J. Appl. Prob. **55,** 1078–1092 (2018)
*doi:*10.1017/jpr.2018.72
© *Applied Probability Trust* 2018

PATHWISE LARGE DEVIATIONS FOR THE ROUGH BERGOMI MODEL

ANTOINE JACQUIER,* ** *Imperial College London*
MIKKO S. PAKKANEN,* *** *Imperial College London and CREATES*
HENRY STONE,* **** *Imperial College London*

Abstract

Introduced recently in mathematical finance by Bayer *et al.* (2016), the rough Bergomi model has proved particularly efficient to calibrate option markets. We investigate some of its probabilistic properties, in particular proving a pathwise large deviations principle for a small-noise version of the model. The exponential function (continuous but superlinear) as well as the drift appearing in the volatility process fall beyond the scope of existing results, and a dedicated analysis is needed.

Keywords: Rough volatility; large deviations; small-time asymptotics; Gaussian measure; reproducing kernel Hilbert space

2010 Mathematics Subject Classification: Primary 60F10
Secondary 60G22; 91G20; 60G15

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 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 [17], the question of the presence of jumps in stock price processes remains controversial [4], [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 a Hurst parameter $H \in (0, 1)$. If $H \in (0, \frac{1}{2})$ then the fractional Brownian motion has negatively correlated increments and 'rough' sample paths, and if $H \in (\frac{1}{2}, 1)$ then it has positively correlated increments and 'smooth' sample paths when compared to a standard Brownian motion, which is recovered by taking $H = \frac{1}{2}$. There has been a resurgent interest in fractional Brownian motion and related processes within the

Received 15 March 2018; revision received 1 August 2018.
* Postal address: Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK.
** Email address: a.jacquier@imperial.ac.uk
*** Email address: m.pakkanen@imperial.ac.uk
**** Email address: henry.stone15@imperial.ac.uk

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

mathematical finance community in recent years. In particular, Gatheral *et al.* [24] carried out an empirical study that suggests that the log volatility behaves at short time scales in a manner similar to a fractional Brownian motion with Hurst parameter $H \approx 0.1$. This finding was corroborated by Bennedsen *et al.* [7], who studied over a thousand individual US equities and found that the Hurst parameter H lies in $(0, \frac{1}{2})$ 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 *et al.* [5] proposed the so-called rough Bergomi model, which they used 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 [7] and [24], as well as fits that observed implied volatility smiles better than traditional Markovian stochastic volatility models, most notably in the close-to-maturity case. In [3], [21], and [22] the authors studied the short-time behaviour of rough volatility models. Despite recent advances in simulation methods for the rough Bergomi model [8], [27], it is necessary to seek a more profound analytical understanding of the properties of this model. Specifically, in the present paper we prove pathwise large deviations for this model, which allow us to characterise its small-time behaviour. Related results have been recently obtained by Bayer *et al.* [6], and large deviations theory is now a common tool for such analysis in standard stochastic volatility models [14], [18], [20], [26], [28], and for their rough counterparts [6], [19].

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 process. 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, and in Section 5 we elucidate the analogous large deviations result for the uncorrelated rough Bergomi model.

Notation. The notation $\mathrm{L}^2 := L^2(\mathcal{T}, \mathbb{R})$ denotes 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 fix $\mathcal{T} = [0, 1]$ for the rest of this paper, although our results can be easily adapted for the general interval $[0, T]$. We further denote by BV the space of paths of finite variation 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, $\mathrm{I}(z_y^x)(t)$ represents the integral (whenever well defined) $\int_0^t \sqrt{x(s)}\,\mathrm{d}y(s)$; the expression $\mathrm{I}(z_y^x)$ is 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 $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, 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 *et al.* [5] introduced a non-Markovian generalisation of Bergomi's 'second generation' stochastic volatility model, which they called the 'rough Bergomi' model. Let Z be the process defined pathwise as

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

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

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

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

for some strictly positive constant η. Note that, for any $t \ge 0$, 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, where X is the log stock price process and v is the volatility process, is then defined as

$$\begin{aligned} X_t &= -\frac{1}{2}\int_0^t v_s \,\mathrm{d}s + \int_0^t \sqrt{v_s}\,\mathrm{d}B_s, \qquad X_0 = 0, \\ v_t &= v_0 \exp\big(Z_t - \tfrac{1}{2}\eta^2 t^{2\alpha+1}\big), \qquad v_0 > 0, \end{aligned} \tag{2.2}$$

where the Brownian motion B is defined as $B := \rho W + \sqrt{1-\rho^2}W^\perp$ for $\rho \in [-1,1]$ and some standard Brownian motion $W^\perp$ independent of W. The filtration $(\mathcal{F}_t)_{t\ge 0}$ can here be taken as the one generated by the two-dimensional Brownian motion $(W, W^\perp)$.

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

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

$$\begin{aligned} \mathbb{E}(Z_s Z_t) &= \int_0^{s\wedge t} K_\alpha(u,s)K_\alpha(u,t)\,\mathrm{d}u \\ &= \frac{\eta^2(2\alpha+1)}{\alpha+1}(s\wedge t)^{1+\alpha}(s\vee t)^\alpha \,{}_2F_1\left(1, -\alpha, 2+\alpha, \frac{s\wedge t}{s\vee t}\right) \quad \textit{for any } s,t \ge 0, \end{aligned}$$

with $\varrho := \rho\eta\sqrt{2\alpha+1}/(\alpha+1)$ and ${}_2F_1$ the Gauss hypergeometric function; see [35, Chapter 5, Section 9].

Proof. Without loss of generality, assume first that $s < t$. This then implies that

$$\mathbb{E}(Z_s Z_t) = \eta^2(2\alpha+1)\int_0^s (t-u)^\alpha (s-u)^\alpha \,\mathrm{d}u = t^\alpha s^{1+\alpha}\int_0^1 (1-v)^\alpha \left(1 - \frac{sv}{t}\right)^\alpha \mathrm{d}v,$$

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

$$\int_0^s (t-u)^\alpha (s-u)^\alpha \,\mathrm{d}u = \frac{1}{\alpha+1}\,{}_2F_1\left(-\alpha, 1;\ \alpha+2;\ \frac{s}{t}\right),$$

from which the proposition follows. □

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

Corollary 2.1. *The process Z is $(\alpha + 1/2)$ self-similar: for any $a > 0$, the processes $(Z_{at})_{t\ge 0}$ and $(a^{\alpha+1/2}Z_t)_{t\ge 0}$ are equal in distribution.*

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

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

Remark 2.1. The process Z in (2.1) is the Holmgren–Riemann–Liouville fractional Brownian motion introduced by Lévy [32], modulo some constant scaling, rather than the more commonly known fractional Brownian motion characterised by Mandelbrot and Van Ness [33, 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. The Mandelbrot–Van Ness representation of W_t^H requires the knowledge of $\widetilde{W}$ from $-\infty$ to t; in comparison we need to know only W from 0 to t to compute the value of Z. Also both Z and W^H are self-similar, but W^H has stationary increments whereas the increments of Z are nonstationary.

