Asymptotics for rough stochastic volatility models

Martin Forde*   Hongzhong Zhang†

March 17, 2021

Abstract

Using the large deviation principle (LDP) for a re-scaled fractional Brownian motion B_t^H where the rate function is defined via the reproducing kernel Hilbert space, we compute small-time asymptotics for a correlated fractional stochastic volatility model of the form $dS_t = S_t\sigma(Y_t)(\bar{\rho}dW_t + \rho dB_t)$, $dY_t = dB_t^H$ where σ is α-Hölder continuous for some $\alpha \in (0,1]$ but need not be bounded; In particular, we show that $\log S_t/t^{\frac{1}{2}-H}$ satisfies the LDP as $t \to 0$ and the model has a well-defined implied volatility smile as $t \to 0$, when the log-moneyness $k(t) = xt^{\frac{1}{2}-H}$. Thus the smile steepens to infinity or flattens to zero depending on whether $H \in (0, \frac{1}{2})$ or $H \in (\frac{1}{2}, 1)$.[1]

Keywords: fractional stochastic volatility; fractional Brownian motion; large deviations; implied volatility asymptotics; rough paths.

1 Introduction

The last few years has seen renewed interest in stochastic volatility models driven by fractional Brownian motion or other self-similar Gaussian processes (see [6, 20, 24, 25, 26]). Recall that fractional Brownian motion B^H (fBM) is a centered self-similar Gaussian process with stationary increments, which depends on a parameter $H \in (0,1)$ called the Hurst index, and B^H is persistent (i.e. more likely to keep a trend than to break it) when $H > \frac{1}{2}$ and anti-persistent when $H < \frac{1}{2}$ (i.e. if B^H was increasing in the past, B^H is more likely to decrease in the future, and vice versa). An earlier application of fractional Brownian motion in finance can be seen in Comte&Renault[11], who introduced a long-memory mean reverting extension of the Hull-White stochastic volatility model, where the log volatility is an Ornstein-Uhlenbeck process but driven by a fractionally integrated Brownian motion process, to capture the (much-documented) effect of volatility persistence. Comte et al.[9] also introduced a long-memory extension of the Heston model, via fractional integration of the usual square root volatility process, which has the desirable feature that the autocovariance function of the volatility process has power decay in the large-time limit (as opposed to the usual exponential decay for the standard CIR process, which has short-memory).

Gatheral et al.[24] provide strong empirical justification for such models; in particular they argue that log-volatility in practice behaves essentially as fBM with Hurst exponent $H \approx 0.1$, at any reasonable time scale (see also Gatheral[22, 23]). In particular, Gatheral et al.[24] advocate a model where the volatility is the exponential of a fractional Ornstein-Uhlenbeck process with small mean-reversion parameter. Alós et al.[4] examine the short-time behaviour of the derivative of the implied volatility with respect to the current log

*Dept. Mathematics, King's College London, Strand, London, WC2R 2LS (Martin.Forde@kcl.ac.uk)

†Dept. IEOR, Columbia University, New York, NY 10027 (hz2244@columbia.edu)

[1]We would like to thank Elisa Alósa, Amir Dembo, Jin Feng, Masaaki Fukasawa, Claudia Klüppelberg, David Nualart, Cheng Ouyang, Chris Rogers, H.M.C. Stone, Mike Tehranchi and Srinivasa Varadhan for helpful discussions.

stock price, for a mean reverting fractional stochatic volatility model driven by a Riemann-Liouville process using Malliavin calculus techniques and an extention of the well known Hull-White decomposition formula to the non-correlated case. They show that this derivative is $O(t^{H-\frac{1}{2}})$ as $t \to 0$ when $H < \frac{1}{2}$ and tends to zero for $H > \frac{1}{2}$ (if the model has no jumps). Their result follows from a novel *anticipative* Ito formula applied to the process $\frac{1}{T-t}\int_t^T \sigma_s^2 ds$ which is clearly not $\mathcal{F}_t$-measurable (see also Alós et al.[3]). Alós&León[2, Theorems 11 and 12] give expressions for the first and second derivative of the implied volatility with respect to the log strike, in terms of a quantity $D_t^+\sigma_t$ defined in their hypothesis H2, which can be seen to depend on the the Malliavin derivative $D_r^W \sigma_u^2$, which is easily computed as $2\sigma(B_u^H)\sigma'(B_u^H)K_H(r,u)$ for our model in (10) where $K_H(s,t)$ is defined in (3). In Alós&León[2, Section 5], they compute this quantity explicitly for a conventional diffusion stochastic volatility model.

Fukusawa[19] derives a small-noise expansion for a general correlated stochastic volatility model driven by fBM, using Yoshida's martingale expansion theory and Edgeworth expansions. More recently, Fukusawa[20] computes a small-time asymptotic expansion for implied volatility for a local-stochastic volatility model driven by fractional Brownian motion, using the (lesser known) Muravlev representation of fractional Brownian motion to capture the effect of correlation, which is non-trivial issue for fractional models as we shall see in this article. In another recent article, Bayer et al.[6] analyze the rough Bergomi variance curve model, which is shown to fit SPX option prices significantly better than conventional Markovian stochastic volatility models, and with less parameters.

Mijatović&Tankov[32] introduce a new parametrization for close-to-the-money options where the log-moneyness $k(t) = \theta\sqrt{t\log(\frac{1}{t})}$ as $t \to 0$, and for exponential Lévy models in this regime, they prove the surprising result that the implied volatility smile converges to a non-degenerate limiting shape. For jumps of infinite variation, the asymptotic smile is a piecewise linear function with three pieces, which depends on three numbers - the constant diffusion coefficient σ and the positive and negative jump activities as measured by the Blumenthal-Getoor index. For jumps of finite variation, the asymptotic smile is constant and equal to σ. These ideas are developed further in Figueroa-López&Olafsson[16], who derive a second-order approximation for at-the-money option prices for a large class of exponential Lévy models, and also when the continuous Brownian component is replaced by an independent stochastic volatility process with leverage. We also mention Baudoin&Ouyang[5] who consider small-time asymptotics for the density of the solution of a rough differential equation driven by enhanced fBM, using Rough paths theory and Malliavin calculus. fBM for $H > \frac{1}{4}$ admits a (step-3) lift as a geometric rough path of order p for any $p > \frac{1}{H}$ and one can prove an LDP for the lifted fBM with small noise (see e.g. Friz&Victoir[18, Theorem 15.59] and Millet&Sanz-Sol[33]), and then in turn for an RDE driven by fBM, using the continuity of the Itô map established in Lyon's celebrated universal limit theorem (see Friz&Victoir[18, Proposition 19.14]). Baudoin&Ouyang[5] also make use of a joint LDP for the terminal level of the small-noise diffusion and the Malliavian covariance matrix associated with the diffusion.

More recently, Gulisashvili et al.[25] compute a sharp small-time density estimate for a model with volatility equal to the absolute value of general self-similar Gaussian process, which includes fBM as a special case. For out-of-the-money strikes, they express the asymptotics explicitly using the self-similarity parameter H of the volatility process, and its first Karhunen-Loéve eigenvalue at time 1, and the multiplicity of this eigenvalue. The Karhunen-Loéve decompsition is essentially an eigenfunction expansion for the path: $X_t = \sum_{n=1}^{\infty}\sqrt{\lambda_n}e_n(t)Z_n$, where Z_n is an i.i.d. sequence of standard normal variates, and λ_n, e_n are the respective eigenvalues and eigenfunctions of the covariance operator $\mathcal{K}f = \int_0^T f(s)Q(s,t)ds$ of the Gaussian process with covariance function $Q(s,t)$, which is a compact self-adjoint linear operator on $L^2[0,T]$. When X is cenrted, the integrated variance $\int_0^T X_s^2 ds$ can then be written simply as $\sum_{n=1}^{\infty}\lambda_n Z_n^2$. Unfortunately, fBM and OU processes driven by fBM do not fall in the class of Gaussian processes for which the Karhunen-Loéve expansion is known explicitly, but efficient numerical techniques exist to compute the eigenfunctions and eigenvalues in these cases.

In Sections 2 and 3 of this article, we give a brief overview of Gaussian processes and in particular fBM, and we recall the classical LDP for a re-scaled Gaussian process and the meaning and construction of the associated reproducing kernel Hilbert space (RKHS). In Section 4, we introduce the fractional stochastic volatility model $dS_t = S_t\sigma(Y_t)(\bar{\rho}dW_t + \rho dB_t), dY_t = dB_t^H$, and using the LDP for a re-scaled fBM, we show that $t^{H-\frac{1}{2}}X_t$ satisfies the LDP as $t \to 0$ with speed $\frac{1}{t^{2H}}$. As a corollary, we show that there is a non-trivial small-maturity implied volatility smile when the log-moneyness of the call option $k(t) = xt^{\frac{1}{2}-H}$ as $t \to 0$; hence the short-maturity smile steepens to infinity or flattens to zero depending on the sign of $H - \frac{1}{2}$. The Hurst exponent H affords us greater flexibility in fitting small-maturity smiles than the Mijatović&Tankov[32] parametrization which only gives a small-time smile for one particular $k(t)$ function, and their asymptotic smile has to be piecewise linear with at most three pieces).

2 Background on Gaussian processes

A zero-mean real-valued Gaussian process $(Z_t)_{t\geq 0}$ is a stochastic process such that on any finite subset $\{t_1, \dots, t_n\} \subset \mathbb{R}$, $(Z_{t_1}, \dots, Z_{t_n})$ has a multivariate normal distribution with mean zero. The law of a Gaussian process is entirely determined by the covariance function $K(s,t) = \mathbb{E}(Z_t Z_s)$ and Z induces a Gaussian probability measure μ on $(E, \mathcal{B}(E))$, where E denotes the Banach space $C_0[0,1]$ with the usual sup norm topology (see e.g. Carmona&Tehranchi[7, Section 3.1.1] for details). Let Z denote the restriction of $(Z_t)_{t\geq 0}$ to $t \in [0,1]$ and let $M(\theta)$ be the moment generating function of Z:

$$M(\theta) \;=\; \mathbb{E}(e^{\langle \theta, Z\rangle}) = e^{\frac{1}{2}Q(\theta)},$$

defined for $\theta \in E^*$, where $Q(\theta) = \int \langle \theta, x\rangle^2 \mu(dx) = \langle \theta, \rho\theta\rangle$ and the covariance functional $\rho : E^* \to E$ is a bounded linear operator given by $\rho\theta = \int_E \langle \theta, x\rangle \, x\mu(dx)$. Using Fubini's theorem, we can re-write $\rho\theta$ as [2] (see Carmona&Tehranchi[7, p.86]).

$$\begin{aligned}\rho\theta(t) &= \int_E \langle \theta, x\rangle \, x(t)\mu(dx) = \int_E \int_{[0,1]} x(u)\theta(du)\, x(t)\mu(dx) = \int_{[0,1]} [\int_E x(u)x(t)\,\mu(dx)]\,\theta(du) \\ &= \int_0^1 K(u,t)\theta(du).\end{aligned}$$

2.1 The reproducing kernel Hilbert space for a Gaussian measure

The reproducing kernel Hilbert space (RKHS) associated with the Gaussian measure μ is defined as the completion of the image $\rho(E^*) \subset E$, using the inner product

$$\langle \rho x^*, \rho y^*\rangle \;=\; \int_E x^*(x)y^*(x)\mu(dx)\,,$$

(see Carmona&Tehranchi[7] for further details, and Subsection 3.3 below for the structure of the RKHS for the specific case of fBM).

