PATHWISE LARGE DEVIATIONS FOR THE ROUGH BERGOMI MODEL

ANTOINE JACQUIER, MIKKO S. PAKKANEN, AND HENRY STONE

Abstract. We study the small-time behaviour of the rough Bergomi model, introduced by Bayer, Friz and Gatheral [4], and prove a large deviations principle for a rescaled version of the normalised log stock price process, which then allows us to characterise the small-time behaviour of the implied volatility.

1. Introduction

The extension of the Black-Scholes model, in which volatility is assumed to be constant, to the case where the volatility is stochastic has proved to be successful in explaining certain phenomena observed in option price data, in particular the implied volatility smile. The main shortcoming of such stochastic volatility models, however, is that they are unable to capture the true steepness of the implied volatility smile close to maturity. While choosing to add jumps to stock price models, for example modelling the stock price process as an exponential Lévy process, does indeed produce steeper implied volatility smiles, see for example [17], the question of the presence of jumps in stock price processes remains controversial [7, 12].

As an alternative to exponential Lévy and classical stochastic volatility models, one may choose a fractional Brownian motion, or a process with similar fine properties, to drive the volatility process, rather than a standard Brownian motion. Since volatility is neither directly observable nor tradable, the issue of arbitrage that is sometimes associated to fractional Brownian motion does not arise in this case. A fractional Brownian motion is a centred Gaussian process, whose covariance structure depends on the Hurst parameter $H \in (0,1)$. If $H \in (0,1/2)$, then the fractional Brownian motion has negatively correlated increments and "rough" sample paths, and if $H \in (1/2,1)$ then it has positively correlated increments and "smooth" sample paths, when compared with a standard Brownian motion, which is recovered by taking $H = 1/2$. There has been a resurgent interest in fractional Brownian motion and related processes within the mathematical finance community in recent years. In particular, Gatheral, Jaisson and Rosenbaum [24] carry out an empirical study that suggests that the log volatility behaves at short time scales in a manner similar to a fractional Brownian motion, in terms of its covariance structure, with Hurst parameter $H \approx 0.1$. This finding is corroborated by Bennedsen, Lunde and Pakkanen [9], who study over a thousand individual US equities and find that the Hurst parameter H lies in $(0,1/2)$ for each equity. In addition, such so-called "rough" volatility models are able to capture the observed steepness of small-time implied volatility smiles and the term structure of at-the-money skew much better than classical stochastic volatility models.

Following [24], Bayer, Friz and Gatheral [4] propose the so-called rough Bergomi model, which they then use to price options on integrated volatility and on the underlying itself. The advantage of their model is that it captures the "rough" behaviour of log volatility, in accordance with [9, 24], as well as fits observed implied volatility smiles better than traditional Markovian stochastic volatility models, most notably in the close-to-maturity case. At the moment, the only known method for pricing mere vanilla options, and computing the corresponding implied volatility smiles, in this setting is Monte Carlo simulation. Despite recent advances in simulation methods for the rough Bergomi model [8, 28], it is

---

Date: June 16, 2017.

2010 *Mathematics Subject Classification.* Primary 60F10, 60G22; Secondary 91G20, 60G15.

Key words and phrases. Rough volatility, large deviations, implied volatility, small-time asymptotics, Gaussian measure, reproducing kernel Hilbert space.

AJ acknowledges financial support from the EPSRC First Grant EP/M008436/1. MSP acknowledges partial support from CREATES (DNRF78), funded by the Danish National Research Foundation.

clearly fruitful to seek more profound analytical understanding of option pricing and implied volatility under this model. Specifically, in the present paper we characterise the small-time behaviour of implied volatility using large deviations theory. We note that related results, concerning the small-time near-the-money skew, have been recently obtained by Bayer, Friz, Gulisashvili, Horvath and Stemper [5]. Large deviations theory is a commonly used tool for analysing small-time and large-time implied volatility smiles, and has been used both for classical stochastic volatility models [13, 18, 26, 27, 29], and for rough volatility models [3, 5, 19, 20, 21].

The structure of the paper is the following. In Section 2, we present the correlated rough Bergomi model, together with its main properties, and lay out the main results of the paper; specifically a small-time large deviations principle for a rescaled version of the normalised log stock price process, and the corresponding small-time implied volatility behaviour for the rough Bergomi model where the log-moneyness is time dependent. In Section 3, we present several elements from the theory of Gaussian measures and large deviations that are required to prove the main results of the paper. In Section 4, we give the proofs of the main results. Finally, Section 5 elucidates the analogous large deviations result for the uncorrelated rough Bergomi model.

Notations: The notation $\mathrm{L}^2 := L^2(\mathcal{T},\mathbb{R})$ stands for the space of real-valued square integrable functions on some index set $\mathcal{T}$. and $\mathcal{C}^d := \mathcal{C}(\mathcal{T},\mathbb{R}^d)$ the space of $\mathbb{R}^d$-valued continuous functions on $\mathcal{T}$. We shall further denote BV the space of paths of finite variations on $\mathcal{T}$, and $\mathbb{R}_+ := [0,\infty)$. For two paths x and y belonging to $\mathcal{C} = \mathcal{C}^1$, we denote by z_y^x the two-dimensional path $(x,y)^\top$. Now, $\mathbb{I}(\mathrm{z}_y^x)(t)$ represents the integral (whenever well-defined) $\int_0^t \sqrt{x(s)}\mathrm{d}y(s)$; the expression $\mathbb{I}(\mathrm{z}_y^x)$ shall be used whenever the integral is taken over the whole time period $[0,1]$, and $x \cdot y := \int_0^1 x(s)\mathrm{d}y(s)$.

2. Model and main results

We assume a given filtered probability space $(\Omega, \mathscr{A}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, where the filtration satisfies the usual conditions, and all stochastic processes here will be assumed to live on this probability space.

2.1. Rough Bergomi Model and Main Properties.

Bayer, Friz and Gatheral [4] introduce a non-Markovian generalisation of Bergomi's "second generation" stochastic volatility model, which they dub the "rough Bergomi" model. Let Z be the process defined pathwise as

$$Z_t := \int_0^t K_\alpha(s,t)\mathrm{d}W_s, \qquad \text{for any } t \geq 0, \tag{2.1}$$

where $\alpha \in \left(-\frac{1}{2}, 0\right)$, W a standard Brownian motion, and where the kernel $K_\alpha : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ reads

$$K_\alpha(s,t) := \eta\sqrt{2\alpha+1}(t-s)^\alpha, \qquad \text{for all } 0 \leq s < t, \tag{2.2}$$

for some strictly positive constant η. Note that, for any $t \geq 0$, the map $s \mapsto K_\alpha(s,t)$ belongs to L^2, so that the stochastic integral (2.1) is well defined. The rough Bergomi model is then defined as

$$\begin{aligned} S_t &= S_0 \exp\left(\int_0^t \sqrt{v_s}\mathrm{d}B_s - \frac{1}{2}\int_0^t v_s \mathrm{d}s\right), && S_0 > 0, \\ v_t &= v_0 \exp\left(Z_t - \frac{\eta^2}{2}t^{2\alpha+1}\right), && v_0 > 0, \end{aligned} \tag{2.3}$$

where the Brownian motion B is defined as $B := \rho W + \sqrt{1-\rho^2}W^\perp$ for $\rho \in [-1,1]$, for some standard Brownian motion $W^\perp$ independent of W. The filtration $(\mathscr{F}_t)_{t\geq 0}$ can here be taken as the one generated by the two-dimensional Brownian motion $(W, W^\perp)$. The stock price process S is clearly then a local $(\mathscr{F}_t)_{t\geq 0}$-martingale, and by Novikov's condition is a true martingale, and hence no arbitrage, in the NFLVR sense [15, Theorem 7.2], is guaranteed.

Proposition 2.1. *The two-dimensional Gaussian process (Z, B) is centred with covariance structure*

$$\mathrm{cov}\left(\begin{pmatrix} Z_t \\ B_t \end{pmatrix}, \begin{pmatrix} Z_t \\ B_t \end{pmatrix}\right) = \begin{pmatrix} \eta^2 t^{2\alpha+1} & \varrho t^{\alpha+1} \\ \varrho t^{\alpha+1} & t \end{pmatrix},$$

$$\begin{aligned}\mathbb{E}(Z_s Z_t) &= \int_0^{s\wedge t} K_\alpha(u,s) K_\alpha(u,t)\mathrm{d}u \\ &= \frac{\eta^2(2\alpha+1)}{\alpha+1}(s\wedge t)^{1+\alpha}(s\vee t)^\alpha {}_2F_1\left(1, -\alpha, 2+\alpha, \frac{s\wedge t}{s\vee t}\right),\end{aligned}$$

for any $s, t \geq 0$, where $\varrho := \frac{\rho\eta\sqrt{2\alpha+1}}{\alpha+1}$ and ${}_2F_1$ is the Gauss hypergeometric function [35, Chapter 5, Section 9].

Proof. Without loss of generality, let us begin by assuming that $s < t$. This then implies that $\mathbb{E}(Z_s Z_t) = \eta^2(2\alpha+1)\int_0^s (t-u)^\alpha (s-u)^\alpha \mathrm{d}u = t^\alpha s^{1+\alpha}\int_0^1 (1-v)^\alpha (1-sv/t)^\alpha \mathrm{d}v$, where the second equality follows from a change of variables. Using Euler's integral representation of the Gauss hypergeometric function ${}_2F_1$, it follows that $\int_0^s (t-u)^\alpha (s-u)^\alpha \mathrm{d}u = \frac{1}{\alpha+1}{}_2F_1(-\alpha, 1; \alpha+2; s/t)$, from which the result then clearly follows. □

Proposition 2.1 implies in particular that the process Z is not stationary, and that the following holds:

Corollary 2.2. *The process Z is $(\alpha + \frac{1}{2})$ self-similar: for any $a > 0$, the processes $(Z_{at})_{t\geq 0}$ and $(a^{\alpha+\frac{1}{2}} Z_t)_{t\geq 0}$ are equal in distribution.*

Note then that the parameter α determines both the local and long-term behaviour of Z.

