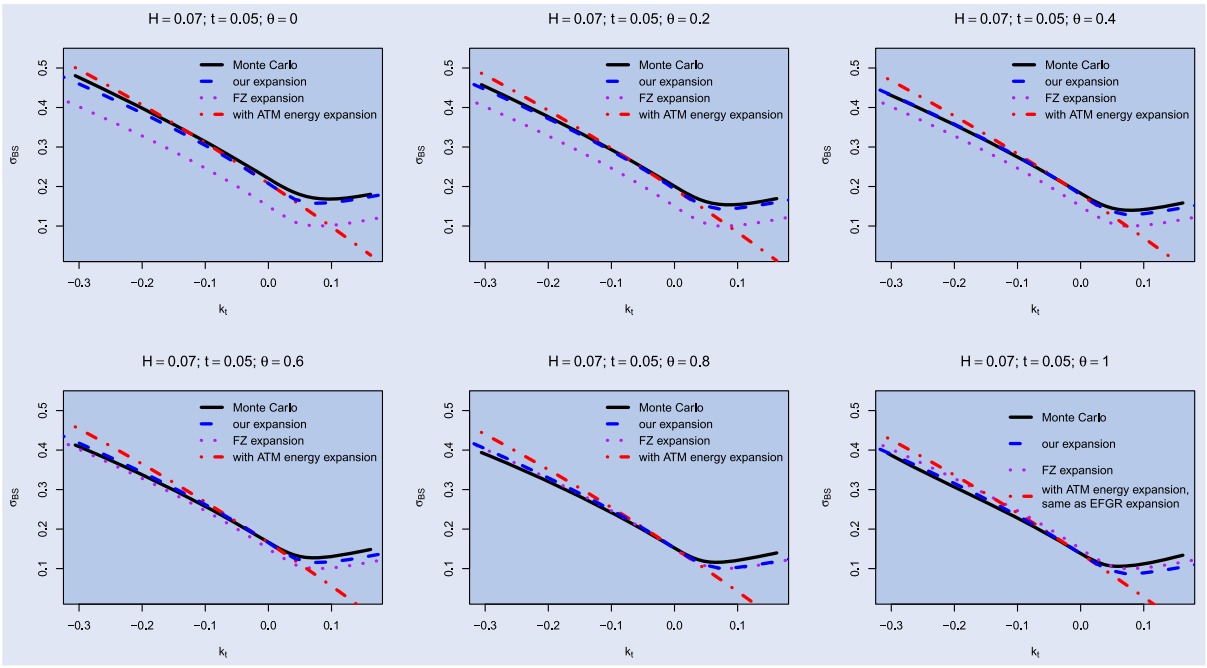

Figure 2. Implied volatility smile approximation for the rBergomi model with parameters $\sigma_0 = 0.15, \eta = 1.8, \rho = -0.78, H = 0.07$, for expiry $t = 0.05$. The Monte Carlo price is computed via the hybrid scheme for rBergomi in Bennedsen *et al.* (2017) with $\kappa = 2$, with 10^9 simulations and 500 time steps. The rate function is computed using the Ritz method with $N = 9$ Fourier basis functions.

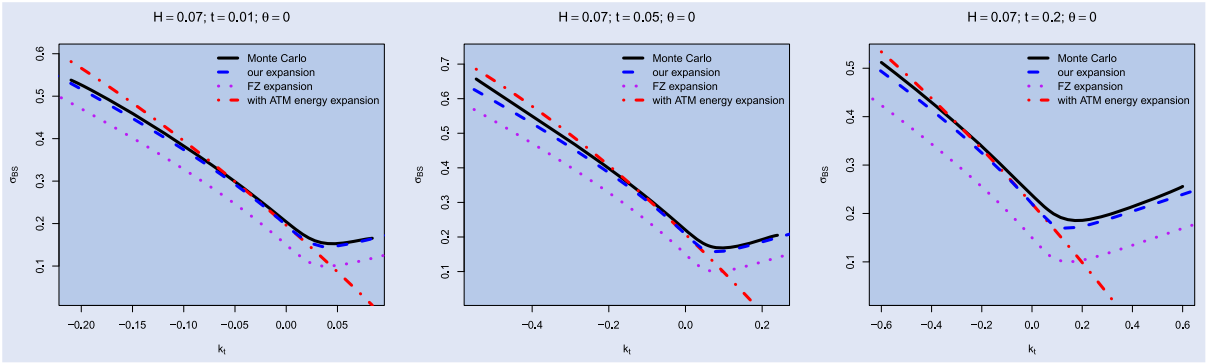

Figure 3. Implied volatility smile approximation for the rBergomi model with parameters $\theta = 0, \sigma_0 = 0.15, \eta = 1.8, \rho = -0.78, H = 0.07$, for expiry $t = 0.01, 0.05, 0.2$. The Monte Carlo price is computed via the hybrid scheme for rBergomi in Bennedsen *et al.* (2017) with $\kappa = 2$, with 10^9 simulations and 500 time steps of length $t/500$. The rate function is computed using the Ritz method with $N = 9$ Fourier basis functions.

case $\theta = 1$. This is always the case: a_0 in (19) is always positive for $\theta = 0$, always negative for $\theta = 1$ (cf. Remark 3.10). Also note that if the coefficient σ_0 were taken non-constant, the slope of the term structure would also be affected.

Finally, as in Remark 3.14, we consider moderate deviations. Figure 5 is as in Bayer *et al.* (2019, Figure 1), the "very rough" case $H = 0.1$ (which was the most problematic case in Bayer *et al.* (2019)). We are plotting, with $k_t = xt^{1/2-H+\beta}$, where $\beta = 0.06$, the Monte Carlo implied volatility and its approximation

$$\sigma(t, k_t) \approx \Sigma(0) + \Sigma'(0)xt^{\beta} + \frac{\Sigma''(0)}{2}x^2t^{2\beta} + \frac{a_0}{2\sigma_0}t^{2H},$$

considering terms up to the first order moderate deviation t^{β}, then up to the second-order moderate deviation $t^{2\beta}$, and finally considering also the term structure t^{2H}. We see how the term structure term improves the moderate deviation pricing. This also explains why, in Bayer *et al.* (2019), the moderate deviation pricing gets worse as $H \downarrow 0$, since the distance of such price from the real (Monte Carlo) one is of order t^{2H}. We also see that using the second order moderate deviation actually does not improve much, and this follows from the fact that the curvature is almost 0 with such choice of parameters (cf. Remark 3.14). As for the term structure, the accuracy of the approximation formula based on moderate deviations gets worse as η increases, for fixed H.

REMARK 4.2 As mentioned above, Monte Carlo pricing is implemented using the hybrid scheme, which introduces a bias in the volatility process, while this process could be simulated exactly. However, in extensive simulations we find that the exact simulation scheme is more unstable for very short maturities, even with a 10^9 trajectories and 500 time steps. This is most likely due to the singularity of the kernel at 0, which is what the hybrid scheme takes care of. On

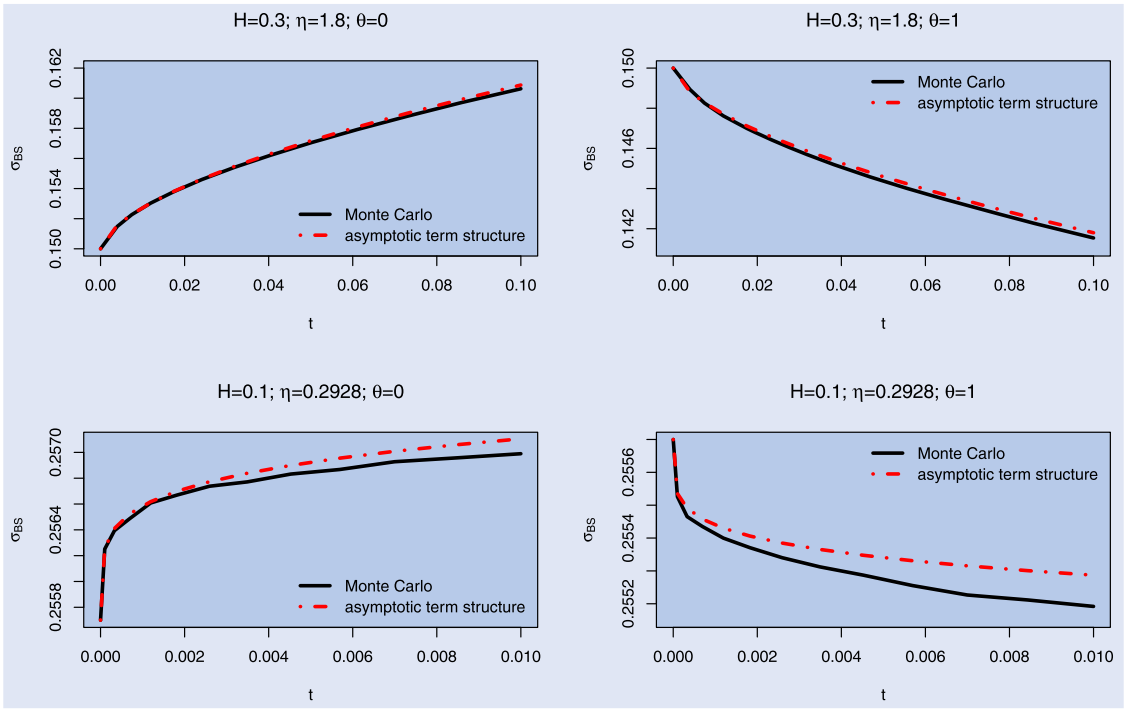

Figure 4. Term structure of volatility for the rBergomi model with parameters $\sigma_0 = 0.15, \eta = 1.8, \rho = -0.78, H = 0.3$ (above) and with parameters $\sigma_0 = 0.2557, \eta = 0.2928, \rho = -0.7571, H = 0.1$ (below). We plot ATM implied volatility as expiration time increases. We consider shorter expiries in the case of rougher trajectories (smaller Hurst parameter H; however, in this case we also take a smaller vol-of-vol parameter η). The Monte Carlo prices are computed via the hybrid scheme in Bennedsen *et al.* (2017) with $\kappa = 2$, with 10^9 simulations and 500 time steps.