3 Fractional Brownian motion

Fractional Brownian motion (fBM) is a natural generalization of standard Brownian motion which preserves the properties of stationary increments, self-similarity and Gaussian finite-dimensional distributions, but it has a more complex dependence structure which exhibits long-range dependence when $H > \frac{1}{2}$. In this subsection, we recall the definition and summarize the basic properties of fBM.

A zero-mean Gaussian process B_t^H is called standard *fractional Brownian motion* (fBM) with Hurst parameter $H \in (0,1)$ if it has covariance function

$$R_H(s,t) \;=\; \mathbb{E}(B_t^H B_s^H) \,-\, \mathbb{E}(B_t^H)\mathbb{E}(B_s^H) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})\,. \qquad (1)$$

[2] Recall that $\langle \theta, f\rangle = \int_{[0,1]} f(u)\theta(du)$ for some signed measure $\theta(du)$, by the Riesz representation theorem.

In order to specify the distribution of a Gaussian process, it is enough to specify its mean and covariance function; therefore, for each H, the law of of B^H is uniquely determined by $R_H(s,t)$. However, this definition by itself does not guarantee the existence of fBM; to show that fBM exists, one needs to verify that the covariance function is non-negative definite.

We now recall some fundamental properties of fBM:

- fBM is continuous a.s. and H-self-similar (H-ss), i.e. for $a > 0$, $(B_{at})_{t\geq 0} \overset{(d)}{=} a^H (B_t)_{t\geq 0}$ where $\overset{(d)}{=}$ means both processes have the same finite-dimensional distributions. For $H \neq \frac{1}{2}$, B^H does not have independent increments; for $H = \frac{1}{2}$, B^H_t is the standard Brownian motion.
- From (1), we see that

$$\begin{aligned}\mathbb{E}((B^H_t - B^H_s)^2) &= \mathbb{E}((B^H)^2_t) + \mathbb{E}((B^H)^2_s) - 2\mathbb{E}(B^H_s B^H_t) \\ &= t^{2H} + s^{2H} - (t^{2H} + s^{2H} - |t-s|^{2H}) = |t-s|^{2H}\,,\end{aligned}$$

  so $B^H_t - B^H_s \sim N(0, |t-s|^{2H})$; thus B^H has stationary increments.
- If we set $X_n = B^H_n - B^H_{n-1}$; then X_n is a discrete-time Gaussian process with covariance function

$$\begin{aligned}\rho_n &= \mathbb{E}(X_{k+n}X_n) = \mathbb{E}((B^H_{k+n} - B^H_{k+n-1})(B^H_k - B^H_{k-1})) \\ &= R_H(k+n,k) + R_H(k+n-1,k-1) - R_H(k+n,k-1) - R_H(k+n-1,k) \\ &= \frac{1}{2}[(n+1)^{2H} + (n-1)^{2H} - 2n^{2H}] \sim H(2H-1)n^{2H-2}, \qquad (n \to \infty),\end{aligned}$$

  and thus (by convexity of the function $g(n) = n^{2H}$), we see that two increments off the form $B_k - B_{k-1}$ and $B_{k+n} - B_{k+n-1}$ are positively correlated if $H \in (\frac{1}{2}, 1)$ and negatively correlated if $H \in (0, \frac{1}{2})$. Thus B^H is *persistent* (i.e. it is more likely to keep a trend than to break it) when $H > \frac{1}{2}$ and anti-persistent when $H < \frac{1}{2}$ (i.e. if B^H was increasing in the past, it is more likely to decrease in the future, and vice versa).
- If $H \in (\frac{1}{2}, 1)$, we can show that $\sum_{n=1}^{\infty} \rho_n = \infty$ which means that the process exhibits long-range dependence, but if $H \in (0, \frac{1}{2})$ then $\sum_{n=1}^{\infty} \rho_n < \infty$.
- Using that $\mathbb{E}((B^H_t - B^H_s)^2) = (t-s)^{2H}$ we can show that sample paths of B^H are α-Hölder-continuous, for all $\alpha \in (0, H)$.
- fBM is the only self-similar Gaussian process with stationary increments (see e.g. Marquardt[31]), and for $H \neq \frac{1}{2}$, B^H_t is neither a Markov process nor a semimartingale (see e.g. Nualart[34]).

3.1 Construction of fractional Brownian motion

Various integral/moving average representations of fBM in terms of a standard Brownian motion have been devised over the years, which we now briefly review:

- Mandelbrot&VanNess[30] give the following moving-average stochastic integral representation of fBM for $t \geq 0$:

$$B^H_t = c_H[\int_{-\infty}^{t} (t-s)^{-\gamma} dB_s - \int_{-\infty}^{0} (-s)^{-\gamma} dB_s],$$

  where B is standard Brownian motion, $\gamma = \frac{1}{2} - H$ and $c_H = (\int_0^{\infty}[(1+s)^{\gamma} - s^{\gamma}]^2 ds + \frac{1}{2H})^{\frac{1}{2}}$,[3] and note that $B^H_0 = 0$.

[3] See Mandelbrot&VanNess[30, Corollary 3.4] for the choice of the normalizing factor c_H.

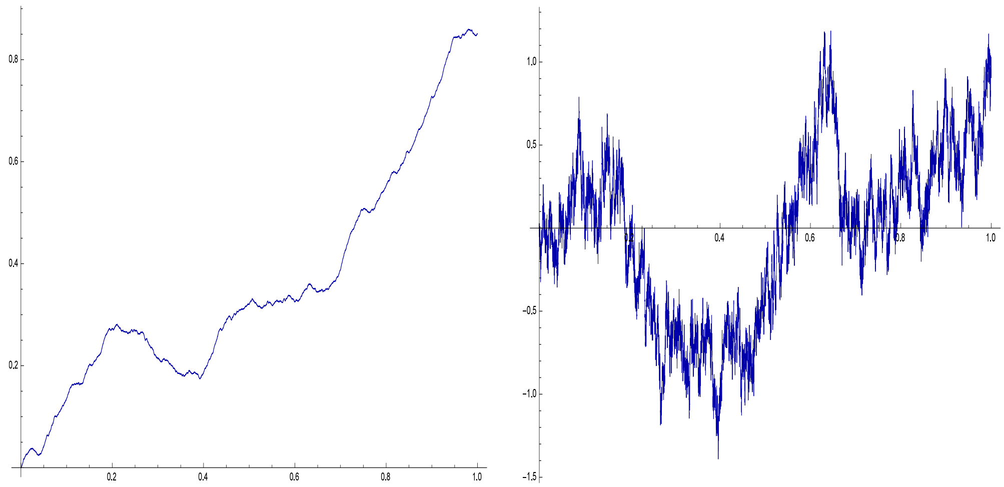

Figure 1: Here we have plotted a Monte Carlo simulation of fBM for $H = 0.9$ (left) and $H = 0.3$ (right), using the command: "data = RandomFunction[FractionalBrownianMotionProcess[H], {0, 1, 0.0001}] ; ListLinePlot[data]" in Mathematica.

- We also have the following Volterra-type representation of fBM on the interval $[0, t]$:

$$B_t^H \quad = \quad \int_0^t K_H(s,t) dB_s, \tag{2}$$

where

$$\begin{aligned} & K_H(s,t) \\ & \quad = \begin{cases} c_+ \, s^{\frac{1}{2}-H} \int_s^t (u-s)^{H-\frac{3}{2}} u^{H-\frac{1}{2}} du \,, & \text{if} \quad H \in (\frac{1}{2}, 1), \\ c_- \, [(\frac{t}{s})^{H-\frac{1}{2}} (t-s)^{H-\frac{1}{2}} - (H-\frac{1}{2}) s^{\frac{1}{2}-H} \int_s^t u^{H-\frac{3}{2}} (u-s)^{H-\frac{1}{2}} du], & \text{if} \quad H \in (0, \frac{1}{2}), \end{cases} \end{aligned} \tag{3}$$

and $c_+ = [\frac{H(2H-1)}{\beta(2-2H, H-\frac{1}{2})}]^{\frac{1}{2}}$, $c_- = [\frac{2H}{(1-2H)\beta(1-2H, H+\frac{1}{2})}]^{\frac{1}{2}}$ and $\beta(\cdot, \cdot)$ denotes the beta function (see Nualart[34, Eq. (5.8) and Proposition 5.1.3]).

- In 1953, Lévy introduced the following variant of fBM with a simpler kernel, also known as the *Riemann-Liouville process* (see Mandelbrot&VanNess[30] for related discussion)

$$\hat{B}_t^H \quad = \quad \frac{1}{\Gamma(H+\frac{1}{2})} \int_0^t (t-s)^{H-\frac{1}{2}} dB_s,$$

for $H \in (0, 1)$, which preserves the self-similarity feature of fBM (but not the stationarity of increments, see e.g. paragraph below Definition 1 in Comte&Renault[10] and Chen et al.[8]). This process is used in the fractional stochastic volatility model introduced in Comte&Renault[10, 11] and a similar representation is used for the fractional Heston model in Comte et al.[9].

3.2 The reproducing kernel Hilbert space for fBM

We first re-prove a well known result, which we later adapt for the proof of the main theorem below.

Lemma 3.1 *(see also [12, 33]). The reproducing kernel Hilbert space of fBM is* $\mathcal{H}_H = \mathbf{K}_H L^2[0, 1]$ *with scalar product given by*

$$\langle f, g \rangle_{\mathcal{H}_H} \quad = \quad \langle \dot{h}_1, \dot{h}_2 \rangle_{L^2[0,1]} \,,$$

for $f = \mathbf{K}_H \dot{h}_1, g = \mathbf{K}_H \dot{h}_2$*, where the operator* $\mathbf{K}_H$ *is defined by* $(\mathbf{K}_H f)(t) = \int_0^t K_H(s,t)f(s)ds$ *for* $t \in [0,1]$*.*

Proof. See Appendix A. ■

3.3 The small-noise large deviation principle for Gaussian processes

A classical result for any centered $\mathbb{R}^n$-valued Gaussian process Z states that $\sqrt{\varepsilon}Z$ satisfies the LDP on $C_0([0,1],\mathbb{R}^n)$ in the uniform topology as $\varepsilon \to 0$ with speed $\frac{1}{\varepsilon}$ and good rate function given by

$$\Lambda(f) \quad = \quad \begin{cases} \frac{1}{2}\|f\|^2_{\mathcal{H}}, & f \in \mathcal{H}, \\ \infty, & \text{otherwise}, \end{cases} \tag{4}$$

where $\mathcal{H}$ is the RKHS associated with the process (see e.g. Millet&Sanz-Sol[33, p.3] or Deuschel&Stroock[14, Theorem 3.4.12] for an elegant proof). fBM is a particular type of centered Gaussian process so $\sqrt{\varepsilon}B^H$ satisfies the LDP with good rate function $\Lambda(f)$, and in this case we know the structure of the RKHS from the Lemma in the previous subsection. From here on, we will use Λ_H in place Λ to signify that we are only working with fBM.