Remark 2.3. The process Z in (2.1) is the Holmgren-Riemann-Liouville fractional Brownian motion introduced by Lévy [33], modulo some constant scaling, rather than the more commonly known fractional Brownian motion characterised by Mandelbrot and Van Ness [34, Definition 2.1] as

$$W_t^H = \frac{1}{\Gamma(H+1/2)}\left(\int_{-\infty}^0 ((t-s)^{H-1/2} - (-s)^{H-1/2})\mathrm{d}\widetilde{W}_s + \int_0^t (t-s)^{H-1/2}\mathrm{d}\widetilde{W}_s\right),$$

where $\widetilde{W}$ is a standard Brownian motion, and Γ the standard Gamma function. Note that the Mandelbrot-Van Ness representation of W_t^H requires the knowledge of $\widetilde{W}$ from $-\infty$ to t; in comparison we only need to know W from 0 to t to compute the value of Z. Finally, both Z and W^H are self-similar, but W^H has stationary increments whereas the increments of Z are non-stationary.

Proposition 2.4. *The process $\log v$ has a modification whose trajectories are almost surely locally γ-Hölder continuous, for all $\gamma \in \left(0, \alpha + \frac{1}{2}\right)$.*

Proof. We first prove that Z has a modification whose trajectories are γ-Hölder continuous, for all $\gamma \in (0, \alpha + \frac{1}{2})$. By the self-similarity of the process Z, standard results of Gaussian expectations yield

$$\mathbb{E}\left(|Z_t - Z_s|^p\right) = \mathbb{E}\left(|Z_1|^p\right)\left|t^{\alpha+1/2} - s^{\alpha+1/2}\right|^p = \frac{2^{p/2}\eta^p}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right)\left|t^{\alpha+1/2} - s^{\alpha+1/2}\right|^p,$$

for all $p > 0$. Since the map $t \mapsto t^{\alpha+1/2}$ is locally γ-Hölder continuous for all $\gamma \in (0, \alpha + \frac{1}{2})$, there exists $K > 0$ such that $|t^{\alpha+1/2} - s^{\alpha+1/2}| \leq K|t-s|^{\alpha+1/2-\varepsilon}$, for all $\varepsilon \in (0, \alpha + \frac{1}{2})$, and therefore

$$\mathbb{E}\left(|Z_t - Z_s|^p\right) \leq K_p|t-s|^{p(\alpha+1/2-\varepsilon)} =: K_p|t-s|^{1+p\gamma_p}.$$

Kolmogorov's continuity theorem [31, Theorem 3.22] then yields that Z has a modification whose trajectories are locally γ-Hölder continuous where $\gamma \in (0, \gamma_p)$; since γ_p converges to $\alpha + 1/2 - \varepsilon$ as p tends to infinity, this proves the claim, as ε is taken arbitrarily small. Now, for the process $\log v$, we have

$$\begin{aligned}|\log v_t - \log v_s| &= \left|Z_t - \frac{\eta^2}{2}t^{2\alpha+1} - \left(Z_s - \frac{\eta^2}{2}s^{2\alpha+1}\right)\right| \\ &\leq |Z_t - Z_s| + \frac{\eta^2}{2}\left|t^{2\alpha+1} - s^{2\alpha+1}\right| \leq C|t-s|^\gamma + \frac{\eta^2}{2}\left|t^{2\alpha+1} - s^{2\alpha+1}\right|,\end{aligned}$$

where C is a strictly positive constant, and $\gamma \in (0, \alpha + 1/2)$. Since the map $t \mapsto t^{2\alpha+1}$ is also locally γ-Hölder continuous for all $\gamma \in (0, 2\alpha + 1]$ and in particular for all $\gamma \in (0, \alpha + 1/2)$, it follows that the process $\log v$ has a modification with locally γ-Hölder continuous trajectories, for all $\gamma \in (0, \alpha + \frac{1}{2})$. □

As a comparison, the fractional Brownian motion has sample paths that are γ-Hölder continuous for any $\gamma \in (0, H)$ [6, Theorem 1.6.1], so that the rough Bergomi model also captures this roughness by identification $\alpha = H - 1/2$; in particular these trajectories are rougher than those of the standard Brownian motion, for which $H = 1/2$.

2.2. **Main Results.** From now on, without loss of generality, we shall fix $\mathcal{T} = [0, 1]$. For any functions $\varphi_1, \varphi_2 : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$, introduce the operators $\mathcal{I}^{\varphi_1} : \mathrm{L}^2 \to \mathbb{R}$ and $\mathcal{I} : \mathrm{L}^2 \to \mathbb{R}^2$ as

$$\mathcal{I}^{\varphi_1} f := \int_0^{\cdot} \varphi_1(u, \cdot) f(u) \mathrm{d}u \qquad \text{and} \qquad \mathcal{I}^{\varphi_1}_{\varphi_2} f := \begin{pmatrix} \mathcal{I}^{\varphi_1} f \\ \mathcal{I}^{\varphi_2} f \end{pmatrix}, \tag{2.4}$$

Whenever the function φ is constant, say equal to c, we shall write $\mathcal{I}^c$ without ambiguity. We also define the space $\mathscr{H}^{\varphi_1}_{\varphi_2} := \left\{\mathcal{I}^{\varphi_1}_{\varphi_2} f : f \in \mathrm{L}^2\right\}$, which is clearly a Hilbert space once endowed with the inner product $\left\langle \mathcal{I}^{\varphi_1}_{\varphi_2} f_1, \mathcal{I}^{\varphi_1}_{\varphi_2} f_2 \right\rangle_{\mathscr{H}^{\varphi_1}_{\varphi_2}} := \langle f_1, f_2 \rangle_{\mathrm{L}^2}$. Now let X be the normalised log stock price process $X_t := \log(S_t/S_0)$, where the stock price process S is the rough Bergomi model in (2.3). For $t, \varepsilon \geq 0$, let us now define the rescaled processes as follows:

$$X^\varepsilon_t := \varepsilon^\beta X_{\varepsilon t}, \qquad Z^\varepsilon_t := \varepsilon^{\beta/2} Z_t \stackrel{\mathrm{d}}{=} Z_{\varepsilon t}, \qquad v^\varepsilon_t := \varepsilon^{1+\beta} v_0 \exp\left(Z^\varepsilon_t - \frac{\eta^2}{2}(\varepsilon t)^\beta\right), \qquad B^\varepsilon_t := \varepsilon^{\beta/2} B_t, \tag{2.5}$$

where $\beta := 2\alpha + 1 \in (0, 1)$. Note in particular that, for any $t, \varepsilon \geq 0$, Z^ε_t and $Z_{\varepsilon t}$ are equal in law, and so are v^ε_t and $\varepsilon^{1+\beta} v_{\varepsilon t}$, which in turn implies that the following representation holds for any $t \geq 0$:

$$\begin{aligned} X^\varepsilon_t := \varepsilon^\beta X_{\varepsilon t} &\stackrel{\mathrm{d}}{=} \varepsilon^\beta \left(\int_0^{\varepsilon t} \sqrt{v_s} \mathrm{d}B_s - \frac{1}{2} \int_0^{\varepsilon t} v_s \mathrm{d}s\right) \stackrel{\mathrm{d}}{=} \varepsilon^\beta \left(\int_0^t \sqrt{v_{\varepsilon s}} \mathrm{d}B_{\varepsilon s} - \frac{\varepsilon}{2} \int_0^t v_{\varepsilon s} \mathrm{d}s\right) \\ &\stackrel{\mathrm{d}}{=} \int_0^t \sqrt{\varepsilon^{1+2\beta} v_{\varepsilon s}} \mathrm{d}B_s - \frac{1}{2} \int_0^t \varepsilon^{1+\beta} v_{\varepsilon s} \mathrm{d}s \stackrel{\mathrm{d}}{=} \int_0^t \sqrt{v^\varepsilon_s} \mathrm{d}B^\varepsilon_s - \frac{1}{2} \int_0^t v^\varepsilon_s \mathrm{d}s. \end{aligned} \tag{2.6}$$

We now state the main result of this section, namely a pathwise large deviations principle for the sequence of rescaled processes $(X^\varepsilon)_{\varepsilon \geq 0}$. We recall first some facts about large deviations on a real, separable Banach space $(\mathscr{E}, \|\cdot\|_{\mathscr{E}})$, following [16] as our guide.

Definition 2.5. A rate function $\Lambda : \mathscr{E} \to [0, +\infty]$ is a lower semi-continuous function if, for all $\mathrm{x}_0 \in \mathscr{E}$,

$$\liminf_{\mathrm{x} \to \mathrm{x}_0} \Lambda(\mathrm{x}) \geq \Lambda(\mathrm{x}_0).$$

Definition 2.6. A family of probability measures $(\mu_\varepsilon)_{\varepsilon \geq 0}$ on $(\mathscr{E}, \mathscr{B}(\mathscr{E}))$ is said to satisfy a large deviations principle (LDP) as ε tends to zero with speed ε^{-1} and rate function Λ if, for any $B \in \mathscr{B}(\mathscr{E})$,

$$-\inf_{\mathrm{x} \in B^\circ} \Lambda(\mathrm{x}) \leq \liminf_{\varepsilon \downarrow 0} \varepsilon \log \mu_\varepsilon(B) \leq \limsup_{\varepsilon \downarrow 0} \varepsilon \log \mu_\varepsilon(B) \leq -\inf_{\mathrm{x} \in \overline{B}} \Lambda(\mathrm{x}), \tag{2.7}$$

where $\overline{B}$ and B° denote respectively the closure and the interior of B.