Proposition 2.2. *The process* $\log v$ *has a modification whose trajectories are almost surely locally* γ*-Hölder continuous for all* $\gamma \in (0, \alpha + \frac{1}{2})$.

Proof. We first prove that Z has a modification whose trajectories are γ-Hölder continuous for all $\gamma \in (0, \alpha + \frac{1}{2})$. First,

$$\begin{aligned}\mathbb{E}(|Z_t - Z_s|^2) &\le \eta^2(2\alpha+1)\left(\int_s^t |t-u|^{2\alpha}\,\mathrm{d}u + \int_0^s |(t-u)^\alpha - (s-u)^\alpha|^2\,\mathrm{d}u\right)\\ &= \eta^2|t-s|^{2\alpha+1} + \eta^2(2\alpha+1)\int_0^s |(t-u)^\alpha - (s-u)^\alpha|^2\,\mathrm{d}u.\end{aligned}$$

Following the change of variables $s-u=(t-s)y$, the integral becomes

$$|t-s|^{2\alpha+1}\int_0^{s/(t-s)} |(y+1)^\alpha - y^\alpha|^2\,\mathrm{d}y,$$

and, hence, $\int_0^{s/(t-s)}|(y+1)^\alpha - y^\alpha|^2\,\mathrm{d}y$ is finite since $\alpha \in (-\frac{1}{2}, 0)$. Therefore, there exists $K > 0$ such that $\mathbb{E}(|Z_t - Z_s|^2) \le K|t-s|^{2\alpha+1}$ for any $s, t \ge 0$. Applying the Kolmogorov continuity theorem [30, Theorem 3.22] then yields that the Gaussian process Z has a modification whose trajectories are locally γ-Hölder continuous where $\gamma \in (0, \alpha + \frac{1}{2})$, thus proving the claim. Now, for the process $\log v$, we have

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

where C is a strictly positive constant, and $\gamma \in (0, \alpha + \frac{1}{2})$. 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 $\gamma \in (0, \alpha+\frac{1}{2})$, 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 $\gamma \in (0, H)$ [9, Theorem 1.6.1], so that the rough Bergomi model also captures this roughness by identification $\alpha = H - \frac{1}{2}$; in particular, these trajectories are rougher than those of the standard Brownian motion, for which $H = \frac{1}{2}$.

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

2.2. Main results

For any functions $\varphi_1, \varphi_2 \colon \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$, introduce the L^2 operators $\mathcal{I}^{\varphi_1}$ and $\mathcal{I}^{\varphi_1}_{\varphi_2}$ defined as

$$\mathcal{I}^{\varphi_1} f \coloneqq \int_0^{\cdot} \varphi_1(u, \cdot) f(u)\, \mathrm{d}u \quad \text{and} \quad \mathcal{I}^{\varphi_1}_{\varphi_2} f \coloneqq \begin{pmatrix} \mathcal{I}^{\varphi_1} f \\ \mathcal{I}^{\varphi_2} f \end{pmatrix}. \tag{2.3}$$

Whenever the function φ is constant, say equal to c, we write $\mathcal{I}^c$ without ambiguity. We also define the space $\mathcal{H}^{\varphi_1}_{\varphi_2} \coloneqq \{\mathcal{I}^{\varphi_1}_{\varphi_2} f \colon f \in \mathrm{L}^2\}$ that is clearly a Hilbert space, once endowed with the inner product $\langle \mathcal{I}^{\varphi_1}_{\varphi_2} f_1, \mathcal{I}^{\varphi_1}_{\varphi_2} f_2 \rangle_{\mathcal{H}^{\varphi_1}_{\varphi_2}} \coloneqq \langle f_1, f_2 \rangle_{\mathrm{L}^2}$, where φ_1, φ_2 are such that $\mathcal{I}^{\varphi_1}_{\varphi_2}$ is injective, making the inner product well defined. For $t, \varepsilon \geq 0$, now define the rescaled processes

$$\begin{gathered} X^\varepsilon_t \coloneqq \varepsilon^\beta X_{\varepsilon t}, \qquad Z^\varepsilon_t \coloneqq \varepsilon^{\beta/2} Z_t \overset{\mathrm{D}}{=} Z_{\varepsilon t}, \qquad v^\varepsilon_t \coloneqq \varepsilon^{1+\beta} v_0 \exp\!\big(Z^\varepsilon_t - \tfrac{1}{2}\eta^2 (\varepsilon t)^\beta\big), \\ B^\varepsilon_t \coloneqq \varepsilon^{\beta/2} B_t, \end{gathered} \tag{2.4}$$

where $\beta \coloneqq 2\alpha + 1 \in (0, 1)$. 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 $t \geq 0$:

$$\begin{aligned} X^\varepsilon_t &\coloneqq \varepsilon^\beta X_{\varepsilon t} \\ &\overset{\mathrm{D}}{=} \varepsilon^\beta \left(\int_0^{\varepsilon t} \sqrt{v_s}\, \mathrm{d}B_s - \frac{1}{2} \int_0^{\varepsilon t} v_s\, \mathrm{d}s \right) \\ &\overset{\mathrm{D}}{=} \varepsilon^\beta \left(\int_0^{t} \sqrt{v_{\varepsilon s}}\, \mathrm{d}B_{\varepsilon s} - \frac{\varepsilon}{2} \int_0^{t} v_{\varepsilon s}\, \mathrm{d}s \right) \\ &\overset{\mathrm{D}}{=} \int_0^t \sqrt{\varepsilon^{1+2\beta} v_{\varepsilon s}}\, \mathrm{d}B_s - \frac{1}{2} \int_0^t \varepsilon^{1+\beta} v_{\varepsilon s}\, \mathrm{d}s \\ &\overset{\mathrm{D}}{=} \int_0^t \sqrt{v^\varepsilon_s}\, \mathrm{d}B^\varepsilon_s - \frac{1}{2} \int_0^t v^\varepsilon_s\, \mathrm{d}s. \end{aligned} \tag{2.5}$$

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 $(\mathcal{E}, \|\cdot\|_{\mathcal{E}})$, following [13] as our guide.