4 Small-time asymptotics for a fractional stochastic volatility model

We work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $(\mathcal{F}_t)_{t\geq 0}$ throughout, supporting two independent standard Brownian motions and satisfying the usual conditions.

4.1 The fBM model - the uncorrelated case

We first consider the following stochastic volatility model for a stock price process S_t driven by fBM:

$$\begin{cases} dS_t \;= S_t\sigma(Y_t)dW_t\,, \\ dY_t \;=\; dB^H_t\,, \end{cases} \tag{5}$$

where σ is is continuous and W and B^H_t denote an independent Brownian and fractional Brownian motion respectively. We set $X_t = \log S_t$ and $X_0 = Y_0 = 0$ without loss of generality[4], and it will be convenient to introduce the corresponding small-noise process

$$\begin{cases} dX^\varepsilon_t \;=\; -\frac{1}{2}\varepsilon\sigma(Y^\varepsilon_t)^2 dt + \sqrt{\varepsilon}\sigma(Y^\varepsilon_t)dW_t, \\ dY^\varepsilon_t \;=\; \varepsilon^H dB^H_t, \end{cases} \tag{6}$$

with $X^\varepsilon_0 = 0, Y^\varepsilon_0 = 0$.

Theorem 4.1 $X_t/t^{\frac{1}{2}-H}$ *satisfies the LDP as* $t \to 0$ *with speed* $\frac{1}{t^{2H}}$ *and good rate function given by*

$$I(x) \quad = \quad \inf_{f\in C_0[0,1]}[\frac{x^2}{2F(f)} + \Lambda_H(f)] \tag{7}$$

where $F(f) = \int_0^1 \sigma(f(s))^2 ds$*, and* $I(\cdot)$ *attains its minimum value of zero at* $x = 0$*.*

[4]Because the law of $X_t - X_0$ is independent of X_0, and if we want $Y_0 \neq 0$ we can just adjust σ accordingly.

Proof. From Itô's formula, we know that $X_t = \log S_t$ satisfies $dX_t = -\frac{1}{2}\sigma(Y_t)^2 dt + \sigma(Y_t)dW_t$. Now let

$$d\tilde{X}^\varepsilon_t \quad = \quad \sqrt{\varepsilon}\sigma(Y^\varepsilon_t)dW_t\,,$$

with $\tilde{X}^\varepsilon_0 = 0$, i.e. the same SDE as in (6) with the same volatility process but without the drift term. Then we have

$$\begin{aligned} \tilde{X}^\varepsilon_1 \quad &\overset{\text{(law)}}{=} \quad \tilde{X}^1_\varepsilon \overset{\text{(law)}}{=} W_{\int_0^\varepsilon \sigma(B^H_s)^2 ds} = W_{\varepsilon\int_0^1 \sigma(B^H_{\varepsilon u})^2 du} \quad \overset{\text{(law)}}{=} \quad W_{\varepsilon\int_0^1 \sigma(\varepsilon^H B^H_u)^2 du} \\ &\overset{\text{(law)}}{=} \quad \sqrt{\varepsilon}\, W_{F(Y^\varepsilon)} \quad \overset{\text{(law)}}{=} \quad F(Y^\varepsilon)^{\frac{1}{2}}\sqrt{\varepsilon}W_1. \end{aligned} \tag{8}$$

From Subsection 3.3, we know that $\varepsilon^H B^H$ satisfies the LDP on $C_0[0,1]$ with speed $\frac{1}{\varepsilon^{2H}}$ and good rate function Λ_H. By the Gärtner-Ellis theorem, we also know that $\varepsilon^H W_1$ satisfies the LDP as $\varepsilon \to 0$ with speed $\frac{1}{\varepsilon^{2H}}$ and good rate function $\frac{1}{2}x^2$; thus (by independence) $(\varepsilon^H W_1, \varepsilon^H B^H)$ satisfies a joint LDP on $\mathbb{R}\times C_0[0,1]$ with speed $\frac{1}{\varepsilon^{2H}}$ and good rate function $\frac{1}{2}x^2 + \Lambda_H(f)$. From (8) we see that

$$\tilde{X}^\varepsilon_1 \quad \overset{\text{(law)}}{=} \quad \varepsilon^{\frac{1}{2}-H}F(Y^\varepsilon)^{\frac{1}{2}}\varepsilon^H W_1 = \varepsilon^{\frac{1}{2}-H}\varphi(\varepsilon^H W_1, Y^\varepsilon),$$

where $\varphi : \mathbb{R}\times C_0[0,1] \to \mathbb{R}$ is given by $\varphi(x,f) = xF(f)^{\frac{1}{2}}$.

F is continuous in the sup norm topology, so (from the contraction principle) $F(Y^\varepsilon)$ satisfies the LDP with speed ε^{-2H} and good rate function $I_F(y) = \inf_{f:F(f)=y}\Lambda_H(f)$ as $\varepsilon \to 0$, and hence (from Remark a) on page 8 in [13]) we know that F is also exponentially tight, so

$$\mathbb{P}(|\frac{X^\varepsilon_1}{\varepsilon^{\frac{1}{2}-H}} - \frac{\tilde{X}^\varepsilon_1}{\varepsilon^{\frac{1}{2}-H}}| > \delta) \quad = \quad \mathbb{P}(\frac{1}{2}\frac{1}{\varepsilon^{\frac{1}{2}-H}}\varepsilon F(Y^\varepsilon) > \delta) \quad \le \quad \mathbb{P}(F(Y^\varepsilon) > \frac{2\delta}{\varepsilon^{\frac{1}{2}+H}}) \quad \le \quad \mathbb{P}(F(Y^\varepsilon) > R)$$

for any $R > 0$ and ε sufficiently small. Hence (by exponential tightness) we see that

$$\limsup_{\varepsilon\to 0}\varepsilon^{2H}\log\mathbb{P}(|\frac{X^\varepsilon_1}{\varepsilon^{\frac{1}{2}-H}} - \frac{\tilde{X}^\varepsilon_1}{\varepsilon^{\frac{1}{2}-H}}| > \delta) \quad \le \quad -I_F(R)\,.$$

Then (by exponential tightness), $\lim_{R\to\infty} I_F(R) = \infty$, so we can letting $R \to \infty$ we see that $\lim_{\varepsilon\to 0}\varepsilon^{2H}\log\mathbb{P}(|\frac{X^\varepsilon_1}{\varepsilon^{\frac{1}{2}-H}} - \frac{\tilde{X}^\varepsilon_1}{\varepsilon^{\frac{1}{2}-H}}| > \delta) = -\infty$. Hence $X^\varepsilon_1/\varepsilon^{\frac{1}{2}-H}$ and $\tilde{X}^\varepsilon_1/\varepsilon^{\frac{1}{2}-H}$ are *exponentially equivalent* in the sense of Dembo&Zeitouni[13, Definition 4.2.10]. Thus (by Dembo&Zeitouni[13, Theorem 4.2.13]), $X^\varepsilon_1/\varepsilon^{\frac{1}{2}-H}$ also satisfies the LDP with speed $\frac{1}{\varepsilon^{2H}}$ and rate function $I(x)$. But $X^\varepsilon_1 \overset{\text{(law)}}{=} X_\varepsilon$ and thus $X_\varepsilon/\varepsilon^{\frac{1}{2}-H}$ also satisfies the LDP with rate $I(x)$, and $I(x)$ simplifies to the expression given in the statement of the theorem (where we now also replace ε with t). The fact that $I(x) = 0$ follows from setting $x = 0$ and $f = 0$ and using that $\Lambda_H(0) = 0$. ■

Remark 4.1 Note that the proof does not use the stationarity of fBM anywhere, so we can actually replace Y with any Gaussian H-self-similar process, as in Gulisashvili et al.[25, 26], e.g. the Riemann-Liouville process $Y_t = \sqrt{2H}\int_0^t (t-s)^{H-\frac{1}{2}}dW_s$.

4.2 The correlated case

We now add correlation to the fractional stochastic volatility model in (5) and assume that $S_t = e^{X_t}$ evolves as

$$\begin{cases} dS_t = S_t\sigma(Y_t)(\bar{\rho}dW_t + \rho dB_t)\,, \\ dY_t \;=\; dB^H_t\,, \end{cases} \tag{9}$$

for $\rho \in (-1,1)$ with $\bar{\rho} = \sqrt{1-\rho^2}$, and σ is α-Hölder continuous. Again it will be convenient to introduce the small-noise process

$$\begin{cases} dX^\varepsilon_t \;=\; -\frac{1}{2}\varepsilon\sigma(Y^\varepsilon_t)^2 dt + \sqrt{\varepsilon}\,\sigma(Y^\varepsilon_t)[\bar{\rho}dW_t + \rho dB_t], \\ dY^\varepsilon_t \;=\; \varepsilon^H dB^H_t\,, \end{cases} \tag{10}$$

with $X^\varepsilon_0 = 0, Y^\varepsilon_0 = 0$.

Theorem 4.2 *$t^{H-\frac{1}{2}}X_t$ satisfies the LDP as $t \to 0$ with speed $\frac{1}{t^{2H}}$ and good rate function given by*

$$I(x) \quad = \quad \inf_{f \in H_1} \Big[\frac{(x-\rho G(f))^2}{2\bar{\rho}^2 F(\mathbf{K}_H f')} + \frac{1}{2}\|f\|^2_{H_1} \Big] \leq \frac{x^2}{2\bar{\rho}^2 \sigma(0)^2} \tag{11}$$

where $F(f) = \int_0^1 \sigma((\mathbf{K}_H f')(s))^2 ds$, $G(f) = \int_0^1 \sigma((\mathbf{K}_H f')(s)) f'(s) ds$ and $H_1 = \{\int_0^{\cdot} \dot{h}(s)ds, \dot{h} \in L^2[0,1]\}$ is the usual Cameron-Martin space for Brownian motion with the Hilbert structure $\langle f, g \rangle_{H_1} = \langle f'(s) g'(s) \rangle_{L^2[0,1]}$. $I(x)$ attains its minimum value of zero at $x = \rho G(0) = 0$.

Proof. See Appendix B. ■

Remark 4.2 Note that all the Theorems above (and we suspect most or all of the other published/unpublished results on fractional stochastic volalility models) are no longer true if we condition on the history of B^H at finite or infinitely many points in $[-\tau, 0]$ for some $\tau > 0$ fixed (for us the problem is that a conditioned fBM is no longer self-similar, which is what is needed in the proof of Theorem 4.1 to translate small noise asymptotics into small-time asymptotics). This is a non-trivial and important issue, which will hopefully be addressed in future work. On this theme, we also recall the prediction formula for fBM of Nuzman&Poor[35]: $\mathbb{E}(B^H_{t+\Delta}|\mathcal{F}_t) = \frac{1}{\pi}\cos(H\pi)\Delta^{H+\frac{1}{2}} \int_{-\infty}^t \frac{B^H_s}{(t-s+\Delta)(t-s)^{H+\frac{1}{2}}} ds$, but what we really want is the conditional covariance structure of fBM.