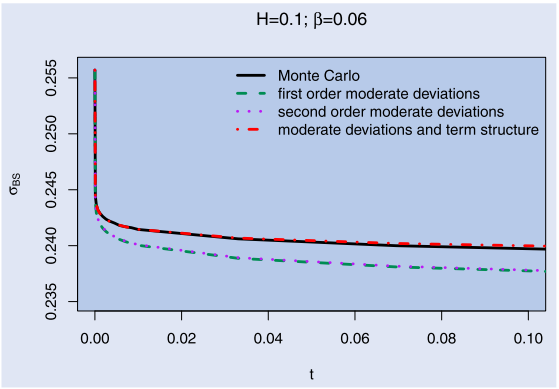

Figure 5. Moderate deviation with $\beta = 0.06$ and $x = 0.4$ (time varying log-strike $k_t = xt^{1/2-H+\beta}$) of implied volatility in rBergomi model with $\sigma_0 = 0.2557, \eta = 0.2928, \rho = -0.7571, H = 0.1, \theta = 0$. Simulation parameters: 10^8 simulation paths, 500 time steps. Time interval $[0, 0.1]$.

the other hand, with such a large number of paths and fine discretisation, for larger maturities the two schemes display no visible difference. Following these considerations, we used for our figures Monte Carlo prices simulated via the hybrid scheme in Bennedsen *et al.* (2017).

REMARK 4.3 In Forde and Zhang (2017, Section 4.5) asymptotics for model (3) with volatility driven by a Mandelbrot–Van Ness fBm (A2) are implemented. Without being completely rigorous, we have applied our expansion also in this case. We computed the K-functional numerically, as in this case no explicit formulas are available. Also in this case the term $a_0 t^{2H}$ lifts the smile, which gets closer to the real (Monte Carlo) implied volatility, for small $|x|$, with respect to the sole FZ expansion.

5. Computing the coefficients via projections

5.1. *Computing $\Sigma(x)$ using the Ritz method*

In order to use (20), the first challenge is the computation of the rate function. A numerical approximation to Λ can be obtained as described in Gelfand and Fomin (2000, Section 40), using the Ritz method, as is done in Forde and Zhang (2017). Natural choices for the orthonormal basis (ONB) $\{e_i\}_{i\geq 1}$ of H^1 are the Fourier basis,

$$\dot{e}_1(s) = 1, \quad \dot{e}_{2n}(s) = \sqrt{2}\cos(2\pi n s),$$
$$\dot{e}_{2n+1}(s) = \sqrt{2}\sin(2\pi n s), \text{ for } n \in \mathbb{N} \setminus \{0\}, \tag{31}$$

or the Haar basis,

$$\dot{e}_1(s) = 1,$$
$$\dot{e}_{2^k+l}(s) = 2^{k/2}\left(1_{\left[\frac{2l-2}{2^{k+1}}, \frac{2l-1}{2^{k+1}}\right]}(s) - 1_{\left[\frac{2l-1}{2^{k+1}}, \frac{2l}{2^{k+1}}\right]}(s)\right) \text{ for } k \geq 0,$$
$$1 \leq l \leq 2^k. \tag{32}$$

We consider functions $h \in H^1$ with $h(0) = 0$ so that

$$\dot{h}(s) = \sum_{n=1}^{N} a_n \dot{e}_n(s),$$

for $N \in \mathbb{N}$ fixed. Then we minimize

$$\Lambda(x) = \frac{(x - \rho G(h))^2}{2\overline{\rho}^2 F(h)} + \frac{\langle \dot{h}, \dot{h} \rangle}{2},$$

with

$$F(h) = \langle \sigma^2(\widehat{h}, 0), 1 \rangle, \quad G(h) = \langle \sigma(\widehat{h}, 0), \dot{h} \rangle, \tag{33}$$

over the Fourier coefficients $(a_n)_n$. This representation of the energy function is also taken from Forde and Zhang (2017) (see notation in Bayer *et al.* 2019, Proposition 5.1). The minimizing value for $\Lambda(x)$ is therefore our approximation for the energy and the corresponding function $\widehat{h}^x$ is the approximate most likely path for the fBm W^H associated with final condition x.

5.2. *A stochastic Taylor development*

The following stochastic Taylor expansion is sketched in Friz *et al.* (2021, Section 7.2) for $\sigma(\omega, t) = \sigma(W_t^H)$. As discussed in Section 2 and Friz *et al.* (2021, Section 7.3), our expansions can actually be carried out in the more general setting $\sigma(\omega, t) = \sigma(W_t^H, t^{2H})$. Under such volatility dynamics, the (rescaled) log-price process is as in (9). As in Friz *et al.* (2021, Section 7.2), we can shift the dynamics via $\widehat{\varepsilon}(W, \overline{W}) \mapsto (\widehat{\varepsilon} W + h, \widehat{\varepsilon}\overline{W} + \overline{h})$, and apply Girsanov theorem in order to center Brownian fluctuations in the minimizer. Then, a stochastic Taylor expansion gives

$$\begin{aligned} &\int_0^1 \sigma\left(\widehat{\varepsilon}\widehat{W}_t + \widehat{h}_t, \widehat{\varepsilon}^2 t^{2H}\right) d[\widehat{\varepsilon}\widetilde{W} + \widetilde{h}]_t \\ &\quad - \frac{\varepsilon\,\widehat{\varepsilon}}{2} \int_0^1 \sigma^2\left(\widehat{\varepsilon}\widehat{W}_t + \widehat{h}_t, \widehat{\varepsilon}^2 t^{2H}\right) \mathrm{d}t \\ &\equiv g_0 + \widehat{\varepsilon} g_1(\omega) + \widehat{\varepsilon}^2 g_2(\omega) + r_3(\omega), \end{aligned}$$

where $r_3(\omega)$ is small† , with

$$g_1 = \int_0^1 \sigma'(\widehat{h}_s^x, 0)\widehat{W}_s \mathrm{d}\widetilde{h}_s^x + \int_0^1 \sigma(\widehat{h}_s^x, 0)\mathrm{d}\widetilde{W}_s, \tag{34}$$

† The precise control of this remainder is detailed in Friz *et al.* (2021) and requires the sophisticated mathematical framework of regularity structures, that we do not intend to introduce in this paper. The interested reader is referred to Bayer *et al.* (2020) and Friz *et al.* (2021).

(cf. Friz *et al.* 2021, Section 7.2) and

$$g_2 = \begin{cases} \frac{1}{2}\int_0^1 \sigma''(\widehat{h}_s^x, 0)\widehat{W}_s^2 \mathrm{d}\widetilde{h}_s^x \\ \quad + \int_0^1 \sigma'(\widehat{h}_s^x, 0)\widehat{W}_s \mathrm{d}\widetilde{W}_s \\ \quad + \int_0^1 \dot{\sigma}(\widehat{h}_s^x, 0) s^{2H} \mathrm{d}\widetilde{h}_s^x & \text{if } H < 1/2, \\ \frac{1}{2}\int_0^1 \sigma''(\widehat{h}_s^x, 0)\widehat{W}_s^2 \mathrm{d}\widetilde{h}_s^x \\ \quad + \int_0^1 \sigma'(\widehat{h}_s^x, 0)\widehat{W}_s \mathrm{d}\widetilde{W}_s \\ \quad + \int_0^1 \dot{\sigma}(\widehat{h}_s^x, 0) s^{2H} \mathrm{d}\widetilde{h}_s^x \\ \quad - \frac{1}{2}\int_0^1 \sigma^2(\widehat{h}_s^x, 0)\mathrm{d}s & \text{if } H = 1/2. \end{cases} \tag{35}$$

The following formula for Δ_2 follows as Friz *et al.* 2021, Equation 7.5

$$\Delta_2 = \begin{cases} \frac{1}{2}\int_0^1 \sigma''(\widehat{h}_s^x, 0)\widehat{V}_s^2 \mathrm{d}\widetilde{h}_s^x \\ \quad + \int_0^1 \sigma'(\widehat{h}_s^x, 0)\widehat{V}_s \mathrm{d}\widetilde{V}_s, \\ \quad + \int_0^1 \dot{\sigma}(\widehat{h}_s^x, 0) s^{2H} \mathrm{d}\widetilde{h}_s^x & \text{if } H < 1/2, \\ \frac{1}{2}\int_0^1 \sigma''(\widehat{h}_s^x, 0)\widehat{V}_s^2 \mathrm{d}\widetilde{h}_s^x \\ \quad + \int_0^1 \sigma'(\widehat{h}_s^x, 0)\widehat{V}_s \mathrm{d}\widetilde{V}_s \\ \quad + \int_0^1 \dot{\sigma}(\widehat{h}_s^x, 0) s^{2H} \mathrm{d}\widetilde{h}_s^x \\ \quad - \frac{1}{2}\int_0^1 \sigma^2(\widehat{h}_s^x, 0)\mathrm{d}s & \text{if } H = 1/2. \end{cases} \tag{36}$$

where we write

$$\begin{aligned} &v_t = E[W_t g_1]/E[g_1^2], \quad \overline{v}_t = E[\overline{W}_t g_1]/E[g_1^2], \\ &\widetilde{v}_t = \rho v_t + \overline{\rho v}_t, \ \widehat{v} = K\dot{v} \end{aligned}$$

and

$$\widetilde{V}_t = \widetilde{W}_t - \widetilde{v}_t g_1, \quad \widehat{V}_t = \widehat{W}_t - g_1 \widehat{v}_t.$$