Definition 2.1. A function $\Lambda \colon \mathcal{E} \to [0, +\infty]$ is said to be a *rate function* if it is lower semi-continuous, i.e. if, for all $x_0 \in \mathcal{E}$,

$$\liminf_{x \to x_0} \Lambda(x) \geq \Lambda(x_0).$$

Definition 2.2. A family of probability measures $(\mu_\varepsilon)_{\varepsilon \geq 0}$ on $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$ is said to satisfy a *large deviations principle* (LDP) as $\varepsilon \to 0$ with speed ε^{-1} and rate function Λ if, for any $B \in \mathcal{B}(\mathcal{E})$,

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

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

Correspondingly, a stochastic process $(Y_\varepsilon)_{\varepsilon \geq 0}$ is said to satisfy an LDP as $\varepsilon \to 0$ if the family of probability measures $(\mathbb{P}(Y_\varepsilon \in \cdot))_{\varepsilon \geq 0}$ satisfies an LDP as $\varepsilon \to 0$. Unless otherwise stated, all LDPs here are as $\varepsilon \to 0$, so we drop this mention for simplicity. To state our results, we now define the operator $\mathcal{M} \colon \mathcal{C}^2 \to \mathcal{C}(\mathcal{T} \times \mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R})$ as

$$(\mathcal{M} z^x_y)(t, \varepsilon) \coloneqq \begin{pmatrix} (\mathfrak{m} x)(t, \varepsilon) \\ y(t) \end{pmatrix} \quad \text{for all } t \in \mathcal{T},\ \varepsilon > 0, \tag{2.6}$$

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

where the operator $\mathfrak{m}\colon \mathcal{C} \to \mathcal{C}(\mathcal{T} \times \mathbb{R}_+, \mathbb{R}_+)$ is defined by

$$(\mathfrak{m}x)(t, \varepsilon) := v_0 \varepsilon^{1+\beta} \exp\big(x(t) - \tfrac{1}{2}\eta^2 (\varepsilon t)^\beta\big), \tag{2.7}$$

as well as the functions $\Lambda^*, \Lambda\colon \mathcal{C}^2 \to \mathbb{R}_+$ defined by

$$\Lambda^*(z_y^x) := \tfrac{1}{2}\|z_y^x\|^2_{\mathcal{H}_\rho^{K_\alpha}} \quad \text{and} \quad \Lambda(z_{y_1}^{x_1}) := \inf\{\Lambda^*(z_{y_2}^{x_2})\colon z_{y_1}^{x_1} = \mathcal{M} z_{y_2}^{x_2}\}.$$

Theorem 2.1. *The sequence $(X^\varepsilon)_{\varepsilon \geq 0}$ satisfies an LDP on $\mathcal{C}$ with speed $\varepsilon^{-\beta}$ and rate function $\Lambda^X\colon \mathcal{C} \to [0, +\infty]$ defined as $\Lambda^X(\varphi) := \inf\{\Lambda(z_y^x)\colon \varphi = \sqrt{x} \cdot y,\ y \in \mathrm{BV} \cap \mathcal{C}\}$.*

Corollary 2.2. *The rescaled log stock price process $(t^\beta X_t)_{t \geq 0}$ satisfies an LDP on $\mathbb{R}$ with speed $t^{-\beta}$ and rate function $\Lambda_1^X(u) := \inf\{\Lambda^X(\varphi)\colon \varphi(1) = u\}$, $u \in \mathbb{R}$.*

Proof. Since X_1^ε and $\varepsilon^\beta X_\varepsilon$ are equal in law, $(\varepsilon^\beta X_\varepsilon)_{\varepsilon \geq 0}$ satisfies an LDP with speed $\varepsilon^{-\beta}$ and rate function Λ_1^X by Theorem 2.1 together with the contraction principle (Proposition 3.2 below). Substituting ε with t completes the proof. □

Remark 2.2. 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 considered is of the following form, for the log stock price process:

$$\mathrm{d}X_t = -\tfrac{1}{2}\sigma(Y_t)^2\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}B_t, \qquad Y_t = W_t^H,$$

where B is a standard Brownian motion and W^H is a (possibly correlated) fractional Brownian motion. In order to prove an LDP, they considered a small-noise version of the stochastic differential equation above; namely,

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

It is, of course, tempting to apply their results to the rough Bergomi model. Unfortunately, the following intricacies make this impossible: first, they assumed 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.

Theorem 2.1 and Corollary 2.2 have several potential applications within mathematical finance, two of which we now outline. One application is to establish the small-time behaviour of implied volatility under the rough Bergomi model. This would require, however, first showing that the stock price process in the rough Bergomi model is a true martingale, which is far from trivial and beyond the scope of the present paper. Another potential application is to variance reduction of Monte Carlo pricing of path-dependent options, following the importance sampling approach employed in [34], [36], and [37].

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 $\gamma := \alpha/(\frac{1}{2}\alpha + \frac{5}{4})$. 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 an LDP with speed $\varepsilon^{-2\gamma(\alpha+1/2)}$ and rate function identical to that in Theorem 2.1. This essentially falls into 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 2.3. The structure of the Hilbert space $\mathcal{H}_\rho^{K_\alpha}$ in Corollary 3.2 below precisely determines the rate function Λ^X. In the uncorrelated case $\rho = 0$, $\mathcal{H}_\rho^{K_\alpha}$ (and its inner product) is

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

degenerate, and clearly Λ^* does not make sense. This case needs to be treated separately and we do so in Section 5.

From (2.3), every $z_y^x \in \mathcal{H}_\rho^{K_\alpha}$ has the representation $z_y^x = \mathcal{I}_\rho^{K_\alpha} f$ for some $f \in \mathrm{L}^2$. Therefore, the rate function in Theorem 2.1 can be written as

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

With this formulation, it is easy to see that $\Lambda^X(0) = 0$: denoting $z_y^{\widetilde{x}} := \mathcal{M} z_y^x$ and using that $\widetilde{x} > 0$, it follows that if $\mathrm{I}(\widetilde{x}, y) = 0$ 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.