Correspondingly, a stochastic process $(Y_\varepsilon)_{\varepsilon \geq 0}$ is said to satisfy a LDP as ε tends to zero if the family of probability measures $(\mathbb{P}(Y_\varepsilon \in \cdot))_{\varepsilon \geq 0}$ satisfies a LDP as ε tends to zero. Unless otherwise stated, all LDP here shall be as ε tends to zero, so we shall drop this mention for simplicity.

To state our results, we now define the operator $\mathcal{M} : \mathcal{C}^2 \to \mathcal{C}(\mathcal{T}^2, \mathbb{R}_+ \times \mathbb{R})$ as

$$(\mathcal{M}\mathrm{z}^x_y)(t, \varepsilon) := \begin{pmatrix} (\mathfrak{m}x)(t, \varepsilon) \\ y(t) \end{pmatrix} \qquad \text{for all } t \in \mathcal{T}, \varepsilon > 0, \tag{2.8}$$

where the operator $\mathfrak{m} : \mathcal{C} \to \mathcal{C}$ is defined by

$$(\mathfrak{m}x)(t, \varepsilon) := v_0 \varepsilon^{1+\beta} \exp\left(x(t) - \frac{\eta^2}{2}(\varepsilon t)^\beta\right), \tag{2.9}$$

as well as the functions $\Lambda^*, \Lambda : \mathcal{C}(\mathcal{T}^2, \mathbb{R}_+ \times \mathbb{R}) \to \mathbb{R}_+$ defined by

$$\Lambda^*(\mathrm{z}_y^x) := \frac{1}{2}\|\mathrm{z}_y^x\|^2_{\mathscr{H}_\rho^{K_\alpha}} \qquad \text{and} \qquad \Lambda(\mathrm{z}_{y_1}^{x_1}) := \inf\left\{\Lambda^*(\mathrm{z}_{y_2}^{x_2}) : \mathrm{z}_{y_1}^{x_1} = \mathcal{M}\mathrm{z}_{y_2}^{x_2}\right\}.$$

Theorem 2.7. *The sequence $(X^\varepsilon)_{\varepsilon\geq 0}$ satisfies a LDP on $\mathcal{C}$ as ε tends to zero, with speed $\varepsilon^{-\beta}$ and rate function $\Lambda^X : \mathcal{C} \to [0, +\infty]$ defined as $\Lambda^X(\varphi) := \inf\left\{\Lambda(\mathrm{z}_y^x) : \varphi = \sqrt{x}\cdot y, y \in \mathrm{BV}\cap\mathcal{C}\right\}$.*

Corollary 2.8. *The rescaled log stock price process $\left(t^\beta X_t\right)_{t\geq 0}$ satisfies a LDP on $\mathbb{R}$ as t tends to zero with speed $t^{-\beta}$ and rate function Λ^X.*

Proof. Since X_1^ε and $\varepsilon^\beta X_\varepsilon$ are equal in law, $(\varepsilon^\beta X_\varepsilon)_{\varepsilon\geq 0}$ satisfies a LDP with speed $\varepsilon^{-\beta}$ and rate function Λ^X by Theorem 2.7; mapping ε to t completes the proof. □

Remark 2.9. Recently, Forde and Zhang [19] derived pathwise large deviations for rough volatility models, with application (by scaling) to small-time asymptotics of the corresponding implied volatility. The model they consider is of the following form, for the log stock price process:

$$\begin{cases} \mathrm{d}X_t = -\frac{1}{2}\sigma(Y_t)^2\mathrm{d}t + \sigma(Y_t)\mathrm{d}B_t, \\ Y_t = W_t^H, \end{cases}$$

where B is a standard Brownian motion, W^H a (possibly correlated) fractional Brownian motion. In order to prove LDP, they consider a small-noise version of the SDE above, namely:

$$\begin{cases} \mathrm{d}X_t^\varepsilon = -\frac{1}{2}\varepsilon\sigma(Y_t)^2\mathrm{d}t + \sqrt{\varepsilon}\sigma(Y_t)\mathrm{d}B_t, \\ Y_t^\varepsilon = \varepsilon^H W_t^H. \end{cases}$$

It is of course tempting to apply their results to the rough Bergomi model. Unfortunately, the following intricacies make this impossible: First, they assume the function σ to have at most linear growth, whereas it is of exponential growth in rough Bergomi. Second, their scaling assumption, allowing them to translate small-noise into small-time estimates crucially relies on the volatility process Y being driftless [19, Equation (4.4)], which does not hold in rough Bergomi.

There is a degree of flexibility when defining the rescaled process X^ε. For example, we may define $X_t^\varepsilon := \varepsilon^\alpha X_{\varepsilon^\gamma t}$, where $\gamma := \alpha/(\alpha/2 + 5/4)$. In this case define $(Z^\varepsilon, B^\varepsilon) := \varepsilon^{\gamma(\alpha+1/2)}(Z, B)$, and $v_t^\varepsilon := \varepsilon^{\alpha+\gamma}v_{\varepsilon^\gamma t}$, so that X^ε satisfies a LDP with speed $\varepsilon^{-2\gamma(\alpha+1/2)}$ and rate function identical to that in Theorem 2.7. This essentially falls in the category of moderate deviations, within the context of [25], for the original process X, which is scaled by $1/(h(t)\sqrt{t})$, where $h(t) \in [1, t^{-1/2}]$ for small enough t.

Remark 2.10. The structure of the Hilbert space $\mathscr{H}_\rho^{K_\alpha}$, see Corollary 3.11, precisely determines the rate function Λ^X. In the uncorrelated case $\rho = 0$, $\mathscr{H}_\rho^{K_\alpha}$ (and its inner product) is degenerate, and clearly Λ^* does not make sense. This case needs to be treated separately and is analysed in Section 5.

From (2.4), every $\mathrm{z}_y^x \in \mathscr{H}_\rho^{K_\alpha}$ has the representation $\mathrm{z}_y^x = \mathcal{I}_\rho^{K_\alpha} f$, for some $f \in \mathrm{L}^2$. Therefore the rate function in Theorem 2.7 can be rewritten as

$$\Lambda^X(\varphi) = \inf_{f\in\mathrm{L}^2}\left\{\frac{1}{2}\int_0^1 f^2(u)\mathrm{d}u : \varphi = \mathbb{I}\left(\mathcal{M}\left(\mathcal{I}_\rho^{K_\alpha} f\right)\right)\right\}. \tag{2.10}$$

Using this formulation, it is then easy to see that $\Lambda^X(0) = 0$: denoting $\mathrm{z}_y^{\widetilde{x}} := \mathcal{M}\mathrm{z}_y^x$ and using that $\widetilde{x} > 0$ it follows that if $\mathbb{I}(\widetilde{x}, y) = 0$ then $y \equiv 0$, which in turn implies that $f \equiv 0$, and hence $\Lambda^X(0) = 0$. Furthermore, since clearly Λ^X cannot take negative values, its minimum value is attained at the origin.

2.3. Asymptotic Behaviour of the Implied Volatility.

Let $\widehat{\sigma}$ denote the implied volatility, that is, for a given log-moneyness $x \in \mathbb{R}$ and maturity $t \geq 0$, the unique non-negative solution to the equation $\mathrm{BS}(x, t, \widehat{\sigma}(x, t)) = \mathrm{C}(x, t)$, where BS denotes the Black-Scholes price of a vanilla Call price, and C the corresponding Call price in a given (here the rough Bergomi) model. Following the methodology developed in [18], or more generally in [22], it is possible to translate the asymptotic behaviour of the log stock price in Corollary 2.8 into small-time behaviour of the implied volatility, as follows:

Corollary 2.11. *The following holds for all $x \neq 0$:*

$$\lim_{t\downarrow 0} t^{1+\beta}\widehat{\sigma}\left(xt^{-\beta}, t\right)^2 = \begin{cases} \dfrac{x^2}{2\inf_{y\geq x}\Lambda^X(y)}, & \text{if } x \geq 0, \\ \dfrac{x^2}{2\inf_{y\leq x}\Lambda^X(y)}, & \text{if } x < 0. \end{cases} \tag{2.11}$$

3. Gaussian Measures on Banach Spaces and Large Deviations

In this section, we gather several elements from the theory of Gaussian measures and large deviations in order to prove Theorem 2.7. This proof shall require a certain number of steps, in particular the precise characterisation of the reproducing kernel Hilbert spaces associated to the different processes under consideration.

3.1. **Gaussian measures on Banach spaces.** Let $\mathcal{T} \subseteq \mathbb{R}$ be some index set. A centred process $(Z_t)_{t\in\mathcal{T}}$ is called Gaussian if for all $n \in \mathbb{N}$ and any $t_1, \ldots, t_n \in \mathcal{T}$, the random variables $Z_{t_1}, \ldots, Z_{t_n}$ are jointly Gaussian; any such process is then completely characterised by its covariance function. We recall some basic facts, needed later, on Gaussian measures on Banach spaces, mostly following Carmona and Tehranchi [10, Chapter 3]. Let $(\mathscr{E}, \|\cdot\|_{\mathscr{E}})$ be a real, separable Banach space, and $\mathscr{E}^*$ its topological dual (i.e., the space of all linear functionals on $\mathscr{E}$), with duality relationship $\langle\cdot,\cdot\rangle_{\mathscr{E}^*\mathscr{E}}$. The bilinear functional $\langle\cdot,\cdot\rangle_{\mathscr{E}^*\mathscr{E}} : \mathscr{E}^* \times \mathscr{E} \to \mathbb{R}$ is such that if $\langle x^*, x\rangle_{\mathscr{E}^*\mathscr{E}} = 0$ for all $x^* \in \mathscr{E}^*$ (resp. $x \in \mathscr{E}$) then $x = 0$ (resp. $x^* = 0$) [2, Page 195]. We shall further let $\mathscr{B}(\mathscr{E})$ denote the Borel σ-algebra of $\mathscr{E}$.