5.3. *Computing $a(x)$ using Karhunen–Loeve decomposition*

Assume we are given h^x computed by the Ritz method. Note then that $\overline{h}^x$ is obtained from h^x via the following formula

$$\dot{\overline{h}}^x(s) = \frac{x - \rho G(h^x)}{\overline{\rho} F(h^x)} \sigma(\widehat{h}^x(s)), \tag{37}$$

with $G(h), F(h)$ as in (33), as can be seen by optimizing over $\overline{h}$ for fixed h in the definition (10) of the rate function. Then we

assume a Karhunen–Loeve (KL) decomposition of $(W, \overline{W})$:

$$W = \sum_i \gamma_i e_i, \quad \overline{W} = \sum_i \overline{\gamma}_i e_i,$$

where $\{e_i\}_i$ is the ONB in (31) or (32) and $\gamma_i, \overline{\gamma}_i$ are i.i.d. standard Gaussians. This implies

$$\widehat{W} = \sum_i \gamma_i \widehat{e}_i,$$

with $\widehat{e}_i(t) = (K * \dot{e}_i)_t$. This yields

$$g_1 = \sum_i \mathrm{g}_i \gamma_i + \overline{\mathrm{g}_i \gamma_i},$$

where

$$\mathrm{g}_i = \int_0^1 \sigma'(\widehat{h}^x_s, 0)\widehat{e}_i(s)\mathrm{d}\widetilde{h}^x(s) + \rho \int_0^1 \sigma(\widehat{h}^x_s, 0)\mathrm{d}e_i(s),$$
$$\overline{\mathrm{g}_i} = \overline{\rho} \int_0^1 \sigma(\widehat{h}^x_s, 0)\mathrm{d}e_i(s).$$

In particular

$$\sigma_x^2 = \sum_i \mathrm{g}_i^2 + \overline{\mathrm{g}_i}^2.$$

Note then that

$$v(t) = \sum_i \mathrm{g}_i e_i(t)/\sigma_x^2, \quad \overline{v}(t) = \sum_i \overline{\mathrm{g}_i} e_i(t)/\sigma_x^2,$$
$$\widehat{v}(t) = \sum_i \mathrm{g}_i \widehat{e}_i(t)/\sigma_x^2,$$

We then can write all the terms in Δ_2 as follows. We denote

$$\alpha_{ij} = \int_0^1 \sigma''(\widehat{h}^x_s, 0)\widehat{e}_i(s)\widehat{e}_j(s)\mathrm{d}\widetilde{h}^x_s, \quad \beta_{ij} = \int_0^1 \sigma'(\widehat{h}^x_s, 0)\widehat{e}_i \mathrm{d}e_j,$$

$\delta_{ij} = \mathbf{1}_{i=j}$ and $\widetilde{\mathrm{g}}_i = \rho \mathrm{g}_i + \overline{\rho}\overline{\mathrm{g}}_i$. Now, expanding (36) with some long but standard computations we get to

$$\Delta_2 = \sum_{ij} (\gamma_i\gamma_j - \delta_{ij})\eta_{0;ij} + \gamma_i\overline{\gamma}_j\eta_{1;ij} + (\overline{\gamma_i\gamma_j} - \delta_{ij})\eta_{2;ij} + C =: \Delta_2^{(2)} + C,$$

where

$$\eta_{0;ij} = \frac{1}{2}\alpha_{ij} - \frac{1}{\sigma_x^2}\mathrm{g}_i \sum_k \mathrm{g}_k\alpha_{jk} + \frac{1}{2\sigma_x^4}\mathrm{g}_i\mathrm{g}_j\left(\sum_{kl}\mathrm{g}_k\mathrm{g}_l\alpha_{kl}\right) + \rho\beta_{ij} - \frac{1}{\sigma_x^2}\mathrm{g}_i\sum_k \widetilde{\mathrm{g}}_k\beta_{jk} - \frac{\rho}{\sigma_x^2}\mathrm{g}_i\sum_k \mathrm{g}_k\beta_{kj} + \frac{1}{\sigma_x^4}\mathrm{g}_i\mathrm{g}_j\left(\sum_{k,l}\mathrm{g}_k\widetilde{\mathrm{g}}_l\beta_{kl}\right)$$

$$\eta_{1;ij} = \overline{\rho}\beta_{ij} - \frac{1}{\sigma_x^2}\overline{\mathrm{g}}_j\sum_k \mathrm{g}_k\alpha_{ik} + \frac{1}{2\sigma_x^4}\mathrm{g}_i\overline{\mathrm{g}}_j\left(\sum_{k,l}\mathrm{g}_k\mathrm{g}_l\alpha_{k,l}\right) - \frac{1}{\sigma_x^2}\overline{\mathrm{g}}_j\sum_k \widetilde{\mathrm{g}}_k\beta_{ik} - \frac{\overline{\rho}}{\sigma_x^2}\mathrm{g}_i\sum_k \mathrm{g}_k\beta_{kj} - \frac{\rho}{\sigma_x^2}\overline{\mathrm{g}}_j\sum_k \mathrm{g}_k\beta_{ki} + \frac{1}{\sigma_x^4}\mathrm{g}_i\overline{\mathrm{g}}_j\left(\sum_k\sum_l \mathrm{g}_k\widetilde{\mathrm{g}}_l\beta_{kl}\right)$$

$$\eta_{2;ij} = \frac{1}{2\sigma_x^4}\overline{\mathrm{g}_i\mathrm{g}_j}\left(\sum_{k,l}\mathrm{g}_k\mathrm{g}_l\alpha_{k,l}\right) - \frac{\overline{\rho}}{\sigma_x^2}\overline{\mathrm{g}_i}\sum_k \mathrm{g}_k\beta_{kj} + \frac{1}{\sigma_x^4}\overline{\mathrm{g}_i\mathrm{g}_j}\left(\sum_{k,l}\mathrm{g}_k\widetilde{\mathrm{g}}_l\beta_{k,l}\right)$$

and

$$C = \int_0^1 \dot{\sigma}(\widehat{h}^x_s, 0)s^{2H}\mathrm{d}\widetilde{h}^x_s + \frac{1}{2}\sum_i \alpha_{ii} - \frac{1}{2\sigma_x^2}\sum_{i,k}\mathrm{g}_i\mathrm{g}_k\alpha_{ik} - \frac{1}{\sigma_x^2}\sum_{i,k}\mathrm{g}_i\widetilde{\mathrm{g}}_k\beta_{ik}.$$

Recall that one has

$$A(x) = e^{\Lambda'(x)C}\mathbb{E}\exp(\Lambda'(x)\Delta_2^{(2)}),$$

where

$$\Lambda'(x) = sgn(x)\sqrt{\frac{2\Lambda(x)}{\sigma_x^2}},$$

and since $\Delta_2^{(2)}$ is an element of the homogeneous Wiener chaos of order 2, the expectation above can be computed as the Carleman–Fredholm determinant $\det_2(I - 2M)^{-1/2}$, where M is the symmetric matrix

$$M = \Lambda'(x)\begin{pmatrix} \frac{1}{2}(\eta_0 + \eta_0^t) & \eta_1 \\ (\eta_1)^t & \frac{1}{2}(\eta_2 + \eta_2^t) \end{pmatrix}.$$

Namely one has

$$E\exp(\Lambda'(x)\Delta_2^{(2)}) = \Pi_{k\geq 0}(1 - 2\lambda_k)^{-1/2}e^{-\lambda_k} \tag{38}$$

where $(\lambda_k)_k$ are the eigenvalues of M (note that the fact that all $\lambda_k < 1/2$ comes from the non-degeneracy assumption). This formula is a simple integral computation if M is diagonal, and the general case follows by diagonalization, cf e.g. Inahama (2013, Remark 5.5) or Janson (1997, p.78).

Of course, in practice we consider approximations $W^N, \overline{W}^N$ obtained by truncating the sums to only keep indices $i \leq N$, where N is fixed, so that all the sums above are then replaced by finite sums. One also needs to compute numerically the integrals appearing in the definition of the coefficients g, α, β. We have found the Haar basis to be more convenient than the Fourier basis for this purpose since the $\widehat{e}_i$'s have explicit expressions in that case.

6. Proofs

6.1. *Energy expansion*

LEMMA 6.1 Fourth order energy expansion *Consider a stochastic volatility model following dynamics* (3) *and the*

associated energy function in (10). *Let* $\Lambda(x)$ *be the energy function in* (10). *Then*

$$\Lambda(x) = \frac{\Lambda''(0)}{2}x^2 + \frac{\Lambda'''(0)}{3!}x^3 + \frac{\Lambda^{(4)}(0)}{4!}x^4 + O(x^5)$$

where

$$\Lambda''(0) = \frac{1}{\sigma_0^2}, \quad \Lambda'''(0) = -6\frac{\rho\sigma_0'}{\sigma_0^4}\langle K1, 1\rangle, \tag{39}$$

and

$$\begin{aligned}\Lambda^{(4)}(0) = 12\frac{(\sigma_0')^2}{\sigma_0^6}\big\{&9\rho^2\langle K1,1\rangle^2 - \rho^2\langle (K1)^2, 1\rangle \\ &- \langle(\overline{K}1)^2, 1\rangle - 2\rho^2\langle K1, \overline{K}1\rangle\big\} \\ &- 12\frac{\sigma_0''}{\sigma_0^5}\rho^2\langle (K1)^2, 1\rangle.\end{aligned}$$

Remark 6.2 In this lemma we expand the rate function $\Lambda(x)$, which has been studied first in Forde and Zhang (2017). The second- and third-order terms in (39) have been computed in Bayer *et al.* (2019, Theorem 3.4). In both these papers, the volatility function is supposed to be $\sigma(W_t^H)$, but adding the dependence $\sigma(W_t^H, t^{2H})$ does not change the large deviations behavior, meaning that the rate function is the same as the one of the model given by $\sigma(W_t^H, 0)$.