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 2.1. This proof requires 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

A centred process $(Z_t)_{t \in \mathcal{T}}$ is called Gaussian if, for all $n \in \mathbb{N}$ and any $t_1, \dots, t_n \in \mathcal{T}$, the random variables $Z_{t_1}, \dots, Z_{t_n}$ 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 $(\mathcal{E}, \|\cdot\|_{\mathcal{E}})$ be a real, separable Banach space, and $\mathcal{E}^*$ its topological dual (i.e. the space of all linear functionals on $\mathcal{E}$), with duality relationship $\langle \cdot, \cdot \rangle_{\mathcal{E}^*\mathcal{E}}$. The bilinear functional $\langle \cdot, \cdot \rangle_{\mathcal{E}^*\mathcal{E}} : \mathcal{E}^* \times \mathcal{E} \to \mathbb{R}$ is such that if $\langle x^*, x \rangle_{\mathcal{E}^*\mathcal{E}} = 0$ for all $x^* \in \mathcal{E}^*$ (respectively, $x \in \mathcal{E}$) then $x = 0$ (respectively, $x^* = 0$); see [2, p. 195]. We further let $\mathcal{B}(\mathcal{E})$ denote the Borel σ-algebra of $\mathcal{E}$.

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

In the following proposition [10, Proposition 3.1] we characterise Gaussian measures on Banach spaces.

Proposition 3.1. *Any (centred) Gaussian measure μ on $\mathcal{E}$ 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 taking values in $\mathcal{E}$ induces some measure on $\mathcal{C}$, the space of continuous functions from $\mathcal{T}$ to $\mathbb{R}$. By Proposition 3.1, one can construct a centred Gaussian probability measure μ on $\mathcal{E}$ 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 $\mathcal{E} = \mathcal{C}^d$. For a Gaussian measure μ on $\mathcal{E}$, we introduce the bounded, linear operator $\Gamma : \mathcal{E}^* \to \mathcal{E}$ as

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

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

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

Definition 3.2. (*Carmona and Tehranchi [10, Definition 3.3].*) The reproducing kernel Hilbert space (RKHS) $\mathcal{H}_\mu$ of μ is defined as the completion of $\Gamma(\mathcal{E}^*)$ with the inner product

$$\langle \Gamma(f^*), \Gamma(g^*) \rangle_{\mathcal{H}_\mu} := \int_{\mathcal{E}} \langle f^*, f \rangle_{\mathcal{E}^*\mathcal{E}} \langle g^*, f \rangle_{\mathcal{E}^*\mathcal{E}} \mu(\mathrm{d}f).$$

For the inclusion map $\iota\colon \mathcal{H}_\mu \to \mathcal{E}$, the space $\iota(\mathcal{H}_\mu)$ is dense in $\mathcal{E}$; it follows then for the adjoint map $\iota^*\colon \mathcal{E}^* \to \mathcal{H}_\mu^*$ that $\iota^*(\mathcal{E}^*)$ 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 $\mathcal{E}$, by Definition 3.1, it follows that

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

This yields the following equivalent form of Definition 3.2 for the RKHS of μ; see [15, p. 88].

Definition 3.3. A real, separable Hilbert space $\mathcal{H}_\mu$ such that $\mathcal{H}_\mu \subset \mathcal{E}$ is the RKHS of μ if the following two conditions hold:

- there exists an embedding $I : \mathcal{H}_\mu \to \mathcal{E}$, i.e. an injective continuous map whose image is dense in $\mathcal{E}$;
- any $f^* \in \mathcal{E}^*$ is a centred Gaussian random variable on $\mathcal{E}$ with variance $\|I^* f^*\|^2_{\mathcal{H}_\mu^*}$.

Remark 3.1. The embedding I need not necessarily be the inclusion map.

Remark 3.2. Given a triplet $(\mathcal{E}, \mathcal{H}_\mu, \mu)$, consider the inclusion map $I^*\colon \mathcal{E}^* \to L^2(\mathcal{E}, \mu)$ (we think of $\mathcal{E}^*$ as a dense subset in $\mathcal{H}_\mu^* \simeq \mathcal{H}_\mu$ by ι^*). Since I^* preserves the Hilbert space structure of $L^2(\mathcal{E}, \mu)$, it can be extended to an isometric embedding $\bar{I}^*\colon \mathcal{H}_\mu^* \to L^2(\mathcal{E}, \mu)$ such that $\|\bar{I}^* f^*\|_{\mathcal{H}_\mu^*} = \|f^*\|_{L^2(\mathcal{E},\mu)}$.

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 (2.3). We introduce the following assumption on the function φ defining the operator $\mathcal{I}^\varphi$.

Assumption 3.1. *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 3.1. *Let φ satisfy Assumption 3.1, which makes $\mathcal{I}^\varphi$ 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 $\mathcal{H}^\varphi = \{\mathcal{I}^\varphi f\colon f \in \mathrm{L}^2\}$, with inner product $\langle \mathcal{I}^\varphi f_1, \mathcal{I}^\varphi f_2 \rangle_{\mathcal{H}^\varphi} := \langle f_1, f_2 \rangle_{\mathrm{L}^2}$.*

Proof. See Section 4. □

Corollary 3.1. *The RKHS of the Gaussian measure induced (on $\mathcal{C}$) by $(Z_t)_{t\in\mathcal{T}}$ in (2.1) is $\mathcal{H}^{K_\alpha}$.*

We need to extend Theorem 3.1 (and Corollary 3.1) 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 (2.2), respectively.

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

Theorem 3.2. *Let φ_1, φ_2 satisfy Assumption 3.1, which makes $\mathcal{I}_{\varphi_2}^{\varphi_1}$ injective on L^2. Introduce the $\mathbb{R}^2$-valued process (Y^1, Y^2) as $Y^i := \int_0 \varphi_i(s, \cdot)\,\mathrm{d}W_s^i$ for $i = 1, 2$, where W^1 and W^2 are two standard Brownian motions with correlation $\rho \in [-1, 1] \setminus \{0\}$. Then the RKHS of the measure induced by (Y^1, Y^2) on $\mathcal{C}^2$ is $\mathcal{H}_{\varphi_2}^{\varphi_1} = \{\mathcal{I}_{\varphi_2}^{\varphi_1} f : f \in \mathrm{L}^2\}$, with inner product $\langle \mathcal{I}_{\varphi_2}^{\varphi_1} f_1, \mathcal{I}_{\varphi_2}^{\varphi_1} f_2 \rangle_{\mathcal{H}_{\varphi_2}^{\varphi_1}} := \langle f_1, f_2 \rangle_{\mathrm{L}^2}$.*

Proof. See Section 4. □

Corollary 3.2. *The RKHS of the measure induced (on $\mathcal{C}^2$) by the process $((Z_t, B_t))_{t\in\mathcal{T}}$ is $\mathcal{H}_\rho^{K_\alpha}$.*

3.2. Large deviations for Gaussian measures

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

$$\int_{\mathcal{E}} \mathrm{e}^{\mathrm{i}\langle y, x\rangle_{\mathcal{E}^*\mathcal{E}}}\,\mu(\mathrm{d}x) = \exp\left(-\frac{1}{2}C_\mu(y, y)\right),$$

where $C_\mu : \mathcal{E}^* \times \mathcal{E}^* \to [0, +\infty)$ is a bilinear, symmetric map. We define $\Lambda_\mu^* : \mathcal{E} \to \mathbb{R}$ as $\Lambda_\mu^*(x) := \sup\{\langle y, x\rangle_{\mathcal{E}^*\mathcal{E}} - \frac{1}{2}C_\mu(y, y) : y \in \mathcal{E}^*\}$ on $\mathcal{E}$. The following lemma was proved in [15, Lemma 3.4.2].