Definition 3.1. [10, Definition 3.1] A measure μ on $(\mathscr{E}, \mathscr{B}(\mathscr{E}))$ is (centred) Gaussian if every $f^* \in \mathscr{E}^*$, when viewed as a random variable via the dual pairing $f \mapsto \langle f^*, f\rangle_{\mathscr{E}^*\mathscr{E}}$, is a (centred) real Gaussian random variable on $(\mathscr{E}, \mathscr{B}(\mathscr{E}), \mu)$.

The following proposition [10, Proposition 3.1] characterises Gaussian measures on Banach spaces.

Proposition 3.2. *Any (centred) Gaussian measure μ on $\mathscr{E}$ is the law of some (centred) Gaussian process with continuous paths, indexed by some compact metric space.*

Note that every real-valued, centred Gaussian process on $\mathscr{E}$ induces some measure on $\mathcal{C}$, the space of continuous functions from $\mathcal{T}$ to $\mathbb{R}$. By Proposition 3.2, one can construct a centred Gaussian probability measure μ on $\mathscr{E}$ by constructing the corresponding Gaussian process. The above argument may be extended to a d-dimensional centred Gaussian process, thereby inducing a Gaussian measure on $\mathscr{E} = \mathcal{C}^d$. For a Gaussian measure μ on $\mathscr{E}$, we introduce the bounded, linear operator $\Gamma : \mathscr{E}^* \to \mathscr{E}$ as

$$\Gamma(f^*) := \int_{\mathscr{E}} \langle f^*, f\rangle_{\mathscr{E}^*\mathscr{E}} f\mu(\mathrm{d}f), \tag{3.1}$$

and note in particular that $\langle f^*, f\rangle_{\mathscr{E}^*\mathscr{E}} f$ is an $\mathscr{E}$-valued random variable on $(\mathscr{E}, \mathscr{B}(\mathscr{E}), \mu)$.

Definition 3.3. [10, Definition 3.3] The reproducing kernel Hilbert space (RKHS) $\mathcal{H}_\mu$ of μ is defined as the completion of $\Gamma(\mathscr{E}^*)$ with the inner product $\langle\Gamma(f^*), \Gamma(g^*)\rangle_{\mathcal{H}_\mu} := \int_{\mathscr{E}} \langle f^*, f\rangle_{\mathscr{E}^*\mathscr{E}}\langle g^*, f\rangle_{\mathscr{E}^*\mathscr{E}}\mu(\mathrm{d}f)$.

For the inclusion map $\iota : \mathcal{H}_\mu \to \mathscr{E}$, the space $\iota(\mathcal{H}_\mu)$ is dense in $\mathscr{E}$; it follows then for the adjoint map $\iota^* : \mathscr{E}^* \to \mathcal{H}_\mu^*$ that $\iota^*(\mathscr{E}^*)$ is dense in $\mathcal{H}_\mu^*$. Recall also that $\mathcal{H}_\mu$ and $\mathcal{H}_\mu^*$ are isometrically isomorphic, which we denote by $\mathcal{H}_\mu^* \simeq \mathcal{H}_\mu$, (by the Riesz representation theorem, as $\mathbb{R}$ is the underlying field). Now, for a centred Gaussian random variable f^* on $\mathscr{E}$, by Definition 3.1, it follows that

$$\mathbb{E}\left(\langle f^*, f\rangle^2_{\mathscr{E}^*\mathscr{E}}\right) = \int_{\mathscr{E}} \langle f^*, f\rangle_{\mathscr{E}^*\mathscr{E}}\langle f^*, f\rangle_{\mathscr{E}^*\mathscr{E}}\mu(\mathrm{d}f) = \|f\|^2_{\mathcal{H}_\mu} = \|\iota^* f^*\|^2_{\mathcal{H}_\mu^*}.$$

This yields the following equivalent form of Definition 3.3 for the RKHS of μ [14, Page 88].

Definition 3.4. A real, separable Hilbert space $\mathcal{H}_\mu$ such that $\mathcal{H}_\mu \subset \mathscr{E}$ is the RKHS of μ if the following two conditions hold:

- there exists an embedding $I : \mathscr{H}_\mu \to \mathscr{E}$, i.e. an injective continuous map whose image is dense in $\mathscr{E}$;

- any $f^* \in \mathscr{E}^*$ is a centred Gaussian random variable on $\mathscr{E}$ with variance $\|I^* f^*\|^2_{\mathcal{H}^*_\mu}$.

Remark 3.5. The embedding I need not necessarily be the inclusion map.

Remark 3.6. Given a triplet $(\mathscr{E}, \mathcal{H}_\mu, \mu)$, consider the inclusion map $I^* : \mathscr{E}^* \to L^2(\mathscr{E}, \mu)$ (we think of $\mathscr{E}^*$ as a dense subset in $\mathcal{H}^*_\mu \simeq \mathcal{H}_\mu$ by ι^*). Since I^* preserves the Hilbert space structure of $L^2(\mathscr{E}, \mu)$, it can be extended to an isometric embedding $\bar{I}^* : \mathcal{H}^*_\mu \to L^2(\mathscr{E}, \mu)$ such that $\|\bar{I}^* f^*\|_{\mathcal{H}^*_\mu} = \|f^*\|_{L^2(\mathscr{E},\mu)}$.

We now explicitly characterise the RKHS of the measures induced by $(Z_t)_{t\in\mathcal{T}}$ (introduced in (2.1)) on $\mathcal{C}$, and by $((Z_t, B_t))_{t\in\mathcal{T}}$ on $\mathcal{C}^2$. In fact, we first prove a more general result, using the operators in (2.4). Introduce the following assumption:

Assumption 3.7. There exists $\phi \in L^2(\mathcal{T}, \mathbb{R})$ such that $\int_0^\varepsilon |\phi(s)|\mathrm{d}s > 0$ for some $\varepsilon > 0$ and $\varphi(\cdot, t) = \phi(t - \cdot)$ for any $t \in \mathcal{T}$.

Theorem 3.8. *Let φ satisfy Assumption 3.7 such that $\mathcal{I}^\varphi$ is injective on* L^2. *The RKHS of the measure induced by the process $\int_0^\cdot \varphi(u, \cdot)\mathrm{d}W_u$ on $\mathcal{C}$ is given by $\mathscr{H}^\varphi = \{\mathcal{I}^\varphi f : f \in \mathrm{L}^2\}$, with inner product $\langle \mathcal{I}^\varphi f_1, \mathcal{I}^\varphi f_2\rangle_{\mathscr{H}^\varphi} := \langle f_1, f_2\rangle_{\mathrm{L}^2}$.*

Corollary 3.9. *The RKHS of the Gaussian measure induced (on $\mathcal{C}$) by $(Z_t)_{t\in\mathcal{T}}$ in* (2.1) *is $\mathscr{H}^{K_\alpha}$.*

We need to extend Theorem 3.8 (and Corollary 3.9) to find the RKHS of the Gaussian measure on the space $\mathcal{C}^2$ induced by the two-dimensional process $((Z_t, B_t))_{t\in\mathcal{T}}$, where Z and B are defined in (2.1) and (2.3) respectively.

Theorem 3.10. *Let φ_1, φ_2 satisfy Assumption 3.7 such that $\mathcal{I}^{\varphi_1}_{\varphi_2}$ is bijective on* L^2. *Introduce the $\mathbb{R}^2$-valued process (Y^1, Y^2) as $Y^i := \int_0^\cdot \varphi_i(s, \cdot)\mathrm{d}W^i_s$ for $i = 1, 2$, where W^1 and W^2 are two standard Brownian motions with correlation $\rho \in [-1, 1]\backslash\{0\}$. Then the RKHS of the measure induced by (Y^1, Y^2) on $\mathcal{C}^2$ is $\mathscr{H}^{\varphi_1}_{\varphi_2} = \left\{\mathcal{I}^{\varphi_1}_{\varphi_2} f : f \in \mathrm{L}^2\right\}$, with inner product $\left\langle \mathcal{I}^{\varphi_1}_{\varphi_2} f_1, \mathcal{I}^{\varphi_1}_{\varphi_2} f_2\right\rangle_{\mathscr{H}^{\varphi_1}_{\varphi_2}} := \langle f_1, f_2\rangle_{\mathrm{L}^2}$.*

Corollary 3.11. *The RKHS of the measure induced (on $\mathcal{C}^2$) by the process $((Z_t, B_t))_{t\in\mathcal{T}}$ is $\mathscr{H}^{K_\alpha}_\rho$.*

3.2. **Large deviations for Gaussian measures.** We now concentrate on large deviations for Gaussian measures. As before, $\mathscr{E}$ denotes a real, separable Banach space with norm $\|\cdot\|_{\mathscr{E}}$, and we introduce a centred Gaussian measure μ on $(\mathscr{E}, \mathscr{B}(\mathscr{E}))$ such that, for any $\mathrm{y} \in \mathscr{E}^*$,

$$\int_{\mathscr{E}} \mathrm{e}^{-\mathrm{i}\langle \mathrm{y},\mathrm{x}\rangle_{\mathscr{E}^*\mathscr{E}}} \mu(\mathrm{dx}) = \exp\left(-\frac{C_\mu(\mathrm{y},\mathrm{y})}{2}\right),$$

where $C_\mu : \mathscr{E}^* \times \mathscr{E}^* \to [0, +\infty)$ is a bilinear, symmetric map. We define $\Lambda^*_\mu : \mathscr{E} \to \mathbb{R}$ as $\Lambda^*_\mu(\mathrm{x}) := \sup\left\{\langle \mathrm{y}, \mathrm{x}\rangle_{\mathscr{E}^*\mathscr{E}} - \frac{1}{2}C_\mu(\mathrm{y}, \mathrm{y}) : \mathrm{y} \in \mathscr{E}^*\right\}$ on $\mathscr{E}$. The following lemma is proved in [14, Lemma 3.4.2].