Corollary 4.3 *The rate function $I(x)$ in* (11) *is continuous.*

Proof. Let $\mathcal{I}(x,f) = \frac{(x-\rho G(f))^2}{2\bar{\rho}^2 F(\mathbf{K}_H f')} + \frac{1}{2}\|f\|^2_{H_1}$. Then $\mathcal{I}(\cdot, f)$ is continuous (and hence USC), and $I(x) = \inf_f \mathcal{I}(x,f)$. The pointwise supremum of a family of LSC functions is LSC (see e.g. Aliprantis&Border[1, Lemma 2.41]), hence the pointwise infimum of a family of USC functions is USC, so $I(x)$ is USC. But $I(x)$ is also a rate function, hence I is also LSC. ■

Corollary 4.4 *Let $\gamma = \frac{1}{2} - H$ as before. Then using the continuity of $I(x)$, we have the following small-time behaviour for digital put/call options*

$$\begin{aligned} \lim_{t \to 0} t^{2H} \log \mathbb{P}(X_t > x t^{\gamma}) &= -\Lambda^*(x)\,, \qquad && (x > 0)\,, \\ \lim_{t \to 0} t^{2H} \log \mathbb{P}(X_t < x t^{\gamma}) &= -\Lambda^*(x)\,, \qquad && (x < 0)\,, \end{aligned}$$

where $\Lambda^*(x) = \inf_{y > x} I(y)$ if $x \geq 0$ and $\Lambda^*(x) = \inf_{y<x} I(y)$ if $x \leq 0$.

4.3 The martingale property and asymptotics for call options and implied volatility

Lemma 4.5 *Assume that $\sigma : \mathbb{R} \mapsto (0, \infty)$ is α-Hölder continuous for some $\alpha \in (0,1]$, then there exists some $A_1 > 0$, such that*

$$\sigma(y)^2 \quad \leq \quad A_1(1 + |y|^2), \quad \forall y \in \mathbb{R}. \tag{12}$$

Proof. Let $L > 0$ be the Hölder constant of σ, we have

$$|\sigma(y) - \sigma(0)| \leq L|y|^{\alpha} \Rightarrow 0 < \sigma(y) < \sigma(0) + L|y|^{\alpha} \leq \sigma(0) + L 1_{|y|<1} + L|y| 1_{|y| \geq 1} \leq \sigma(0) + L + L|y|.$$

Hence, $\sigma(y)^2 \leq (\sigma(0) + L + L|y|)^2 \leq (2[(\sigma(0)+L) \vee L|y|])^2 \leq 4(\sigma(0)+L)^2 + 4L^2|y|^2$. It follows that (12) holds for $A_1 = 4(\sigma(0)+L)^2$. ■

We now prove that the stock price process in (9) is a martingale:

Proposition 4.6 *If σ satisfies* (12)*, then S_t is a martingale.*

Proof. For any $s > 0$, $(B^H_s)^2/s^{2H}$ follows a chi-squared distribution with degree of freedom 1, so

$$\mathbb{E}(e^{C(B^H_s)^2}) = \frac{1}{\sqrt{1-2s^{2H}C}}, \quad \forall C < \frac{1}{2s^{2H}}. \tag{13}$$

Hence, if $\varepsilon, C > 0$ and $s \geq 0$ satisfy

$$\varepsilon C(s+\varepsilon)^{2H} \leq \frac{1}{4}, \tag{14}$$

then by Jensen's inequality and Fubini's theorem, we have that

$$\mathbb{E}(e^{\int_s^{s+\varepsilon} C|B^H_u|^2 du}) = \mathbb{E}(e^{\frac{1}{\varepsilon}\int_s^{s+\varepsilon} \varepsilon C|B^H_u|^2 du}) \leq \frac{1}{\varepsilon}\int_s^{s+\varepsilon} \mathbb{E}(e^{\varepsilon C|B^H_u|^2})du = \frac{1}{\varepsilon}\int_s^{s+\varepsilon} \frac{du}{\sqrt{1-2u^{2H}\varepsilon C}}.$$

Since for all $u \in (s, s+\varepsilon]$, $1-2u^{2H}\varepsilon C \geq 1-2(s+\varepsilon)^{2H}\varepsilon C \geq 1-\frac{1}{2} = \frac{1}{2}$, we know that

$$\mathbb{E}(e^{\int_s^{s+\varepsilon} C|B^H_u|^2 du}) \leq \frac{1}{\varepsilon}\int_s^{s+\varepsilon} \sqrt{2}du = \sqrt{2}. \tag{15}$$

We use the above estimate to prove that $S_t = S_0 \exp(\int_0^t \sigma(B^H_s)(\bar{\rho}dW_s + \rho dB_s) - \frac{1}{2}\int_0^t \sigma(B^H_s)^2 ds)$ is a martingale for all $t > 0$. To this end, we define $s_0 = 0$ and ε_0 be such that $\varepsilon_0^{2H+1}A_1 = \frac{1}{4}$, and for $n \geq 1$, define $s_n = s_{n-1} + \varepsilon_{n-1}$ and $\varepsilon_n > 0$ be such that $\varepsilon_n(s_n+\varepsilon_n)^{2H}A_1 = \varepsilon_n s_{n+1}^{2H}A_1 = \frac{1}{4}$. We claim that s_n increases to ∞ as n tends to ∞; indeed, if we assume to the contrary that $\lim_{n\to\infty} s_n = M < \infty$, then $\varepsilon_n \geq 1/(4A_1(\sup_{n\geq 1} s_n)^{2H}) = 1/(4A_1M^{2H}) > 0$, which is a contradiction.

From (15) and (12), we know that for all $n \geq 1$,

$$\begin{aligned}\mathbb{E}(e^{\frac{1}{2}\int_{s_{n-1}}^{s_n} \sigma(B^H_s)^2 ds}) &\leq \mathbb{E}(e^{\frac{1}{2}\int_{s_{n-1}}^{s_n} A_1(1+|B^H_s|^2)ds}) = e^{\frac{1}{2}A_1\varepsilon_{n-1}}\mathbb{E}(e^{\frac{1}{2}A_t\int_{s_{n-1}}^{s_n} |B^H_s|^2 ds}) \\ &\leq \sqrt{2}e^{\frac{1}{2}A_1\varepsilon_{n-1}} < \infty.\end{aligned}$$

Hence by Karatzas&Shreve[27, Corollary 5.5.14], S_t is a true martingale. ■

4.4 Asymptotics for call options and implied volatility

Corollary 4.7 *Consider the model in* (9) *and assume σ satisfies* (12)*. Then we have the following small-time behaviour for out-of-the-money put/call options on $S_t = e^{X_t}$ with $S_0 = 1$:*

$$-\lim_{t\to 0} t^{2H}\log\mathbb{E}(S_t - e^{xt^\gamma})^+ = \Lambda^*(x), \qquad (x \geq 0), \tag{16}$$

$$-\lim_{t\to 0} t^{2H}\log\mathbb{E}(e^{xt^\gamma} - S_t)^+ = \Lambda^*(x), \qquad (x \leq 0), \tag{17}$$

where $x = \log K$ is the log-moneyness.

Proof. See Appendix C. ■

Remark 4.3 Note that this is a small-time, small log-moneyness parametrization if $H \in (0, \frac{1}{2})$, and a small-time, large log-moneyness parametrization for $H \in (\frac{1}{2}, 1)$. Put differently, we expect the implied volatility smile to steepen to an infinite V-shape as the maturity $t \to 0$ if $H \in (0, \frac{1}{2})$ (similar to jump models) and to flatten when $H \in (\frac{1}{2}, 1)$. The empirical findings in Gatheral et al.[24] report that $H = 0.1$ is realistic, but historically H is usually found or chosen to be greater than $\frac{1}{2}$ to capture long-memory dependence. Recall that $\gamma = \frac{1}{2} - H$, hence the limit $H \to 0$ here is consistent with the parameterization used in FX option markets where the log moneyness scales as $x\sqrt{t}$ as $t \to 0$.

Corollary 4.8 *Let $\hat{\sigma}(x,t)$ denote the implied volatility at log-moneyness x and maturity t. Then for the model in* (9)*, we have*

$$\hat{\sigma}_0(x) = \lim_{t\to 0}\hat{\sigma}(xt^\gamma,t) \quad = \quad \frac{|x|}{\sqrt{2\Lambda^*(x)}} \tag{18}$$

where $\gamma = \frac{1}{2}-H$ as before.

Proof. Setting $k = xt^\gamma = xt^{\frac{1}{2}-H}$, we know that the log absolute call price $L = L(k) = |\log \mathbb{E}(S_t - e^{xt^\gamma})^+| \sim \frac{\Lambda^*(x)}{t^{2H}}$ as $t \to 0$. Then by in Gao&Lee[21, Corollary 7.1], we have the following small-time behaviour for the dimensionless implied volatility squared $V^2 = \hat{\sigma}(x,t)^2 t$:

$$\begin{aligned} V &= G[k, L - \frac{3}{2}\log L + \log(\frac{k}{4\sqrt{\pi}})] \,+\, o(\frac{k}{L^{\frac{3}{2}}}) \\ &= \sqrt{2}[\frac{\Lambda^*(x)}{t^{2H}} + t^{\frac{1}{2}-H}x - \frac{3}{2}\log(\frac{\Lambda^*(x)}{t^{2H}}) + \log\frac{t^{\frac{1}{2}-H}x}{4\sqrt{\pi}}]^{\frac{1}{2}} \\ &\quad -\sqrt{2}[\frac{\Lambda^*(x)}{t^{2H}} - \frac{3}{2}\log(\frac{\Lambda^*(x)}{t^{2H}}) + \log\frac{t^{\frac{1}{2}-H}x}{4\sqrt{\pi}}]^{\frac{1}{2}} \,+\, o(\frac{k}{L^{\frac{3}{2}}}) \\ &= \frac{\sqrt{2\Lambda^*(x)}}{t^H}[\frac{1}{2}\frac{t^{2H}}{\Lambda^*(x)}t^{\frac{1}{2}-H}x \,+\, o(\frac{xt^{\frac{1}{2}-H}}{(\frac{\Lambda^*(x)}{t^{2H}})^{\frac{3}{2}}}) = \frac{x\sqrt{t}}{\sqrt{2\Lambda^*(x)}}\,[1+o(1)], \end{aligned}$$

where $G(\cdot,\cdot)$ is defined in Gao&Lee[21], and the result follows by dividing by $\sqrt{t}$. ■

4.5 Numerical implementation and computing the most likely path

We can use the Ritz method described in Gelfand&Fomin[17, Section 40] to provide an approximate numerical solution to the rate function $I(x)$ in Theorem 4.2. More specifically, using that $(\cos(2\pi ns), \sin(2\pi ns))_{n=0}^\infty$ is an orthogonal basis for $L^2[0,1]$, we consider f functions such that $f(0) = 0$ and $f'(s) = a_0 + \sum_{n=1}^N [a_n\cos(2\pi ns) + b_n\sin(2\pi ns)]$ for some finite N, and we then minimize $\frac{(x-\eta_1 f(1))^2}{2\int_0^t \sigma(y(\int_0^t K_H(s,t)f'(s)ds))^2 dt} \,+\, \frac{1}{2}\|f\|^2_{H_1}$ over the N Fourier coefficients. The optimal f then gives the "most likely path" for B^H given that the log stock price $X_t = x$.