Lemma 3.1. *The following three statements hold for the measure μ:*

(i) *there exists $\alpha > 0$ such that $\int_{\mathcal{E}} \exp(\alpha\|x\|_{\mathcal{E}}^2)\mu(\mathrm{d}x)$ is finite;*

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

(iii) *Λ_μ^* defines a rate function on $\mathcal{E}$ and satisfies $\Lambda_\mu^*(ay) = a^2\Lambda_\mu^*(y)$ for all $a \in \mathbb{R}$.*

The reader may recognise statement (i) in Lemma 3.1 above as Fernique's theorem [16]. For an $\mathcal{E}$-valued Gaussian random variable X with distribution μ, define $X^\varepsilon := \varepsilon^{1/2}X$, with law μ_ε. Then the following theorem holds; see [15, Theorem 3.4.5].

Theorem 3.3. *The sequence of probability measures $(\mu_\varepsilon)_{\varepsilon\ge 0}$ satisfies an LDP on $\mathcal{E}$ with speed ε^{-1} and rate function Λ_μ^*.*

Remark 3.3. Theorem 3.3 implies, in particular, that the standard Brownian motion $(W_t)_{t\ge 0}$ satisfies an LDP on $\mathbb{R}$ with speed t^{-1}, since W_t and $\sqrt{t}W_1$ are equal in law.

Corollary 3.3. *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 an LDP on $\mathbb{R}$ with speed $t^{-\beta}$ and rate function $\Lambda_\mu^*(x) := x^2/2\eta^2$ for $x \in \mathbb{R}$.*

Proof. Here, $\mathcal{E} = \mathbb{R}$ and $\langle u, v\rangle_{\mathcal{E}^*\mathcal{E}} = uv$. Since Z_t and $t^{\beta/2}Z_1$ are equal in law, and $\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}yx}\mathbb{P}(Z_1 \in \mathrm{d}x) = \exp(-\frac{1}{2}y^2\eta^2)$, taking $C_\mu(x, y) \equiv xy\eta^2$, the proof follows from Theorem 3.3 and Remark 4.1. □

The following two results will be essential for establishing an 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.

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

Proposition 3.2. (Contraction principle; [13, Theorem 4.2.1].) *Let $\mathcal{E}$ and $\widetilde{\mathcal{E}}$ be two Hausdorff topological spaces and let $f\colon \mathcal{E} \to \widetilde{\mathcal{E}}$ be a continuous mapping. Let $(\nu_\varepsilon)_{\varepsilon\geq 0}$ and $(\widetilde{\nu}_\varepsilon)_{\varepsilon\geq 0}$ be two families of probability measures on $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$ and $(\widetilde{\mathcal{E}}, \mathcal{B}(\widetilde{\mathcal{E}}))$, respectively, such that $\widetilde{\nu}_\varepsilon = \nu_\varepsilon \circ f^{-1}$ for each $\varepsilon > 0$. If $(\nu_\varepsilon)_{\varepsilon\geq 0}$ satisfies an LDP on $\mathcal{E}$ with rate function Λ then $(\widetilde{\nu}_\varepsilon)_{\varepsilon\geq 0}$ satisfies an LDP on $\widetilde{\mathcal{E}}$ with rate function*

$$\widetilde{\Lambda}(y) := \inf\{\Lambda(x)\colon y = f(x)\}.$$

Theorem 3.4. (Deuschel and Stroock [15, Theorem 3.4.12].) *Let B be a d-dimensional Gaussian process, inducing a measure μ on $(\mathcal{C}^d, \mathcal{B}(\mathcal{C}^d))$ with RKHS $\mathcal{H}_\mu$. Then $(\varepsilon\mu)_{\varepsilon\geq 0}$ satisfies an LDP with speed ε^{-1} and rate function*

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

In order to deal with the stochastic differential equation (2.5), we have to consider the stochastic integral $\int_0^\cdot \sqrt{v^\varepsilon_s}\,\mathrm{d}B^\varepsilon_s$. Assuming that the sequence $(\sqrt{v^\varepsilon_s}, B^\varepsilon_s)$ converges weakly as $\varepsilon \to 0$ yields, under some conditions, a weak convergence for the stochastic integral [29], [31]. However, in order to state a large deviations principle, we need a stronger result, proved by Garcia [23]. Before stating it (Theorem 3.5 below), though, we introduce the following class of sequences of stochastic processes.

Definition 3.4. (*Garcia [23, Definition 1.1].*) Let $\mathcal{U}$ 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

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

Theorem 3.5. (Garcia [23, Theorem 1.2].) *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 an LDP with rate function Λ then the sequence of stochastic integrals $(X^\varepsilon \cdot Y^\varepsilon)_{\varepsilon\geq 0}$ satisfies an LDP with rate function $\widehat{\Lambda}(\varphi) := \inf\{\Lambda(z^x_y)\colon \varphi = x\cdot y,\ y \in \mathrm{BV}\}$.*

4. Proofs of the main results

Proof of Theorem 3.1. Let Assumption 3.1 hold for a given function $\varphi \in \mathrm{L}^2$. The operator $\mathcal{I}^\varphi$ in (2.3) is surjective on $\mathcal{H}^\varphi$. 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 [38, 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 $\langle \mathcal{I}^\varphi f_1, \mathcal{I}^\varphi f_2\rangle_{\mathcal{H}^\varphi} := \langle f_1, f_2\rangle_{\mathrm{L}^2}$ defines an inner product on $\mathcal{H}^\varphi$ and, hence, $(\mathcal{H}^\varphi, \langle\cdot,\cdot\rangle_{\mathcal{H}^\varphi})$ is a real inner product space. In order for $\mathcal{H}^\varphi$ to satisfy Definition 3.3, 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 $\mathcal{H}^\varphi$. Therefore, $\|\mathcal{I}^\varphi f_n - \mathcal{I}^\varphi f_m\|_{\mathcal{H}^\varphi} = \|f_n - f_m\|_{\mathrm{L}^2} \to 0$ as $n, m \to \infty$. Since L^2 is a complete (Hilbert) space, there exists a

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

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

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

which converges to 0 as $n \to \infty$. Therefore, $f = \widetilde{f}$, $\mathcal{I}^\varphi f \in \mathcal{H}^\varphi$, and $\mathcal{H}^\varphi$ 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 $\mathcal{H}^\varphi$, which is then separable.