Lemma 3.12. *The following three statements hold for the measure μ:*

(1) There exists $\alpha > 0$ such that $\int_{\mathscr{E}} \exp\left(\alpha\|\mathrm{x}\|^2_{\mathscr{E}}\right)\mu(\mathrm{dx})$ is finite;

(2) For all $\mathrm{y} \in \mathscr{E}^*$, $C_\mu(\mathrm{y}, \mathrm{y}) = \int_{\mathscr{E}} \langle \mathrm{y}, \mathrm{x}\rangle^2_{\mathscr{E}^*\mathscr{E}}\mu(\mathrm{dx}) \le \|\mathrm{y}\|^2_{\mathscr{E}^*} \int_{\mathscr{E}} \|\mathrm{x}\|^2_{\mathscr{E}}\mu(\mathrm{dx}) \in (0, +\infty)$;

(3) Λ^*_μ defines a rate function on $\mathscr{E}$ and satisfies $\Lambda^*_\mu(a\mathrm{y}) = a^2\Lambda^*_\mu(\mathrm{y})$ for all $a \in \mathbb{R}$.

For a Gaussian random variable X on $\mathscr{E}$ with distribution μ, define $X^\varepsilon := \varepsilon^{1/2}X$, with law μ_ε. Then the following holds [14, Theorem 3.4.5]:

Theorem 3.13. *The sequence of probability measures $(\mu_\varepsilon)_{\varepsilon\ge 0}$ satisfies a LDP on $\mathscr{E}$ with speed ε^{-1} and rate function Λ^*_μ.*

Remark 3.14. Theorem 3.13 implies in particular that the standard Brownian motion $(W_t)_{t\ge 0}$ satisfies a LDP on $\mathbb{R}$ with speed t^{-1}, since W_t and $\sqrt{t}W_1$ are equal in law.

Corollary 3.15. *For any $t \in \mathcal{T}$, let ν_t be the law of Z_t defined in* (2.1). *Then the sequence $(\nu_t)_{t>0}$ satisfies a LDP on $\mathbb{R}$ as t tends to zero with speed $t^{-\beta}$ and rate function $\Lambda^*_\mu(x) := \frac{x^2}{2\eta^2}$, for $x \in \mathbb{R}$.*

Proof. Here, $\mathscr{E} = \mathbb{R}$ and $\langle u, v\rangle_{\mathscr{E}^*\mathscr{E}} = uv$. Since Z_t and $t^{\beta/2}Z_1$ are equal in law, and $\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}yx}\mathbb{P}(Z_1 \in \mathrm{d}x) = \exp(-y^2\eta^2/2)$, taking $C_\mu(x, y) \equiv xy\eta^2$, the proof follows from Theorem 3.13 and Remark 4.1. □

The following two results will be essential for establishing a LDP for the rough Bergomi model. The first one, the contraction principle, states that continuous mappings preserve large deviations principles, while the second one is a universal LDP result for general Gaussian measures on Banach spaces.

Proposition 3.16 (Theorem 4.2.1. in [16] **(Contraction Principle)**)**.** *Let $\mathscr{E}$ and $\widetilde{\mathscr{E}}$ be two Hausdorff topological spaces and let $f : \mathscr{E} \to \widetilde{\mathscr{E}}$ be a continuous mapping. Let $(\nu_\varepsilon)_{\varepsilon\geq 0}$, $(\widetilde{\nu}_\varepsilon)_{\varepsilon\geq 0}$ be two families of probability measures on $(\mathscr{E}, \mathscr{B}(\mathscr{E}))$ and $(\widetilde{\mathscr{E}}, \mathscr{B}(\widetilde{\mathscr{E}}))$ respectively, such that $\widetilde{\nu}_\varepsilon = \nu_\varepsilon \circ f^{-1}$ for each $\varepsilon > 0$. If $(\nu_\varepsilon)_{\varepsilon\geq 0}$ satisfies a LDP on $\mathscr{E}$ with rate function Λ, then $(\widetilde{\nu}_\varepsilon)_{\varepsilon\geq 0}$ satisfies a LDP on $\widetilde{\mathscr{E}}$ with rate function*

$$\widetilde{\Lambda}(\mathrm{y}) := \inf\left\{\Lambda(\mathrm{x}) : \mathrm{y} = f(\mathrm{x})\right\}.$$

Theorem 3.17 (Theorem 3.4.12 in [14])**.** *Let B be a d-dimensional Gaussian process, inducing a measure μ on $(\mathcal{C}^d, \mathscr{B}(\mathcal{C}^d))$ with RKHS $\mathcal{H}_\mu$. Then $(\varepsilon\mu)_{\varepsilon\geq 0}$ satisfies a LDP with speed ε^{-1} and rate function*

$$\Lambda^*_\mu(\mathrm{x}) := \begin{cases} \dfrac{1}{2}\|\mathrm{x}\|^2_{\mathcal{H}_\mu}, & \textit{if } \mathrm{x} \in \mathcal{H}_\mu,\\ +\infty, & \textit{otherwise.}\end{cases}$$

In order to deal with the stochastic differential equation (2.6), we have to consider the stochastic integral $\int_0^\cdot \sqrt{v_s^\varepsilon}\mathrm{d}B_s^\varepsilon$. Assuming that the sequence $(\sqrt{v_s^\varepsilon}, B_s^\varepsilon)$ converges weakly as ε tends to zero yields, under some conditions, a weak convergence for the stochastic integral [30, 32]. However, in order to state a large deviations principle, we need a stronger result, proved by Garcia [23]. Before stating it (Theorem 3.19 below), though, we introduce the following class of sequences of stochastic processes:

Definition 3.18 (Definition 1.1 in [23])**.** Let $\mathscr{U}$ denote the space of simple, real-valued, adapted processes Z such that $\sup_{t\geq 0}|Z_t| \leq 1$. A sequence of semi-martingales $(Y^\varepsilon)_{\varepsilon\geq 0}$ is said to be uniformly exponentially tight (UET) if, for every $c, t > 0$, there exists $K_{c,t} > 0$ such that

$$\limsup_{\varepsilon\downarrow 0} \varepsilon \log\left(\sup_{Z\in\mathscr{U}} \mathbb{P}\left(\sup_{s\leq t}\left|\int_0^s Z_{u-}\mathrm{d}Y_u^\varepsilon\right| \geq K_{c,t}\right)\right) \leq -c.$$

Theorem 3.19 (Theorem 1.2 in [23])**.** *Let $(X^\varepsilon)_{\varepsilon\geq 0}$ be a sequence of adapted, càdlàg stochastic processes, and $(Y^\varepsilon)_{\varepsilon\geq 0}$ a sequence of uniformly exponentially tight semi-martingales. If the sequence $((X^\varepsilon, Y^\varepsilon))_{\varepsilon\geq 0}$ satisfies a LDP with rate function Λ, then the sequence of stochastic integrals $(X^\varepsilon \cdot Y^\varepsilon)_{\varepsilon\geq 0}$ satisfies a LDP with rate function $\widehat{\Lambda}(\varphi) := \inf\left\{\Lambda(\mathrm{z}_y^x) : \varphi = \mathrm{x}\cdot\mathrm{y}, \mathrm{y} \in \mathrm{BV}\right\}$.*

4. Proof of the main results

Proof of Theorem 3.8. Let Assumption 3.7 hold for a given function $\varphi \in \mathrm{L}^2$. The operator $\mathcal{I}^\varphi$ in (2.4) is surjective on $\mathscr{H}^\varphi$. Let $f_1, f_2 \in \mathrm{L}^2$ be such that $\mathcal{I}^\varphi f_1 = \mathcal{I}^\varphi f_2$. Then $\int_0^t \varphi(u,t)[f_1(u) - f_2(u)]\mathrm{d}u = 0$ for any $t \in \mathcal{T}$. Titchmarsh's convolution theorem [36, Theorem VII] then implies that $f_1 = f_2$ almost everywhere, so that $\mathcal{I}^\varphi$ is a bijection. The linearity of $\mathcal{I}^\varphi$ implies that $\langle \mathcal{I}^\varphi f_1, \mathcal{I}^\varphi f_2\rangle_{\mathscr{H}^\varphi} := \langle f_1, f_2\rangle_{\mathrm{L}^2}$ defines an inner product on $\mathscr{H}^\varphi$, and hence $(\mathscr{H}^\varphi, \langle\cdot,\cdot\rangle_{\mathscr{H}^\varphi})$ is a real inner product space. In order for $\mathscr{H}^\varphi$ to satisfy Definition 3.4, we first need to show that it is a separable Hilbert space. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of functions in L^2 such that $\{\mathcal{I}^\varphi f_n\}_{n\in\mathbb{N}}$ converges to $\mathcal{I}^\varphi f$ in L^2. Therefore $\|\mathcal{I}^\varphi f_n - \mathcal{I}^\varphi f_m\|_{\mathscr{H}^\varphi} = \|f_n - f_m\|_{\mathrm{L}^2}$ tends to zero as n and m tend to infinity. Since L^2 is a complete (Hilbert) space, there exists a function $\widetilde{f} \in \mathrm{L}^2$ such that the sequence $\{f_n\}_{n\in\mathbb{N}}$ converges to $\widetilde{f}$. Assume for a contradiction that $f \neq \widetilde{f}$, then, since $\mathcal{I}^\varphi$ is a bijection, the triangle inequality yields

$$0 < \left\|\mathcal{I}^\varphi f - \mathcal{I}^\varphi \widetilde{f}\right\|_{\mathscr{H}^\varphi} \leq \|\mathcal{I}^\varphi f - \mathcal{I}^\varphi f_n\|_{\mathscr{H}^\varphi} + \left\|\mathcal{I}^\varphi \widetilde{f} - \mathcal{I}^\varphi f_n\right\|_{\mathscr{H}^\varphi},$$

which converges to zero as n tends to infinity. Therefore $f = \widetilde{f}$, $\mathcal{I}^\varphi f \in \mathscr{H}^\varphi$ and $\mathscr{H}^\varphi$ is complete, hence a real Hilbert space. Since L^2 is separable with countable orthonormal basis $\{\phi_n\}_{n\in\mathbb{N}}$, then $\{\mathcal{I}^\varphi\phi_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for $\mathscr{H}^\varphi$, which is then separable.