Proof We have the following development for the minimizer $h_\cdot^x$ in (10), for $x \to 0$:

$$h_t^x = \alpha_t x + \beta_t\frac{x^2}{2} + \gamma_t\frac{x^3}{6} + O(x^4), \tag{40}$$

with

$$\begin{aligned}\alpha_t &= \frac{\rho}{\sigma_0}t, \\ \beta_t &= 2\frac{\sigma_0'}{\sigma_0^3}[\rho^2\langle K1, 1_{[0,t]}\rangle + \langle K1_{[0,t]}, 1\rangle - 3\rho^2 t\langle K1, 1\rangle],\end{aligned}$$

where α, β have been also computed in Bayer *et al.* (2019). We make here the ansatz that the expansion goes on one more order with γ, that we do not actually need to compute. The existence of such γ follows from the smoothness of $\sigma(\cdot,\cdot)$ (cf. Friz *et al.* 2021 and Bayer *et al.* 2019, Section 5.2). We can compute, using $\langle K(K1), 1\rangle = \langle K1, \overline{K}1\rangle$ and $\langle K(\overline{K}1), 1\rangle = \langle(\overline{K}1)^2, 1\rangle$,

$$\begin{aligned}K\dot\beta &= 2\frac{\sigma_0'}{\sigma_0^3}[\rho^2 K(K1) + K(\overline{K}1) - 3\rho^2\langle K1, 1\rangle K1], \\ \langle K\dot\beta, 1\rangle &= 2\frac{\sigma_0'}{\sigma_0^3}[\rho^2\langle K1, \overline{K}1\rangle + \langle(\overline{K}1)^2, 1\rangle - 3\rho^2\langle K1, 1\rangle^2], \\ \langle K1, \dot\beta\rangle &= 2\frac{\sigma_0'}{\sigma_0^3}[\rho^2\langle (K1)^2, 1\rangle + \langle K1, \overline{K}1\rangle - 3\rho^2\langle K1, 1\rangle^2].\end{aligned}$$

We also have

$$\sigma(\widehat{h}_s^x, 0) = \sigma_0 + x\frac{\sigma_0'}{\sigma_0}\rho K1(s) + \left(\frac{\sigma_0''}{\sigma_0^2}\rho^2(K1)^2(s) + \sigma_0' K\dot\beta(s)\right)\frac{x^2}{2} + O(x^3). \tag{41}$$

We use now (33) and compute

$$\begin{aligned}F(h^x) &= \sigma_0^2 + x2\rho\sigma_0'\langle K1, 1\rangle \\ &\quad + x^2\left\{\left(\left(\frac{\sigma_0'}{\sigma_0}\right)^2 + \frac{\sigma_0''}{\sigma_0}\right)\rho^2\langle (K1)^2, 1\rangle\right. \\ &\quad \left. + \sigma_0\sigma_0'\langle K\dot\beta, 1\rangle\right\} + O(x^3), \\ G(h^x) &= \rho x + x^2\left(\frac{\sigma_0'}{\sigma_0^2}\rho^2\langle K1, 1\rangle + \frac{\sigma_0}{2}\beta_1\right) \\ &\quad + x^3\left(\frac{\sigma_0}{6}\gamma_1 + \frac{\sigma_0''}{2\sigma_0^3}\rho^3\langle (K1)^2, 1\rangle\right. \\ &\quad + \rho\frac{(\sigma_0')^2}{\sigma_0^4}\big[(\rho^2+1)\langle K1, \overline{K}1\rangle + \rho^2\langle (K1)^2, 1\rangle \\ &\quad \left. + \langle(\overline{K}1)^2, 1\rangle - 6\rho^2\langle K1, 1\rangle^2\big]\right) + O(x^3),\end{aligned} \tag{42}$$

from which we get

$$\begin{aligned}x - \rho G(h^x) &= (1-\rho^2)x - x^2\rho\frac{\sigma_0'}{\sigma_0^2}(1-\rho^2)\langle K1, 1\rangle \\ &\quad - x^3\rho\left(\frac{\sigma_0}{6}\gamma_1 + \frac{\sigma_0''}{2\sigma_0^3}\rho^3\langle (K1)^2, 1\rangle\right. \\ &\quad + \rho\frac{(\sigma_0')^2}{\sigma_0^4}\big[(\rho^2+1)\langle K1, \overline{K}1\rangle \\ &\quad + \rho^2\langle (K1)^2, 1\rangle + \langle(\overline{K}1)^2, 1\rangle \\ &\quad \left. - 6\rho^2\langle K1, 1\rangle^2\big]\right) + O(x^3), \\ \frac{1}{F(h^x)} &= \frac{1}{\sigma_0^2} - x2\rho\frac{\sigma_0'}{\sigma_0^4}\langle K1, 1\rangle \\ &\quad - x^2\left\{\frac{\sigma_0''}{\sigma_0^5}\rho^2\langle (K1)^2, 1\rangle\right. \\ &\quad + \left(\frac{\sigma_0'}{\sigma_0^3}\right)^2\big(\rho^2\langle (K1)^2, 1\rangle + 2\langle(\overline{K}1)^2, 1\rangle \\ &\quad \left. + 2\rho^2\langle K1, \overline{K}1\rangle - 10\rho^2\langle K1, 1\rangle^2\big)\right\} \\ &\quad + O(x^3), \\ \frac{(x-\rho G(h^x))^2}{1-\rho^2} &= (1-\rho^2)x^2 - x^3 2\rho\frac{\sigma_0'}{\sigma_0^2}(1-\rho^2)\langle K1, 1\rangle \\ &\quad + x^4\left[\rho^2\frac{(\sigma_0')^2}{\sigma_0^4}(1+11\rho^2)\langle K1, 1\rangle^2\right. \\ &\quad - \frac{\sigma_0''}{\sigma_0^3}\rho^4\langle (K1)^2, 1\rangle - \frac{\rho\sigma_0}{3}\gamma_1 - 2\rho^2\frac{(\sigma_0')^2}{\sigma_0^4} \\ &\quad \times\big[(\rho^2+1)\langle K1, \overline{K}1\rangle + \rho^2\langle (K1)^2, 1\rangle \\ &\quad \left. + \langle(\overline{K}1)^2, 1\rangle\big]\right] + O(x^5),\end{aligned}$$

$$\frac{(x-\rho G(h^x))^2}{2(1-\rho^2)F(h^x)} = (\dots)x^2 + (\dots)x^3 - x^4\frac{\rho}{6\sigma_0}\gamma_1 + x^4\left\{-\frac{\rho^2}{2}\frac{\sigma_0''}{\sigma_0^5}\langle (K1)^2, 1\rangle - 2\rho^2\frac{(\sigma_0')^2}{\sigma_0^6}\langle K1, \overline{K}1\rangle - \frac{(\sigma_0')^2}{\sigma_0^6}\frac{\rho^4+\rho^2}{2}\langle (K1)^2, 1\rangle - \frac{(\sigma_0')^2}{\sigma_0^6}\langle (\overline{K}1)^2, 1\rangle + \frac{(\sigma_0')^2}{\sigma_0^6}\frac{3}{2}\rho^2(5-\rho^2)\langle K1, 1\rangle^2\right\} + O(x^5).$$

We also have, from (40)

$$\langle \dot{h}^x, \dot{h}^x\rangle = \cdots + x^4\left(\frac{\rho}{3\sigma_0}\gamma_1 + \left(\frac{\sigma_0'}{\sigma_0^3}\right)^2\left[\rho^4\langle (K1)^2, 1\rangle + \langle (\overline{K}1)^2, 1\rangle + 2\rho^2\langle K1, \overline{K}1\rangle + 3\rho^4\langle K1, 1\rangle^2 - 6\rho^2\langle K1, 1\rangle\langle \overline{K}1, 1\rangle\right]\right) + O(x^5).$$

Now we write, from Bayer *et al.* (2019, Proposition 5.1),

$$\Lambda(x) = \frac{(x-\rho G(h^x))^2}{2\overline{\rho}^2 F(h^x)} + \frac{\langle \dot{h}^x, \dot{h}^x\rangle}{2}$$

and use the expansions above for the two summands. The fourth-order expansion of $\Lambda(x)$ follows. ■

6.2. *Proof of Lemma 3.4*

Let us take $x \neq 0$.

STEP 1: We first need to expand $\overline{h}_t^x$ in (10), for small x (an expansion of h^x was computed in Bayer *et al.* 2019). We write

$$\Phi_1(W, \overline{W}) = X_1 \tag{43}$$

for the Itô map associated with the RoughVol model (8). Computing the Frechet derivative of Φ_1 with respect to the second component at $\mathrm{h} = (h, \overline{h})$ in the direction f we get (cf. (34))

$$\langle D\Phi_1(\mathrm{h}), (0, f)\rangle = \langle D_2\Phi_1(\mathrm{h}), f\rangle = \frac{\mathrm{d}}{\mathrm{d}\delta}\Phi_1(h, \overline{h} + \delta f) = \overline{\rho}\int_0^1 \sigma(\widehat{h}, 0)\,\mathrm{d}f. \tag{44}$$

From the first order optimality condition (Friz *et al.* 2021, Appendix B), we get that for h^x minimizer and any f in the Cameron–Martin space H^1,

$$\langle \mathrm{h}_t^x, \mathrm{f}_t\rangle_{H^1} = \Lambda'(x)\langle D\Phi_1, \mathrm{f}\rangle.$$

Let f be the second component of f. Using (44) we get

$$\int_0^1 \dot{\overline{h}}_t^x \dot{f}_t \mathrm{d}t = \langle \overline{h}_t^x, f_t\rangle_{H^1} = \Lambda'(x)\langle D_2\Phi_1, f\rangle = \overline{\rho}\Lambda'(x)\int_0^1 \sigma(\widehat{h}_t^x, 0)\dot{f}_t\mathrm{d}t.$$

Now, from (39) we derive that, for $x \to 0$,

$$\Lambda'(x) = \frac{x}{\sigma_0^2} - 3x^2\frac{\rho\sigma_0'}{\sigma_0^4}\langle K1, 1\rangle + O(x^3). \tag{45}$$