We now find a dense embedding $I\colon \mathcal{H}^\varphi \to \mathcal{C}$ as in Definition 3.3. Since $\mathcal{H}^\varphi \subset \mathcal{C}$, take the embedding to be the inclusion map $I = \iota$. By [11, Lemma 2.1], the conditions on ϕ in Assumption 3.1 imply that $\mathcal{H}^\varphi$ 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 $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$, and f^* is a centred, real Gaussian random variable on $(\mathcal{E}, \mathcal{B}(\mathcal{E}), \mu)$ by Definition 3.1 . In turn, Remark 3.2 implies that I^*, the dual of I, admits an isometric embedding $\bar{I}^*$ such that $\|\bar{I}^* f^*\|^2_{(\mathcal{H}^\varphi)^*} = \|f^*\|^2_{L^2(\mathcal{E},\mu)} = \int_{\mathcal{E}} (f^*)^2\,\mathrm{d}\mu = \mathcal{V}(f^*)$; hence, $\mathcal{H}^\varphi$ is the RKHS of μ. □

Proof of Theorem 3.2. We proceed in a similar manner to the proof of Theorem 3.1. Let φ_1, φ_2 satisfy Assumption 3.1. Clearly, the operator $\mathcal{I}^{\varphi_1}_{\varphi_2}$ in (2.3) is surjective on $\mathcal{H}^{\varphi_1}_{\varphi_2} \subset \mathcal{C}^2$. By Titchmarsh's convolution theorem [38, 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, $\langle \mathcal{I}^{\varphi_1}_{\varphi_2} f_1, \mathcal{I}^{\varphi_1}_{\varphi_2} f_2\rangle_{\mathcal{H}^{\varphi_1}_{\varphi_2}} := \langle f_1, f_2\rangle_{\mathrm{L}^2}$ is a well-defined inner product and, following the proof of Theorem 3.1, $(\mathcal{H}^{\varphi_1}_{\varphi_2}, \langle\cdot,\cdot\rangle_{\mathcal{H}^{\varphi_1}_{\varphi_2}})$ is a real, separable Hilbert space. To find a dense embedding $I\colon \mathcal{H}^{\varphi_1}_{\varphi_2} \to \mathcal{C}^2$, take I as the inclusion map ι; then the conditions on ϕ_1, ϕ_2 in Assumption 3.1 imply that $\mathcal{H}^{\varphi_1}_{\varphi_2}$ 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 $(\mathcal{C}^2, \mathcal{B}(\mathcal{C}^2), \mu)$, where μ denotes the measure induced by (Y^1, Y^2), and $\mathcal{V}(f^*) = \int_{\mathcal{C}^2} (f^*)^2\,\mathrm{d}\mu_2 = \|f^*\|^2_{L^2(\mathcal{C}^2,\mu_2)} = \|\iota^* f^*\|^2_{(\mathcal{H}^{\varphi_1}_{\varphi_2})^*}$, so that by Definition 3.3, $\mathcal{H}^{\varphi_1}_{\varphi_2}$ is the RKHS of the measure induced by (Y^1, Y^2) on $\mathcal{C}^2$. □

Proof of Theorem 2.1. Let $(Z^\varepsilon, B^\varepsilon)$ be as in (2.4). From Theorem 3.4 and Corollary 3.2, the sequence $((Z^\varepsilon, B^\varepsilon))_{\varepsilon\geq 0}$ satisfies an LDP with speed $\varepsilon^{-\beta}$ and rate function (with $\mathcal{H}^{K_\alpha}_\rho$ given in Corollary 3.2)

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

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 (2.6) is continuous with respect to the $\mathcal{C}(\mathcal{T}\times\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R})$ 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

$$\left\|\mathcal{M}\begin{pmatrix} f+\delta^f \\ g+\delta^g \end{pmatrix} - \mathcal{M}\begin{pmatrix} f \\ g \end{pmatrix}\right\|_\infty = \sup_{t\in\mathcal{T},\,\varepsilon>0} \{|(\mathfrak{m}(f+\delta^f))(t,\varepsilon) - (\mathfrak{m} f)(t,\varepsilon)| + |\delta^g(t)|\}$$

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

$$\leq \sup_{t\in\mathcal{T},\,\varepsilon>0}\{v_0^{1+\beta}\exp\bigl(-\tfrac{1}{2}\eta^2(\varepsilon t)^\beta\bigr)|\mathrm{e}^{f(t)}||\mathrm{e}^{\delta^f(t)}-1|\}+\sup_{t\in\mathcal{T}}|\delta^g(t)|$$
$$\leq C\sup_{t\in\mathcal{T}}|\mathrm{e}^{\delta^f(t)}-1|+\sup_{t\in\mathcal{T}}|\delta^g(t)|$$

for some strictly positive constant C. The right-hand side clearly tends to 0 as $(\delta^f,\delta^g)\to 0$ with respect to the sup norm on $\mathcal{C}^2$ and, hence, $\mathcal{M}$ is a continuous operator. The contraction principle (Proposition 3.2) therefore implies that the sequence $((v^\varepsilon, B^\varepsilon))_{\varepsilon\geq 0}$ satisfies an LDP on $\mathcal{C}(\mathcal{T}\times\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R})$, with speed $\varepsilon^{-\beta}$ and rate function Λ. Since $\mathcal{M}$ is clearly a bijection, the rate function Λ may then be expressed as $\Lambda(z_{y_1}^{x_1})=\Lambda^*(\mathcal{M}^{-1}(z_{y_1}^{x_1}))$ for any $(x_1,y_1)\in\mathcal{C}^2$.