In the following tables (see also Figure 2 and Figure 3) we calculate the rate function $I(x)$ in (11) and the asymptotic implied volatility $\hat{\sigma}_0(x)$ for the uncorrelated model in (5) and the correlated model in (9) respectively, using the Ritz method with the NMinimize command in Mathematica with $N = 4$ and $\sigma(y) = 0.1 + .05\tanh(y)$, $H = 0.25$ (Mathematica code available on request).

Table 1: Implied volatility in the uncorrelated model (5).

x	$\hat{\sigma}_0(x)$ Ritz method	$\hat{\sigma}_t(x)$ Monte Carlo $t = .005$
0.001	10.0000%	10.0179%
0.02	10.0183%	10.0364%
0.04	10.0716%	10.0832%
0.06	10.1551%	10.1594%
0.08	10.2625%	10.2589%
0.10	10.3866%	10.3778%

Table 2: Implied volatility in the correlated model (9) with $\rho = -0.1$.

x	$\hat{\sigma}_0(x)$ Ritz method	$\hat{\sigma}_t(x)$ Monte Carlo $t = .005$
-0.06	10.3064%	10.3127%
-0.04	10.1769%	10.1674%
-0.02	10.0724%	10.0774%
0.001	9.99731%	10.0067%
0.02	9.96412%	9.97780%
0.04	9.96607%	9.97724%
0.06	10.0036%	10.0115%
0.08	10.0714%	10.0740%

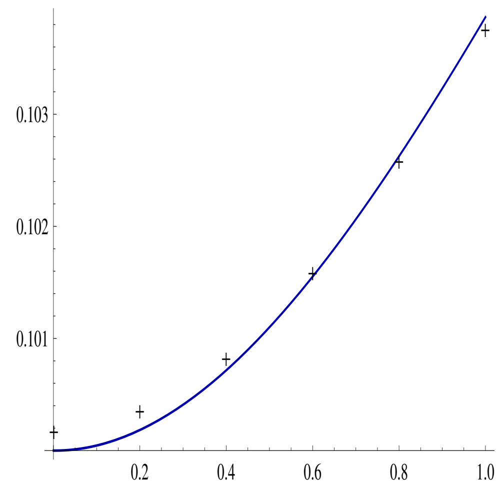

Figure 2: Here we have plotted the right half of the (symmetric) small-maturity implied volatility smile for the model in (9) for $\rho = 0$, $\sigma(y) = 1 + .05\tanh(y)$, $H = 0.25$ and $t = .005$. verses the values obtained by Monte Carlo using the well known Willard[37] conditioning method with 500,000 simulations and 100 time steps in Mathematica. We use a discretization of the Volterra formula in (2) to generate the fBM.

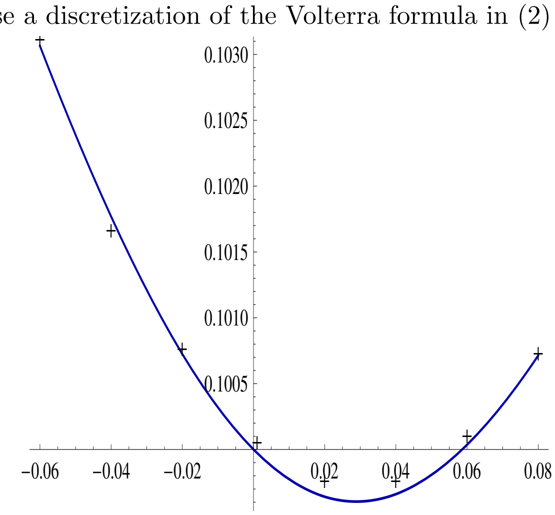

Figure 3: Here we have plotted the small-maturity implied volatility smile for the model in (9) for $\rho = -0.1$, $\sigma(y) = 1 + .05\tanh(y)$, $H = 0.25$ and $t = .002$. verses the values obtained by Monte Carlo using the Willard[37] conditioning method with 500,000 simulations and 100 time steps in Mathematica.

References

[1] C.D. Aliprantis and K. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, 3rd ed., 2007.

[2] E. Alós and J.A. León, *On the second derivative of the at-the-money implied volatility in stochastic volatility models*. Available at https://ideas.repec.org/p/upf/upfgen/1458.html, 2015.

[3] E. Alós, J.A. León, and J. Vives, *A generalization of the Hull and White formula with applications to option pricing approximation*, Finance Stoch., 10 (2006), pp. 353–365.

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

[5] F. Baudoin and C. Ouyang, *On small time asymptotics for rough differential equations driven by fractional Brownian motions*, in Large Deviations and Asymptotic Methods in Finance, Springer, 2015, pp. 413–438.

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

[7] R. Carmona and M. Tehranchi, *Interest Rate Models: an Infinite Dimensional Stochastic Analysis Perspective*, Springer, 2006.

[8] X. Chen, W.V. Li, J. Rosiński, and Q.-M. Shao, *Large deviations for local times and intersection local times of fractional Brownian motions and Riemann-Liouville processes*, Ann. Probab., 39 (2011), pp. 729–778.

[9] F. Comte, L. Coutin, and E. Renault, *Affine fractional stochastic volatility models*, Ann. Finance, 8 (2012), pp. 337–378.

[10] F. Comte and E. Renault, *Long memory continuous time models*, J. Econometrics, 73 (1996), pp. 101–149.

[11] ——, *Long memory in continuous-time stochastic volatility models*, Math. Finance, 8 (1998), pp. 291–323.

[12] L. Decreusefond and A.S. Ustunel, *Stochastic analysis of the fractional Brownian motion*, Potential Anal., 10 (1999), pp. 177–214.

[13] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Bartlet publishers, Boston, 1998.

[14] J.D. Deuschel and D.W. Stroock, *Large deviations*, Academic Press, New York, 1989.

[15] M.D. Donsker and S.R.S. Varadhan, *Asymptotic evaluation of markov process expectations for large time, III*, Comm. Pure Appl. Math., 29 (1976), pp. 389–461.

[16] J.E. Figueroa-López and S. Olafsson, *Short-time expansions for close-to-the-money options under a Lévy jump model with stochastic volatility*, Finance Stoch., 20 (2016), pp. 219–265.

[17] S.V. Fomin and I.M. Gelfand, *Calculus of Variations*, Dover Publication, 2000.

[18] P. Friz and N. Victoir, *Multidimensional Dimensional Processes Seen as Rough Paths*, Cambridge University Press, 2010.

[19] M. Fukasawa, *Asymptotic analysis for stochastic volatility: martingale expansion*, Finance Stoch., 15 (2011), pp. 635–654.

[20] ——, *Short-time at-the-money skew and rough fractional volatility*. Forthcoming Quant. Finance. Available at http://arxiv.org/abs/1501.06980., 2016.

[21] K. Gao and R. Lee, *Asymptotics of implied volatility to arbitrary order*, Finance Stoch., 18 (2014), pp. 349–392.

[22] J. Gatheral, *Fractional volatility models.* BBQ seminar, 2014.

[23] ———, *Volatility is rough, part 2: Pricing.* Workshop on Stochastic and Quant. Finance, Imperial College London, 2014.

[24] J. Gatheral, T. Jaisson, and M. Rosenbaum, *Volatility is rough.* Available at http://arxiv.org/abs/1410.3394, 2014.

[25] A. Gulisashvili, F. Viens, and X. Zhang, *Small-time asymptotics for Gaussian self-similar stochastic volatility models.* Available at http://arxiv.org/abs/1505.05256, 2015.

[26] ———, *Extreme-strike asymptotics for general Gaussian stochastic volatility models.* Available at http://arxiv.org/abs/1502.05442, 2015.

[27] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, 1991.

[28] M.A. Lifshits, *Gaussian Random Functions*, Springer, 1995.

[29] K. Majewski, *Large deviations for multi-dimensional reflected fractional Brownian motion*, Stochastics and Stochastic Reports, 75 (2003), pp. 233–257.

[30] B. Mandelbrot and J.W. Van Ness, *Fractional Brownian motions, fractional noises and applications*, SIAM Rev., 10 (1968), pp. 422–437.

[31] T. Marquardt, *Fractional Lévy processes with an application to long memory moving average processes*, Bernoulli, 12 (2006), pp. 1099–1126.

[32] A. Mijatović and P. Tankov, *A new look at short-term implied volatility in asset price models with jumps*, Math. Finance, 26 (2016), pp. 149–183.

[33] A. Millet and M. Sanz-Sol, *Large deviations for rough paths of the fractional Brownian motion*, Ann. Inst. Henri Poincaré Probab. Stat., 42 (2006), pp. 245–271.

[34] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer, Berlin, 2006.

[35] C.J. Nuzman and H. Vincent Poor, *Linear estimation of self-similar processes via Lamperti's transformation*, J. Appl. Probab., 37 (2000), pp. 429–452.

[36] M. Veraar, *The stochastic Fubini theorem revisited*, Stochastics, 84 (2012), pp. 543–551.

[37] G. Willard, *Calculating prices and sensitivities for path-independent derivatives securities in multifactor models*, J. Derivatives, 5 (1997), pp. 45–61.