We get

$$\overline{h}_t^x = x\frac{\overline{\rho}}{\sigma_0}t + O(x^2).$$

We also have

$$h_t^x = x\frac{\rho}{\sigma_0}t + O(x^2), \quad \widetilde{h}_t^x = \rho h_t^x + \overline{\rho}\overline{h}_t^x = x\frac{t}{\sigma_0} + O(x^2),$$
$$\widehat{h}_t^x = (K\dot{h}^x)_t = x\frac{\rho}{\sigma_0}K1(t), \tag{46}$$

and

$$\sigma(\widehat{h}^x, 0) = \sigma_0 + x\frac{\rho\sigma_0'}{\sigma_0}K1 + O(x^2),$$
$$\sigma'(\widehat{h}^x, 0) = \sigma_0' + x\frac{\rho\sigma_0''}{\sigma_0}K1 + O(x^2),\ \dot{\sigma}(\widehat{h}^x, 0) = \dot{\sigma}_0 + O(x).$$

STEP 2: We recall here, from Friz *et al.* (2021), the definition of some quantities needed to compute $A(x)$. Let g_1 be as in (34) and let us write $\sigma_x^2 = Var(g_1)$ for its variance. We recall, again from Friz *et al.* (2021, Equation (6.3)), $\sigma_x^2 = 2\Lambda(x)/\Lambda'(x)^2$, from which we get

$$\sigma_x^2 = \sigma_0^2 + 4\rho\sigma_0'\langle K1, 1\rangle x + O(x^2). \tag{47}$$

From (34) we define and compute

$$v_t = \frac{E[W_t g_1]}{E[g_1^2]} = \frac{1}{\sigma_x^2}\left(\rho\int_0^t \sigma(\widehat{h}_s^x, 0)\mathrm{d}s + \int_0^1 \sigma'(\widehat{h}_s^x, 0)K1_{[0,t]}(s)\mathrm{d}\widetilde{h}_s^x\right),$$
$$\overline{v}_t = \frac{E[\overline{W}_t g_1]}{E[g_1^2]} = \frac{1}{\sigma_x^2}\overline{\rho}\int_0^t \sigma(\widehat{h}_s^x, 0)\mathrm{d}s. \tag{48}$$

(Note that $v, \overline{v}$ are in the Cameron–Martin space). From (11) we have that $A(x)$ in Theorem 3.2 is

$$A(x) = E\left[\exp(\Lambda'(x)\Delta_2)\right],$$

where Δ_2 is given in (36).

STEP 3: We can expand now such quantity, for $x \to 0$ and we get

$$A(x) = 1 + x\Lambda''(0)E\left[\Delta_2^0\right] + x^2\left(\frac{\Lambda'''(0)E\left[\Delta_2^0\right] + \Lambda''(0)^2E\left[(\Delta_2^0)^2\right]}{2} + \Lambda''(0)E\left[\partial_x\big|_{x=0}\Delta_2\right]\right) + O(x^3), \tag{49}$$

where Δ_2^0 denotes $\Delta_2|_{x=0}$. The statement of the theorem follows from the computation of the quantities in (49).

STEP 4: We compute

$$\dot{v}_t = \frac{1}{\sigma_x^2}\left(\rho\sigma(\widehat{h}_t^x, 0) + \int_t^1 \sigma'(\widehat{h}_s^x, 0)K(s,t)\mathrm{d}\widetilde{h}_s^x\right),$$

and we obtain, also using (48),

$$\begin{aligned}
\sigma_x^2 v_t &= \rho\sigma_0 t + x\frac{\sigma_0'}{\sigma_0}\big(\rho^2\langle K1, 1_{[0,t]}\rangle + \langle K1_{[0,t]}, 1\rangle\big) + O(x^2),\\
\sigma_x^2\widetilde{v}_t &= \sigma_0 t + x\frac{\sigma_0'}{\sigma_0}\rho\big(\langle K1, 1_{[0,t]}\rangle + \langle K1_{[0,t]}, 1\rangle\big) + O(x^2),\\
\sigma_x^2\widehat{v}_t &= \rho\sigma_0 K1(t) + x\frac{\sigma_0'}{\sigma_0}\big(\rho^2 K(K1)(t) + K(\overline{K}1)(t)\big) + O(x^2),
\end{aligned} \tag{50}$$

where we have used $\langle K1_{[0,t]}, 1\rangle = \int_0^t \overline{K}1(u)\,\mathrm{d}u$. We have

$$\begin{aligned}
\sigma_x^4\widehat{v}_t\mathrm{d}\widetilde{v}_t = \rho\sigma_0^2 K1(t)\mathrm{d}t + x\sigma_0'\Big(&\rho^2 K(K1)(t) + K(\overline{K}1)(t)\\
&+ \rho^2 K1(t)\big(K1(t) + \overline{K}1(t)\big)\Big)\mathrm{d}t.
\end{aligned}$$

Putting together the previous expressions and using $\langle K(K1), 1\rangle = \langle K1, \overline{K}1\rangle$ and $\langle K(\overline{K}1), 1\rangle = \langle(\overline{K}1)^2, 1\rangle$ we get

$$\begin{aligned}
\sigma_x^4\int_0^1 \widehat{v}_t\mathrm{d}\widetilde{v}_t &= \rho\sigma_0^2\langle K1, 1\rangle + x\sigma_0'\big(\rho^2\langle(K1)^2, 1\rangle\\
&\quad + \langle(\overline{K}1)^2, 1\rangle + 2\rho^2\langle K1, \overline{K}1\rangle\big) + O(x^2),\\
\sigma_x^4\int_0^1 K1(t)\widehat{v}_t\mathrm{d}\widetilde{v}_t &= \rho\sigma_0^2\langle(K1)^2, 1\rangle + O(x),\\
\sigma_x^4\int_0^1 \widehat{v}_t^2\mathrm{d}t &= \rho^2\sigma_0^2\langle(K1)^2, 1\rangle + O(x).
\end{aligned} \tag{51}$$

This implies, together with (47),

$$\begin{aligned}
\partial_x\left(\sigma_x^2\int_0^1 \widehat{v}_t\mathrm{d}\widetilde{v}_t\right) = \frac{\sigma_0'}{\sigma_0^2}\big(&2\rho^2\langle K1, \overline{K}1\rangle + \rho^2\langle(K1)^2, 1\rangle\\
&+ \langle(\overline{K}1)^2, 1\rangle - 4\rho^2\langle K1, 1\rangle^2\big).
\end{aligned} \tag{52}$$

We can now compute

$$\begin{aligned}
E\int_0^1(\widehat{V}_s\mathrm{d}\widetilde{V}_s) &= E\int_0^1 \widehat{W}_s\mathrm{d}\widetilde{W}_s + E[g_1^2]\int_0^1 \widehat{v}_s\mathrm{d}\widetilde{v}_s\\
&\quad - E\left[g_1\left(\int_0^1 \widehat{W}_s\mathrm{d}\widetilde{v}_s + \widehat{v}_s\mathrm{d}\widetilde{W}_s\right)\right],
\end{aligned}$$

where

$$\begin{aligned}
&E\left[g_1\int_0^1 \widehat{W}_s\mathrm{d}\widetilde{v}_s\right]\\
&\quad = \int_0^1\int_0^s K(s,u)\mathrm{d}E\big[g_1 W_u\big]\mathrm{d}\widetilde{v}_s = \sigma_x^2\int_0^1 \widehat{v}_s\mathrm{d}\widetilde{v}_s,\\
&E\big[g_1\int_0^1 \widehat{v}_s\mathrm{d}\widetilde{W}_s\big] = \int_0^1 \widehat{v}_s\mathrm{d}E\big[g_1\widetilde{W}_s\big] = \sigma_x^2\int_0^1 \widehat{v}_s\mathrm{d}\widetilde{v}_s,
\end{aligned}$$

so that

$$E\int_0^1\big(\widehat{V}_s\mathrm{d}\widetilde{V}_s\big) = -\sigma_x^2\int_0^1 \widehat{v}_s\mathrm{d}\widetilde{v}_s.$$

We also compute

$$\begin{aligned}
\int_0^1 E[\widehat{V}_s^2]\mathrm{d}s &= \langle K^2 1, 1\rangle - \sigma_x^2\int_0^1 \widehat{v}_s^2\mathrm{d}s,\\
\int_0^1 K1(s)E[\widehat{V}_s\mathrm{d}\widetilde{V}_s] &= -\sigma_x^2\int_0^1 K1(s)\widehat{v}_s\mathrm{d}\widetilde{v}_s,
\end{aligned}$$

and all these quantities can be expanded in x using (51). Now we use (36) to write, in the case $H < 1/2$

$$E\Delta_2^0 = \sigma_0' E\int_0^1 \widehat{V}_s^0\mathrm{d}\widetilde{V}_s^0 = -\rho\sigma_0'\langle K1, 1\rangle. \tag{53}$$

Moreover, using (46),

$$\partial_x\big|_{x=0}\int_0^1 \dot{\sigma}(\widehat{h}_s^x, 0)s^{2H}\mathrm{d}\widetilde{h}_s^x = \frac{\dot{\sigma}_0}{(2H+1)\sigma_0}.$$

Now, also using (36) and (51) we get

$$\begin{aligned}
\partial_x E\Delta_2\big|_{x=0} &= \frac{\sigma_0''}{2\sigma_0}\int_0^1 E[(\widehat{V}_s^0)^2]\mathrm{d}s\\
&\quad + \rho\frac{\sigma_0''}{\sigma_0}\int_0^1 K1(s)E[\widehat{V}_s^0\mathrm{d}\widetilde{V}_s^0]\\
&\quad + \sigma_0'\int_0^1 \partial_x E[\widehat{V}_s\mathrm{d}\widetilde{V}_s]\big|_{x=0} + \frac{\dot{\sigma}_0}{(2H+1)\sigma_0}\\
&= \frac{\sigma_0''}{\sigma_0}\left(\frac{\langle K^2 1, 1\rangle}{2} - \frac{3}{2}\rho^2\langle(K1)^2, 1\rangle\right)\\
&\quad - \left(\frac{\sigma_0'}{\sigma_0}\right)^2\big(2\rho^2\langle K1, \overline{K}1\rangle\\
&\quad + \rho^2\langle(K1)^2, 1\rangle + \langle(\overline{K}1)^2, 1\rangle - 4\rho^2\langle K1, 1\rangle^2\big)\\
&\quad + \frac{\dot{\sigma}_0}{(2H+1)\sigma_0}.
\end{aligned}$$