In the second step we will apply Theorem 3.5 to prove that the sequence of stochastic integrals $(\mathrm{I}(v^\varepsilon,B^\varepsilon)(\cdot))_{\varepsilon\geq 0}:=(\int_0^\cdot\sqrt{v_s^\varepsilon}\,\mathrm{d}B_s^\varepsilon)_{\varepsilon\geq 0}$ satisfies an LDP. Since $B^\varepsilon=\varepsilon^{\alpha+1/2}B$ by (2.5), we can write the stochastic integral $\mathrm{I}(v^\varepsilon,B^\varepsilon)(\cdot)=\mathrm{I}(\varepsilon^{2\alpha}v^\varepsilon,\sqrt{\varepsilon}B)(\cdot)$, which holds almost surely, and, therefore (see [23, Example 2.1]), the sequence of (semi)-martingales $(\sqrt{\varepsilon}B)_{\varepsilon\geq 0}$ is UET in the sense of Definition 3.4. Since the sequence $(\sqrt{\varepsilon^{2\alpha}v^\varepsilon})_{\varepsilon\geq 0}$ consists of càdlàg, $(\mathcal{F}_t)$-adapted processes, Theorem 3.5 implies that the sequence of stochastic integrals $(\mathrm{I}(v^\varepsilon,B^\varepsilon)(\cdot))_{\varepsilon\geq 0}$ satisfies an LDP with speed $\varepsilon^{-\beta}$ and rate function

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

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 $(\mathrm{I}(v^\varepsilon,B^\varepsilon))_{\varepsilon\geq 0}$ are exponentially equivalent. For any $\delta>0$, it follows that

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

where $b_\varepsilon:=\delta/v_0\varepsilon^{1+\beta}$. Using the fact 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

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

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

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

This then implies that

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

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

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

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

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

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

and, therefore,

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

which is precisely the definition of exponential equivalence [13, Definition 4.2.10]. Then, by [13, Theorem 4.2.13], the sequence $(X^{\varepsilon})_{\varepsilon\geq 0}$ satisfies an LDP with speed $\varepsilon^{-\beta}$ and rate function Λ^X. □

Proof of Theorem 3.3. Let $\underline{X} \coloneqq (X^1, \dots, X^n)$ be an n-dimensional random vector taking values on $\mathcal{E}^n$, where each X^k has distribution μ, so that the average $(1/n)\sum_{k=1}^n X^k$ has distribution $\mu_{1/n}$. Lemma 3.1 implies that $\int_{\mathcal{E}} \exp(\alpha\|x\|_{\mathcal{E}}^2)\mu_{1/n}(\mathrm{d}x)$ is finite for some $\alpha > 0$, and [15, Theorem 3.3.11] yields an LDP for the sequence $(\mu_{1/n})_{n\geq 1}$, with rate function Λ_μ^*. Define now $n(\varepsilon) \coloneqq \lfloor 1/\varepsilon\rfloor \vee 1$ and $\ell(\varepsilon) \coloneqq \varepsilon n(\varepsilon)$ for $\varepsilon > 0$, noting that $\ell(\varepsilon) \in [1-\varepsilon, 1)$ for $\varepsilon \in (0, \frac{1}{2})$ and in $[\frac{1}{2}, 1]$ for $\varepsilon \in [\frac{1}{2}, 1]$; 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 $\mathcal{E}$, we define the dilated set $\widetilde{B} \coloneqq \{\ell^{-1/2}x\colon \text{ for all } \ell \in [\frac{1}{2}, 1],\ x \in B\}$, so that

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

The large deviations upper bound then follows from the obvious equalities

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

Now, for any x in any open set $C \subset \mathcal{E}$, 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 $\varepsilon_0 \in (0, \frac{1}{2}]$. The large deviations lower bound then follows from

$$\begin{aligned}
\liminf_{\varepsilon\downarrow 0} \varepsilon\log\mu_\varepsilon(C) &= \liminf_{\varepsilon\downarrow 0} \frac{\ell(\varepsilon)}{n(\varepsilon)}\log\mu_{1/n(\varepsilon)}(\ell(\varepsilon)^{-1/2}C)\\
&\geq \liminf_{n\uparrow\infty} \frac{1}{n}\log\mu_{1/n}(\mathcal{O}_x)\\
&\geq -\inf_{y\in\mathcal{O}_x} \Lambda_\mu^*(y)\\
&\geq -\Lambda_\mu^*(x).
\end{aligned}$$

□

Remark 4.1. The proof of Theorem 3.3 still 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.

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

5. Large deviations for the uncorrelated rough Bergomi model

We now treat the special case of (2.2), where the Brownian motions W and B are independent ($\rho = 0$). Following arguments similar to Corollary 3.2, and mimicking (2.3), we introduce the L^2 operator $\mathcal{I}^0$ as

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

so that the RKHS (on $\mathcal{C}^2$) of the measure induced by (Z, B) is

$$\mathcal{H} \coloneqq \{\mathcal{I}^0(f_1, f_2) \colon f_1, f_2 \in \mathrm{L}^2\},$$

with inner product

$$\langle \mathcal{I}^0(f_1, f_2), \mathcal{I}^0(g_1, g_2)\rangle_{\mathcal{H}} \coloneqq \langle f_1, g_1\rangle_{\mathrm{L}^2} + \langle f_2, g_2\rangle_{\mathrm{L}^2} \quad \text{for any } f_1, f_2, g_1, g_2 \in \mathrm{L}^2.$$

Similarly to Theorem 2.1, [15, Theorem 3.4.12] yields an LDP on $\mathcal{C}^2$ for $((Z^\varepsilon, B^\varepsilon))_{\varepsilon\geq 0}$ with speed $\varepsilon^{-\beta}$ and rate function

$$\Lambda(z_y^x) \coloneqq \begin{cases} \frac{1}{2}\|z_y^x\|_{\mathcal{H}}^2 & \text{if } (x, y)^\top \in \mathcal{H}, \\ +\infty & \text{otherwise.} \end{cases}$$

This, in turn, yields an LDP for $((v^\varepsilon, B^\varepsilon))_{\varepsilon\geq 0}$ in (2.4) on $\mathcal{C}^2$ with speed $\varepsilon^{-\beta}$ and rate function $\widetilde{\Lambda}(z_y^x) \coloneqq \inf\{\Lambda(z_{y^*}^{x^*}) \colon z_y^x = \mathcal{M} z_{y^*}^{x^*}\}$, where the operator $\mathcal{M}$ is defined in (2.6). In the same vein as Theorem 2.1, Theorem 3.5 yields an LDP for $(\int_0^\cdot \sqrt{v_s^\varepsilon}\,\mathrm{d}B_s^\varepsilon)_{\varepsilon\geq 0}$ on $\mathcal{C}$ with speed $\varepsilon^{-\beta}$ and rate function $\widehat{\Lambda}^X$, defined as

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

with $\mathfrak{m}$ introduced in (2.7). Following an argument identical to the one presented in the proof of Theorem 2.1, we conclude that $(X^\varepsilon)_{\varepsilon>0}$ satisfies an LDP with speed $\varepsilon^{-\beta}$ and rate function $\widehat{\Lambda}^X$.