A Proof of Lemma 3.1

We first note that $\mathbf{K}_H$ is a bijection from $L^2[0,1]$ into $\mathcal{H}_H$.[5] We now construct the adjoint of $\mathbf{K}^H$ (see also Deuschel&Stroock[14, Section 3.4] where $S = \mathbf{K}^H$ in their notation), loosely following the arguments in Decreusefond&Ustunel[12, Theorem 3.3] (which contains some minor errors). For $f \in H := L^2[0,1]$, $\mathbf{K}_H f \in \mathcal{H}_H \subset E$ and for $\theta \in E^*$ let $Af = \langle \theta, \mathbf{K}_H f\rangle$. Then $A : L^2[0,1] \to \mathbb{R}$ is a continuous linear functional,

[5] The kernel is continuous and positive, so it induces an injection. Suppose this is not the case; then a nonzero cádlág function f is mapped to 0 (because $\mathbf{K}_H$ is linear). Then pick the first interval where the sign of f is either $+$ or $-$, without loss of generality, say $f(t) > 0$ for all $t \in (0,u)$, then the image of f will also have positive value for argument in $(0,u)$, due to positivity of the kernel. This is a contradiction.

i.e. an element of $H^* \simeq H$, so we have a continuous linear map $\mathbf{K}_H^* : E^* \to H$. More explicitly, using Fubini's theorem, we have

$$\langle \theta, \mathbf{K}_H f \rangle = \int_0^1 [\int_0^t K_H(s,t) f(s) ds]\, \theta(dt) = \int_0^1 [\int_s^1 K_H(s,t)\theta(dt)]\, f(s) ds,$$

so $(\mathbf{K}_H^*\theta)(s) = \int_s^1 K_H(s,t)\theta(dt)$. In order to verify that $\mathcal{H}_H$ is the RKHS, we have to verify that $\mathcal{H}_H$ is dense in E and that

$$\int e^{i\langle \theta, f\rangle} d\mu = e^{-\frac{1}{2}\|\mathbf{K}_H^*\theta\|^2},$$

where μ is the law of B^H on $C_0[0,1]$. We note that $\mathbf{K}_H t^r = c_{r,H}\, t^{r+H-\frac{1}{2}}$ for all $r > -\frac{1}{2}$ for some $c_{r,H} \neq 0$, so $\mathcal{H}_H$ contains all polynomials null at zero and this is dense in $C_0[0,1]$ (by the Stone-Weierstrauss theorem). Moreover, we also see that

$$\begin{aligned}
\int \langle \theta, f\rangle^2 d\mu = \mathbb{E}(\int_0^1 B_s^H \theta(ds) \int_0^1 B_t^H \theta(dt)) &= \mathbb{E}(\int_0^1 \int_0^1 B_t^H B_s^H \theta(ds)\theta(dt)) \\
&= \int_0^1 \int_0^1 R_H(s,t)\theta(ds)\theta(dt) \\
&= \int_0^1 \int_0^1 \int_0^{s\wedge t} K_H(r,s)K_H(r,t) dr\, \theta(ds)\theta(dt) \\
\text{(if we define } K_H(r,s) \text{ to be zero if } r > s) &= \int_0^1 \int_0^1 \int_0^1 K_H(r,s)K_H(r,t) dr\, \theta(ds)\theta(dt) \\
&= \int_0^1 (\int_0^1 K_H(r,s)\theta(ds))(\int_0^1 K_H(r,t)\theta(dt)) dr \\
&= \int_0^1 (\int_r^1 K_H(r,s)\theta(ds))(\int_r^1 K_H(r,t)\theta(dt)) dr \\
&= \int_0^1 (\mathbf{K}_H^*\theta)(r)^2 dr = \|\mathbf{K}_H^*\theta\|^2_{L^2[0,1]} = \|\mathbf{R}_H\theta\|^2_{\mathcal{H}_H},
\end{aligned}$$

where $\mathbf{R}_H = \mathbf{K}_H \mathbf{K}_H^*$, and we have used that $R_H(s,t) = \int_0^{s\wedge t} K_H(r,s)K_H(r,t)dr$ in the third line (see e.g. Nualart[34, Eq. (5.9)]). This verifies that $\mathcal{H}_H$ is the RKHS for fBM.

B Proof of Theorem 4.2

Given that (B, B^H) is a Gaussian process, applying similar arguments to the proof of Lemma 3.1, we see that the RKHS for (B, B^H) is $\mathcal{H}_H^2 = \{(f,g) \in C_0([0,1],\mathbb{R}^2) : f(t) = \int_0^t \dot{h}(s)ds,\ g(t) = \int_0^t K_H(s,t)\dot{h}(s)ds, \dot{h} \in L^2[0,1]\}$. Using the general LDP for Gaussian processes in Subsection 3.3, we know that $\varepsilon^H(B^H, B)$ satisfies a joint LDP on $C_0([0,1],\mathbb{R}^2)$ as $\varepsilon \to 0$ with speed $\frac{1}{\varepsilon^{2H}}$ and rate function

$$I_H(f,g) = \begin{cases} \frac{1}{2}\int_0^1 \dot{h}(s)^2 ds, & \text{if } (f,g) \in \mathcal{H}_H^2, \\ +\infty, & \text{otherwise}. \end{cases} \tag{B-1}$$

From Itô's formula, we know that $X_t = \log S_t$ satisfies $dX_t = -\frac{1}{2}\sigma(Y_t)^2 dt + \sigma(Y_t)(\bar{\rho} dW_t + \rho dB_t)$. Now let $d\tilde{X}_t^\varepsilon = \sqrt{\varepsilon}\sigma(Y_t^\varepsilon)(\bar{\rho} dW_t + \rho dB_t), dY_t^\varepsilon = \varepsilon^H dB_t^H$, i.e. the same SDE as in (10) but without the drift term.

Then we have

$$
\begin{aligned}
\tilde{X}^1_\varepsilon \overset{\text{(law)}}{=} \bar{\rho}\, W_{\int_0^\varepsilon \sigma(B^H_s)^2 ds} + \rho \int_0^\varepsilon \sigma(B^H_s) dB_s \quad &= \quad \bar{\rho}\, W_{\varepsilon \int_0^1 \sigma(B^H_{\varepsilon u})^2 du} \,+\, \rho \int_0^1 \sigma(B^H_{\varepsilon u}) dB_{\varepsilon u} \\
&\overset{\text{(law)}}{=} \quad \bar{\rho}\, W_{\varepsilon \int_0^1 \sigma(\varepsilon^H B^H_u)^2 du} + \rho \int_0^1 \sigma(\varepsilon^H B^H_u) \sqrt{\varepsilon} dB_u \\
&\overset{\text{(law)}}{=} \quad \sqrt{\varepsilon}\,[\,\bar{\rho}\, W_{F(Y^\varepsilon)} + \rho \int_0^1 \sigma(Y^\varepsilon_u) dB_u] \\
&\overset{\text{(law)}}{=} \quad \sqrt{\varepsilon}\,[\,\bar{\rho} F(Y^\varepsilon)^{\frac{1}{2}}\, W_1 + \rho \int_0^1 \sigma(Y^\varepsilon_u) dB_u] \overset{\text{(law)}}{=} \tilde{X}^\varepsilon_1 \,. \qquad \text{(B-2)}
\end{aligned}
$$

$\varepsilon^H W_1$ satisfies the LDP as $\varepsilon \to 0$ with speed $\frac{1}{\varepsilon^{2H}}$ and rate function $\frac{1}{2}x^2$; thus by independence, $(\varepsilon^H W_1, \varepsilon^H B, \varepsilon^H B^H)$ satisfies a joint LDP on $\mathbb{R} \times C_0[0,1] \times C_0[0,1]$ with speed $\frac{1}{\varepsilon^{2H}}$ and rate function $\frac{1}{2}x^2 + I_H(f,g)$. Loosely following the arguments in Dembo&Zeitouni[13, Lemma 5.6.9] by "freezing" the coefficient $\sigma(\cdot)$ over small time intervals, we now define

$$
\begin{aligned}
Z^\varepsilon_t \quad &= \quad \varepsilon^H \int_0^t \sigma(\varepsilon^H B^H_s) dB_s = \varepsilon^H \int_0^t \sigma(\varepsilon^H \int_0^s K_H(u,s) dB_u) dB_s \,, \\
Z^{m,\varepsilon}_t \quad &= \quad \varepsilon^H \int_0^t \sigma(\varepsilon^H B^H_{\frac{[ms]}{m}}) dB_s = \varepsilon^H \int_0^t \sigma(\varepsilon^H \int_0^{\frac{[ms]}{m}} K_H(u, \frac{[ms]}{m}) dB_u) dB_s \,,
\end{aligned}
$$

Lemma B.1 *$Z^{\varepsilon,m}_t$ is an exponentially good approximation to Z^ε_t for all $t \in [0,1]$:*

$$
\lim_{m\to\infty} \lim_{\varepsilon \to 0} \varepsilon^{2H} \log \mathbb{P}(\sup_{t\in[0,1]} |Z^\varepsilon_t - Z^{\varepsilon,m}_t| > \delta) \quad = \quad -\infty \,.
$$

Proof. See Appendix B.1 below. ■

We can define the following mappings Φ, Φ^m from $\mathcal{H}^2_H$ to $C_0[0,1]$:

$$
\begin{aligned}
(\Phi(f,g))(t) \quad &= \quad \int_0^t \sigma(g(s)) f'(s) ds \\
(\Phi^m(f,g))(t) \quad &= \quad (\Phi^m(f,g))(t) = \int_0^t \sigma(g(\frac{[ms]}{m})) f'(s) ds \qquad \text{for} \quad t \in [0,1], \qquad \text{(B-3)}
\end{aligned}
$$

and recall that for $(f,g) \in \mathcal{H}^2_H$, $f = \mathbf{K}_{\frac{1}{2}} \dot{h}$ and $g = \mathbf{K}_H \dot{h}$ for some $\dot{h} \in L^2[0,1]$. Moreover, we can extend the domain of Φ and Φ^m to $C_0[0,1] \times C_0[0,1]$ by setting $\Phi(f,g) = 0$ for $f, g \notin \mathcal{H}^2_H$ and

$$
\Phi^m(f,g) \quad = \quad \sum_{j=0}^{[mt]-1} \sigma(g(\frac{j}{m}))(f(\frac{j+1}{m}) - f(\frac{j}{m})) \,+\, \sigma(g(\frac{[mt]}{m}))(f(t) - f(\frac{[mt]}{m}))
$$

and we note that Φ is measurable on this extended domain, and Φ^m is continuous on this extended domain and agrees with the original definition of Φ^m in (B-3) for $(f,g) \in \mathcal{H}^2_H$ (where we are using the sup norm topology for both arguments of Φ^m).

Lemma B.2

$$
\lim_{m\to\infty} \sup_{(f,g)\in\mathcal{H}^2_H : I_H(f,g) \le \alpha} \|\Phi((f,g)) - \Phi^m((f,g))\|_\infty = 0. \qquad \text{(B-4)}
$$

Proof. See Appendix B.2 below. ■

Returning now to (B-2), we see that

$$
\tilde{X}^\varepsilon_1 \quad \overset{\text{(law)}}{=} \quad \sqrt{\varepsilon}\,[\,\bar{\rho} F(Y^\varepsilon)^{\frac{1}{2}}\, W_1 + \rho \int_0^1 \sigma(Y^\varepsilon_u) dB_u] \quad = \quad \varepsilon^{\frac{1}{2}-H} [\,\bar{\rho} F(\varepsilon^H B^H)^{\frac{1}{2}}\, \varepsilon^H W_1 + \varepsilon^H \rho \int_0^1 \sigma(\varepsilon^H B^H_u) dB_u].
$$

We now define the functional $\varphi^m : \mathbb{R} \times C_0[0,1] \times C_0[0,1]$ as

$$\varphi^m(x,f,g) \quad = \quad \bar{\rho}F(g)^{\frac{1}{2}}x + \rho\Phi^m(f,g)(1),$$

and the "(1)" just means the function evaluated at the point $t = 1$, and again φ^m is continuous if we use the sup norm for the f and g arguments. Then we have (recall that $\rho \neq 0$)

$$\begin{aligned}
\mathbb{P}(|\frac{\tilde{X}_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}} - \varphi^m(\varepsilon^H W_1, \varepsilon^H B^1, \varepsilon^H B^H)| > \delta) &= \mathbb{P}(\rho(Z_1^\varepsilon - Z_1^{\varepsilon,m})| > \delta) \\
&\leq \mathbb{P}(|\rho||Z_1^\varepsilon - Z_1^{\varepsilon,m}| > \delta) \\
&\leq \mathbb{P}(|Z_1^\varepsilon - Z_1^{\varepsilon,m}| > \frac{1}{2}\delta/|\rho|)\,.
\end{aligned}$$

