Short-time near-the-money skew in rough fractional volatility models

C. Bayer, P. K. Friz, A. Gulisashvili, B. Horvath & B. Stemper

To cite this article: C. Bayer, P. K. Friz, A. Gulisashvili, B. Horvath & B. Stemper (2019) Short-time near-the-money skew in rough fractional volatility models, Quantitative Finance, 19:5, 779-798, DOI: 10.1080/14697688.2018.1529420

To link to this article: https://doi.org/10.1080/14697688.2018.1529420

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 13 Nov 2018.

Submit your article to this journal

Article views: 491

View Crossmark data

Citing articles: 8

View citing articles
Short-time near-the-money skew in rough fractional volatility models

C. BAYER†, P. K. FRIZ‡*, A. GULISASHVILI§, B. HORVATH¶ and B. STEMPER†‡

†WIAS Berlin, Berlin, Germany
‡TU Berlin, Berlin, Germany
§Ohio University, Athens, USA
¶Imperial College London, London, UK

(Received 22 March 2017; accepted 7 September 2018; published online 13 November 2018)

We consider rough stochastic volatility models where the driving noise of volatility has fractional scaling, in the ‘rough’ regime of Hurst parameter $H < 1/2$. This regime recently attracted a lot of attention both from the statistical and option pricing point of view. With focus on the latter, we sharpen the large deviation results of Forde-Zhang [Asymptotics for rough stochastic volatility models. SIAM J. Financ. Math., 2017, 8(1), 114–145] in a way that allows us to zoom-in around the money while maintaining full analytical tractability. More precisely, this amounts to proving higher order moderate deviation estimates, only recently introduced in the option pricing context. This in turn allows us to push the applicability range of known at-the-money skew approximation formulae from CLT type log-moneyness deviations of order $t^{1/2}$ (works of Alòs, León & Vives and Fukasawa) to the wider moderate deviations regime.

Keywords: Rough stochastic volatility model; European option pricing; Small-time asymptotics; Moderate deviations

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

1. Introduction

Since the groundbreaking work of Gatheral et al. (2014), the past two years have brought about a gradual shift in volatility modeling, leading away from classical diffusive stochastic volatility models towards so-called rough volatility models. The term was coined in Gatheral et al. (2014) and Bayer et al. (2016), and it essentially describes a family of (continuous-path) stochastic volatility models where the driving noise of the volatility process has Hölder regularity lower than Brownian motion, typically achieved by modeling the fundamental noise innovations of the volatility process as a fractional Brownian motion with Hurst exponent (and hence Hölder regularity) $H < 1/2$. Here, we would also like to mention pioneering work on asymptotics for rough volatility models in Alòs et al. (2007) and Fukasawa (2011). A major appeal of such rough volatility models lies in the fact that they effectively capture several stylized facts of financial markets both from a statistical (Gatheral et al. 2014; Bennedsen et al. 2016) and an option-pricing point of view (Bayer et al. 2016). In particular, with regards to the latter point of view, a widely observed empirical phenomenon in equity markets is the ‘steepness of the smile on the short end’ describing the fact that as time to maturity becomes small the empirical implied volatility skew follows a power law with negative exponent, and thus becomes arbitrarily large near zero. While standard stochastic volatility models with continuous paths struggle to capture this phenomenon, predicting instead a constant at-the-money implied volatility behavior on the short end (Gatheral 2011), models in the fractional stochastic volatility family (and more specifically so-called rough volatility models) constitute a class, well-tailored to fit empirical implied volatilities for short dated options.

Typically, the popularity of asset pricing models hinges on the availability of efficient numerical pricing methods. In the case of diffusions, these include Monte Carlo estimators, PDE discretization schemes, asymptotic expansions and transform methods. With fractional Brownian motion being the prime example of a process beyond the semimartingale framework, most currently prevalent option pricing methods...
to a non-Markovian case, and using the new expansion, we can high resolution scale. To be more specific, our paper adds more details). On the other hand, it allows us to zoom in on volatility is driven by a rough uncertainty approaches zero, strikes with acceptable bid-ask spreads one hand, it reflects the market reality that as time to matur-
tion above, by rescaling the strike with respect to the time

Rather recently, Friz et al. (2018) introduced another regime called moderately-out-of-the-money (MOTM), which, in a

cial case of the Rough Heston model – on an explicit formula

The paper is organized as follows: In Section 2 we set the

take completeness) a number of works, either based on

large deviations or central limit type expansion regime, that inspired this work: Alòs et al. (2007), Fukasawa (2011), Deuschel et al. (2014a), Deuschel et al. (2014b) and Fuku-
and Tankov (2016) and especially Forde and Zhang (2017).

2. Exposition and assumptions

We consider a rough stochastic volatility model, normalized to \( r = 0 \) and \( S_0 = 1 \), of the form suggested by Forde and Zhang (2017)\)

\[
\frac{dS_t}{S_t} = \sigma (\hat{B}_t) \left( d\tilde{\rho}W_t + \rho B_t \right).
\]

Here \((W, B)\) are two independent standard Brownian motions, \( \rho \in (-1, 1) \) a correlation parameter, and \( \tilde{\rho}^2 = 1 - \rho^2 \). Then \( \tilde{\rho}W + \rho B \) is another standard Brownian motion which has constant correlation \( \rho \) with the factor \( B \), which drives the stochastic volatility

\[
\sigma_{\text{stoch}} (t, \omega) := \sigma (\hat{B}_t (\omega)) \equiv \sigma (\tilde{B}).
\]

Here \( \sigma(\cdot) \) is some real-valued function, typically smooth but not bounded, and we will denote by \( \sigma_0 := \sigma(0) \) the spot volatility, with \( \hat{B} \) a Gaussian (Volterra) process of the form

\[
\hat{B}_t = \int_0^t K (t, s) dB_s,
\]

for some kernel \( K \), which shall be further specified in Assump-
tions 2.1 and 2.5 below. The log-price \( X_t = \log(S_t) \) satisfies

\[
\frac{dX_t}{X_t} = -\frac{1}{2} \sigma^2 (\hat{B}_t) \left( dt + \sigma (\hat{B}_t) \left( d\tilde{\rho}W_t + \rho B_t \right) \right),
\]

\[
X_0 = 0.
\]

Recall that by Brownian scaling, for fixed \( t > 0 \),

\[
(B_{ts}, W_{ts})_{s \geq 0} \xrightarrow{\text{law}} \left( \epsilon (B_t, W_t) \right)_{s \geq 0}, \quad \text{where} \ \epsilon \equiv \epsilon (t) \equiv t^{1/2}.
\]

As a direct consequence, classical short-time SDE problems can be analyzed as small-noise problems on a unit time hori-
zon. For our analysis, it will also be crucial to impose such a scaling property on the Gaussian process \( \hat{B} \) (more precisely, on the kernel \( K \) in (2) driving the volatility process in our model:

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

\[
(\hat{B}_t : 0 \leq s \leq t_0) \xrightarrow{\text{law}} (\tilde{\epsilon} \hat{B}_s : 0 \leq s \leq t_0).
\]
In fact, we will always have
\[ \hat{k} \equiv \hat{k}(t) \equiv t^{H} = \varepsilon^{2H}, \]
which covers the examples of interest, in particular standard fractional Brownian motion \( \hat{B} = B^{H} \) or Riemann-Liouville fBM with explicit kernel \( K(t, s) = \sqrt{2H} |t - s|^{H-1/2} \). (This is very natural, even from a general perspective of self-similar processes, see Lamperti 1962.)

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

Remark 2.2 It should be possible to replace the fractional Brownian motion by a certain fractional Ornstein-Uhlenbeck process in the results obtained in this paper. Intuitively, this replacement creates a negligible perturbation (for \( \varepsilon \ll 1 \)) in the results obtained in this paper. Intuitively, this might be possible to employ a version of this condition in the fBM environment. A similar situation was in fact encountered in Cass and Friz (2010), where fractional scaling at times near zero was important. To quantify the perturbation, the authors of Cass and Friz (2010) introduced an easy to verify coupling condition (see Corollary 2 in Cass and Friz 2010). It should be possible to employ a version of this condition in the present paper to justify the replacement mentioned above. We will however not pursue this point further here.

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

General facts on large deviations of Gaussian measures on Banach spaces (Deuschel and Stroock 1989) such as the path space \( C([0, 1], \mathbb{R}^{3}) \) imply that a large deviation principle holds for the triple \( \{\varepsilon_{\cdot}(W, B, \hat{B}) : \hat{\varepsilon} > 0\} \), with speed \( \hat{\varepsilon}^{2} \) and rate function
\[
\frac{1}{2} \|h\|_{H_{0}^{1}}^{2} + \frac{1}{2} \|f\|_{H_{0}^{1}}^{2}, \quad f, h \in H_{0}^{1} \text{ and } \hat{f} = K\hat{f},
\]
otherwise,
\[ (4) \]
where
\[ K\hat{f}(t) := \int_{0}^{t} K(t, s)\hat{f}(s) \, ds \]
for \( f \in H_{0}^{1} \), the space of absolutely continuous paths with \( L^{2} \) derivative
\[
H_{0}^{1} := \left\{ f : [0, 1] \to \mathbb{R} \text{continuous} \left| \|f\|_{H_{0}^{1}}^{2} = \int_{0}^{1} |f(s)|^{2} \, ds < \infty, \ f(0) = 0 \right. \right\}.
\]
\[ (5) \]
This enables us to derive a large deviations principle for \( X_{\cdot} \) in (3): the (local) small-time self-similarity property of \( \hat{B} \) (Assumption 2.1) implies that \( X_{\cdot} \equiv X_{\cdot}^{\varepsilon} \) where
\[
dX_{\cdot}^{\varepsilon} = \sigma(\hat{\varrho}_{\cdot})\varepsilon \, d(\varrho W_{\cdot} + \rho B_{\cdot}) - \frac{1}{2} \varepsilon^{2} \sigma^{2}(\hat{\varrho}_{\cdot}) \, d\varrho, \quad X_{0}^{\varepsilon} = 0.
\]
For what follows, it will be convenient to consider a rescaled version of (3)
\[
d\hat{X}^{\varepsilon} = d\left(\frac{\hat{\varrho}}{\varepsilon}, \hat{X}^{\varepsilon}\right) = \sigma(\hat{\varrho}_{\cdot})\hat{\varrho} \, dW_{\cdot} + \rho_{\cdot} \, dB_{\cdot} - \frac{1}{2} \varepsilon \sigma^{2}(\hat{\varrho}_{\cdot}) \, d\varrho, \quad X_{0}^{\varepsilon} = 0.
\]
Under a linear growth condition on the function \( \sigma \), Forde and Zhang (2017) use the extended contraction principle to establish a large deviations principle for \( (\hat{X}^{\varepsilon}) \) with speed \( \varepsilon^{2} \). More precisely, with
\[
\varphi_{1}(h, f) := \Phi_{1}(h, f, \hat{f}) = \int_{0}^{1} \sigma(\hat{f}) \, d(\varrho h + \rho f), \quad (6)
\]
the rate function is given by
\[
I(x) = \inf_{h, f \in H_{0}^{1}} \left\{ \frac{1}{2} \int_{0}^{1} h^{2} \, dt + \frac{1}{2} \int_{0}^{1} f^{2} \, dt : \varphi_{1}(h, f) = x \right\}
= \inf_{f \in H_{0}^{1}} \left\{ \frac{1}{2} \left( x - \rho \left(\sigma(\hat{f}), \hat{f}\right) \right)^{2} + \frac{1}{2} \int_{0}^{1} f^{2} \, dt \right\}, \quad (7)
\]
where \( \{., .\} \) denotes the inner product on \( L^{2}([0, 1], dt) \). Several other proofs (under varying assumptions on \( \sigma \)) have appeared since (Jacquier et al. 2017; Bayer et al. 2017; Gulisashvili 2017).

As a matter of fact, this paper relies on moderate – rather than large – deviations, as emphasized in (iiic) below. To this end, let us make

Assumption 2.4

(i) (Positive spot vol) Assume \( \sigma : \mathbb{R} \to \mathbb{R} \) is smooth with \( \sigma_{0} := \sigma(0) > 0 \).
(ii) (Roughness) The Hurst parameter \( H \) satisfies \( H \in (0, 1/2] \).
(iii) (Martingality) The price process \( S = \exp X \) is a martingale.
(iv) (Short-time moments) \( \forall m < \infty \exists \theta > 0 : E(S_{\theta}^{m}) < \infty \).