We now wish to find a dense embedding $I : \mathscr{H}^\varphi \to \mathscr{E}$ as in Definition 3.4. Since $\mathscr{H}^\varphi \subset \mathcal{C}$, take the embedding to be the inclusion map $I = \iota$. By [11, Lemma 2.1], the conditions on ϕ in Assumption 3.7 imply that $\mathscr{H}^\varphi$ is dense in $\mathcal{C}$. Finally, for $f^* \in \mathcal{C}^*$, the measure μ induced by the process $\int_0^\cdot \varphi(u,\cdot)\mathrm{d}W_s$ is a Gaussian probability measure on $(\mathscr{E}, \mathscr{B}(\mathscr{E}))$, and f^* is a centred, real Gaussian random variable on $(\mathscr{E}, \mathscr{B}(\mathscr{E}), \mu)$ by Definition 3.1 . In turn, Remark 3.6 implies that I^*, the dual of I, admits an isometric embedding $\bar{I}^*$ such that $\|\bar{I}^* f^*\|^2_{(\mathscr{H}^\varphi)^*} = \|f^*\|^2_{L^2(\mathscr{E},\mu)} = \int_{\mathscr{E}} (f^*)^2 \mathrm{d}\mu = \mathbb{V}(f^*)$, and hence $\mathscr{H}^\varphi$ is the RKHS of μ. □

Proof of Theorem 3.10. We proceed in a similar manner to the proof of Theorem 3.8. Let φ_1, φ_2 satisfy Assumption 3.7. Clearly the operator $\mathcal{I}^{\varphi_1}_{\varphi_2}$ in (2.4) is surjective on $\mathscr{H}^{\varphi_1}_{\varphi_2} \subset \mathcal{C}^2$. By Titchmarsh's convolution Theorem [36, Theorem VII], if $\mathcal{I}^{\varphi_1}_{\varphi_2} f_1 = \mathcal{I}^{\varphi_1}_{\varphi_2} f_2$ on $\mathcal{T}$, then $f_1 = f_2$ and $\mathcal{I}^{\varphi_1}_{\varphi_2}$ is a bijection. Furthermore $\langle \mathcal{I}^{\varphi_1}_{\varphi_2} f_1, \mathcal{I}^{\varphi_1}_{\varphi_2} f_2 \rangle_{\mathscr{H}^{\varphi_1}_{\varphi_2}} := \langle f_1, f_2 \rangle_{\mathrm{L}^2}$ is a well-defined inner product and, following the proof of Theorem 3.8, $(\mathscr{H}^{\varphi_1}_{\varphi_2}, \langle \cdot, \cdot \rangle_{\mathscr{H}^{\varphi_1}_{\varphi_2}})$ is a real, separable Hilbert space. To find a dense embedding $I : \mathscr{H}^{\varphi_1}_{\varphi_2} \to \mathcal{C}^2$, take I as the inclusion map ι; then the conditions on ϕ_1, ϕ_2 in Assumption 3.7 imply that $\mathscr{H}^{\varphi_1}_{\varphi_2}$ is dense in $\mathcal{C}^2$ by [11, Lemma 2.1]. Finally, $f^* \in (\mathcal{C}^2)^*$ is a real, centred Gaussian random variable on $(\mathcal{C}^2, \mathscr{B}(\mathcal{C}^2), \mu)$, where μ denotes the measure induced by (Y^1, Y^2), and $\mathbb{V}(f^*) = \int_{\mathcal{C}^2} (f^*)^2 \mathrm{d}\mu_2 = \|f^*\|^2_{L^2(\mathcal{C}^2,\mu_2)} = \|\iota^* f^*\|^2_{(\mathscr{H}^{\varphi_1}_{\varphi_2})^*}$, so that by Definition 3.4, $\mathscr{H}^{\varphi_1}_{\varphi_2}$ is the RKHS of the measure induced by (Y^1, Y^2) on $\mathcal{C}^2$. □

Proof of Theorem 2.7. Let the two-dimensional rescaled process $(Z^\varepsilon, B^\varepsilon)$ be as in (2.5). From Theorem 3.17 and Corollary 3.11 the sequence $((Z^\varepsilon, B^\varepsilon))_{\varepsilon \geq 0}$ satisfies a LDP with speed $\varepsilon^{-\beta}$ and rate function (with $\mathscr{H}^{K_\alpha}_\rho$ given in Corollary 3.11)

$$\Lambda^*(\mathrm{z}^x_y) = \begin{cases} \dfrac{1}{2} \left\|\mathrm{z}^x_y\right\|^2_{\mathscr{H}^{K_\alpha}_\rho}, & \text{if } \mathrm{z}^x_y \in \mathscr{H}^{K_\alpha}_\rho, \\ +\infty, & \text{otherwise.} \end{cases}$$

Pathwise, we may view the map $t \mapsto (Z^\varepsilon_t, B^\varepsilon_t)^\top$ as an element of $\mathcal{C}^2$, and write

$$\begin{pmatrix} v^\varepsilon_t \\ B^\varepsilon_t \end{pmatrix} = \mathcal{M} \begin{pmatrix} Z^\varepsilon \\ B^\varepsilon \end{pmatrix} (t, \varepsilon).$$

We first verify that the operator $\mathcal{M}$ in (2.8) is continuous with respect to the $\mathcal{C}(\mathcal{T}^2, \mathbb{R}_+ \times \mathbb{R})$ norm $\|\cdot\|_\infty$. For any $(f,g)^\top \in \mathcal{C}^2$, introduce a small perturbation $(\delta^f, \delta^g) \in \mathcal{C}^2$. Then

$$\begin{aligned} \left\| \mathcal{M} \begin{pmatrix} f + \delta^f \\ g + \delta^g \end{pmatrix} - \mathcal{M} \begin{pmatrix} f \\ g \end{pmatrix} \right\|_\infty &= \sup_{t \in \mathcal{T}, \varepsilon > 0} \left\{ \left| (\mathfrak{m}(f + \delta^f))(t, \varepsilon) - (\mathfrak{m} f)(t, \varepsilon) \right| + |\delta^g(t)| \right\} \\ &\leq \sup_{t \in \mathcal{T}, \varepsilon > 0} \left\{ v_0^{1+\beta} \exp\left(-\frac{\eta^2}{2} (\varepsilon t)^\beta \right) \left| \mathrm{e}^{f(t)} \right| \left| \mathrm{e}^{\delta^f(t)} - 1 \right| \right\} + \sup_{t \in \mathcal{T}} |\delta^g(t)| \\ &\leq C \sup_{t \in \mathcal{T}} \left| \mathrm{e}^{\delta^f(t)} - 1 \right| + \sup_{t \in \mathcal{T}} |\delta^g(t)|, \end{aligned}$$

for some strictly positive constant C. The right-hand side clearly tends to zero as (δ^f, δ^g) tends to zero with respect to the sup norm on $\mathcal{C}^2$, and hence $\mathcal{M}$ is a continuous operator. The Contraction Principle (Proposition 3.16) therefore implies that the sequence $(v^\varepsilon, B^\varepsilon)_{\varepsilon \geq 0}$ satisfies a LDP on $\mathcal{C}(\mathcal{T}^2, \mathbb{R}_+ \times \mathbb{R})$, with speed $\varepsilon^{-\beta}$ and rate function Λ. Since $\mathcal{M}$ is clearly a bijection, the rate function Λ may then be expressed as $\Lambda(\mathrm{z}^{x_1}_{y_1}) = \Lambda^* \left(\mathcal{M}^{-1}(\mathrm{z}^{x_1}_{y_1}) \right)$, for any $(x_1, y_1) \in \mathcal{C}^2$.