Combining this with Lemma B.1, we see that

$$\lim_{m\to\infty}\lim_{\varepsilon\to 0}\varepsilon^{2H}\log\mathbb{P}(|\frac{\tilde{X}_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}} - \varphi^m(\varepsilon^H W_1, \varepsilon^H B^1, \varepsilon^H B^H)| > \delta) = -\infty\,.$$

Thus $\varphi^m(\varepsilon^H W_1, \varepsilon^H B^1, \varepsilon^H B^H)$ is an exponentially good approximation to $\tilde{X}_1^\varepsilon/\varepsilon^{\frac{1}{2}-H}$. Moreover, we see that

$$\begin{aligned}
&\lim_{m\to\infty}\sup_{(z,f,g)\in\mathbb{R}\times\mathcal{H}_H^2:\frac{1}{2}z^2+I_H(f,g)\leq\alpha}|\rho(\Phi((f,g))(1) - \Phi^m((f,g))(1))| \\
\leq\ &\lim_{m\to\infty}\sup_{(f,g)\in\mathcal{H}_H^2:I_H(f,g)\leq\alpha}||\rho||\Phi((f,g))(1) - \Phi^m((f,g))(1)| \\
\leq\ &\lim_{m\to\infty}\sup_{(f,g)\in\mathcal{H}_H^2:I_H(f,g)\leq\alpha}|\rho||\Phi((f,g))(1) - \Phi^m((f,g))(1)| = 0,
\end{aligned}$$

where the final equality follows from Lemma B.2. Thus by the extended contraction principle in Dembo&Zeitouni[13, Theorem 4.2.23], $\tilde{X}_1^\varepsilon/\varepsilon^{\frac{1}{2}-H}$ satisfies the LDP with speed $\frac{1}{\varepsilon^{2H}}$ and rate function

$$\begin{aligned}
\inf_{w,f,g\,:\,\bar{\rho}F(g)^{\frac{1}{2}}w+\rho\Phi(f)=x}[\frac{1}{2}w^2 + I_H(f,g)] &= \inf_{(f,g)\in\mathcal{H}_H^2}[\frac{(x-\rho\Phi(f))^2}{2\bar{\rho}^2F(g)} + I_H(f,g)] \\
&= \inf_{f\in H_1}\Big[\frac{(x-\rho G(f))^2}{2\bar{\rho}^2F(\mathbf{K}_H f')} + \frac{1}{2}\|f\|_{H_1}^2\Big]\,,
\end{aligned}$$

as required. The rest of the proof just involves dealing with the drift term and proceeds as for Theorem 4.1.

B.1 Proof of Lemma B.1

We need to prove that for any given $\delta > 0$

$$\lim_{m\to\infty}\lim_{\varepsilon\to 0}\varepsilon^{2H}\log\mathbb{P}(\sup_{t\in[0,1]}|Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) = -\infty\,.$$

To this end, let $\sigma_t = \sigma(\varepsilon^H B_t^H) - \sigma(\varepsilon^H B^H_{\frac{[mt]}{m}})$, for $t \in [0,1]$. Fix a $\rho > 0$ and consider

$$\tau_1^m \quad = \quad \inf\{t > 0 : \varepsilon^H|B_s^H - B^H_{\frac{[ms]}{m}}| > \rho\} \wedge 1,$$

then for all $t \in [0, \tau_1^m]$, by the α-Hölder continuity of $\sigma(\cdot)$, we have

$$-L\rho^\alpha < \sigma_t < L\rho^\alpha\,, \tag{B-5}$$

where $L > 0$ is the α-Hölder continuity coefficient for $\sigma(\cdot)$. For any $\lambda > 0$ fixed, we have

$$\mathbb{P}(\varepsilon^H\sup_{t\in[0,\tau_1^m]}\int_0^t\sigma_s dB_s > \delta) \quad = \quad \mathbb{P}(\sup_{t\in[0,\tau_1^m]}\exp(\varepsilon^H\lambda\int_0^t\sigma_s dB_s) > e^{\lambda\delta})\,.$$

Since $\exp(\varepsilon^H\lambda\int_0^t \sigma_s dB_s)$ is a submartingale, by the maximum inequality (see e.g. Karatzas&Shreve[27, Theorem 1.3.8]), we have

$$\mathbb{P}(\sup_{t\in[0,\tau_1^m]}\exp(\varepsilon^H\lambda\int_0^t \sigma_s dB_s) > e^{\lambda\delta}) \quad \le \quad e^{-\lambda\delta}\,\mathbb{E}[\exp(\varepsilon^H\lambda\int_0^{\tau_1^m}\sigma_s dB_s)]\,.$$

Introducing the supermartingale $M_t = \exp(\varepsilon^H\lambda\int_0^{t\wedge\tau_1^m}\sigma_s dB_s - \frac{1}{2}\varepsilon^{2H}\lambda^2\int_0^{t\wedge\tau_1^m}\sigma_s^2 ds)$, then

$$\mathbb{E}(\exp(\varepsilon^H\lambda\int_0^{\tau_1^m}\sigma_s dB_s)) = \mathbb{E}(M_{\tau_1^m}\cdot\exp(\frac{1}{2}\varepsilon^{2H}\lambda^2\int_0^{\tau_1^m}\sigma_s^2 ds)) \quad \le \quad \exp(\frac{1}{2}\varepsilon^{2H}\lambda^2 L^2\rho^{2\alpha}),$$

where the last inequality is due to (B-5). Hence,

$$\mathbb{P}(\sup_{t\in[0,\tau_1^m]}\exp(\varepsilon^H\lambda\int_0^t \sigma_s dB_s) > e^{\lambda\delta}) \quad \le \quad \exp(\frac{1}{2}\varepsilon^{2H}\lambda^2 L^2\rho^{2\alpha} - \lambda\delta).$$

Letting $\lambda = \frac{\delta}{\varepsilon^{2H}L^2\rho^{2\alpha}}$,

$$\mathbb{P}(\sup_{t\in[0,\tau_1^m]}\varepsilon^H\int_0^t \sigma_s dB_s > \delta) \quad \le \quad \exp(-\frac{\delta^2}{2\varepsilon^{2H}L^2\rho^{2\alpha}})\,.$$

That is,

$$\varepsilon^{2H}\log\mathbb{P}(\sup_{t\in[0,\tau_1^m]}|Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) \le -\frac{\delta^2}{2L^2\rho^{2\alpha}} + \varepsilon^{2H}\log 2\,, \tag{B-6}$$

where we have used that $\mathbb{P}(\sup_{t\in[0,\tau_1^m]}|Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) \le \mathbb{P}(\sup_{t\in[0,\tau_1^m]}(Z_t^\varepsilon - Z_t^{m,\varepsilon}) > \delta) + \mathbb{P}(\sup_{t\in[0,\tau_1^m]}(Z_t^\varepsilon - Z_t^{m,\varepsilon}) < -\delta)$. Hence

$$\lim_{\rho\to 0}\sup_{m\ge 1}\limsup_{\varepsilon\to 0}\varepsilon^{2H}\log\mathbb{P}(\sup_{t\in[0,\tau_1^m]}|Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) = -\infty\,.$$

On the other hand,

$$\mathbb{P}(\tau_1^m < 1) \quad = \quad \mathbb{P}(\sup_{t\in[0,1]}\varepsilon^H|B_t^H - B^H_{\frac{[mt]}{m}}| > \rho) \le m\mathbb{P}(\sup_{t\in[0,\frac{1}{m}]}\varepsilon^H|B_t^H| > \rho), \tag{B-7}$$

because fBM has stationary increments. By the scaling property of fBM we have

$$\begin{aligned}\mathbb{P}(\sup_{t\in[0,\frac{1}{m}]}\varepsilon^H B_t^H > \rho) = \mathbb{P}(\sup_{t\in[0,1]}\varepsilon^H B^H_{\frac{t}{m}} > \rho) \quad &= \quad \mathbb{P}(\sup_{t\in[0,1]}(\frac{\varepsilon}{m})^H B_t^H > \rho) \\ &= \quad \mathbb{P}(\sup_{t\in[0,1]}B_t^H > \rho(\frac{m}{\varepsilon})^H)\,.\end{aligned} \tag{B-8}$$

By Theorem 1 of Lifshits[28, p.139], there exists a constant d such that

$$\lim_{r\to\infty} r^{-1}[\log\mathbb{P}(\sup_{t\in[0,1]}B_t^H > r) + (r+d)^2/2] = 0. \tag{B-9}$$

Thus for all $c > 0$, there exists an $r^* = r^*(c) > 0$ sufficiently large such that for all $r > r^*$ we have

$$\mathbb{P}(\sup_{t\in[0,1]}B_t^H > r) \le \exp(-\frac{(r+d)^2}{2} + rc)\,.$$

Thus for the range of B_t^H we have for all $r > 0$ sufficiently large,

$$\mathbb{P}(\sup_{t\in[0,1]}|B_t^H| > r) \le 2\exp(-\frac{(r+d)^2}{2} + rc).$$

Now setting $r = \rho(\frac{m}{\varepsilon})^H$ and using (B-9) we have

$$\lim_{\varepsilon\to 0} \varepsilon^{2H} \log \mathbb{P}(\sup_{t\in[0,\frac{1}{m}]} |\varepsilon^H B_t^H| > \rho) \le \lim_{\varepsilon\to 0} \varepsilon^{2H}[-\frac{(\rho(m/\varepsilon)^H + d)^2}{2} + \rho(m/\varepsilon)^H c + \log 2] = -\frac{\rho^2 m^2}{2}.$$

Thus from (B-7) we have

$$\limsup_{m\to\infty} \lim_{\varepsilon\to 0} \varepsilon^{2H} \mathbb{P}(\tau_1^m < 1) \le \limsup_{m\to\infty} \lim_{\varepsilon\to 0} \varepsilon^{2H}[\log m + \log \mathbb{P}(\sup_{t\in[0,\frac{1}{m}]} |\varepsilon^H B_t^H| > \rho)] = -\infty, \tag{B-10}$$

and the right hand side is $-\infty$ so we can replace the lim sup's in m with a genuine limit.