While condition (iii) hardly needs justification, we emphasize that conditions (iiiia-b) are only used to the extent that they imply condition (iiic) given below (which thus may replace (iiiia-b) as an alternative, if more technical, assumption). The reason we point this out explicitly is that all the conditions (iiia-c) are implicit (growth) conditions on the function \( \sigma(\cdot) \). For instance, (iiiia-b) was seen to hold under a linear growth assumption (Forde and Zhang 2017; Gulisashvili 2017), whereas the log-normal volatility case (think of \( \sigma(x) = e^{x} \)) is complicated. Martingality, for instance, requires \( \rho \leq 0 \) and there is a critical moment \( m^{*} = m^{*}(\rho) \), even when \( \rho < 0 \). See Sin (1998), Jourdain (2004) and Lions and Musiela (2007) for the case \( H = 1/2 \) and the forthcoming work (Friz and Gassiat 2018) for the general rough case \( H \in (0, 1] \). We view (iiic) simply as a more flexible condition that can hold in situations where (iiib) fails.
(iii) (Call price upper moderate deviation bound) For every \( \beta \in (0, H) \), and every fixed \( x > 0 \), and \( \hat{x} := \frac{x}{\epsilon} \),
\[
E[I(e^{\epsilon x_i^\beta} - e^{\hat{x}})^+] \leq \exp \left(-\frac{x^2 + o(1)}{2\sigma_0^2 e^{4H-2\beta}}\right).
\]
This condition is reminiscent of the ‘upper part’ of the large deviation estimate obtained in Forde and Zhang (2017)

\[
E[I(e^{\epsilon x_i^\beta} - e^{\epsilon x_i^{2-\gamma}})^+] \leq \exp \left(-\frac{I(x) + o(1)}{\epsilon^{4H}}\right).
\]

If fact, if one formally applies this with \( x \) replaced by \( x e^{2\beta} \), followed by Taylor expanding the rate function,
\[
I(xe^{2\beta}) \sim \frac{1}{2} I''(0)x^2 e^{4\beta} = \frac{1}{2\sigma_0^2} x^2 e^{4\beta},
\]
one readily arrives at the estimate (iii). Unfortunately, \( o(1) = \alpha_0(1) \) in (8), which is a serious obstacle in making this argument rigorous. Instead, we will give a direct argument (Lemma 7.1) to see how (iiia-b) implies (iii). In the sequel, we will use another mild assumption on the kernel.

**Assumption 2.5** The kernel \( K \) has the following properties

(i) \( \hat{B}_t = \int_0^t K(t,s) dB_s \) has a continuous (in \( t \)) version on \([0,1]\).

(ii) \( \forall t \in [0,1] : \int_0^t K(t,s) ds < \infty \).

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

**Remark 2.6** Assumption 2.5 implies that the Cameron-Martin space \( \mathcal{H} \) of \( \hat{B} \) is generated by the image of \( H_1^0 \) under \( K \), i.e.

\( \mathcal{H} = \{ f | f \in H_1^0 \} \).

See Lemma 5.3 and Remark 5.4 for more details. A reference and also a sufficient condition for Assumption 2.5 (i) can be found e.g. in Decreusefond (2005, Section 3).

3. Main results

The following result can be seen as a non-Markovian extension of work by Osajima (2015). The statement here is a combination of Theorem 5.10 and Proposition (5.14) below. Recall that \( \sigma_0 = \sigma(0) \) represents spot-volatility. We also set \( \sigma_0^2 \equiv \sigma(0) \).

**Theorem 3.1** Energy expansion The rate function (or energy) \( I \) in (7) is smooth in a neighborhood of \( x = 0 \) (at-the-money) and it is of the form

\[
I(x) = \frac{1}{\sigma_0^2} x^2 - \left(6\rho - \frac{\sigma_0^2}{\sigma_0^2} \int_0^t \int_0^t K(t,s) ds dt\right) \frac{x^3}{3!} + O(x^4).
\]

The next result is an exact representation of call prices, valid in a non-Markovian generality, and amenable to moderate- and large-deviation analysis (Theorem 3.4 below).

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

\[
c(\hat{x},t) = E \left[ \exp(\epsilon X_t) - \exp(\hat{x}) \right] = e^{-\frac{1}{2}(\epsilon)^2} e^\alpha f(\epsilon, x),
\]

where

\[
J(\epsilon, x) := E \left[ e^{-(\epsilon^2/2)t^2} \left( \exp \left( \frac{\epsilon}{\hat{x}} \right)^2 - 1 \right) \right] e^{\frac{\epsilon}{\epsilon^2 \sigma_0^2} \int_0^t \frac{1}{\epsilon} \right] \exp(\epsilon X_t) - \exp(\hat{x}) \right]
\]

\( \hat{U}^\epsilon \) is a random variable of the form

\[
\hat{U}^\epsilon = \hat{g}_1 + \hat{\epsilon}^2 \hat{R}_2^\epsilon
\]

with \( g_1 \), a centred Gaussian random variable, explicitly given in equation (38) below, and \( R_2^\epsilon \) is a (random) remainder term, in the sense of a stochastic Taylor expansion in \( \hat{\epsilon} \), see Lemma 6.2 for more details.

**Example 3.3** Black-Scholes model We fix volatility \( \sigma(\cdot) \equiv \sigma > 0 \), and \( H = 1/2 \) so that \( \hat{\epsilon} = \epsilon \) and all \( \hat{\epsilon} \) can be omitted. Energy is given by

\[
I(x) = x^2/2\sigma^2 \text{ and } U^\epsilon = \epsilon^2 \hat{R}_2^\epsilon \equiv \epsilon^2 W_1 - \epsilon^2 \sigma^2/2
\]

with \( \hat{R}_2^\epsilon \equiv -\sigma^2/2 \) independent of \( \epsilon \). Moreover,

\[
J(\epsilon, x) := E \left[ e^{-(\epsilon^2/2)t^2} \left( \exp \left( \frac{\epsilon}{\epsilon} \right)^2 - 1 \right) \right] e^{\frac{\epsilon}{\epsilon^2 \sigma_0^2} \int_0^t \frac{1}{\epsilon} \right] \exp(\epsilon X_t) - \exp(\hat{x}) \right]
\]

\( \hat{U}^\epsilon \) is a random variable of the form

\[
\hat{U}^\epsilon = \hat{g}_1 + \hat{\epsilon}^2 \hat{R}_2^\epsilon
\]

and \( \hat{U}/\epsilon \) has a centred Gaussian random variable, explicitly given in equation (38) below, and \( \hat{R}_2^\epsilon \) is a (random) remainder term, in the sense of a stochastic Taylor expansion in \( \hat{\epsilon} \), see Lemma 6.2 for more details.

Using the expansion \( \Phi(-y) = (1/y \sqrt{2\pi}) e^{-y^2/2} \) \( (1 - y^2 + \cdots) \), as \( y \to \infty \) one deduces, for fixed \( x > 0 \), the asymptotic relation, as \( \epsilon \to 0 \),

\[
J(\epsilon, x) \sim \frac{e^{-x^2/2} e^{3\sigma^3}}{\sqrt{2\pi} x^2}.
\]

We will be interested (cf. Theorem 3.4) in replacing \( x \) by \( x e^{2\beta} \to 0 \) for \( \beta > 0 \). This gives \( \hat{\alpha} = (1/\sigma)(\epsilon^2 + 2\beta) \) and the above analysis, now based on \( \hat{\alpha} \to \infty \), remains valid for \( \beta \) in the ‘moderate’ regime \( \beta \in [0, 1/2) \) and we obtain

\[
\forall x > 0, \quad \beta \in [0, 1/2] : J(\epsilon, x e^{2\beta}) \sim \frac{1}{\sqrt{2\pi} \epsilon^2 x^2} e^{-3\sigma^3}. \]

Let us point out, for the sake of completeness, that a similar expansion is not valid for \( \beta > 1/2 \). To see this, first note

\[\text{More terms in the expansion of } \Phi \text{ are needed.}\]
Moreover, (9) implies that \( J(\varepsilon, x)|_{x=0} \) is precisely the ATM call price with time \( t = \varepsilon^2 \) from expiration. Well-known ATM asymptotics then imply that \( J(\varepsilon, x)|_{x=\varepsilon} \sim (1/\sqrt{2\pi})e^{\sigma^2} \) as \( \varepsilon \to 0 \). These asymptotics are unchanged in case of \( o(t^{1/2}) = o(\varepsilon) \) out-of-moneyness (‘almost-at-the-money’ in the terminology of Friz et al. 2018), which readily implies

\[
\forall x > 0, \quad \beta > 1/2 : J(\varepsilon, xe^{2\beta}) \sim \frac{1}{\sqrt{2\pi}}e^{\sigma^2} = \text{const} \times \varepsilon.
\]

At last, we have the borderline case \( \beta = 1/2 \), or \( \tilde{x} = x\varepsilon \). From e.g. Muhrle-Karbe and Nutz (2011, Theorem 3.1), we see that \( c(x, \varepsilon^2) - a(x, \sigma)\varepsilon \) with positive constant \( a(x, \sigma) \). A look at (9) then reveals

\[
\forall x > 0 : J(\varepsilon, xe^{2\beta}) \sim a(x, \sigma)\varepsilon e^{2\beta}\varepsilon = \text{const} \times \varepsilon.
\]

For the call price expansion in the large / moderate deviations regime, \( \beta \in [0, 1/2) \), the polynomial in \( \varepsilon \) behavior of (13) implies that the \( J \)-term in the pricing formula will be negligible on the moderate / large deviation scale, in the sense for any \( \theta > 0 \), we have \( e^{\theta}J(\varepsilon, xe^{2\beta}) \to 0 \) as \( \varepsilon \to 0 \). Consequently, with \( k_t = k\beta, \) for \( t = \varepsilon^2, k > 0, \beta \in [0, 1/2) \), we get the ‘moderate’ Black-Scholes call price expansion,

\[
- \log c_k(k, t) = \frac{1}{t^{\beta-2\beta}} \frac{k^2}{2\sigma^2} (1 + o(1)) \quad \text{as} \quad t \downarrow 0.
\]

While the above can be confirmed by elementary analysis of the Black–Scholes formula, the following theorem exhibits it as an instance of a general principle. See Friz et al. (2018) for a general diffusion statement.

**Theorem 3.4** Moderate Deviations

*In the rough volatility regime \( H \in (0, 1/2) \), consider log-strikes of the form

\( k_t = k t^{1/2-H+\beta} \) for a constant \( k \geq 0 \). |

(i) For \( \beta \in (0, H) \), and every \( \theta > 0 \), we have

\[
- \log c(k_t, t) = \frac{I''(0)}{t^{2H-2\beta}} \frac{k^2}{2} + O(t^{3\beta-2H}) + O(t^{-\theta}) \quad \text{as} \quad t \downarrow 0.
\]

(ii) For \( \beta \in (0, \frac{1}{2}H) \), and every \( \theta > 0 \), we have

\[
- \log c(k_t, t) = \frac{I''(0)}{t^{2H-2\beta}} \frac{k^2}{2} + \frac{I''(0)}{t^{2H-2\beta}} \frac{k^3}{6} + O(t^{4\beta-2H}) + O(t^{-\theta}) \quad \text{as} \quad t \downarrow 0.
\]

Moreover,

\[
I''(0) = \frac{1}{\sigma_0^2},
\]

\[
I'''(0) = -6\rho \sigma_0' \int_0^1 \int \mathcal{K}(t, s) \, ds \, dt = -6\rho \sigma_0' \mathcal{K}(1, 1),
\]

where \((\cdot, \cdot)\) is the inner product in \( L^2([0, 1])\).

**Remark 3.5** In principle, further terms (of order \( t^{i\beta-2H}, \ i = 4, 5, \ldots \)) can be added to this expansion of log call prices, given that the energy has sufficient regularity, see Theorem 3.6. We also note that, for small enough \( \beta \), the error term \( O(t^{-\theta}) \) can be omitted. In any case, one can replace the additive error bounds by (cruder) ones, where the right-most term in the expansion is multiplied with \((1 + o(1))\), as was done in Friz et al. (2018).

**Proof of Theorem 3.4** We apply Theorem 3.2 with \( \tilde{x} = k_t = k t^{1/2-H+\beta} \), i.e. with \( x = k_t \beta = k\beta \). In particular, we so get, with \( \tilde{\varepsilon} = t^{\beta} \) and \( \varepsilon = t^{1/2} \),

\[
c(k_t, t) = e^{-t^{(\beta^2)} } e^{t^{(\beta)}} \, J(\varepsilon, k\beta \varepsilon).
\]

The technical Proposition 7.3 asserts that, for fixed \( k > 0 \), the factor \( J \) is negligible in the sense that, for every \( \theta > 0 \),

\[
\varepsilon^{\theta} \log J(\varepsilon, k\beta \varepsilon) \to 0 \quad \text{as} \quad \varepsilon \to 0.
\]

The theorem now follows immediately from the Taylor expansion of \( I(x) \) around \( x = 0 \) (see Theorem 3.1), plugging in \( x = k \beta \). Indeed, replacing \( I(x) \) by the Taylor-jet seen in (i),(ii), leads exactly to an error term \( O(t^{3\beta-2H}) \), resp. \( O(t^{4\beta-2H}) \).

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

\[
k_t = k t^{1/2-H+\beta},
\]

and denote

\[
\phi_{n,H,\beta}(t) = \max \left\{ t^{2H-2\beta-\theta}, t^{(n-1)\beta} \right\}.
\]

Here, \( \theta > 0 \) can be arbitrarily small. It is clear that for all small \( t \) and \( \theta \) small enough,

\[
\phi_{n,H,\beta}(t) = t^{2H-2\beta-\theta} \Leftrightarrow 2H - 2\beta \leq (n - 1)\beta
\]

\[
\Leftrightarrow \frac{2H}{n + 1} \leq \beta,
\]

while

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

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

**Theorem 3.6** Suppose \( 0 \beta < 2H/n \) and \( \theta > 0 \) small enough. Then as \( t \to 0 \) (and for \( k > 0 \),

\[
\sigma_{\text{impl}}(k_t, t)^2 = \sum_{j=0}^{n-2} \frac{(-1)^{j/2}}{t^{(j+1)!}} \left( \sum_{i=3}^{n} \frac{I''(0)}{i!} k^{i-2} t^{i-2\beta} \right)^j
\]

\[
+ O\left( \phi_{n,H,\beta}(t) \right).
\]

The \( O \)-estimate in (14) depends on \( n, H, \beta, \theta, \) and \( k \). It is uniform on compact subsets of \([0, \infty)\) with respect to the variable \( k \).