STEP 5: We need now to compute

$$E[(\Delta_2^0)^2] = (\sigma_0')^2 E\left(\int_0^1 \widehat{V}_s^0\mathrm{d}\widetilde{V}_s^0\right)^2,$$

where (using definitions and (50))

$$\widehat{V}_s^0 = \widehat{W}_s - \rho K1(s)\widetilde{W}_1 \text{ and } \mathrm{d}\widetilde{V}_s^0 = \mathrm{d}\widetilde{W}_s - \widetilde{W}_1\mathrm{d}s.$$

We can rewrite

$$\begin{aligned}
\int_0^1 \widehat{V}_s^0\mathrm{d}s\widetilde{W}_1 &= \overline{\rho}\int_0^1 \overline{K}1(u)\widetilde{W}_u\mathrm{d}B_u\\
&\quad + \int_0^1\Big(\rho\big(\overline{K}1(u) - 2\langle K1, 1\rangle\big)\widetilde{W}_u\\
&\quad + \int_0^u \overline{K}1(s)\,\mathrm{d}W_s\Big)\mathrm{d}\widetilde{W}_u.
\end{aligned}$$

and, differentiating the product $\int_0^1 K1(u)\mathrm{d}\widetilde{W}_u, \widetilde{W}_1$

$$\begin{aligned}
&\int_0^1 \widehat{V}_s^0 \mathrm{d}\widetilde{V}_s^0 \\
&\quad = -\overline{\rho}\int_0^1 \overline{K}1(u)\widetilde{W}_u dB_u \\
&\qquad + \int_0^1 \left(\widehat{V}_u^0 - \int_0^u \overline{K}1(s)\,\mathrm{d}W_s \right. \\
&\qquad \left. + \rho\big(2\langle K1,1\rangle - \overline{K}1(u)\big)\widetilde{W}_u\right)\mathrm{d}\widetilde{W}_u \\
&\quad = -\overline{\rho}\int_0^1 \overline{K}1(u)\widetilde{W}_u dB_u - \rho\int_0^1 K1(u)\,\mathrm{d}u \\
&\qquad + \int_0^1 \left(\widehat{W}_u - \int_0^u \overline{K}1(s)\,\mathrm{d}W_s \right. \\
&\qquad + \rho\big(2\langle K1,1\rangle - \overline{K}1(u) - K1(u)\big)\widetilde{W}_u \\
&\qquad \left. - \rho\int_0^u K1(s)\mathrm{d}\widetilde{W}_s\right)\mathrm{d}\widetilde{W}_u
\end{aligned}$$

with $\widetilde{W}$ independent of B. Therefore, by Itô isometry,

$$\begin{aligned}
&E\left[\left(\int_0^1 \widehat{V}_s^0 \mathrm{d}\widetilde{V}_s^0\right)^2\right] \\
&\quad = \overline{\rho}^2\int_0^1 \overline{K}1(u)^2 u\mathrm{d}u + \rho^2\left(\int_0^1 K1(u)\,\mathrm{d}u\right)^2 \\
&\qquad + E\int_0^1 \left(\widehat{W}_u - \int_0^u \overline{K}1(s)\,\mathrm{d}W_s + \rho\big(2\langle K1,1\rangle\right. \\
&\qquad \left. -\overline{K}1(u) - K1(u)\big)\widetilde{W}_u - \rho\int_0^u K1(s)\mathrm{d}\widetilde{W}_s\right)^2 \mathrm{d}u.
\end{aligned}$$

We can apply again Itô isometry to compute the last expectations, and

$$\begin{aligned}
&E\left[\left(\int_0^1 \widehat{V}_s^0 \mathrm{d}\widetilde{V}_s^0\right)^2\right] \\
&\quad = \rho^2\int_0^1\int_0^u \big(K(u,s) - \overline{K}1(s) + 2\langle K1,1\rangle - \overline{K}1(u) \\
&\qquad - K1(u) - K1(s)\big)^2\,\mathrm{d}s\,\mathrm{d}u \\
&\qquad + \rho^2\left(\int_0^1 K1(u)\,\mathrm{d}u\right)^2 \\
&\qquad + \overline{\rho}^2\int_0^1\int_0^u (K(u,s) - \overline{K}1(s))^2\mathrm{d}s\,\mathrm{d}u \\
&\qquad + \overline{\rho}^2\int_0^1 \overline{K}1(u)^2 u\mathrm{d}u.
\end{aligned}$$

At this point it is a (long) calculus excercise (noting $\langle\overline{K}1,1\rangle = \langle K1,1\rangle$) to show that

$$\begin{aligned}
&E\left[\left(\int_0^1 \widehat{V}_s^0 \mathrm{d}\widetilde{V}_s^0\right)^2\right] \\
&\quad = \rho^2(3\langle K1,1\rangle^2 - \langle (K1)^2,1\rangle - 2\langle K1,\overline{K}1\rangle) \\
&\qquad + \langle K^2 1,1\rangle - \langle(\overline{K}1)^2,1\rangle
\end{aligned} \tag{54}$$

STEP 6: Substituting in (49) we get

$$\begin{aligned}
A(x) &= 1 - x\frac{\rho\sigma_0'}{\sigma_0^2}\langle K1,1\rangle + x^2 \\
&\quad \left\{\frac{(\sigma_0')^2}{\sigma_0^4}\left(3\rho^2\langle K1,1\rangle^2 + \frac{1}{2}E\left[\left(\int_0^1 \widehat{V}_s^0\mathrm{d}\widetilde{V}_s^0\right)^2\right]\right)\right. \\
&\quad + \frac{\sigma_0''}{\sigma_0^3}\left(\frac{\langle K^2 1,1\rangle}{2} - \frac{3}{2}\rho^2\langle (K1)^2,1\rangle\right) \\
&\quad - \frac{(\sigma_0')^2}{\sigma_0^4}\big(2\rho^2\langle K1,\overline{K}1\rangle + \rho^2\langle (K1)^2,1\rangle + \langle(\overline{K}1)^2,1\rangle \\
&\quad \left. -4\rho^2\langle K1,1\rangle^2\big) + \frac{\dot{\sigma}_0}{(2H+1)\sigma_0^3}\right\} + O(x^3) \\
&= 1 - x\frac{\rho\sigma_0'}{\sigma_0^2}\langle K1,1\rangle + x^2 \\
&\quad \left\{\frac{(\sigma_0')^2}{\sigma_0^4}\left(\frac{\langle K^2 1,1\rangle}{2} - \frac{3}{2}\langle(\overline{K}1)^2,1\rangle\right.\right. \\
&\quad \left.\left. + \rho^2\left(\frac{17}{2}\langle K1,1\rangle^2 - \frac{3}{2}\langle (K1)^2,1\rangle - 3\langle K1,\overline{K}1\rangle\right)\right)\right. \\
&\quad \left. + \frac{\sigma_0''}{\sigma_0^3}\left(\frac{\langle K^2 1,1\rangle}{2} - \frac{3}{2}\rho^2\langle (K1)^2,1\rangle\right) + \frac{\dot{\sigma}_0}{(2H+1)\sigma_0^3}\right\} \\
&\quad + O(x^3)
\end{aligned}$$

and we get Theorem 12.

STEP 7: When $H = 1/2$, Δ_2 in (36) has an additional summand. Let us write

$$\begin{aligned}
\widetilde{\Delta}_2 &= \frac{1}{2}\int_0^1 \sigma''(\widehat{h}_s^x,0)\widehat{V}_s^2\mathrm{d}\widetilde{h}_s^x + \int_0^1 \sigma'(\widehat{h}_s^x,0)\widehat{V}_s\mathrm{d}\widetilde{V}_s \\
&\quad + \int_0^1 \dot{\sigma}(\widehat{h}_s^x,0)\mathrm{d}\widetilde{h}_s^x,
\end{aligned}$$

so that $\widetilde{\Delta}_2$ has the same expression as Δ_2 in the rough case $H < 1/2$. For $H = 1/2$ we can write

$$\Delta_2 = \widetilde{\Delta}_2 - \frac{1}{2}\int_0^1 \sigma^2(\widehat{h}_s^x,0)\mathrm{d}s, \tag{55}$$

so that

$$\Delta_2^0 = \widetilde{\Delta}_2^0 - \frac{\sigma_0^2}{2} \tag{56}$$

and, using (41)

$$\partial_x\big|_{x=0}\Delta_2 = \partial_x\big|_{x=0}\widetilde{\Delta}_2 - \sigma_0'\rho\langle K1,1\rangle \tag{57}$$

Now, $A(x)$ in Theorem 3.2 is

$$A(x) = e^x E\left[\exp(\Lambda'(x)\Delta_2)\right]$$

with Δ_2 as above. Expanding in x we find

$$A(x) = 1 + x(\Lambda''(0)E[\Delta_2^0] + 1)$$

$$+x^2\left(\frac{\Lambda'''(0)E[\Delta_2^0]+E\left[(\Lambda''(0)\Delta_2^0+1)^2\right]}{2}\right.$$
$$\left.+\Lambda''(0)E\left[\partial_x\big|_{x=0}\Delta_2\right]\right)+O(x^3)$$
$$=1+x\Lambda''(0)E[\widetilde{\Delta}_2^0]$$
$$+x^2\left(\frac{\Lambda'''(0)E[\widetilde{\Delta}_2^0]+\Lambda''(0)^2E[(\widetilde{\Delta}_2^0)^2]}{2}\right.$$
$$\left.+\Lambda''(0)E[\partial_x\big|_{x=0}\widetilde{\Delta}_2]\right)+\frac{x}{2}+\frac{x^2}{8}+O(x^3) \qquad (58)$$

(we have used (39) and (53)).