In the second step we will apply Theorem 3.19 to prove that the sequence $(\mathbb{I}(v^\varepsilon, B^\varepsilon)(\cdot))_{\varepsilon \geq 0} := (\int_0^\cdot \sqrt{v^\varepsilon_s} \mathrm{d}B^\varepsilon_s)_{\varepsilon \geq 0}$ satisfies a LDP. Since $B^\varepsilon = \varepsilon^{\alpha + 1/2} B$ by (2.6), the equality $\mathbb{I}(v^\varepsilon, B^\varepsilon)(\cdot) = \mathbb{I}(\varepsilon^{2\alpha} v^\varepsilon, \sqrt{\varepsilon} B)(\cdot)$ holds almost surely, and therefore [23, Example 2.1] the sequence of (semi)-martingales $(\sqrt{\varepsilon} B)_{\varepsilon \geq 0}$ is UET in the sense of Definition 3.18. Since the sequence $(\sqrt{\varepsilon^{2\alpha} v^\varepsilon})_{\varepsilon \geq 0}$ consists of càdlàg, $(\mathscr{F}_t)$-adapted processes, Theorem 3.19, implies that the sequence $(\mathbb{I}(v^\varepsilon, B^\varepsilon)(\cdot))_{\varepsilon \geq 0}$ satisfies a LDP with speed $\varepsilon^{-\beta}$ and rate function

$$\Lambda^X(z) = \inf\{\Lambda(\mathrm{z}^x_y), z = \mathbb{I}(x,y) \text{ and } y \in \mathrm{BV} \cap \mathcal{C}\}.$$

The final step is to prove the LDP for $X^\varepsilon = \int_0^\cdot \sqrt{v_s^\varepsilon}\mathrm{d}B_s^\varepsilon - \frac{1}{2}\int_0^\cdot v_s^\varepsilon \mathrm{d}s$. To do this we show that the sequences $(X^\varepsilon)_{\varepsilon\geq 0}$ and $(\mathbb{I}(v^\varepsilon, B^\varepsilon))_{\varepsilon\geq 0}$ are exponentially equivalent. For any $\delta > 0$ it follows that

$$\mathbb{P}\left(\sup_{t\in[0,1]} |X_t^\varepsilon - \mathbb{I}(v^\varepsilon, B^\varepsilon)(t)| > \delta\right) \leq \mathbb{P}\left(\int_0^1 v_s^\varepsilon \mathrm{d}s > \delta\right) \leq \mathbb{P}\left(\int_0^1 \exp(Z_s^\varepsilon)\mathrm{d}s > b_\varepsilon\right),$$

where $b_\varepsilon := \delta/v_0\varepsilon^{1+\beta}$. Using that $\int_0^1 \exp(Z_s^\varepsilon)\mathrm{d}s \leq \exp(\sup_{t\in[0,1]} Z_t^\varepsilon)$ almost surely, it follows that

$$\mathbb{P}\left(\int_0^1 \exp(Z_s^\varepsilon)\mathrm{d}s > b_\varepsilon\right) \leq \mathbb{P}\left(\sup_{t\in[0,1]} Z_t^\varepsilon > \log b_\varepsilon\right) = \mathbb{P}\left(\sup_{t\in[0,1]} Z_t > \frac{\log b_\varepsilon}{\varepsilon^{\beta/2}}\right).$$

The process $(Z_t)_{t\in[0,1]}$ is almost surely bounded [1, Theorem 1.5.4], and so we may apply the Borell-TIS inequality; a consequence of which [1, Theorem 2.1.1 and discussion thereafter], implies that

$$\mathbb{P}\left(\sup_{t\in[0,1]} Z_t > \frac{\log b_\varepsilon}{\varepsilon^{\beta/2}}\right) \leq \exp\left(-\frac{1}{2}\left\{\frac{\log b_\varepsilon}{\varepsilon^{\beta/2}} - \mathbb{E}\left(\sup_{t\in[0,1]} Z_t\right)\right\}^2\right).$$

This then implies that

$$\varepsilon^\beta \log \mathbb{P}\left(\int_0^1 \exp(Z_s)\mathrm{d}s > b_\varepsilon\right) \leq \varepsilon^\beta\left(-\frac{(\log b_\varepsilon)^2}{2\varepsilon^\beta} + \frac{\log b_\varepsilon}{\varepsilon^{\beta/2}}\mathbb{E}\left(\sup_{t\in[0,1]} Z_t\right) - \frac{1}{2}\mathbb{E}\left(\sup_{t\in[0,1]} Z_t\right)^2\right).$$

Note that $\varepsilon^{\beta/2}\log b_\varepsilon \to 0$ as $\varepsilon \downarrow 0$, which in turn implies that

$$\limsup_{\varepsilon\downarrow 0} \varepsilon^{\beta/2}\log b_\varepsilon \mathbb{E}\Big(\sup_{t\in[0,1]} Z_t\Big) = 0.$$

Similarly, $\limsup_{\varepsilon\downarrow 0}\varepsilon^\beta \mathbb{E}(\sup_{t\in[0,1]} Z_t)^2 = 0$. Furthermore, it follows that

$$\limsup_{\varepsilon\downarrow 0}\varepsilon^\beta\left(-\frac{(\log b_\varepsilon)^2}{2\varepsilon^\beta}\right) = -\infty.$$

Therefore

$$\limsup_{\varepsilon\downarrow 0}\varepsilon^\beta \log \mathbb{P}\left(\sup_{t\in[0,1]} |X_t^\varepsilon - \mathbb{I}(v^\varepsilon, B^\varepsilon)(t)| > \delta\right) = -\infty,$$

which is precisely the definition of exponential equivalence [16, Definition 4.2.10]. Then, by [16, Theorem 4.2.13], the sequence $(X^\varepsilon)_{\varepsilon\geq 0}$ satisfies a LDP with speed $\varepsilon^{-\beta}$ and rate function Λ^X. □

Proof of Theorem 3.13. Let $\underline{X} := (X^1, \ldots, X^n)$ be an n-dimensional random vector taking values on $\mathscr{E}^n$, where each X^k has distribution μ, so that the average $\frac{1}{n}\sum_{k=1}^n X^k$ has distribution $\mu_{1/n}$. Lemma 3.12 implies that $\int_{\mathscr{E}} \exp(\alpha\|x\|_{\mathscr{E}}^2)\mu_{1/n}(\mathrm{d}x)$ is finite for some $\alpha > 0$, and [14, Theorem 3.3.11] yields a LDP for the sequence $(\mu_{1/n})_{n\geq 1}$, with rate function Λ_μ^*. Define now $n(\varepsilon) := \left\lfloor \frac{1}{\varepsilon}\right\rfloor \vee 1$ and $\ell(\varepsilon) := \varepsilon n(\varepsilon)$ for $\varepsilon > 0$, noting that $\ell(\varepsilon) \in [1-\varepsilon, 1)$ for $\varepsilon \in (0, 1/2)$ and in $[\frac{1}{2}, 1]$ for $\varepsilon \in [1/2, 1]$; for a Gaussian random variable X with distribution $\mu_{1/n(\varepsilon)}$, it follows that $\ell(\varepsilon)^{1/2}X$ has distribution μ_ε. For a closed subset B of $\mathscr{E}$, we define the dilated set $\widetilde{B} := \left\{\ell^{-1/2}x :\ \text{for all } \ell \in \left[\frac{1}{2}, 1\right],\ x \in B\right\}$, so that

$$\begin{aligned}\limsup_{\varepsilon\downarrow 0}\varepsilon \log \mu_\varepsilon(B) &= \limsup_{\varepsilon\downarrow 0}\frac{\ell(\varepsilon)}{n(\varepsilon)}\log \mu_{1/n(\varepsilon)}\left(\ell(\varepsilon)^{-1/2}B\right)\\ &\leq \limsup_{\varepsilon\downarrow 0}\frac{1}{n(\varepsilon)}\log\mu_{1/n(\varepsilon)}(\widetilde{B}) = \limsup_{n\uparrow\infty}\frac{1}{n}\log\mu_{1/n}(\widetilde{B}) \leq -\inf_{x\in\widetilde{B}}\Lambda_\mu^*(x).\end{aligned}$$

The large deviations upper bound then follows from the obvious equalities

$$\inf_{x\in\widetilde{B}}\Lambda_\mu^*(x) = \inf_{\ell\in\left[\frac{1}{2},1\right]}\inf_{x\in B}\Lambda_\mu^*(\ell^{-1/2}x) = \inf_{\ell\in\left[\frac{1}{2},1\right]}\ell^{-1}\inf_{x\in B}\Lambda_\mu^*(x) = \inf_{x\in B}\Lambda_\mu^*(x).$$

Now for any x in any open set $C \subset \mathscr{E}$, we can find an open neighbourhood $\mathcal{O}_x$ such that $\mathcal{O}_x \subseteq \ell(\varepsilon)^{-1/2}C$ for all $0 < \varepsilon < \varepsilon_0$ with $\varepsilon_0 \in \left(0, \frac{1}{2}\right]$. The large deviations lower bound then follows from the inequalities below:

$$\liminf_{\varepsilon\downarrow 0} \varepsilon \log \mu_\varepsilon(C) = \liminf_{\varepsilon\downarrow 0} \frac{\ell(\varepsilon)}{n(\varepsilon)} \log \mu_{\frac{1}{n(\varepsilon)}} \left(\ell(\varepsilon)^{-1/2}C\right) \geq \liminf_{n\uparrow\infty} \frac{1}{n} \log \mu_{\frac{1}{n}}(\mathcal{O}_x) \geq -\inf_{y\in\mathcal{O}_x} \Lambda^*_\mu(y) \geq -\Lambda^*_\mu(x).$$

□

Remark 4.1. The proof of Theorem 3.13 stills holds for the case where $t^{\beta/2}X \sim \mu_t$ with speed $t^{-\beta}$, and the proof can be easily adapted to confirm this case.