**Remark 3.7** Using the multinomial formula, we can represent the expression on the left-hand side of (14) in terms of...
certain powers of \( t \). However, the coefficients become rather complicated.

**Remark 3.8** Let an integer \( n \geq 2 \) be fixed, and suppose we would like to use only the derivatives \( I^{(i)}(0) \) for \( 2 \leq i \leq n \) in formula (14) to approximate \( \sigma_{\text{impl}}(k_1, t)^2 \). Then, the optimal range for \( \beta \) is the following: \( 2H/(n+1) \leq \beta < 2H/n \). On the other hand, if \( \beta \) is outside of the interval \([2H/(n+1), 2H/n]\), more derivatives of the energy function at zero may be needed to get a good approximation of the implied variance in formula (14).

We will next derive from Theorem 3.6 several asymptotic formulas for the implied volatility. In the next corollary, we take \( n = 2 \).

**Corollary 3.9** As \( t \to 0 \),
\[
\sigma_{\text{impl}}(k_1, t) = \sigma_0 + O(\phi_2 H, \beta, \sigma(\phi_2 H, \beta, \sigma)(t)).
\] (15)

Corollary 3.9 follows from Theorem 3.6 with \( n = 2 \), the equality
\[
I''(0) = \sigma_0^{-2}
\] (16)
given in Theorem 3.4, and the Taylor expansion \( \sqrt{1 + h} = 1 + O(h) \) as \( h \to 0 \).

In the next corollary, we consider the case where \( n = 3 \).

**Corollary 3.10** Suppose \( \beta < 2H/3 \). Then, as \( t \to 0 \),
\[
\sigma_{\text{impl}}(k_1, t) = \sigma_0 + \rho \frac{\sigma_0'}{\sigma_0} (K_1, 1) kt^\beta + O(\phi_3 H, \beta, \sigma(\phi_3 H, \beta, \sigma)(t)).
\] (17)

Corollary 3.10 follows from Theorem 3.6 with \( n = 3 \), formula (16), the equality
\[
I'''(0) = -6\rho \frac{\sigma_0'}{\sigma_0}(K_1, 1)
\] (18)
(see Theorem 3.4), and the expansion \( \sqrt{1 + h} = 1 + \frac{1}{2} h + O(h^2) \) as \( h \to 0 \).

Using Corollary 3.10, we establish the following implied volatility skew formula in the moderate deviation regime.

**Corollary 3.11** Let \( 0 < H < \frac{1}{2} \), \( 0 < \beta < \frac{2}{3} H \), and fix \( y, z > 0 \) with \( y \neq z \). Then as \( t \to 0 \),
\[
\frac{\sigma_{\text{impl}}(y^{1/2 - H + \beta}, t) - \sigma_{\text{impl}}(z^{1/2 - H + \beta}, t)}{(y - z)^{1/2 - H + \beta}} \sim \rho \frac{\sigma_0'}{\sigma_0}(K_1, 1) t^{H - 1/2}.
\] (19)

**Remark 3.12** Corollary 3.11 complements earlier works of Alòs et al. (2007) and Fukasawa (2011, 2017). For instance, the following formula can be found in Fukasawa (2017, p. 6), see also Fukasawa (2011, p. 14):
\[
\frac{\sigma_{\text{impl}}(y^{1/2}, t) - \sigma_{\text{impl}}(z^{1/2}, t)}{(y - z)^{1/2}} \sim \rho C(H) \frac{\sigma_0'}{\sigma_0} t^{H - 1/2}.
\] (20)

In formula (20), we employ the notation used in the present paper. Our analysis shows that the applicability range of skew approximation formulas is by no means restricted to the Central Limit Theorem type log-moneyness deviations of order \( t^{1/2} \). It also includes the moderate deviations regime of order \( t^{1/2 - H + \beta} \). The previous rate is clearly \( \gg t^{1/2} \) as \( t \to 0 \).

**Remark 3.13** Symmetry Write \( \Phi_1(W, B, \beta; \rho; \sigma) \) for the ‘Itô-type map’
\[
\Phi_1(W, B, \beta) := \int_0^1 \sigma(\beta, t) d(\beta W + \rho B).
\]
It equals, in law, \( \Phi_1(W, -B, -\beta; -\rho; \sigma(-)) \), and indeed all our formulae are invariant under this transformation. In particular, the skew remains unchanged when the pair \((\rho, \sigma_0')\) is replaced by \(( -\rho, -\sigma_0')\).

## 4. Simulation results

We verify our theoretical results numerically with a variant of the rough Bergomi model (Bayer et al. 2016) which fits nicely into the general rough volatility framework considered in this paper. As before, the model has been normalized such that \( \beta_0 = 1 \) and \( r = 0 \). We let \((W, B)\) be two independent Brownian motions and \( \rho \in (-1, 1) \) with \( \beta^2 = 1 - \rho^2 \) such that \( Z = \rho W + \beta B \) is another Brownian motion having constant correlation \( \rho \) with \( B \). For some spot volatility \( \sigma_0 \) and volatility of volatility parameter \( \eta \), we then assume the following dynamics for some asset \( S \):
\[
\frac{dS_t}{S_t} = \sigma(\beta_t) dZ_t
\] (21)
\[
\sigma(x) = \sigma_0 \exp \left( \frac{1}{2} \eta x^2 \right)
\] (22)
where \( \hat{B} \) is a Riemann-Liouville fBm given by
\[
\hat{B}_t = \sqrt{2H} \int_0^t |s|^{-H-1/2} dB_s.
\]

The approach taken for the Monte Carlo simulations of the quantities we are interested in is the one initially explored in the original rough Bergomi pricing paper (Bayer et al. 2016). That is, exploiting their joint Gaussianity, where we use the well-known Cholesky method to simulate the joint paths of \((Z, \hat{B})\) on some discretization grid \( D \). With (22) being an explicit function in terms of the rough driver, an Euler discretization of the Ito SDE (21) on \( D \) then yields estimates for the price paths.

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

**Lemma 4.1** For convenience, define constants \( y = \frac{1}{2} - H \in [0, \frac{1}{2}] \) and \( D_H = \sqrt{2H}/(H + \frac{1}{2}) \) and define an auxiliary

\(\dagger\) Note that expressions for the exact same scenario have have been computed before in the original pricing paper (Bayer et al. 2016), yet in that version the expression for the autocorrelation of the IBM \( B \) was incorrect. We compute and state here all the relevant terms for the sake of completeness.
Short-time near-the-money skew in RFV models

Figure 1. Illustration of the term structure of implied volatility of the Modified Rough Bergomi model in the Moderate deviations regime with time-varying log-strike \( k_t = 0.4 t^\beta \). Depicted are the asymptotic formula (equation (17), dashed line) and an estimate based on \( N = 10^6 \) samples of a MC Cholesky Option Pricer (solid line) with 500 time steps. Model parameters are given by spot vol \( \sigma_0 \approx 0.2557 \), vvol \( \eta = 0.2928 \) and correlation parameter \( \rho = -0.7571 \).

\[
G(x) = 2H \left( \frac{1}{1 - \gamma} x^{-\gamma} + \frac{\gamma}{1 - \gamma} x^{-(1+\gamma)} \right) \times \frac{1}{2 - \gamma} _2F_1(1, 1 + \gamma, 3 - \gamma, x^{-1})
\]

where \(_2F_1\) denotes the Gaussian hypergeometric function (Olver et al. 2010). Then the joint process \((Z, \hat{B})\) has zero mean and covariance structure governed by

\[
\text{Var}[\hat{B}^2_t] = t^{2H}, \quad \text{for } t \geq 0, \\
\text{Cov}[\hat{B}_s, \hat{B}_t] = t^{2H} G(s/t), \quad \text{for } s > t \geq 0, \\
\text{Cov}[\hat{B}_s, Z_t] = \rho D_H(s^{H+1/2} - (s - \min(t, s))^H), \quad \text{for } t, s \geq 0, \\
\text{Cov}[Z_s, Z_t] = \min(t, s), \quad \text{for } t, s \geq 0.
\]

Numerical simulations† confirm the theoretical results obtained in the last section. In particular – as can be seen in figure 1 – the asymptotic structure for the implied volatility (17) captures very well the geometry of the term structure of implied volatility, with particularly good results for higher \( H \) and worsening results as \( H \downarrow 0 \). Quite surprisingly, despite being an asymptotic formula, it seems to be fairly accurate over a wide array of maturities extending up to a single year.

5. Proof of the energy expansion

Consider

\[
dX = -\frac{1}{2} \sigma^2(Y) dt + \sigma(Y) d(\tilde{\rho} dW + \rho dB), \quad X_0 = 0 \\
dY = d\hat{B}, \quad Y_0 = 0
\]

where \( \hat{B}_t = \int_0^t K(t, s) dB_s \) for a fixed Volterra kernel (recall (3) in the previous section). We study the small noise problem \((X', Y')\) where \((W, B, \hat{B})\) is replaced by \((\varepsilon W, \varepsilon B, \varepsilon \hat{B})\). The following proposition roughly says that

\[
\mathbb{P} \left( X'_1 \approx \frac{\varepsilon}{\varepsilon}\hat{X}_1 \right) \approx \exp \left( -\frac{I(x)}{2\varepsilon^2} \right).
\]

Proposition 5.1 Forde and Zhang 2017 Under suitable assumptions (cf. Section 2), the rescaled process \((\varepsilon X'_1): \)

† The Python 3 code used to run the simulations can be found at github.com/RoughStochVol.
\( \varepsilon \geq 0 \) satisfies an LDP (with speed \( \hat{\varepsilon}^2 \)) and rate function

\[
I(x) = \inf_{f \in H^0_1} \left[ \frac{(x - \rho G(f))^2}{2 \hat{\rho}^2 F(f)} + \frac{1}{2} E(f) \right]
\]

\[
\equiv \inf_{f \in H^0_1} \mathcal{I}_x(f) \equiv \mathcal{I}_x(f^*) ,
\]

(24)

where

\[
G(f) = \int_0^1 \sigma \left( (K \hat{f}) (s) \right) \hat{f}_s ds = \{ \sigma (K \hat{f}) , \hat{f} \} = \{ \sigma (\hat{f}) , \hat{f} \}
\]

\[
F(f) = \int_0^1 \sigma \left( (K \hat{f}) (s) \right) ds = \{ \sigma^2 (K \hat{f}) , 1 \} \equiv \{ \sigma \hat{f} , 1 \}
\]

\[
E(f) = \int_0^1 \vec{\sigma}^2 (s) ds = \{ \vec{f} , \vec{f} \}
\]

The rest of this section is devoted to analysis of the function \( I \) as defined in (24). First, we derive the first order optimality condition for the above minimization problem.

**Proposition 5.2** First order optimality condition For any \( x \in \mathbb{R} \) we have at any local minimizer \( f = f^* \) of the functional \( \mathcal{I}_x \) in (24) that

\[
f^*_t = \frac{\rho (x - \rho G(f^*)) \left( \{ \sigma (K \hat{f}^*) , 1_{[0,1]} \} + \{ \sigma^\prime (K \hat{f}^*) , \hat{f}^* , 1_{[0,1]} \} \right)}{\hat{\rho}^2 F(f^*)} + \frac{(x - \rho G(f^*))^2}{\hat{\rho}^2 F^2(f^*)} \left( \{ \sigma \sigma^\prime (K \hat{f}^*) , 1_{[0,1]} \} \right).
\]

(25)

for all \( t \in [0,1] \).