6.3. *Proof of Theorem 3.7*

A Taylor expansion gives

$$\Sigma(x)=\frac{x}{\sqrt{2\Lambda(x)}}$$
$$=\frac{1}{\sqrt{\Lambda''(0)}}\left(1-\frac{\Lambda'''(0)}{6\Lambda''(0)}x\right.$$
$$\left.+\frac{\Lambda'''(0)^2-\Lambda''(0)\Lambda^{(4)}(0)}{24\Lambda''(0)^2}x^2\right)+O(x^3).$$

The explicit expressions for the three terms now follow from Lemma 6.1. Let us compute $a(x)$. The rate function is quadratic and $\Lambda(0)=\Lambda'(0)=0$. Then, using Taylor developments of Λ and $x\to\frac{1}{1+x}$ we get

$$\frac{2\Lambda(x)}{x\Lambda'(x)}=1-\frac{x}{6}\frac{\Lambda'''(0)}{\Lambda''(0)}+\frac{x^2}{12}\left\{\left(\frac{\Lambda'''(0)}{\Lambda''(0)}\right)^2-\frac{\Lambda^{(4)}(0)}{\Lambda''(0)}\right\}$$
$$+O(x^3)=1+\sqrt{\Lambda''(0)}\left(x\Sigma'(0)+x^2\Sigma''(0)\right)$$
$$+O(x^3)$$

From Lemma 6.1,

$$\frac{2\Lambda(x)}{x\Lambda'(x)}=1+x\frac{\rho\sigma_0'\langle K1,1\rangle}{\sigma_0^2}+x^2\sqrt{\Lambda''(0)}\Sigma''(0)+O(x^3)$$

and, with $A(x)$ given in Lemma 12, we have when $H<1/2$

$$\frac{2A(x)\Lambda(x)}{x\Lambda'(x)}=1+x^2\overline{a}_0+O(x^3)$$

with

$$\overline{a}_0=\frac{(\sigma_0')^2}{\sigma_0^4}C_{K,\rho}+\frac{\sigma_0''}{\sigma_0^3}\overline{C}_{K,\rho}+\frac{\dot\sigma_0}{(2H+1)\sigma_0^3}+\sqrt{\Lambda''(0)}v''(0)$$
$$-\frac{\rho^2(\sigma_0')^2\langle K1,1\rangle^2}{\sigma_0^4}$$
$$=\frac{(\sigma_0')^2}{\sigma_0^4}\frac{D_{K,\rho}}{2}+\frac{\sigma_0''}{\sigma_0^3}\frac{\overline{D}_{K,\rho}}{2}+\frac{\dot\sigma_0}{(2H+1)\sigma_0^3},$$

with $D_{K,\rho},\overline{D}_{K,\rho}$, defined in (17). Now, as a consequence of Lemma 6.1, we have

$$\frac{x^2}{2\Lambda(x)^2}\sim\frac{2}{\Lambda''(0)^2x^2}=\frac{2\sigma_0^4}{x^2}$$

and the expansion of $a(x)$ follows. When $H=1/2$,

$$\frac{2A(x)\Lambda(x)}{x\Lambda'(x)\exp(x/2)}=1+x^2\overline{a}_0+O(x^3)$$

with

$$\overline{a}_0=\frac{(\sigma_0')^2}{\sigma_0^4}\frac{D_{K,\rho}}{2}+\frac{\sigma_0''}{\sigma_0^3}\frac{\overline{D}_{K,\rho}}{2}+\frac{\dot\sigma_0}{(2H+1)\sigma_0^3}+\frac{\rho\sigma_0'}{2\sigma_0^2}\langle K1,1\rangle.$$

We conclude as in the case $H<1/2$.

6.4. *Proof of Theorem 3.13*

The call asymptotics is a corollary of Theorem 3.2, taking into consideration that

$$\Lambda(x_t)=\sum_{i=2}^n\frac{\Lambda^{(i)}(0)}{i!}x^it^{i\beta}+O(t^{(n+1)\beta})$$

and that $O(t^{(n+1)\beta-2H})\to0$ under $\beta\in(\frac{2H}{n+1},\frac{2H}{n}]$. Recall $\Lambda''(0)=\sigma_0^{-2}$ and the first statement follows.

Let us write $\alpha=2H-2\beta$, $\delta=1/2+2H-2\beta$, $\gamma=1/2-H+\beta$ and $\mathcal{M}(t,x)=\sum_{i=3}^n\frac{\Lambda^{(i)}(0)}{i!}x^it^{i\beta-2H}$. We intend to apply (Gao and Lee 2014, Corollary 7.1, Equation (7.2)), where $G_-(k,u)$ denotes $\sqrt{2}(\sqrt{u+k}-\sqrt{u})$ and V denotes $\sqrt{t}\sigma_{BS}$. To do so, we notice that

$$G_-^2(k,u)=\frac{k^2}{2u}+o\left(\frac{k^2}{u^2}\right) \qquad (59)$$

when $k\downarrow0$ and $u\to\infty$. In the notation of Gao and Lee (2014), we have

$$L_t=-\log c(t,k_t),$$

and G will be computed for $u=L_t-\frac{3}{2}\log L_t+\log(\frac{k_t}{4\sqrt{\pi}})$, so let us compute

$$L_t-\frac{3}{2}\log L_t+\log\left(\frac{k_t}{4\sqrt{\pi}}\right)$$
$$=\frac{x^2}{2\sigma_0^2t^\alpha}+\mathcal{M}(t,x)-\log\frac{\sigma_0^3}{x^2\sqrt{2\pi}}-\delta\log t-\frac{3}{2}\log L_t$$
$$+\log\left(\frac{k_t}{4\sqrt{\pi}}\right)+o(1)$$

and take care of the logarithmic terms in t. For $t\downarrow0$,

$$-\delta\log t-\frac{3}{2}\log L_t+\log\left(\frac{k_t}{4\sqrt{\pi}}\right)$$
$$=\left(-\delta+\frac{3}{2}\alpha+\gamma\right)\log t-\frac{3}{2}\log\left(\frac{x^2}{2\sigma_0^2}\right)+\log\left(\frac{x}{4\sqrt{\pi}}\right)$$

$$+ o(1)$$
$$= -\frac{3}{2}\log\left(\frac{x^2}{2\sigma_0^2}\right) + \log\left(\frac{x}{4\sqrt{\pi}}\right) + o(1)$$

So

$$L_t - \frac{3}{2}\log L_t + \log\left(\frac{x}{4\sqrt{\pi}}\right) = \frac{x^2}{2\sigma_0^2 t^\alpha} + \mathcal{M}(t,x) + o(1) \tag{60}$$

Equations (59) and (60) tell us that

$$\frac{1}{t}G_-^2\left(k_t, L_t - \frac{3}{2}\log L_t + \log\left(\frac{k_t}{4\sqrt{\pi}}\right)\right) = \sigma_0^2 \frac{1}{1 + \frac{2\sigma_0^2\mathcal{M}(t,x)}{x^2}t^\alpha + o(t^\alpha)} + o(t^\alpha) \tag{61}$$

The proof now boils down to writing the development of this factor using the Taylor developement of $\frac{1}{1+u}$, with $u = 2\sigma_0^2\mathcal{M}(t,x)t^\alpha/x^2 + o(t^\alpha)$. We have, for $j \in \mathbb{N}$,

$$u^j = \left(\frac{2\sigma_0^2\mathcal{M}(t,x)}{x^2}t^\alpha\right)^j + o(t^\alpha)$$

using $\mathcal{M}(t,x)^{p-1}t^{j\alpha} = o(t^\alpha)$ for $j \geq p \geq 1$. Also notice $u^{n-1} = O((\mathcal{M}(t,x)t^\alpha)^{n-1}) = o(t^\alpha)$ because $\beta \in (\frac{2H}{n+1}, \frac{2H}{n}]$. We have

$$\frac{1}{1+u} = \sum_{j=0}^{n-2}(-1)^j u^j + O(u^{n-1}) = \sum_{j=0}^{n-2}(-1)^j\left(\frac{2\sigma_0^2\mathcal{M}(t,x)}{x^2}t^\alpha\right)^j + o(t^\alpha) \tag{62}$$

So from (61) and (62)

$$\frac{1}{t}G_-^2(k_t, L_t - \frac{3}{2}\log L_t + \log\left(\frac{k_t}{4\sqrt{\pi}}\right)) = \sum_{j=0}^{n-2}(-1)^j 2^j \sigma_0^{2(j+1)}\left(\frac{\mathcal{M}(t,x)}{x^2}t^\alpha\right)^j + o(t^\alpha) \tag{63}$$

We apply now (Gao and Lee 2014, Corollary 7.1, Equation (7.2)):

$$\left|\frac{1}{t}G_-^2(k_t, L_t - \frac{3}{2}\log L_t + \log(\frac{k_t}{4\sqrt{\pi}})) - \sigma_{BS}^2(k_t)\right| = o\left(\frac{k_t^2}{tL_t^2}\right) = o\big(t^\alpha\big)$$

and obtain expansion (27).

Acknowledgments

We are grateful to C. Bayer and M. Fukasawa for discussion and to F. Bourgey and M. Pakkanen for the Python and R code for simulating the rough Bergomi model. We thank an anonymous reviewer for several remarks that helped us to improve the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

PKF and PP gratefully acknowledge financial support from European Research Council (ERC) Grant CoG-683164 and German science foundation (DFG) via the cluster of excellence MATH+, project AA4-2. PG acknowledges financial support from the French ANR via the project ANR-16-CE40-0020-01.

References

Ait-Sahalia, Y., Li, C. and Li, C.X., Implied stochastic volatility models. *Rev. Financ. Stud.*, 2020, **34**, 394–450.

Alòs, E. and León, J., On the curvature of the smile in stochastic volatility models. *SIAM J. Financ. Math.*, 2017, **8**(1), 373–399.

Alòs, E., León, J. and Vives, J., On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance Stoch.*, 2007, **11**(4), 571–589.

Bayer, C., Friz, P. and Gatheral, J., Pricing under rough volatility. *Q. Finance*, 2016, **16**(6), 887–904.

Bayer, C., Friz, P.K., Gassiat, P., Martin, J. and Stemper, B., A regularity structure for rough volatility. *Math. Finance*, 2020, **30**(3), 782–832.