Acknowledgements

AJ acknowledges financial support from the EPSRC First Grant EP/M008436/1. HS thanks the EPSRC CDT in Financial Computing and Analytics for financial support. MSP acknowledges partial support from CREATES (DNRF78), funded by the Danish National Research Foundation.

References

[1] Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*. Springer, New York.

[2] Aliprantis, C. D. and Border, K. C. (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd edn. Springer, Berlin.

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

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press

[4] Bajgrowicz, P., Scaillet, O. and Treccani, A. (2015). Jumps in high-frequency data: spurious detections, dynamics, and news. *Manag. Sci.* **62,** 2198–2217.
[5] Bayer, C., Friz, P. and Gatheral, J. (2016). Pricing under rough volatility. *Quant. Finance* **16,** 887–904.
[6] Bayer, C. *et al.* (2017). Short-time near-the-money skew in rough fractional volatility models. Preprint. Available at https://arxiv.org/pdf/1703.05132.pdf.
[7] Bennedsen, M., Lunde, A. and Pakkanen, M. S. (2016). Decoupling the short- and long-term behavior of stochastic volatility. Preprint. Available at https://arxiv.org/pdf/1610.00332.pdf.
[8] Bennedsen, M., Lunde, A. and Pakkanen, M. S. (2017). Hybrid scheme for Brownian semistationary processes. *Finance Stoch.* **21,** 931–965.
[9] Biagini, F., Hu, Y., Øksendal, B. and Zhang, T. (2008). *Stochastic Calculus for Fractional Brownian Motion and Applications*. Springer, London.
[10] Carmona, R. A. and Tehranchi, M. R. (2006). *Interest Rate Models: An Infinite Dimensional Stochastic Analysis Perspective*. Springer, Berlin.
[11] Cherny, A. (2008). Brownian moving averages have conditional full support. *Ann. Appl. Prob.* **18,** 1825–1830.
[12] Christensen, K., Oomen, R. C. A. and Podolskij, M. (2014). Fact or friction: jumps at ultra high frequency. *J. Financial Econom.* **114,** 576–599.
[13] Dembo, A. and Zeitouni, O. (2010). *Large Deviations Techniques and Applications*. Springer, Berlin.
[14] Deuschel, J. D., Friz, P. K., Jacquier, A. and Violante, S. (2014). Marginal density expansions for diffusions and stochastic volatility I: theoretical foundations. *Commun. Pure Appl. Math.* **67,** 40–82.
[15] Deuschel, J.-D. and Stroock, D. W. (1989). *Large Deviations*. Academic Press, Boston.
[16] Fernique, X. (1970). Intégrabilité des vecteurs gaussiens. *C. R. Acad. Sci. Paris A-B* **270,** A1698–A1699.
[17] Figueroa-López, J. E. and Forde, M. (2012). The small-maturity smile for exponential Lévy models. *SIAM J. Financial Math.* **3,** 33–65.
[18] Forde, M. and Jacquier, A. (2009). Small-time asymptotics for implied volatility under the Heston model. *Int. J. Theoret. Appl. Finance* **12,** 861–876.
[19] Forde, M. and Zhang, H. (2017). Asymptotics for rough stochastic volatility models. *SIAM J. Financial Math.* **8,** 114–145.
[20] Friz, P. K. *et al.* (eds) (2015). *Large Deviations and Asymptotic Methods in Finance*. Springer, Cham.
[21] Fukasawa, M. (2011). Asymptotic analysis for stochastic volatility: martingale expansion. *Finance Stoch.* **15,** 635–654.
[22] Fukasawa, M. (2017). Short-time at-the-money skew and rough fractional volatility. *Quant. Finance* **17,** 189–198.
[23] Garcia, J. (2008). A large deviation principle for stochastic integrals. *J. Theoret. Prob.* **21,** 476–501.
[24] Gatheral, J., Jaisson, T. and Rosenbaum, M. (2018). Volatility is rough. *Quant. Finance* **18,** 933–949.
[25] Guillin, A. (2003). Averaging principle of SDE with small diffusion: moderate deviations. *Ann. Prob.* **31,** 413–443.
[26] Jacquier, A., Keller-Ressel, M. and Mijatović, A. (2013). Large deviations and stochastic volatility with jumps: asymptotic implied volatility for affine models. *Stochastics* **85,** 321–345.
[27] Jacquier, A., Martini, C. and Muguruza, A. (2018). On VIX futures in the rough Bergomi model. *Quant. Finance* **18,** 45–61.
[28] Jacquier, A. and Roome, P. (2015). Asymptotics of forward implied volatility. *SIAM J. Financial Math.* **6,** 307–351.
[29] Jakubowski, A., Mémin, J. and Pagès, G. (1989). Convergence en loi des suites d'intégrales stochastiques sur l'espace D^1 de Skorokhod. *Prob. Theory Relat. Fields* **81,** 111–137.
[30] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.
[31] Kurtz, T. G. and Protter, P. E. (1996). Weak convergence of stochastic integrals and differential equations. II. Infinite-dimensional case. In *Probabilistic Models for Nonlinear Partial Differential Equations* (Lecture Notes Math. **1627**). Springer, Berlin, pp. 197–285.
[32] Lévy, P. (1953). Random functions: general theory with special reference to Laplacian random functions. *Univ. California Publ. Statist.* **1,** 331–390.
[33] Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10,** 422–437.
[34] Morse, M. and Spiliopoulos, K. (2018). Importance sampling for slow-fast diffusions based on moderate deviations. Preprint. Available at https://arxiv.org/pdf/1805.10229.pdf.
[35] Olver, F. W. J. (1997). *Asymptotics and Special Functions*. A K Peters/CRC Press, Boca Raton, FL.
[36] Robertson, S. (2010). Sample path large deviations and optimal importance sampling for stochastic volatility models. *Stoch. Process. Appl.* **120,** 66–83.
[37] Spiliopoulos, K. (2013). Large deviations and importance sampling for systems of slow-fast motion. *Appl. Math. Optimization* **67,** 123–161.
[38] Titchmarsh, E. C. (1926). The zeros of certain integral functions. *Proc. London Math. Soc.* **25,** 283–302.

https://doi.org/10.1017/jpr.2018.72 Published online by Cambridge University Press