**Proof** We denote \( a \approx b \) whenever \( a = b + o(\delta) \) for a small parameter \( \delta \). We expand

\[
E(f + \delta g) \approx E(f) + 2 \delta \{ \hat{f} , \hat{g} \}
\]

\[
F(f + \delta g) \approx F(f) + \delta \left( \sigma \hat{f} \right)^\prime (K \hat{f}) \hat{K} \hat{g}
\]

\[
G(f + \delta g) \approx G(f) + \delta \left( \{ \sigma (K \hat{f}) , \hat{g} \} + \{ \sigma^\prime (K \hat{f}) , \hat{f} , \hat{K} \hat{g} \} \right)
\]

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

\[
\mathcal{I}_x(f + \delta g) = \frac{(x - \rho G(f))^2}{2 \hat{\rho}^2 F(f + \delta g)} + \frac{1}{2} E(f + \delta g)
\]

\[
\approx \frac{(x - \rho G(f) - \delta \rho \left( \{ \sigma (K \hat{f}) , \hat{g} \} + \{ \sigma^\prime (K \hat{f}) , \hat{f} , \hat{K} \hat{g} \} \right))}{2 \hat{\rho}^2} \left( \{ \sigma \sigma^\prime (K \hat{f}) , \hat{K} \hat{g} \} \right)^2 + \frac{1}{2} E(f) + \delta \{ \hat{f} , \hat{g} \}
\]

\[
+ \frac{(x - \rho G(f))^2 - \delta \rho (x - \rho G(f))}{2 \hat{\rho}^2 F(f)} \left( \{ \sigma (K \hat{f}) , \hat{g} \} + \{ \sigma^\prime (K \hat{f}) , \hat{f} , \hat{K} \hat{g} \} \right)
\]

\[
\approx \frac{1}{2 \hat{\rho}^2 F(f) \left[ 1 + \frac{\delta}{\hat{\rho}^2} \left( \{ \sigma \sigma^\prime (K \hat{f}) , \hat{K} \hat{g} \} \right) \right]}
\]

As a consequence, we must have, for \( f = f^* \) and every \( \hat{g} \in L^2(0,1) \)

\[
0 = \frac{d}{d\delta} \left( \mathcal{I}_x(f + \delta \hat{g}) \right) \bigg|_{\delta = 0}
\]

\[
= - \frac{\rho \left( x - \rho G(f) \right) \left( \{ \sigma (K \hat{f}) , \hat{g} \} + \{ \sigma^\prime (K \hat{f}) , \hat{f} , \hat{K} \hat{g} \} \right)}{\hat{\rho}^2 F(f)}
\]

\[
- \frac{(x - \rho G(f))^2}{\hat{\rho}^2 F^2(f)} \left( \{ \sigma \sigma^\prime (K \hat{f}) , \hat{K} \hat{g} \} \right) + \{ \hat{f} , \hat{g} \}.
\]

Recall \( f^*_0 = 0 \), any \( x \). We now test with \( \hat{g} = 1_{[0,1]} \) for a fixed \( t \in [0,1] \) and obtain

\[
f^*_t = \frac{\rho \left( x - \rho G(f^*) \right) \left( \{ \sigma (K \hat{f}^*) , 1_{[0,1]} \} + \{ \sigma^\prime (K \hat{f}^*) , \hat{f} , 1_{[0,1]} \} \right)}{\hat{\rho}^2 F(f^*)}
\]

\[
+ \frac{(x - \rho G(f^*))^2}{\hat{\rho}^2 F^2(f^*)} \left( \{ \sigma \sigma^\prime (K \hat{f}^*) , 1_{[0,1]} \} \right).
\]

\]

\[
5.1. \text{Smoothness of the energy}
\]

Having formally identified the first order condition for minimality in (24), we will now show that the energy \( x \mapsto I(x) \) is a smooth function. More precisely, we will use the implicit function theorem to show that the minimizing configuration \( f^* \) is a smooth function in \( x \) (locally at \( x = 0 \)). As \( \mathcal{I}_x \) is a smooth function, too, this will imply smoothness of \( x \mapsto \mathcal{I}_x(f^*) = I(x) \), at least in a neighborhood of 0.

As the Cameron-Martin space \( \mathcal{H} \) of the process \( \hat{B} \) continuously embeds into \( C([0,1], K) \), maps \( H^0_1 \) continuously into \( C([0,1], I) \), i.e. there is a constant \( C > 0 \) such that for any \( f \in H^0_1 \) we have

\[
\| \hat{K}^\prime \|_\infty \leq C \| I(f) \|_{H^0_1}.
\]

(26)

This result will follow from

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

\[
\| h \|_\infty \leq C \| h \|_{\mathcal{H}}.
\]

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

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

On the other hand, a generic element $h \in \mathcal{H}$ can be written as $h_t = E[V_t Z]$ where $Z$ is a centred Gaussian random variable with variance $\|h\|^2_2$, see, e.g. Friz and Hairer (2014, page 150). By Cauchy–Schwarz,

$$|h_t| \leq E[|V_t|](1/2) \|h\|_{\mathcal{H}} \leq \sigma \|h\|_{\mathcal{H}}$$

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

**Remark 5.4** Assume $V$ is of Volterra form, i.e. $V_t = \int_0^t K(t, s) dB_s$. Then it can be shown (see Decreusefond 2005, Section 3) that $\mathcal{H}$ is the image of $L^2$ under the map

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

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

$$\|K\hat{f}\|_\infty \leq C \|K\hat{f}\|_{\mathcal{H}} = C \|\hat{f}\|_{L^2} = C \|f\|_{H^1_0}. \tag{27}$$

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

By Proposition 5.2, any local optimizer $f = f^+ \in \mathcal{H}^+_0$ to the functional $I_x: H^+_0 \to \mathbb{R}$ in the uncorrelated case $\rho = 0$ satisfies for any $t \in [0, 1]$

$$f_t = \frac{x^2}{F^2(f)} \left(\langle \sigma \sigma' \rangle (K \hat{f}), K1_{[0,1]}\right).$$

We define a map $H: H^+_0 \times \mathbb{R} \to H^+_0$ by

$$H(f, x)(t) := f_t - \frac{x^2}{F^2(f)} \left(\langle \sigma \sigma' \rangle (K \hat{f}), K1_{[0,1]}\right). \tag{27}$$

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

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

Note that invertibility should hold for $x$ small enough, as $DH(0, 0) = \text{id}_{H^+_0} - x^2 R$ for some $R$, which is invertible as long as $R$ has a bounded norm for sufficiently small $x$.

**Remark 5.5** The method of proof in this section is purely local in $H^+_0$. Hence, we only really need smoothness of $\sigma$ locally around 0. Note, however, that stochastic Taylor expansions used in Section 6 will actually require global smoothness of $\sigma$.

**Lemma 5.6** The functions $F: H^1_0 \to \mathbb{R}$ and $R_1: H^1_0 \to C([0, 1])$ defined by

$$R_1(f)(t) := \langle (\sigma \sigma') (K \hat{f}), K1_{[0,1]}\rangle, \quad t \in [0, 1],$$

are smooth in the sense of Fréchet.

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

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

By Lemma 5.3, we can bound

$$\|D^N F(f) \cdot (g_1, \ldots, g_N)\| \leq \text{const} \int_0^1 \|K \hat{g}_1(s) \cdots K \hat{g}_N(s)\| \, ds \leq \text{const} \|K \hat{g}_1\|_\infty \cdots \|K \hat{g}_N\|_\infty \leq \text{const} C \|g_1\|_{H^1_0} \cdots \|g_N\|_{H^1_0},$$

for $\text{const} = \|(d^N / ds^N) \sigma^2\|_\infty$. Thus, $D^N F(f)$ is a multi-linear form on $H^1_0$ with operator norm $\|D^N F(f)\| \leq \|(d^N / ds^N) \sigma^2\|_\infty C$ independent of $f$. As $f \mapsto D^N F(f)$ is continuous, we conclude that $D^N F(f)$ as given above is, in fact, a Fréchet derivative.

Let us next consider the functional $R_1$. Note that

$$(D^N R_1(f) \cdot (g_1, \ldots, g_N))(t) = [\hat{s}_N(K \hat{f}) K \hat{g}_1 \cdots K \hat{g}_N, K1_{[0,1]}]$$

for $\hat{s}_N(x) := (d^N / dx^N) \sigma(x) \sigma'(x)$. Hence, Assumption 2.5 implies that

$$\|D^N R_1(f) \cdot (g_1, \ldots, g_N)\|^2_{H^1_0}$$

$$= \int_0^1 \left(\int_0^1 \hat{s}_N (K \hat{f}(s)) \prod_{i=1}^N (K \hat{g}_i(s)) K(s, t) \, ds \right)^2 \, dt$$

$$\leq \|\hat{s}_N\|_\infty^2 \prod_{i=1}^N \|K \hat{g}_i\|_\infty^2 \int_0^1 \int_0^1 \|K(s, t)\|^2 \, ds \, dt$$

$$\leq \|\hat{s}_N\|_\infty^2 C^N \prod_{i=1}^N \|g_i\|_{H^1_0}^2 \int_0^1 \int_0^1 \|K(s, t)\|^2 \, ds \, dt$$

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

$$\|D^N R_1(f)\| \leq \|\hat{s}_N\|_\infty C^N \sqrt{\int_0^1 \int_0^1 \|K(s, t)\|^2 \, ds \, dt},$$

independent of $f$. From continuity of $f \mapsto D^N R_1(f)$, it follows that $D^N R_1(f)$ is the $N$'th Fréchet derivative.
Theorem 5.7 Zero correlation

**Assuming \( \rho = 0 \), the energy \( I(x) \) (as defined in (24)) is smooth in a neighborhood of \( x = 0 \).

**Proof** By construction, we have

\[
DH(f, x) = id_{H_0^1} - x^2A(f)
\]

for \( A : H_0^1 \rightarrow L(H_0^1, H_0^1) \) defined by

\[
A(f) := R_1(f) \otimes DF^{-2}(f) + F^{-2}(f)DR_1(f).
\]

Here,

\[
(R_1(f) \otimes DF^{-2}(f)) \cdot g = (DF^{-2}(f) \cdot g) R_1(f),
\]

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

For the energy, we prove that \( I(x) = \mathcal{I}_0(f^s) \) in a neighborhood of \( x = 0 \). First of all, we show that a minimizer exists. If not, there is a function \( g \in H_0^1 \) with \( \mathcal{I}_0(g) < \mathcal{I}_0(f^s) \). For small enough \( x \) such a \( g \) must be inside a ball with radius \( \epsilon \) around \( 0 \) in \( H_0^1 \), as \( \mathcal{I}_0(g) \geq \frac{1}{2}\|g\|_{H_0^1}^2 \) and \( \lim_{\epsilon \to 0} \mathcal{I}_0(f^s) = 0 \). Then note that for any \( g \in H_0^1 \)

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

where \( D^2\mathcal{I}_0(f) \) denotes the second derivative of \( f \mapsto \mathcal{I}_0(f) \). By continuity, \( D^2\mathcal{I}_0(f) \) stays positive definite for \( (x, f) \) in a neighborhood of \( (0, 0) \). As noted, for \( x \) small enough, both \( g \) and \( f^s \) (and the line connecting them) lie in this neighborhood.

For \( h := g - f^s \), this implies

\[
\mathcal{I}_0(g) - \mathcal{I}_0(f^s) = D\mathcal{I}_0(f^s) \cdot h + \int_0^1 D^2\mathcal{I}_0(f^s + th) \cdot (h, h) \, dt > 0,
\]

since \( D\mathcal{I}_0(f^s) \cdot h = 0 \) and \( D^2\mathcal{I}_0(f^s + th) \cdot (h, h) > 0 \). This contradicts the assumption that \( \mathcal{I}_0(g) < \mathcal{I}_0(f^s) \), and we conclude that \( f^s \) is, indeed, a minimizer of \( \mathcal{I}_0 \), implying that \( I(x) = \mathcal{I}_0(f^s) \) locally.

Finally, as \( x \mapsto f^s \) is smooth and \( (f, x) \mapsto \mathcal{I}_0(f) = x^2/2F(f) + \frac{1}{2}\|f\|_{H_0^1}^2 \) is smooth, we see that \( x \mapsto I(x) = \mathcal{I}_0(f^s) \) is smooth in a neighborhood of \( 0 \). (Note that this argument relies on \( \sigma(0) \neq 0 \), implying that \( F(f) \neq 0 \) for \( f \) in a neighborhood to \( 0 \).)

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

\[
J(u) := \int_0^1 \left[ (u'(s)^2 - 1)^2 + u(s)^2 \right] \, ds
\]

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

### 5.1.2. The general case

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

\[
H(f, x)(t) := f_t - \frac{\rho (x - \rho G(f))}{\rho^2 F(f)} + \frac{(x - \rho G(f))^2}{\rho^2 F(f)} \left( \langle (\sigma \sigma') \, (K^f) \rangle, K[0, 1] \right) - \frac{\rho (x - \rho G(f))}{\rho^2 F(f)} \left( \langle (K^f) \rangle, K[0, 1] \right) + \frac{(x - \rho G(f))^2}{\rho^2 F(f)^2} R_2(f)(t),
\]

where \( R_2, R_3 : H_0^1 \rightarrow H_0^1 \) are defined by

\[
R_2(f)(t) := \langle \sigma(K^f) \rangle \cdot K[0, 1],
\]

\[
R_3(f)(t) := \langle \sigma'(K^f) \rangle \cdot K[0, 1],
\]

\( t \in [0, 1] \).

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

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

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

\[
D^N G(f) \cdot (g_1, \ldots, g_N) = \left\langle \sigma^{(N)}(K^f) \cdot \prod_{i=1}^N K_\delta_1, \sum_{k=1}^N \sigma^{(N-1)}(K^f) \cdot \prod_{i \neq k} K_\delta_1 \right\rangle.
\]

For \( R_2 \) and, respectively, \( R_3 \), we obtain

\[
(D^N R_2(f) \cdot (g_1, \ldots, g_N))(t) = \int_0^t \sigma^{(N)}(\langle (K^f)(s) \rangle) \prod_{i=1}^N (K_\delta_1)(s) \, ds,
\]
and
\[
(D^N R_n(f) \cdot (g_1, \ldots, g_N))(t) = \left(\sigma^{(N+1)}(Kf) \cdot (\prod_{i=0}^{N} K_i) \right) + \sum_{k=1}^{N} \left(\sigma^{(N)}(Kf) \cdot (\prod_{i \neq k} K_i) \right).
\]

**Theorem 5.10** Let \( \sigma \) be smooth with \( \sigma(0) \neq 0 \). Then the energy \( I(x) \) as defined in (24) is smooth in a neighborhood of \( x = 0 \).