Bayer, C., Friz, P.K., Gulisashvili, A., Horvath, B. and Stemper, B., Short-time near-the-money skew in rough fractional volatility models. *Quant. Finance*, 2019, **19**(5), 779–798.

Bayer, C., Hammouda, C.B. and Tempone, R., Hierarchical adaptive sparse grids and quasi-Monte Carlo for option pricing under the rough Bergomi model. *Quant. Finance*, 2020, **20**(9), 1457–1473.

Bayer, C., Harang, F.A. and Pigato, P., Log-modulated rough stochastic volatility models. *SIAM J. Financ. Math.*, 2021, **12**(3), 1257–1284.

Bayer, C., Horvath, B., Muguruza, A., Stemper, B. and Tomas, M., On deep calibration of (rough) stochastic volatility models. arXiv preprint arXiv:1908.08806, 2019,

Bennedsen, M., Lunde, A. and Pakkanen, M.S., Hybrid scheme for Brownian semistationary processes. *Finance Stoch.*, 2017, **21**(4), 931–965.

Bennedsen, M., Lunde, A. and Pakkanen, M.S., Decoupling the short- and long-term behavior of stochastic volatility. *J. Financ. Econ.*, 2021, nbaa049.

Berestycki, H., Busca, J. and Florent, I., Computing the implied volatility in stochastic volatility models. *Commun. Pure. Appl. Math.*, 2004, **57**(10), 1352–1373.

Camara, A., Krehbiel, T. and Lib, W., Expected returns, risk premia, and volatility surfaces implicit in option market prices. *J. Banking Finance*, 2011, **35**(1), 215–230.

El Euch, O., Fukasawa, M., Gatheral, J. and Rosenbaum, M., Short-term at-the-money asymptotics under stochastic volatility models. *SIAM J. Financ. Math.*, 2019, **10**(2), 491–511.

El Euch, O., Fukasawa, M. and Rosenbaum, M., The microstructural foundations of leverage effect and rough volatility. *Finance Stoch.*, 2018, **22**(2), 241–280.

El Euch, O. and Rosenbaum, M., The characteristic function of rough heston models. *Math. Finance*, 2019, **29**(1), 3–38.

Forde, M., Gerhold, S. and Smith, B., Small-time, large-time and $H \to 0$ asymptotics for the rough heston model. *Math. Finance*, 2020.

Forde, M., Jacquier, A. and Lee, R., The small-time smile and term structure of implied volatility under the Heston model. *SIAM J. Financ. Math.*, 2012, **3**(1), 690–708.
Forde, M. and Zhang, H., Asymptotics for rough stochastic volatility models. *SIAM J. Financ. Math.*, 2017, **8**(1), 114–145.
Friz, P.K., Gassiat, P. and Pigato, P., Precise asymptotics: Robust stochastic volatility models. *Ann. Appl. Probab.*, 2021, **31**(2), 896–940.
Friz, P.K., Gerhold, S. and Pinter, A., Option pricing in the moderate deviations regime. *Math. Finance*, 2017, **28**(3), 962–988.
Fukasawa, M., Asymptotic analysis for stochastic volatility: Martingale expansion. *Finance Stoch.*, 2011, **15**(4), 635–654.
Fukasawa, M., Short-time at-the-money skew and rough fractional volatility. *Quant. Finance*, 2017, **17**(2), 189–198.
Fukasawa, M., Volatility has to be rough. *Quant. Finance*, 2021, **21**(1), 1–8.
Fukasawa, M., Takabatake, T. and Westphal, R., Is volatility rough? arXiv preprint arXiv:1905.04852, 2019.
Gao, K. and Lee, R., Asymptotics of implied volatility to arbitrary order. *Finance Stoch.*, 2014, **18**(2), 349–392.
Gatheral, J., Jaisson, T. and Rosenbaum, M., Volatility is rough. *Quant. Finance*, 2018, **18**, 1–17.
Gelfand, I.M. and Fomin, S.V., *Calculus of Variations*, 2000 (Dover Publications: Englewood Cliffs, NJ).
Goudenège, L., Molent, A. and Zanette, A., Machine learning for pricing American options in high-dimensional Markovian and non-Markovian models. *Quant. Finance*, 2020, **20**(4), 573–591.
Gulisashvili, A., Gaussian stochastic volatility models: Scaling regimes, large deviations, and moment explosions. *Stoch. Proc. Appl.*, 2020, **130**(6), 3648–3686.
Gulisashvili, A., Time-inhomogeneous Gaussian stochastic volatility models: Large deviations and super roughness. arXiv preprint arXiv:2002.05143, 2020.
Guo, B., Han, Q. and Zhao, B., The Nelson-Siegel model of the term structure of option implied volatility and volatility components. *J. Futures Mark.*, 2014, **34**(8), 788–806.
Inahama, Y., Laplace approximation for rough differential equation driven by fractional Brownian motion. *Ann. Probab.*, 2013, **41**(1), 170–205.
Jacquier, A., Pakkanen, M.S. and Stone, H., Pathwise large deviations for the rough Bergomi model. *J. Appl. Probab.*, 2018, **55**(4), 1078–1092.
Jacquier, A. and Pannier, A., Large and moderate deviations for stochastic Volterra systems. arXiv preprint arXiv:2004.10571, 2020.
Janson, S., *Gaussian Hilbert Spaces*, Vol. 129, 1997 (Cambridge University Press).
Jost, C., A note on ergodic transformations of self-similar Volterra Gaussian processes. *Electron. Commun. Probab.*, 2007, **12**, 259–266.
Krylova, E., Nikkinen, J. and Vähämaa, S., Cross-dynamics of volatility term structures implied by foreign exchange options. *J. Econ. Bus.*, Sept. 2009, **61**(5), 355–375.
Lee, R.W., Implied volatility: Statics, dynamics, and probabilistic interpretation. In *Recent Advances in Applied Probability*, pp. 241–268, 2005 (Springer: New York).
Mandelbrot, B. and Van Ness, J.W., Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 1968, **10**, 422–437.
McCrickerd, R. and Pakkanen, M.S., Turbocharging Monte Carlo pricing for the rough bergomi model. *Quant. Finance*, 2018, **18**(11), 1877–1886.
Medvedev, A. and Scaillet, O., A simple calibration procedure of stochastic volatility models with jumps by short term asymptotics. Research Paper No. 93, September 2003, FAME – International Center for Financial Asset Management and Engineering, 2003. Available at SSRN 477441, 2003.
Medvedev, A. and Scaillet, O., Approximation and calibration of short-term implied volatilities under jump-diffusion stochastic volatility. *Rev. Financ. Studies*, 2007, **20**(2), 427–459.
Nualart, D., *The Malliavin Calculus and Related Topics*, Vol. 1995, 2006 (Springer).
Osajima, Y., General asymptotics of Wiener functionals and application to implied volatilities. In *Large Deviations and Asymptotic Methods in Finance*, pp. 137–173, 2015 (Springer).
Pham, H., Large deviations in finance. *Third SMAI European Summer School in Financial Mathematics*, 2010.
Vasquez, A., Equity volatility term structures and the cross section of option returns. *J. Financ. Quant. Anal.*, 2017, **52**(06), 2727–2754.

Appendix: Fractional Brownian motion

The fBM is a "rough" continuous-time Gaussian process in that, depending on a parameter $H \in (0,1)$, its trajectories are locally Hölder continuous of any order strictly less than H. Unlike classical BM, the increments of fBm are not independent if $H \neq 1/2$. The fBM was introduced for the first time by Mandelbrot and Van Ness in Mandelbrot and Van Ness (1968) as the following stochastic integral, for $t \geq 0$:

$$Z_t^H = c_H \left[\int_{-\infty}^{t} (t-s)^{H-1/2} \mathrm{d}Z_s - \int_{-\infty}^{0} (-s)^{H-1/2} \mathrm{d}Z_s \right],$$

where Z is a BM and $c_H = (\int_0^\infty [(1+s)^{1/2-H} - s^{1/2-H}]^2 \mathrm{d}s + \frac{1}{2H})^{1/2}$. Such process is Gaussian with covariance

$$E[Z_t^H Z_s^H] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}). \tag{A1}$$

It can also be represented as a Volterra integral on the interval $[0,t]$:

$$Z_t^H = \int_0^t K_H(s,t)\, \mathrm{d}B_s, \tag{A2}$$

with K_H as in Nualart (2006) or Forde and Zhang (2017, Section 3.1)). One can consider the following variant of fBM, known as *Riemann–Liouville process* (Mandelbrot and Van Ness 1968), introduced in 1953 by Lévy. This process is also represented as Volterra integral as

$$\widehat{B}_t^H = \int_0^t K(t,s)\, \mathrm{d}B_s, \tag{A3}$$

with a simpler kernel

$$K(t,s) = \sqrt{2H}(t-s)^{H-1/2}, \text{for} \quad H \in (0,1). \tag{A4}$$

It is still self-similar, but stationarity of increments does not hold. Moreover, the covariance structure is more complicated than (A1). It can be expressed using hypergeometric functions (see Bayer *et al.* 2019, Lemma 4.1). The K-functionals that we find in our expansion can be computed in this case as

$$\begin{aligned}
\langle K1, 1\rangle &= \frac{\sqrt{2H}}{(H+1/2)(H+3/2)} \\
\langle K^2 1, 1\rangle &= \frac{1}{2H+1} \\
\langle (K1)^2, 1\rangle = \langle (\overline{K}1)^2, 1\rangle &= \frac{H}{(H+1)(H+1/2)^2} \\
\langle K1, \overline{K}1\rangle &= \frac{2H}{(H+1/2)^2}\beta(H+3/2, H+3/2)
\end{aligned} \tag{A5}$$

where β is the beta function. In the case, $K \equiv 1$ the fBM driving the volatility is actually a BM and we are back to the classical setting of a diffusive Markovian volatility. In this case, our expansions can be compared e.g. to Medvedev and Scaillet (2003, 2007).