Finally, notice that

$$\{\sup_{t\in[0,1]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta\} \quad \subset \quad \{\tau_1^m < 1\} \quad \cup \quad \{\sup_{t\in[0,\tau_1^m]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta\}\,.$$

Hence

$$\begin{aligned} \mathbb{P}(\sup_{t\in[0,1]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) \quad &\le \quad \mathbb{P}(\tau_1^m < 1) \,+\, \mathbb{P}(\sup_{t\in[0,\tau_1^m]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) \\ &\le \quad 2\max\{\mathbb{P}(\tau_1^m < 1)\,,\, \mathbb{P}(\sup_{t\in[0,\tau_1^m]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta)\}\,. \end{aligned}$$

We can then proceed as

$$\begin{aligned} & \varepsilon^{2H} \log \mathbb{P}(\sup_{t\in[0,1]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) \\ \le \quad & \varepsilon^{2H} \log 2 \,+\, \max\{\varepsilon^{2H} \log \mathbb{P}(\tau_1^m < 1), \varepsilon^{2H} \log \mathbb{P}(\sup_{t\in[0,\tau_1^m]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta)\} \\ \le \quad & \max\{\limsup_{\varepsilon\to 0} \varepsilon^{2H} \log \mathbb{P}(\tau_1^m < 1)\,,\, \limsup_{\varepsilon\to 0} \varepsilon^{2H} \log \mathbb{P}(\sup_{t\in[0,\tau_1^m]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta)\}, \end{aligned} \tag{B-11}$$

as $\varepsilon \to 0$. For any $N < 0$, from (B-6) we can select a $\rho > 0$ sufficiently small such that

$$\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^{2H} \log \mathbb{P}(\sup_{t\in[0,\tau_1^m]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) \quad \le \quad \sup_{m\ge 1} \limsup_{\varepsilon\to 0} \varepsilon^{2H} \log \mathbb{P}(\sup_{t\in[0,\tau_1^m]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) \le N.$$

On the other hand, for this $\rho > 0$, from (B-10) we have that

$$\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^{2H} \log \mathbb{P}(\tau_1^m < 1) = -\infty\,.$$

Thus, from (B-11), we have

$$\lim_{m\to\infty} \limsup_{\varepsilon\to\infty} \varepsilon^{2H} \log \mathbb{P}(\sup_{t\in[0,1]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) \le\, N\,.$$

By the arbitrariness of $N < 0$ we know that

$$\lim_{m\to\infty} \limsup_{\varepsilon\to\infty} \varepsilon^{2H} \log \mathbb{P}(\sup_{t\in[0,1]} |Z_t^\varepsilon - Z_t^{m,\varepsilon}| > \delta) = -\infty\,.$$

B.2 Proof of Lemma B.2

For all $\dot{h} \in L^2[0,1]$ such that $\|\dot{h}\|_{L^2[0,1]} \le \beta$ and assuming that $s < t$ (without loss of generality), from the Cauchy-Schwarz inequality we see that

$$\begin{aligned} |(\mathbf{K}_H \dot{h})(t) - (\mathbf{K}_H \dot{h})(s)| \quad &= \quad |\int_0^t K_H(u,t)\dot{h}(u)du - \int_0^s K_H(u,s)\dot{h}(u)du| \\ &\le \quad [\int_0^t |K_H(u,t) - K_H(u,s) 1_{\{u<s\}}|^2 du]^{\frac{1}{2}} \cdot \beta := \delta(t,s)\cdot\beta\,. \end{aligned} \tag{B-12}$$

But we can re-write $\delta(t,s)$ as

$$\begin{aligned}\delta^2(t,s) &= \int_0^t K_H^2(u,t)du + \int_0^s K_H^2(u,s)du - 2\int_0^s K_H(u,t)K_H(u,s)du \\ &= \mathrm{Var}(B_t^H) + \mathrm{Var}(B_s^H) - 2\mathrm{Cov}(B_t^H B_s^H) = |t-s|^{2H}.\end{aligned}$$

Thus $\delta(t-s) = |t-s|^H$. Moreover, by the α-Hölder continuity of σ and the Cauchy-Schwartz inequality, we have

$$\begin{aligned}|(\Phi((f,g)))(t) - (\Phi^m((f,g)))(t)| &= |\int_0^t [\sigma((\mathbf{K}_H\dot{h})(s)) - \sigma((\mathbf{K}_H\dot{h})(\frac{[ms]}{m}))]\,\dot{h}(s)ds\,| \\ &\leq L\int_0^1 |(\mathbf{K}_H\dot{h})(s) - (\mathbf{K}_H\dot{h})(\frac{[ms]}{m})|^\alpha \cdot |\dot{h}(s)|ds \\ &\leq L\,(\int_0^1 |(\mathbf{K}_H\dot{h})(s) - (\mathbf{K}_H\dot{h})(\frac{[ms]}{m})|^{2\alpha}ds\,)^{\frac{1}{2}} \cdot \beta \\ &\leq L\beta^{1+\alpha}\sup_{s\in[0,1]} \delta(s, \frac{[ms]}{m})^\alpha \leq L\beta^{1+\alpha}\sup_{s\in[0,1]} |s - \frac{[ms]}{m}|^{\alpha H} = \frac{L\beta^{1+\alpha}}{m^{\alpha H}}\,,\end{aligned}$$

where L is the Hölder constant for σ, and we have used (B-13) in the final line. Letting $m\to\infty$ we obtain the result.

C Proof of Corollary 4.7

- (i) Lower bound. Recall that $\gamma = \frac{1}{2} - H$; then for any $\delta > 0$, we have

  $$\begin{aligned}\mathbb{E}(S_t - e^{xt^\gamma})^+ &\geq (e^{x(1+\delta)t^\gamma} - e^{xt^\gamma})\,\mathbb{P}(S_t > e^{x(1+\delta)t^\gamma}) \\ &= e^{xt^\gamma}(e^{\delta t^\gamma} - 1)\,\mathbb{P}(S_t > e^{x(1+\delta)t^\gamma}) \geq e^{xt^\gamma}\,\delta t^\gamma\,\mathbb{P}(S_t > e^{x(1+\delta)t^\gamma})\,.\end{aligned}$$

  By Theorem 4.2 and using that $\lim_{t\to 0} t^{2H}\log t = 0$, we have that

  $$\begin{aligned}&\liminf_{t\to 0} t^{2H}\log\mathbb{E}(S_t - S_0e^{xt^\gamma})^+ \\ \geq\; &\liminf_{t\to 0}[t^{2H}(xt^\gamma + \log\delta + \gamma\log t) + t^{2H}\log\mathbb{P}(S_t > e^{x(1+\delta)}t^\gamma)] \\ =\; &\liminf_{t\to 0}\, t^{2H}\log\mathbb{P}(S_t > e^{x(1+\delta)t^\gamma}) \geq -\Lambda^*(x+\delta)\,.\end{aligned}$$

  Now take $\delta\to 0+$; then by continuity of $\Lambda^*(\cdot)$, we obtain the desired lower bound.

- Upper bound. We note that for $q > 1$, we have

  $$\mathbb{E}(S_t - e^{xt^\gamma})^+ = \mathbb{E}((S_t - e^{xt^\gamma})^+ 1_{S_t\geq e^{xt^\gamma}}) \leq \mathbb{E}[((S_t - e^{xt^\gamma})^+)^q]^{1/q}\,\mathbb{E}(1_{S_t\geq e^{xt^\gamma}})^{1-1/q}.$$

  Thus

  $$\begin{aligned}t^{2H}\log\mathbb{E}(S_t - e^{xt^\gamma})^+ &\leq \frac{t^{2H}}{q}\log\left[\mathbb{E}[((S_t - e^{xt^\gamma})^+)^q]\right] + t^{2H}(1-\frac{1}{q})\log\mathbb{P}(S_t\geq e^{xt^\gamma}) \\ &\leq \frac{t^{2H}}{q}\log\mathbb{E}(S_t^q) + t^{2H}(1-\frac{1}{q})\log\mathbb{P}(S_t\geq e^{xt^\gamma}).\end{aligned} \tag{C-1}$$

  Now recall that for $q > 1$

  $$\mathbb{E}(S_t^q) = \mathbb{E}(\exp(\tfrac{1}{2}(q^2\bar{\rho}^2 - q)\int_0^t \sigma(B_s^H)^2ds + q\rho\int_0^t \sigma(B_s^H)dB_s)) = \mathbb{E}(R_1\cdot R_2),$$

where we define the nonnegative random variables R_1, R_2 as

$$\begin{aligned} R_1 &= \exp(\frac{1}{2}(q^2 - q + q^2\rho^2)\int_0^t \sigma(B_s^H)^2 ds), \\ R_2 &= \exp(-q^2\rho^2 \int_0^t \sigma(B_s^H)^2 ds + q\rho \int_0^t \sigma(B_s^H) dB_s)\,. \end{aligned}$$

We now choose $t > 0$ small enough so that

$$\max\{(q^2 - q + q^2\rho^2), 2q^2\rho^2\} \cdot A_1 t^{2H+1} \leq \frac{1}{4}\,.$$

Then by the Cauchy-Schwarz inequality,

$$\begin{aligned} \mathbb{E}(S_t^q) &\leq \sqrt{\mathbb{E}(R_1^2)\cdot \mathbb{E}(R_2^2)} \\ &= \sqrt{\mathbb{E}(e^{(q^2-q+q^2\rho^2)\int_0^t \sigma(B_s^H)^2 ds}) \cdot \mathbb{E}(e^{-2q^2\rho^2 \int_0^t \sigma(B_s^H)^2 ds + 2q\rho \int_0^t \sigma(B_s^H) dB_s})} \\ &= \sqrt{\mathbb{E}(e^{(q^2-q+q^2\rho^2)\int_0^t \sigma(B_s^H)^2 ds}) \cdot 1}. \end{aligned}$$

where the last step is due to the observation that R_2^2 can be considered as the time-t value of a martingale starting from 1. Moreover,

$$\begin{aligned} \mathbb{E}(S_t^q) &\leq \sqrt{\mathbb{E}(e^{(q^2-q+q^2\rho^2)\int_0^t \sigma(B_s^H)^2 ds})} \leq \sqrt{\mathbb{E}(e^{(q^2-q+q^2\rho^2)\int_0^t A_1(1+|B_s^H|^2) ds})} \\ &\leq \sqrt{e^{(q^2-q+q^2\rho^2)A_1 t}\sqrt{2}} = \sqrt[4]{2} e^{\frac{1}{2}(q^2-q+q^2\rho^2)A_1 t}, \end{aligned}$$

where the last inequality follows from (15) with $s = 0$, $\varepsilon = t$ and $C = (q^2 - q + q^2\rho^2)A_1$ after verifying that (14) is satisfied. As a consequence, we obtain

$$\lim_{q\to\infty} \limsup_{t\to 0} \frac{t^{2H}}{q} \log \mathbb{E}(S_t^q) \quad \leq \quad \lim_{q\to\infty} \limsup_{t\to 0} \frac{t^{2H}}{q} (\frac{1}{2}(q^2 - q + q^2\rho^2)A_1 t + \frac{1}{4}\log 2) = 0.$$

Hence we see that

$$\limsup_{t\to 0} \frac{t^{2H}}{q} \log \mathbb{E}(S_t^q) \quad \leq \quad 0\,.$$

If we then take $\lim_{q\to\infty} \limsup_{t\to 0}$ on both sides of (C-1), we have (by Theorem 4.1) the upper bound

$$\limsup_{t\to 0} t^{2H} \log \mathbb{E}(S_t - S_0 e^{xt^\gamma})^+ \quad \leq \quad -\Lambda^*(x)\,,$$

as required.