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

Note that (28) contains three terms. The derivative of the first term \( f \mapsto f \) is always equal to \( \text{id}_{H^1} \). For the second term, we note that
\[
(x - \rho G(f))|_{x=0, f=0} = 0.
\]
Hence, the only non-vanishing contribution to the derivative of the second term evaluated in direction \( g \in H^1_0 \) at \( x = 0, f = 0 \) and \( t \in [0,1] \) is
\[
\frac{\rho^2 DG(0) \cdot g}{\rho^2 G(0)} (R_2(0) + R_3(0)) = \frac{\rho^2 \sigma_0 g(1)}{\rho^2 \sigma_0} (\sigma_0 t + 0) = \rho^2 \frac{\sigma_0}{\rho^2} g(1)t.
\]
For the same reason, the derivative of the third term at \( f \mapsto f \) is \( (0,0) \), vanishes entirely. Hence,
\[
(DH(0,0) \cdot g)(t) = g(t) + \rho^2 \frac{\sigma_0}{\rho^2} g(1)t.
\]
It is easy to see that \( g \mapsto DH(0,0) \cdot g \) is invertible. Indeed, let us construct the pre-image \( g = DH(0,0)^{-1} \cdot h \) of some \( h \in H^1_0 \). At \( t = 1 \) we have
\[
\frac{\rho^2 + \rho^2}{\rho^2} g(1) = h(1),
\]
implies \( g(1) = \rho^2 h(1) \). For \( 0 \leq t < 1 \), we then get
\[
g(t) + \rho^2 \frac{\sigma_0}{\rho^2} g(1)t = g(t) + \rho^2 \frac{\sigma_0}{\rho^2} \rho^2 h(1)t = g(t) + \rho^2 h(1)t = h(t),
\]
or \( g(t) = h(t) - \rho^2 h(1)t \).

For existence of the minimizer, note that
\[
D^2 J_0(0) \cdot (g, g) = \frac{\rho^2}{\rho^2} g(1)^2 + \|g\|_{H^0}^2,
\]
which is again positive definite.

**Remark 5.11** Though only formulated in terms of ‘smoothness’, it is easy to show that \( \sigma \in C^3 \) implies that \( I \in C^{k-1} \) (locally at 0).

### 5.2. Energy expansion

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

#### 5.2.1. Expansion of the minimizing configuration.

**Theorem 5.12** We have
\[
f_t^x = \alpha_t x + \beta_t x^2 + O(x^3),
\]
\[
\alpha_t = \frac{\rho}{\sigma_0},
\]
\[
\beta_t = 2 \frac{\sigma_0}{\sigma_0^2} \left[ \rho^2 \langle K1, 1 \rangle |_{t=1} + \|K1, 1\|_{1} - 3 \rho^2 t (K1, 1) \right].
\]

**Remark 5.13** Non-Markovian transversality In the RL-fBM case, \( K(t,s) = \sqrt{2H} |t-s|^{\gamma} \) with \( \gamma = H - 1/2 \) one computes
\[
\{1, K1, 1\} = \frac{1}{(1+y)(2+y)} \left[ 1 - (1-t)^{2+y} \right] \in C^1[0,1].
\]
Interestingly, the transversality condition known from the Markovian setting (\( g_1 = 0 \), which readily translates to \( f_t^1 = 0 \) there) remains valid here (for \( \rho = 0 \), at least to order \( x^2 \), in the sense that
\[
\dot{f}_t^x \approx \beta_t x^2 = (\text{const}) (1 - t)^{1+y} |_{t=1} = 0
\]

**Proof of Theorem 5.12** First order expansion:

Up to the order needed in order to get the first order term, we have
\[
f_t^x = \alpha_t x + O(x^2),
\]
\[
\beta_t = \alpha_t x + O(x^2),
\]
\[
\sigma(\dot{K}^x) = \sigma_0 + \sigma_0^2 K \dot{x} + O(x^2),
\]
\[
\sigma'(\dot{K}^x) = \sigma_0^2 + \sigma_0^2 K \dot{x} + O(x^2),
\]
\[
F(f^x) = \langle \sigma(\dot{K}^x), 1 \rangle = \sigma_0^2 + O(x),
\]
\[
G(f^x) = \langle \sigma(\dot{K}^x), \dot{f}^x \rangle = \langle \sigma_0, \dot{x} \rangle x + O(x^2).
\]
Therefore,
\[
\sigma(K \dot{f}^s) \cdot 1_{[0,1]} = \sigma_0 t + O(x),
\]
\[
\sigma'(K \dot{f}^s) \cdot 1_{[0,1]} = \sigma_0 x + O(x^2),
\]
\[
\sigma \sigma'(K \dot{f}^s) \cdot K \cdot 1_{[0,1]} = O(1),
\]
\[
(x - \rho G(f^s))^2 = (1 - \rho \sigma_0 \alpha_1) x + O(x^2),
\]
\[
(x - \rho G(f^s))^2 = O(x^2).
\]
This yields for the first order term in (25)
\[
\alpha_t = \frac{\rho(1 - \rho \sigma_0 \alpha_1)}{\rho^2 \sigma_0^2} t.
\]
Setting \( t = 1 \), we get
\[
\alpha_1 = \frac{\rho}{\rho^2 \sigma_0} - \frac{\rho^2}{\rho^2 \sigma_0^2} \alpha_1,
\]
which is solved by \( \alpha_1 = \rho / \sigma_0 \). Inserting this term back into the equation for \( \alpha_t \), we get
\[
\alpha_t = \frac{\rho}{\sigma_0} t. \quad (31)
\]
Second order expansion:
Using (31) and the ansatz \( f_t^s = \alpha_1 x + \frac{1}{2} \beta x^2 + O(x^3) \), we re-compute the relevant terms appearing in the (25). We have
\[
\sigma(K \dot{f}^s(s)) = \sigma_0 + \sigma_0' \frac{\rho}{\sigma_0} (K11)(s) x + O(x^2)
\]
and analogously for \( \sigma \) replaced by \( \sigma' \), \( \sigma' \). This implies
\[
\sigma(K \dot{f}^s(s)) \cdot 1_{[0,1]} = \sigma_0 t + \sigma_0' \frac{\rho}{\sigma_0} [K11, 1_{[0,1]}] x + O(x^2),
\]
\[
\sigma'(K \dot{f}^s(s)) \cdot K1_{[0,1]} = \rho \sigma_0' [K1_{[0,1]}, 1] x + O(x^2),
\]
\[
\sigma \sigma'(K \dot{f}^s(s)) \cdot K1_{[0,1]} = \sigma_0 \sigma_0' [K1_{[0,1]}, 1] + O(x).
\]
Using the notation introduced earlier, we have
\[
F(f^s) = \sigma_0^2 + 2 \sigma_0' \rho \langle K11, 1 \rangle x + O(x^2),
\]
\[
G(f^s) = \rho x + \left( \frac{1}{2} \sigma_0 \beta_1 + \frac{\rho^2}{\sigma_0^2} \langle K11, 1 \rangle \right) x^2 + O(x^3).
\]
This directly implies
\[
x - \rho G(f^s) = \tilde{\rho}^2 x - \rho \left( \frac{1}{2} \sigma_0 \beta_1 + \frac{\rho^2}{\sigma_0^2} \langle K11, 1 \rangle \right) x^2 + O(x^3),
\]
\[
(x - \rho G(f^s))^2 = \tilde{\rho}^2 x^2 - 2 \tilde{\rho}^2 \rho \]
\[
\times \left[ \frac{1}{2} \sigma_0 \beta_1 + \frac{\rho^2}{\sigma_0^2} \langle K11, 1 \rangle \right] x^3 + O(x^4).
\]
We next compute some auxiliary terms appearing in (25).
\[
N_1 := \rho (x - \rho G(f^s)) \frac{1}{2} \sigma_0 \dot{t} x + \left[ \rho^2 \tilde{\rho}^2 \frac{\sigma_0'}{\sigma_0} \langle [K1, 1_{[0,1]}] + [K1_{[0,1]}], 1 \rangle \right]
\]
\[
- \rho \alpha_0 \frac{\sigma_0'}{\sigma_0} \langle K11, 1 \rangle - \frac{1}{2} \rho^2 \sigma_0^2 t \beta_1 \right] x^2 + O(x^3)
\]
The corresponding denominator is \( \tilde{\rho}^2 F(f^s) \). Using the formula
\[
\frac{a_1 x + a_2 x^2 + O(x^3)}{b_0 + b_1 x + O(x^2)} = \frac{a_1}{b_0} + \frac{a_2 b_0 - a_1 b_1}{b_0} x + O(x^3),
\]
we obtain
\[
\frac{N_1}{\tilde{\rho}^2 F(f^s)} = \tilde{\rho}^2 \frac{\sigma_0'}{\sigma_0} [K11, 1_{[0,1]}] + [K1_{[0,1]}, 1] + \left( \frac{\rho^4}{\tilde{\rho}^2} + 2 \tilde{\rho}^2 \right) \sigma_0^2 \sigma_0' t \langle K11, 1 \rangle - \frac{1}{2} \rho^2 \beta_1 t \right] x^2 + O(x^3)
\]
\[
(32)
\]
For the second term in (25), let
\[
N_2 := (x - \rho G(f^s))^2 \frac{1}{2} \sigma_0 \sigma_0' [K1_{[0,1]}, 1] x^2 + O(x^3).
\]
The corresponding denominator is \( \tilde{\rho}^2 F(f^s) = \tilde{\rho}^2 \sigma_0^2 + O(x) \). Hence,
\[
\frac{N_2}{\tilde{\rho}^2 F(f^s)} = \tilde{\rho}^2 \sigma_0^2 [K1_{[0,1]}, 1] x^2 + O(x^3). \quad (33)
\]
Combining (32) and (33), we get
\[
f_t^s = \frac{\rho}{\sigma_0} t + \left[ \frac{\rho^2}{\sigma_0} \sigma_0' \langle [K11, 1_{[0,1]}] + [K1_{[0,1]}, 1] \rangle \right]
\]
\[
- \frac{\rho^2}{\sigma_0} \sigma_0' \langle [K11, 1_{[0,1]}] + [K1_{[0,1]}, 1] \rangle t \langle K11, 1 \rangle + \tilde{\rho}^2 \sigma_0^2 \sigma_0' [K1_{[0,1]}, 1] \right] x^2
\]
\[
+ O(x^3).
\]
We shall next compute \( \beta_1 \). Taking the second order terms on both sides and letting \( t = 1 \), we obtain
\[
\frac{1}{2} \beta_1 = \frac{\rho}{\rho^2} \sigma_0^2 \langle K11, 1 \rangle + \frac{\rho^4}{\tilde{\rho}^2} \sigma_0^2 \langle K11, 1 \rangle
\]
\[
- \frac{1}{2} \rho^2 \beta_1 - 2 \rho^2 \sigma_0^2 \sigma_0' \langle K11, 1 \rangle + \tilde{\rho}^2 \sigma_0^2 \sigma_0' \langle K1_{[0,1]}, 1 \rangle.
\]
Moving \( \beta_1 \) to the other side with \( 1 + \rho^2 / \tilde{\rho}^2 = 1 / \tilde{\rho}^2 \) and collecting terms on the right hand side, we arrive at
\[
\frac{1}{2} \beta_1 = \frac{\sigma_0^2}{\sigma_0^2} \langle K11, 1 \rangle \left( 2 \rho^2 - \frac{\rho^4}{\tilde{\rho}^2} - 2 \rho^2 + \tilde{\rho}^2 \right)
\]
\[
= \frac{1}{2} \rho^2 - \frac{2}{\tilde{\rho}^2} \sigma_0' \langle K11, 1 \rangle.
\]
We conclude that
\[
\beta_1 = 2 \left( 2 - \rho^2 / \tilde{\rho}^2 \right) \sigma_0' \langle K11, 1 \rangle
\]
Hence, we obtain

\[ \beta_t = 2 \frac{\sigma'_0}{\sigma_0} \left[ \rho^2 \{K_1, 1_{[0,t]} \} + \{K_1, 1_{[0,t]} \} ^2 \right] \]

5.2.2. Energy expansion in the general case. Now we compute the Taylor expansion of \( I(x) \) as defined in Proposition 5.1. We start with the second term. Plugging in the optimal path \( f_1^* = \alpha_x + \frac{1}{2} \beta_1 x^2 + O(x^3) \) (and using \( \{\beta, 1\} = \beta_1 \) as \( \beta_0 = 0 \)) we obtain

\[ \frac{1}{2} \langle f^*, f^* \rangle = \frac{1}{2} \frac{\rho^2}{\sigma_0^2} x^2 + \frac{1}{2} \frac{\rho}{\sigma_0} \beta_1 x^3 + O(x^4). \]

Inserting \( \beta_1 = 2(1 - 2 \rho^2) (\sigma'_0/\sigma_0) \{K_1, 1\} \) into the above formula for \( (x - \rho G(f^*))^2 \), we get

\[ (x - \rho G(f^*))^2 = \tilde{\rho}^3 x^2 - 2 \tilde{\rho}^2 \frac{\sigma'_0}{\sigma_0} \{K_1, 1\} x^3 + O(x^4). \]

Recall the denominator

\[ 2 \tilde{\rho}^2 F(f^*) = 2 \tilde{\rho}^2 \sigma_0^2 + 4 \tilde{\rho}^2 \sigma'_0 \rho \{K_1, 1\} x + O(x^2). \]