5. Large deviations for the uncorrelated Rough Bergomi model

We treat here the special case of (2.3), where the Brownian motions W and B are independent ($\rho = 0$). Following similar arguments to Corollary 3.11, and analogously to (2.4), we introduce the operator $\mathcal{I}^0 : \mathrm{L}^2 \times \mathrm{L}^2 \to \mathbb{R}^2$ as

$$\mathcal{I}^0(f_1, f_2) := \begin{pmatrix} \mathcal{I}^{K_\alpha} f_1 \\ \mathcal{I}^1 f_2 \end{pmatrix}, \qquad \text{for any } f_1, f_2 \in \mathrm{L}^2,$$

so that the RKHS (on $\mathcal{C}^2$) of the measure induced by (Z, B) is $\mathscr{H} := \left\{\mathcal{I}^0(f_1, f_2) : f_1, f_2 \in \mathrm{L}^2\right\}$, with inner product $\left\langle \mathcal{I}^0(f_1, f_2), \mathcal{I}^0(g_1, g_2)\right\rangle_{\mathscr{H}} := \langle f_1, g_1\rangle_{\mathrm{L}^2} + \langle f_2, g_2\rangle_{\mathrm{L}^2}$, for any $f_1, f_2, g_1, g_2 \in \mathrm{L}^2$. Similarly to Theorem 2.7, [14, Theorem 3.4.12] yields a LDP on $\mathcal{C}^2$ for $((Z^\varepsilon, B^\varepsilon))_{\varepsilon\geq 0}$ with speed $\varepsilon^{-\beta}$ and rate function

$$\Lambda(\mathrm{z}^x_y) := \begin{cases} \frac{1}{2}\left\|\mathrm{z}^x_y\right\|^2_{\mathscr{H}}, & \text{if}(x, y)^\top \in \mathscr{H}, \\ +\infty, & \text{otherwise.} \end{cases}$$

This in turn yields a LDP for $((v^\varepsilon, B^\varepsilon))_{\varepsilon\geq 0}$ in (2.5) on $\mathcal{C}^2$ with speed $\varepsilon^{-\beta}$ and rate function $\widetilde{\Lambda}(\mathrm{z}^x_y) := \inf\left\{\Lambda(\mathrm{z}^{x^*}_{y^*}) : \mathrm{z}^x_y = \mathcal{M}\mathrm{z}^{x^*}_{y^*}\right\}$, where the operator $\mathcal{M}$ is defined in (2.8). In the same vein as Theorem 2.7, Theorem 3.19 yields a LDP for $(\int_0^\cdot \sqrt{v^\varepsilon_s}\mathrm{d}B^\varepsilon_s)_{\varepsilon\geq 0}$ on $\mathcal{C}$ with speed $\varepsilon^{-\beta}$ and rate function $\widehat{\Lambda}^X$, defined as

$$\begin{aligned}\widehat{\Lambda}^X(\varphi) &:= \inf\left\{\widetilde{\Lambda}(\mathrm{z}^x_y) : \varphi = x \cdot y, y \in \mathrm{BV} \cap \mathcal{C}\right\} = \inf\left\{\Lambda(\mathrm{z}^{x^*}_{y^*}) : \varphi = x \cdot y, \mathrm{z}^x_y = \mathcal{M}\mathrm{z}^{x^*}_{y^*}, x^*, y^* \in \mathscr{H}\right\} \\ &= \inf\left\{\Lambda(\mathrm{z}^x_y) : \varphi = x \cdot y, \mathrm{z}^x_y = \mathcal{M}(\mathcal{I}^0(f_1, f_2)), f_1, f_2 \in \mathrm{L}^2\right\} \\ &= \inf_{f_1, f_2\in\mathrm{L}^2} \left\{\frac{1}{2}\|f_1\|^2_{\mathrm{L}^2} + \frac{1}{2}\|f_2\|^2_{\mathrm{L}^2} : \varphi = \int_0^\cdot \sqrt{\mathfrak{m}\left((\mathcal{I}^{K_\alpha} f_1)(s)\right)} f_2(s)\mathrm{d}s\right\}.\end{aligned}$$

with $\mathfrak{m}$ introduced in (2.9). Following identical an identical argument to that presented in Theorem 2.7, we conclude that $(X^\varepsilon)_{\varepsilon>0}$ satisfies a LDP with speed $\varepsilon^{-\beta}$ and rate function $\widehat{\Lambda}^X$.

References

[1] R.J. Adler and J.E. Taylor. Random Fields and Geometry. Springer-Verlag, New York, 2007.

[2] C.D. Aliprantis and K.C. Border. Infinite dimensional analysis: a hitchhiker's guide, Third Edition. Springer-Verlag, Berlin, Heidelberg, 2006.

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

[4] C. Bayer, P.K. Friz and J. Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6): 887-904, 2016.

[5] C. Bayer, P.K. Friz, A. Gulisashvili, B. Horvath and B. Stemper. Short-time near-the-money skew in rough fractional volatility models. Preprint available at arXiv:1703.05132, 2017.

[6] F. Biagini, Y. Hu, B. Øksendal and T. Zhang. Stochastic calculus for fractional Brownian motion and applications. Springer, London, 2008.

[7] P. Bajgrowicz, O. Scaillet and A. Treccani. Jumps in high-frequency data: Spurious detections, dynamics, and news. *Management Science*, 62(2): 2198-2217, 2015.

[8] M. Bennedsen, A. Lunde and M.S. Pakkanen. Hybrid scheme for Brownian semistationary processes. *Finance and Stochastics*, to appear.

[9] M. Bennedsen, A. Lunde and M.S. Pakkanen. Decoupling the short- and long-term behavior of stochastic volatility. Preprint available at arXiv:1610.00332, 2016.

[10] R. Carmona and M. Tehranchi. Interest rate models: an infinite dimensional stochastic analysis perspective. Springer-Verlag, Berlin, Heidelberg, 2006.

[11] A. Cherny. Brownian moving averages have conditional full support. *Ann. App. Prob.*, 18(5): 1825-1850, 2008.

[12] K. Christensen, R.C.A. Oomen and M. Podolskij. Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics* 114(3): 576-599, 2014.

[13] J.D. Deuschel, P. Friz, A. Jacquier and S. Violante. Marginal density expansions for diffusions and stochastic volatility, Part I: Theoretical foundations. *Communications on Pure and Applied Mathematics*, 67(1): 40-82, 2014.

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

[15] F. Delbaen, W. Schachermayer. A generalised version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300: 463-520, 1994.

[16] A. Dembo and O. Zeitouni. Large deviations theory and applications, Second Edition. Springer, Berlin, 2010.

[17] J.E. Figueroa-López and M. Forde. The small-maturity smile for exponential Lévy models. *SIAM Journal on Financial Mathematics*, 3(1): 33-65, 2012.

[18] M. Forde and A. Jacquier. Small-time asymptotics for implied volatility under the Heston model. *IJTAF*, 12(6): 861-876, 2009.

[19] M. Forde and H. Zhang. Asymptotics for rough stochastic volatility models. *SIAM Journal on Financial Mathematics*, 8, 114-145, 2017.

[20] M. Fukasawa. Short-time at-the-money skew and rough fractional volatility. *Quant. Finance*, 17(2): 189-198, 2017.

[21] M. Fukasawa. Asymptotic analysis for stochastic volatility: martingale expansion. *Fin. Stoch*, 15: 635-654, 2011.

[22] K. Gao and R. Lee. Asymptotics of implied volatility to arbitrary order. *Finance and Stochastics*, 18(2): 349-392, 2014.

[23] J. Garcia. A large deviations principle for stochastic integrals. *Journal of Theoretical Probability*, 21: 476-501, 2008.

[24] J. Gatheral, T. Jaisson and M. Rosenbaum. Volatility is rough. Preprint available at arXiv:1410.3394, 2014.

[25] A. Guillin. Averaging principle of SDE with small diffusion: moderate deviations. *Ann. Prob.*, 31(1): 413-443, 2003.

[26] A. Jacquier, P. Friz, J. Gatheral, A. Gulisashvili and J. Teichmann. Large Deviations and Asymptotic Methods in Finance Springer Proceedings in Mathematics and Statistics, Volume 110, 2015.

[27] A. Jacquier, M. Keller-Ressel and A. Mijatović. Large deviations and stochastic volatility with jumps: asymptotic implied volatility for affine models. *Stochastics*, 85(2): 321-345, 2013.

[28] A. Jacquier, C. Martini and A. Muguruza. On VIX Futures in the rough Bergomi model. Preprint available at arXiv:1701.04260, 2017.

[29] A. Jacquier and P. Roome. Asymptotics of forward implied volatility. *SIAM Journ. Fin. Math.*, 6(1): 307-351, 2015.

[30] A. Jakubowski, J. Memin and G. Pagès. Convergence en loi des suites d'intégrales stochastique sur l'espace $\mathbb{D}^1$ de Skorokhod. *PTRF*, 81: 111-137, 1989.

[31] O. Kallenberg. Foundations of modern probability, 2nd Edition. Springer, New York, 2002.

[32] T. Kurtz and P. Protter. Weak convergence of stochastic integrals and differential equations II. *Lecture Notes in Mathematics*, 1627: 197-285, 1996.

[33] P. Lévy. Random functions: general theory with special references to Laplacian random functions. *University of California Publications in Statistics*, 1: 331-390, 1953.

[34] B.B. Mandelbrot and J.W. Van Ness. Fractional Brownian motions, fractional noises, and applications. *SIAM Review*, 10: 422-437, 1968.

[35] F.W.J. Olver. Asymptotics and special functions, 2nd Edition. A.K. Peters / CRC Press, 1997.

[36] E.C. Titchmarsh. The zeros of certain integral functions. *Proceedings of the LMS*, 25: 283-302, 1926.

Department of Mathematics, Imperial College London
E-mail address: a.jacquier@imperial.ac.uk

Department of Mathematics, Imperial College London and CREATES, Aarhus University
E-mail address: m.pakkanen@imperial.ac.uk

Department of Mathematics, Imperial College London
E-mail address: henry.stone15@imperial.ac.uk