Using the expansion of a fraction

\[ \frac{a x^2 + a x^3 + O(x^4)}{b_0 + b_1 x + O(x^2)} = \frac{a x^2}{b_0} + \frac{a x^3}{b_0} \frac{a x}{b_0} \frac{a x}{b_0} + O(x^4), \]

we obtain from

\[ \frac{(x - \rho G(f^*))^2}{2 \tilde{\rho}^2 F(f^*)} = \frac{\tilde{\rho}^4}{2 \tilde{\rho}^2 \sigma_0^2} x^2 + \frac{\left( -2 \tilde{\rho}^2 \rho \frac{\sigma'_0}{\sigma_0} \{K_1, 1\} \right)}{4 \tilde{\rho}^2 \sigma'_0} x^3 + O(x^4). \]

We note that

\[ \frac{1}{2} \frac{\rho}{\sigma_0} \beta_1 = \frac{\rho}{\sigma_0} \frac{\sigma'_0}{\sigma_0} \{K_1, 1\} \]

\[ = \left( (1 - 2 \rho^2) - 2(1 - \rho^2) \right) \frac{\sigma'_0}{\sigma_0} \{K_1, 1\} = -\rho \frac{\sigma'_0}{\sigma_0} \{K_1, 1\}. \]

Adding both terms, we arrive at the

**Proposition 5.14.** The energy expansion to third order gives

\[ I(x) = \frac{1}{2} x^2 - \rho \frac{\sigma'_0}{\sigma_0} \{K_1, 1\} x^3 + O(x^4). \]

5.2.3. Energy expansion for the Riemann-Liouville kernel. Let us specialize the energy expansion given in Proposition 5.14 for the Riemann-Liouville fBm. Choose \( \gamma = H - \frac{1}{2} \) and recall that the kernel \( K \) takes the form \( K(t, s) = (t - s)\gamma \). We get

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

The key term \( \{K_1, 1\} \) appearing in the energy expansion now gives

\[ \{K_1, 1\} = \int_0^1 (K(t)) \, dt = \int_0^1 t^{1+\gamma} \, dt = \frac{1}{(1 + \gamma)(2 + \gamma)} \]

\[ = \frac{1}{(H + 1/2)(H + 3/2)}. \]

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

\[ I(x) = \frac{1}{2} \sigma_0^2 x^2 - \frac{\rho}{(H + 1/2)(H + 3/2)} \frac{\sigma'_0}{\sigma_0} x^3 + O(x^4). \]

For completeness, let us also fully describe the time-dependence of the second order term \( \beta_3 \), in the expansion of the optimal trajectory \( f_3^* \). Unlike the first order time, here we do not have a linear movement any more. Indeed

\[ \{K_1, 1_{[0,t]} \} = \frac{1}{(1 + \gamma)(2 + \gamma)} \left[ (1 - (1 - t)^{2+\gamma}) \right]. \]

6. Proof of the pricing formula

Fix \( x \geq 0 \) and \( \tilde{x} = (\epsilon \tilde{\epsilon}) x \) where \( \epsilon = t^{1/2} \) and \( \tilde{\epsilon} = t^{H} = \epsilon^{2H} \). We have

\[ c(\tilde{x}, t) = E \left( \exp \left( X^t \right) - \exp \left( \tilde{x}^t \right) \right)^+ \]

\[ = E \left( \exp \left( \frac{\epsilon}{\epsilon} X^t \right) - \exp \left( \frac{\epsilon}{\epsilon} \tilde{x}^t \right) \right)^+ \]

where we recall

\[ \tilde{X}_t^\epsilon = \frac{\epsilon}{\epsilon} X^t = \int_0^1 \sigma (\tilde{\epsilon} \tilde{B}) \, d(\tilde{\rho} \tilde{W} + \rho B) \]

\[ = \frac{1}{2} \tilde{\epsilon} \int_0^1 \sigma (\tilde{\epsilon} \tilde{B})^2 \, dt. \]
Consider a Cameron–Martin perturbation of $\hat{X}_t$. That is, for a Cameron–Martin path $h = (h, f) \in H_0^1 \times H_0^1$ consider a measure change corresponding to a transformation $\hat{\epsilon}(W, B) \sim \hat{\epsilon}(W, B) + (h, f)$ (transforming the Brownian motions to Brownian motions with drift), we obtain the Girsanov density

$$G_{\hat{\epsilon}} = \exp \left( -\frac{1}{\hat{\epsilon}} \int_0^t \hat{h}_s \, dW_s - \frac{1}{\hat{\epsilon}} \int_0^t \hat{f}_s \, dB_s \right) - \frac{1}{2\hat{\epsilon}^2} \int_0^t (\hat{h}_s + \hat{f}_s)^2 \, ds.$$  \hfill (36)

Under the new measure, $\hat{X}_t$ becomes $Z_{t\hat{\epsilon}}$, where

$$Z_{t\hat{\epsilon}} = \int_0^t \sigma (\hat{\epsilon} \hat{B}_s + \hat{f}_s) \left[ \hat{\epsilon} \, d(\hat{\rho}_{sW} + \rho B_s) + d(\hat{\rho}_sW + \rho f_s) \right]$$

$$- \frac{1}{2} \hat{\epsilon} \int_0^t \sigma (\hat{\epsilon} \hat{B}_s + \hat{f}_s)^2 \, ds.$$

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

$$I (x) = \inf_{h, f \in H_0^1} \left\{ \frac{1}{2} \int_0^t h_s^2 \, ds + \frac{1}{2} \int_0^t f_s^2 \, ds : \Phi_1 (h, f, \hat{f}) = x \right\},$$

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

$$\Phi_1 (h, f, \hat{f}) = \int_0^1 \sigma (\hat{f}_s) \, d(\hat{\rho}_sW + \rho f_s).$$

For any Cameron–Martin path $(h, f)$, the perturbed random variable $Z_{t\hat{\epsilon}}$ admits a stochastic Taylor expansion with respect to $\hat{\epsilon}$.

**Lemma 6.2** Fix $(h, f) \in K^\epsilon$ and define $Z_{t\hat{\epsilon}}$ accordingly. Then

$$Z_{t\hat{\epsilon}} = x + \hat{\epsilon} g_1 + \hat{\epsilon}^2 R_2^g,$$ \hfill (37)

where $g_1$ is a Gaussian random variable, given explicitly by

$$g_1 = \int_0^1 \sigma (\hat{f}_s) \, d(\hat{\rho}_sW + \rho f_s) + \sigma' (\hat{f}_s) \hat{B}_s \, d(\hat{\rho}_sW + \rho f_s).$$ \hfill (38)

and

$$R_2^g = \int_0^1 \sigma' (\hat{f}_s) \hat{B}_s \, d(\hat{\rho}_sW + \rho B_s) - \frac{1}{\hat{\epsilon}} \int_0^1 \sigma (\hat{\epsilon} \hat{B}_s + \hat{f}_s)^2 \, ds + \frac{1}{2\hat{\epsilon}^2} \int_0^t \int_0^t \sigma'' (\hat{\epsilon} \hat{B}_s + \hat{f}_s) \hat{B}_s^2 \times \hat{\epsilon} \, d(\hat{\rho}_sW + \rho B_s) + d(\hat{\rho}_sW + \rho f_s) \left( \hat{\epsilon} - \zeta \right) \, d\zeta.$$ \hfill (39)

**Proof** By a stochastic Taylor expansion for the controlled process $Z_{t\hat{\epsilon}}$ with control $(h, f) \in K^\epsilon$ as in Definition 6.1 and thanks to $\sigma \in C^2$, we have at $t = 1$

$$Z_{t\hat{\epsilon}} = \int_0^1 \sigma (\hat{\epsilon} \hat{B}_s + \hat{f}_s) \left[ \hat{\epsilon} \, d(\hat{\rho}_sW + \rho B_s) + d(\hat{\rho}_sW + \rho f_s) \right]$$

$$- \frac{1}{2} \hat{\epsilon} \int_0^t \sigma (\hat{\epsilon} \hat{B}_s + \hat{f}_s)^2 \, ds.$$

Collecting terms in powers of $\hat{\epsilon}$ and with the random variable $g_1$ as in (38) (recalling that $\hat{\epsilon} \in O(\hat{\epsilon}^2)$), we have

$$\hat{Z}_{t\hat{\epsilon}} = \int_0^1 \sigma (\hat{f}_s) \, d(\hat{\rho}_sW + \rho f_s) + \hat{\epsilon} g_1 + O(\hat{\epsilon}^2),$$

furthermore, since $(h, f) \in K^\epsilon$, by the definition of $\Phi_1$, it holds that

$$\int_0^1 \sigma (\hat{f}_s) \, d(\hat{\rho}_sW + \rho f_s) = x.$$

This proves the statement (37) and the statement that $g_1$ is Gaussian is immediate from the form (38).

Finally, we determine an explicit form of the Girsanov density $G_\epsilon$ for the choice where $(h', f')$ in (36) are chosen the cheapest admissible control (cf. Definition 6.1). Similarly to classical works of Azencott, Ben Arous, and others, see, for instance, Ben Arous (1988), we show that the stochastic integrals in the exponent of $G_{\epsilon}$ are proportional to the first order term $g_1$ (with factor $\Gamma' (x)$) when evaluated at the minimizing configuration $(h', f')$.

**Lemma 6.3** We have

$$\int_0^1 h_t^2 \, dW_t + \int_0^1 f_t^2 \, dB_t = I' (x) \, g_1.$$

**Proof** See Lemma A.2.

With these preparations in place, we are now ready to prove the pricing formula from Section 3.

**Proof of Theorem 3.2** With a Girsanov factor (all integrals on $[0, 1]$)

$$G_{\epsilon} = e^{-1/\hat{\epsilon}} \int h \, dW - \frac{1}{2} \int f \, dB - \int \frac{1}{\hat{\epsilon}} f \, (h' + f') \, dt$$

and (evaluated at the minimizer)

$$G_{\epsilon}|_\ast = e^{-L(\lambda)/2} e^{-\Gamma (xS_0)/(\mu^2/2)},$$
we have, setting \( \hat{U}^x := \hat{Z}_1^x - x = \hat{g}_1 + \hat{e}^2 R_2^x \)

\[
c(\hat{x}, t) = E\left[ \exp\left( \frac{\hat{e}}{\hat{g}} \hat{Z}_1^x \right) - \exp\left( \frac{\hat{e}}{\hat{g}} \hat{U}^x \right) + G_{\hat{z}, l}|x_0 \right]
= e^{-\frac{\hat{e}}{\hat{g}}} \mathcal{E} \left[ \left( \exp\left( \frac{\hat{e}}{\hat{g}} \hat{U}^x \right) - 1 \right)^{+} G_{\hat{z}, l} \right]
= e^{-\frac{\hat{e}}{\hat{g}}} e^{\frac{\hat{e}}{\hat{g}} H} \mathcal{E} \left[ \left( \exp\left( \frac{\hat{e}}{\hat{g}} \hat{U}^x \right) - 1 \right)^{+} e^{-H(x)} \right]
= e^{-\frac{\hat{e}}{\hat{g}}} e^{\frac{\hat{e}}{\hat{g}} H} \mathcal{E} \left[ \left( \exp\left( \frac{\hat{e}}{\hat{g}} \hat{U}^x \right) - 1 \right)^{+} e^{-H(x)} \right]
\]

7. Proof of the moderate deviation expansions

In Section 2, we pointed out that (iiiic) is exactly what one gets from (call price) large deviations (8), if heuristically applied to \( x e^{2\beta} \). We now give a proper derivation based on moderate deviations.

Lemma 7.1 Assume (iiiia-b) from Assumption 2.4. Then an upper moderate deviation estimate holds both for calls and digital calls. That is, we have

(iiic) For every \( \beta \in (0, H) \), and every fixed \( x > 0 \), and \( \hat{x}_\epsilon := x e^{-2H+2\beta} \),

\[
E[|e^{X_1^x} - e^{\hat{x}_\epsilon}|^+] \leq \exp \left( -\frac{x^3 + o(1)}{2\sigma_0^2 e^{4H-4\beta}} \right)
\]

and also

\[
P[X_1^x > \hat{x}_\epsilon] \leq \exp \left( -\frac{x^3 + o(1)}{2\sigma_0^2 e^{4H-4\beta}} \right). \tag{40}
\]

Proof Recall \( \sigma(\cdot) \) smooth but unbounded and recall \( \hat{x}_\epsilon := x e^{-2H+2\beta} \). In case of \( \beta = 0 \) and \( H = 1/2 \) a large deviation principle (LDP) for \( (X_1^x e) \) is readily reduced, via exponential equivalence, to a LDP for the family of stochastic Itô integrals given by \( \sigma(\hat{e}\hat{B}) \hat{e} dZ \) for some Brownian \( Z \), \( \rho \)-correlated with \( B \). There are then many ways to establish a LDP for this family. A particularly convenient one, that requires no growth restriction on \( \sigma \), uses continuity of stochastic integration with respect to the rough path \( (B, Z, f \hat{B}dZ) = (B, Z, f \hat{B}dZ) \) in suitable metrics, for which a LDP is known (Friz and Hairer 2014, Ch 9.3). It was pointed out in Bayer et al. (2017) that a similar reasoning is possible when \( H < 1/2 \), the rough path is then replaced by a ‘richer enhancement’ of \( (B, Z) \), the precise size of which depends on \( H \), for which again one has a LDP. A moderate deviation principle (MDP) for \( (X_1^x e) \) is a LDP for \( (e^{-2H}X_1^x e) \) for \( \beta \in (0, H) \). This can be reduced to a LDP, with \( \hat{e} := e^{-2H} \hat{e} = e^{2H-2\beta} \), for

\[
e^{-2\beta} \int_0^1 \sigma(\hat{e}\hat{B}) \hat{e} dZ = \int_0^1 \sigma(\hat{e}\hat{B}) \hat{e} dZ \equiv \int_0^1 \sigma(\hat{e}\hat{B}) \hat{e} dZ
\]

with speed \( \hat{e}^2 \). Since \( \sigma(\cdot) \equiv \sigma(2\hat{e}^2) \) converges (with all derivatives) locally uniformly to the constant function \( \sigma_0 \), and one checks that the above is exponentially equivalent to the (Gaussian) family given by \( \sigma_0 Z_t \), with law \( N(0, \sigma_0^2 \hat{e}^2) \) which gives (40), even with equality. (By localization, exponential equivalence can again be done for \( \sigma \) without growth restrictions.)

We have not yet used either assumption (iiiia-b). These become important in order to extend estimate (40) to the case of genuine call payoffs. We can follow here a well-known argument (e.g. Forde and Jacquier 2009; Pham 2010; Forde and Zhang 2017) with the ‘moderate’ caveat to carry along a factor \( e^{2\beta} \). In fact, this is follows precisely the argument of Forde and Zhang (2017) where the authors carry along a factor \( \hat{e} \). (This provides a unified view on rough and moderate deviations.) The remaining details then follow essentially Appendix C. Proof of Corollary 4.13., part (ii) upper bound’ of Forde and Zhang (2017), noting perhaps that the authors use their assumptions to show validity of what we simply assumed as condition (iib), and also that one works with the quadratic rate function \( I''(0) \beta^2 = x^2 / 2\sigma_0^2 \) throughout.

Remark 7.2 By an easy argument similar to ‘Appendix C. Proof of Corollary 4.13., part (i) lower bound’ of Forde and Zhang (2017) one sees that validity of the call price upper bound (iiiic) implies the corresponding digital call price upper bound (40). For this reason, we only emphasized (iiic) but not (40) in Section 2.

In a classical work Azencott (1982) (see also Azencott 1985; Ben Arous 1988, Théorème 2) obtained asymptotic expansions of functionals of Laplace type on Wiener space, of the type \( \mathbb{E}[\exp(-F(X')/\epsilon)] \), for small noise diffusions \( X' \). This refines the large deviation (equivalently: Laplace) principle of Freidlin–Wentzell for small noise diffusions. In a nutshell, for fixed \( X_0 = x \), Azencott gets expansions of the form \( e^{-c/\epsilon} \epsilon^2 \). His ideas (used by virtually all subsequent works in this direction) are a Girsanov transform, to make the minimizing path ‘typical’, followed by localization around the minimizer (justified by a good large deviation principle), and finally a local (stochastic Taylor) type analysis near the minimizer. None of these ingredients rely on the Markovian structure (or, relatedly, PDE arguments). As a consequence (and motivation for this work) such expansions were also obtained in the (non-Markovian) context of rough differential equations driven by fractional Brownian motion (Inahama 2013; Baudoin and Ouyang 2015) with \( H < 1/2 \).

And yet, our situation is different in the sense that call price Wiener functionals do not fit the form studied by Azencott and others, nor can we in fact expect a similar expansion: Example 3.3 gives a Black–Scholes call price expansion of the form constant times \( e^{-c/\epsilon} (\epsilon^2 + \cdots) \). Azencott’s ideas are nonetheless very relevant to us: we already used the Girsanov formula.
in Theorem 3.2 in order to have a tractable expression for $J$. It thus ‘only’ remains to carry out the localization and do some local analysis.

**Proposition 7.3** Let $x > 0$ and $\beta \in (0, H)$. Then the factor $J$ is negligible in the sense that, for every $\theta > 0$,

$$ e^{\theta} \log J(\varepsilon, x e^{2\beta}) \to 0 \quad \text{as} \quad \varepsilon \to 0. $$

**Proof** Step 1. Localization Write $x_\varepsilon := x e^{2\beta}, \hat{x}_\varepsilon := x_\varepsilon e^{1-2H}$. By definition,

$$ E[(e^{X_\varepsilon} - e^{\hat{x}_\varepsilon})^2 e^{-\hat{x}_\varepsilon}] = J(\varepsilon, x_\varepsilon). $$

Fix $x, \delta > 0$ and write $\delta_\varepsilon = \delta e^{2\beta}$. We claim that (the positive quantity)

$$ J(\varepsilon, x_\varepsilon) - J_{\delta_\varepsilon}(\varepsilon, x_\varepsilon) = e^{(\theta_\varepsilon)/2} e^{-\hat{x}_\varepsilon} E[(e^{X_\varepsilon} - e^{\hat{x}_\varepsilon})1_{\hat{x}_\varepsilon > x_\varepsilon}] $$

is exponentially small, in the sense that, for some $c > 0$ and $\varepsilon^2 = \varepsilon^{4H-4\beta}$,

$$ J(\varepsilon, x_\varepsilon) - J_{\delta_\varepsilon}(\varepsilon, x_\varepsilon) = O\left(e^{-c\varepsilon^2}\right). $$

There is a battle here between the exploding factor $e^{(\theta_\varepsilon)/2}$, with exponent

$$ \frac{I(\varepsilon)}{\varepsilon^2} \sim \frac{I''(0)(\varepsilon)^2}{2\varepsilon^2} = \frac{I''(0)(\varepsilon^2)}{2\varepsilon^{4H-4\beta}}, $$

and on the other hand

$$ E[(e^{X_\varepsilon} - e^{\hat{x}_\varepsilon})1_{\hat{x}_\varepsilon > x_\varepsilon}] \leq \exp\left(-\frac{(x + \delta)^2 + o(1)}{2\varepsilon^{4H-4\beta}}\right), $$

where the given estimate is an easy consequence of Lemma 7.1. Since $I''(0) = 1/\alpha^2_0$ we see that the last factor ‘exponentially over-compensates’ the rest, so that the difference is indeed exponentially negligible.

**Step 2. Upper bound.** For any $x > 0$, recall that $\hat{U}^{\varepsilon,x} = \hat{U}^{\varepsilon}$ decomposes into a Gaussian random variable $g_1 = g_1^{\varepsilon}$ and remainder $R_2^{\varepsilon} = R_2^{\varepsilon}$. In order to control this remainder without imposing boundedness assumption on $\sigma$, we will crucially used a ‘localized remainder tail estimate’ as given in Proposition 7.4 below. For any $\varepsilon \in (0, 1]$,

$$ J_{\delta}(\varepsilon, x) = E\left[e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}} \exp\left(\frac{E}{\varepsilon} \hat{U}^{\varepsilon}\right) - 1 \right] e^{\left(T(x)/\varepsilon\right)R_2^{\varepsilon, x}} 1_{\hat{U}^{\varepsilon} \in [0, \delta]}. $$

To proceed, recall $e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}} = e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}} 1_{\hat{U}^{\varepsilon} \in [0, \delta]}$, so that, for any $\kappa > 0$,

$$ e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}} = e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}} 1_{\hat{U}^{\varepsilon} \in [0, \delta]} e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon} \in [0, \delta]} 1_{\hat{U}^{\varepsilon} \in [0, \delta]} e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon} \in [0, \delta]} \leq \frac{\kappa}{\varepsilon}. $$

Since $I'(x) > 0$ for small enough $x > 0$, it follows that $-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon} < 0$ on the event $\{\hat{U}^{\varepsilon} \in [0, \delta]\}$, which leads us to

$$ J_{\delta}(\varepsilon, x) \leq (e^{\delta} - 1) E[e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}} 1_{\hat{U}^{\varepsilon} \in [0, \delta]} \geq \kappa] $$

$$ + (e^{\delta} - 1) E[e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}} 1_{\hat{U}^{\varepsilon} \in [0, \delta]} < \kappa] $$

$$ \leq (e^{\delta} - 1) \sqrt{E[(e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}})\varepsilon]} \sqrt{P[\hat{U}^{\varepsilon} \in [0, \delta] \geq \kappa]} $$

$$ + (e^{\delta} - 1) C $$

where, by Proposition 7.4, the constant $C = C(\kappa)$ is uniform in small $\varepsilon$ and $x$. The square-root terms are computed resp. (Fernique) estimated by

$$ \exp\left(\frac{(I'(x))^2\varepsilon}{\varepsilon^2}\right) \times \exp(-c\varepsilon^2/\varepsilon^2) $$

for some $c > 0$ which depends on the law of $B$ (hence $H$), but is uniform in $\varepsilon$ and $x$. Hence, for $x$ small enough, the resulting exponent $(I'(x))^2\varepsilon - c\varepsilon^2$ is negative, which is more than enough to conclude the upper bound.

**Step 3. Lower bound.** Write $E_{\delta,x} \{|x| = E[1_{\hat{U}^{\varepsilon} \in [0, \delta]} 1_{\hat{U}^{\varepsilon} \in [0, \delta]}]$$ \leq \kappa \} \}$ and estimate

$$ E_{\delta,x} \left[ e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}/2} \left(\exp\left(\frac{E}{\varepsilon} \hat{U}^{\varepsilon}\right) - 1\right)^{1/2} \right] $$

$$ = E_{\delta,x} \left[ e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}/2} \left(\exp\left(\frac{E}{\varepsilon} \hat{U}^{\varepsilon}\right) - 1\right)^{1/2} e^{\left(T(x)/\varepsilon\right)R_2^{\varepsilon}/2} \right] $$

$$ \leq J_{\delta}(\varepsilon, x)^{1/2} E_{\delta,x} \left[ e^{\left(T(x)/\varepsilon\right)R_2^{\varepsilon}/2} \right]^{1/2} $$

where we used Cauchy–Schwarz and discarded the event $\{\hat{U}^{\varepsilon} \in [0, \delta] \geq \kappa\}$, the localized remainder estimate provides an upper bound on $E_{\delta,x} \{e^{-\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}}\}$, uniformly over small (enough) $\varepsilon$ and $x$.

It then suffices to get a suitable lower bound of the left-hand side above. Indeed, for $u \in [0, \varepsilon^2/\varepsilon^2] = [0, 0, \varepsilon^2]$ in $\varepsilon$ small enough, not dependent on $x$.

$$ u \mapsto (e^{(u/\varepsilon)}u - 1) 1/2 e^{-\left(T(x)/\varepsilon\right)u/2} \geq \gamma \left(\frac{E}{\varepsilon} u\right)^{1/2} \frac{u}{\varepsilon} $$

for a constant $\gamma > 0$ which can also be taken uniformly in small $x, \varepsilon$. Then estimate

$$ E_{\delta,x} \left[ e^{\left(T(x)/\varepsilon\right)\hat{U}^{\varepsilon}/2} \left(\exp\left(\frac{E}{\varepsilon} \hat{U}^{\varepsilon}\right) - 1\right)^{1/2} e^{\left(T(x)/\varepsilon\right)R_2^{\varepsilon}/2} \right] $$

$$ \geq \gamma \varepsilon^{1/2} H E[\hat{U}^{\varepsilon} \in [0, \varepsilon^2]) $$

As a quick sanity check, pretend zero remainder so that $\hat{U}^{\varepsilon} = \frac{\varepsilon}{\varepsilon} g_1$: dropping further the (exponentially close to probability one) event $\{|\varepsilon B| \in [0, \varepsilon^2/\varepsilon^2] \}$, a Gaussian computation then
shows that we are left with \((\gamma \varepsilon^{1/2-H} \text{ times } \hat{\varepsilon}^{1/2} \text{ times})\)

\[ E[|g|^{1/2}; g \in [0, \hat{\varepsilon}]] \sim (\text{const}) \varepsilon^{3/2}. \]

In general, set \(V^e = \hat{U}^e / \hat{\varepsilon} = g + \varepsilon R_2 \varepsilon^b\), so that \(E_r \left[ \hat{U}^e |^1/2; \hat{U}^e \in [0, \hat{\varepsilon}^2 \eta] \right] = \hat{\varepsilon}^{1/2} E_r \left[ V^e |^1/2; V^e \in [0, \hat{\varepsilon} \eta] \right].\)

At this stage, it is difficult to treat \(\hat{r} R^e\) as perturbation of \(g\) since, on the given event \(\{V^e \in [0, \hat{\varepsilon} \eta]\}\), all terms are of order \(\hat{\varepsilon}\). We can solve this issue by realizing that we can replace, throughout, \(x \times \varepsilon^{2 \beta}\). Since \(I^e(x, \varepsilon) \sim (\text{const}) x \varepsilon^{2 \beta}\), with see from (43), that in the above estimate the event \(\hat{U}^e \in [0, \hat{\varepsilon}^2 \eta] = [0, \hat{\varepsilon}^{2H} \eta] \text{ (resp. } V^e \in [0, \hat{\varepsilon} \eta] \text{ is of the same order as } E[|g|^{1/2}; g \in [0, \varepsilon^{2H} \eta]]\text{)},\) possibly with an insignificantly modified constant \(\varepsilon\). It is now straightforward to show that the behavior of \(E_r[|V^e|^{1/2}; V^e \in [0, \varepsilon^{2H-2p} \eta]]\) is of the same order as \(E[|g|^{1/2}; g \in [0, \varepsilon^{2H-2p} \eta]]\), the correct behavior (i.e. positive power of \(\varepsilon\)) is obtained by spelling out the (Gaussian) integral. ■

PROPOSITION 7.4 Localized remainder tail estimate  For every \(\kappa > 0\), there exists \(c_1, c_2 > 0\) such that, for all \(r\) and uniformly in small \(\varepsilon, x\) we have

\[ P \left( |R_2^e| > r, |\hat{\varepsilon} B|_{\infty[0,1]} < \kappa \right) \leq c_1 \exp \left( -c_2 r \right). \]

Proof We decompose \(\hat{\varepsilon}^2 R_2^e = M^e + N^e\) in terms of the (local) martingale

\[ M^e := \hat{\varepsilon} \int_0 \left[ \sigma \left( \hat{\varepsilon} \tilde{B} + \hat{\jmath} \right) - \sigma \left( \tilde{\jmath} \right) \right] \, d[\tilde{B}W + \rho B] \]

and the (bounded variation) process

\[ N^e := \int_0 \left[ \sigma \left( \hat{\varepsilon} \tilde{B} + \hat{\jmath} \right) - \sigma \left( \tilde{\jmath} \right) - \sigma' \left( \tilde{\jmath} \right) \hat{\varepsilon} \right] \, d[\tilde{\jmath} h + \rho h] \]

\[ - \frac{1}{\varepsilon} \hat{\varepsilon} \int_0 \sigma \, d\left[ \hat{\varepsilon} \tilde{B} + \hat{\jmath} \right] \, dt. \]

Let \(\tau^{\varepsilon, x}\) be the stopping time when \(\hat{\varepsilon} \tilde{B}\) first leaves the uniform ball of radius \(\kappa\). Then

\[ M^e_t := M^e_{t, \tau^{\varepsilon, x}} \]

still yields a (local) martingale. The point is that \(|[\hat{\varepsilon} \tilde{B}]|_{\infty[0,1]} < \kappa\) \(\Rightarrow \tau^{\varepsilon, x} > 1\). On this event, \(M^e |[0,1] = M^e |[0,1]\) and we can thus replace \(M^e\), in the definition of the remainder, by \(M^\varepsilon, x\). Let \(K = K^\varepsilon, x\) be the \(\varepsilon\)-fattening of \( f (t) : 0 \leq t \leq 1\), recall \(f = f^e\), then, for \(t \in [0, 1]\),

\[ d [M^\varepsilon]_t = \varepsilon^2 (\hat{\varepsilon} \tilde{B} + \hat{\jmath}f) - \sigma(f_t) \right)^2 \leq \hat{\varepsilon}^2 \| \sigma' \|^2 \varepsilon^{2H} |\hat{\varepsilon} B|^2. \]

Clearly, we can replace \(K\) by \(K\), which contains all \(K^\varepsilon, x\) for small \(\varepsilon\). To summarize, we have, on the event \(|[\hat{\varepsilon} B]|_{\infty[0,1]} < \kappa\),

\[ R^e \left( \hat{\varepsilon}^{2} |M^\varepsilon + \hat{\varepsilon}^{-2} N^e\right) \]

with \(\hat{\varepsilon}^{-2} M^{\varepsilon} \in O(\hat{\varepsilon}^{2} |B|_{\infty[0,1]}\), and, as seen by a similar (but easier) reasoning, \(\hat{\varepsilon}^{-2} N^e = O(|B|_{\infty[0,1]}\), always for fixed \(\kappa > 0\), but uniformly in small \(\varepsilon\) (equivalently, \(\hat{\varepsilon}\)) and small \(x > 0\). This clearly shows that \(\varepsilon^{-2} N^e\) has exponential tails. The same is true for the martingale part, whose bracket is \(O(\text{Gaussian})\). This is exactly the situation for the ‘model’ martingale increment 2 \(I^e_0 BDB = B^2 - 1\) which clearly has exponential tails. To make this rigorous, recall that Gaussian resp. exponential tails are characterized by \(O(\sqrt{\varepsilon})\) resp. \(O(p)\)-growth of the \(L^p\)-norms. The statement is then an easy consequence of the sharp (upper) BDG constant (Carlen and Kree 1991), known to be \(O(\sqrt{\varepsilon})\). ■

8. Proof of the implied volatility expansion

With Theorem 3.2 in place, we now turn to the proof of the implied volatility expansion, formulated in Theorem 3.6.

Proof of Theorem 3.6  We will use an asymptotic formula for the dimensionless implied variance

\[ V^e_i = t \sigma_{\text{empir}} (k_i, t)^2, \quad t > 0, \]

obtained in Gao and Lee (2014). It follows from the first formula in Remark 7.3 in Gao and Lee (2014) that

\[ V^e_i - \frac{k_i^2}{2 L_i} = O \left( \frac{k_i^2}{L_i} (k_i + |\log k_i| + \log L_i) \right), \quad t \to 0, \tag{44} \]

where \(L_i = -\log c(k_i, t), t > 0\).

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

\[ L_e = \frac{I(k r^e)}{2 \beta} + O(t^{\theta}) \tag{45} \]

as \(t \to 0\), for all \(x \geq 0\) and \(\beta \in [0, H]\) and any \(\theta > 0\). Let us first assume \(2H/(n + 1) < \beta < 2H/n\). Using the energy expansion, we obtain from (45) that

\[ L_e = \sum_{i=2}^n \frac{I^{(i)}(0)}{i!} k_i \beta^{i-2H} + O(t^{\theta}) \quad \frac{I^{(i)}(0)}{2} k_i \beta^{i-2H} \cdot \left( 1 + \sum_{i=3}^n 2 \frac{I^{(i)}(0)}{i!} k_i \beta^{i-2H} + O(t^{2H-2\beta-\theta}) \right) \tag{46} \]

as \(t \to 0\). The second term in the brackets on the right-hand side of (46) disappears if \(n = 2\).

Remark 8.1 Suppose \(n \geq 2\) and \(2H/(n + 1) < \beta < 2H/n\). Then formula (46) is optimal. Next, suppose \(n \geq 2\) and \(0 < \beta < 2H/(n + 1)\). In this case, there exists \(m \geq n + 1\) such that \(2H/(m + 1) \leq \beta < 2H/m\), and hence (46) holds with \(m\) instead of \(n\). However, we can replace \(m\) by \(n\), by making the error term worse. It is not hard to see that the following
Next, using the Taylor formula for the function $u \mapsto 1/(1 + u)$, and setting

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

we obtain from (46) that

$$(2L_t)^{-1} = \frac{2H-2\beta}{k^2I'(0)} \left[ \sum_{j=0}^{n-2} (-1)^j u^j + O(u^{n-1}) \right]$$

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

$$(2L_t)^{-1} = \frac{2H-2\beta}{k^2I'(0)} \left[ \sum_{j=0}^{n-2} (-1)^j u^j \right] + O(t^{4H-4\beta-\theta})$$

as $t \to 0$. Now, (48) gives

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

as $t \to 0$. Finally, by canceling a factor of $t$ in the previous formula, we obtain formula (14) for $2H/(n+1) \leq \beta < 2H/n$. The proof in the case where $\beta \leq 2H/(n+1)$ is similar. Here we take into account Remark 8.1. This completes the proof of Theorem 3.6.

Acknowledgments

Two referees are thanked for their useful comments. We further thank Martin Forde for valuable feedback.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

We gratefully acknowledge financial support through DFG research grants FR2943/2 and BA5484/1 (C. Bayer, P.K. Friz, B. Stemper), European Research Council Grant CoG-683164 (P.K. Friz), and SNF Early Postdoc Mobility Grant 165248 (B. Horvath) respectively.

References


### Appendix. Auxiliary lemmas

In this section we provide and prove some auxiliary lemmas, which are used in the preparations to the proof of Theorem 3.2. We start with a technical Lemma, that justifies the derivation.

**Lemma A.1** Assume $\sigma(.) > 0$ and $|\rho| < 1$. Then $K^3$ is a Hilbert manifold near any $h^0 := (h^0_1, h^0_2) \in K^3 \ni \delta h^0_0 \times \delta h^0_1$.

**Proof** Similar to Bismut (1984, p. 25) we need to show that $D\varphi_1(h)$ is surjective where $\varphi_1(h) : \delta h \to \mathbb{R}$ with

$$
\varphi_1(h) = \varphi_1(h,f) = \int_0^1 \sigma(\hat{y}) \, d(\hat{\rho} + \rho f).
$$

From

$$
\varphi_1(h + \delta h') = \int_0^1 \sigma(\hat{y}' + \delta) \, d(\hat{\rho} + \rho f + \delta(\hat{\rho} h' + \rho f'))
$$

$$
= \varphi_1(h) + \delta \int_0^1 \sigma(\hat{y}') \, d\hat{\rho}' + \rho f'
$$

$$
+ \delta \int_0^1 \sigma(\hat{y}') \, d(\hat{\rho} + \rho f) + o(\delta).
$$

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

$$(D\varphi_1)(h)(h', 0) = \hat{\rho} \int_0^1 \sigma(\hat{y}) \, d\hat{\rho}'$$

is sufficient to guarantee surjectivity of $D\varphi_1(h)$. \endProof

We now give the proof of Lemma 6.3, which determines the form of the Girsanov measure change (36) for the minimizing configuration.

**Lemma A.2** (i) Any optimal control $h^0 = (h^0_1, h^0_2) \in K^3$ is a critical point of

$$
h = (h,f) \mapsto -I\left(\varphi_1(h)\right) + \frac{1}{2} \|h\|_{L^2}\text{.}
$$

(ii) it holds that

$$
\int_0^1 h^2 \, dW + \int_0^1 \hat{y}^2 \, dB = I'(x) g_1.
$$

**Proof** (Step 1) Write $h = (h,f)$ and

$$
\varphi_1(h) = \varphi_1(h,f) = \int_0^1 \sigma(\hat{y}) \, d(\hat{\rho} + \rho f).
$$

Let $h^0 = (h^0_1, h^0_2) \in K^3$ an optimal control. Then

$$
\text{Rect}D\varphi_1\left(h^0\right) = T_{h^0} K^3 = \left\{h \in \mathcal{F}^1 : D\varphi_1(h) = 0\right\}.
$$

(There is $K^3$ to be a Hilbert manifold near $h^0$, as was seen in the last lemma.)

(Step 2) For fixed $h^0 \in \mathcal{F}^1$, define

$$
u(t) := -I\left(\varphi_1(h^0 + \nu(t))\right) + \frac{1}{2} \left\|\hat{y}^0 + \rho \nu(t)\right\|_{L^2}^2 \geq 0
$$

with equality at $t = 0$ (since $x = \varphi_1(h^0)$ and $I(x) = \frac{1}{2} \left\|\hat{y}^0\right\|_{L^2}^2$) and non-negativity for all $t$ because $\hat{y}^0 + \rho \nu(t)$ is an admissible control for reaching $x = \varphi_1(h^0 + \nu(t))$ (so that $I(x) = \text{inf}\{\cdots \leq \frac{1}{2} \left\|\hat{y}^0\right\|_{L^2}^2 + \rho \nu(t)\}$).

(Step 3) We note that $\nu(t) = 0$ is a consequence of $u \in C^1$ near 0, $\nu(0) = 0$ and $u \geq 0$. In other words, $h^0_1$ is a critical point for

$$
\mathcal{F}^1 \ni \nu(t) \mapsto -I\left(\varphi_1(h^0)\right) + \frac{1}{2} \left\|\nu(t)\right\|_{L^2}^2
$$

(Step 4) The functional derivative of this map at $h^0$ must hence be zero. In particular, for all $h \in \mathcal{F}^1$,
\[ 0 = -I'(\varphi^h) \langle D\varphi_1 (h^0), h \rangle + \langle h^0, h \rangle \]
\[ = -I'(x) \langle D\varphi_1 (h^0), h \rangle + \langle h^0, h \rangle. \]

(Step 5) With \( h^0 = (h^t, f^t) \) and \( h = (h, f) \)

\[ \langle D\varphi_1 (h^0), h \rangle = \frac{d}{dx} \int_0^1 \sigma (\hat{\xi}^t + \hat{\eta}) \ d\]
\[ \times (\hat{\rho} h^t + \hat{\rho} f^t + \epsilon (\hat{\rho} h + \hat{\rho} f)) \]
\[ = \int_0^1 \sigma (\hat{\xi}^t) d(\hat{\rho} h + \hat{\rho} f) + \int_0^1 \sigma (\hat{\xi}^t) d (\hat{\rho} h^t + \hat{\rho} f^t) \]

By continuous extension, replace \( h = (h, f) \) by \((W, B)\) above and note that
\[ \langle D\varphi_1 (h^0), (W, B) \rangle = g_1 \]

since indeed
\[ g_1 = \int_0^1 \sigma (\hat{\xi}_t) d(\hat{\rho} W_t + \rho B_t) + \sigma (\hat{\eta}_t) \hat{B}_t d(\hat{\rho} h_t + \rho f_t). \]

Hence
\[ \int_0^1 \dot{h}^t dW + \int_0^1 \dot{f}^t dB = I'(x) g_1. \]

\[ \square \]