Received: 12 December 2022 | Accepted: 13 December 2022

DOI: 10.1111/mafi.12374

ORIGINAL ARTICLE

MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics WILEY

Reconstructing volatility: Pricing of index options under rough volatility

Peter K. Friz | **Thomas Wagenhofer**

Institut für Mathematik, Technische Universität Berlin, Berlin, Germany

Correspondence
Peter K. Friz, Institut für Mathematik, Technische Universität Berlin, Berlin, Germany.
Email: friz@math.tu-berlin.de

To the memory of Marco Avellaneda

Funding information
DFG, Grant/Award Number: EXC 2046

Abstract

Avellaneda et al. (2002, 2003) pioneered the pricing and hedging of index options – products highly sensitive to implied volatility and correlation assumptions – with large deviations methods, assuming local volatility dynamics for all components of the index. We present an extension applicable to non-Markovian dynamics and in particular the case of rough volatility dynamics.

1 | INTRODUCTION

Given N assets, whose discounted price has risk-neutral dynamics

$$dS_t^i/S_t^i = \sigma_i(t,\omega)\,dB_t^i,$$

with Brownian driving noise, we consider an index of the form

$$I_t := \sum_{i=1}^{N} w_i S_t^i$$

where the w_i are positive constants. From standard Itô rules, assuming $\langle B^i, B^j\rangle_t/dt = \rho_{ij}$,

$$\frac{d\langle I\rangle_t/dt}{I_t^2} = \sum_{i,j=1}^{N} p_t^i p_t^j \rho_{ij}\sigma_i(t,\omega)\sigma_j(t,\omega) =: \sigma_I^2(t,\omega),$$

with $p_t^i = w_i S_t^i / I_t$. Taking $t = 0$, we have today's (deterministic) spot volatilities $\sigma_{i,0} := \sigma_i(0,\omega)$, with similar notation being used for the index. One then has the standard formula that relates spot volatilities of the index and its components:

$$\sigma_{I,0}^2 = \sum_{i,j=1}^{N} \rho_{ij}\sigma_{i,0}\sigma_{j,0}p_0^i p_0^j. \tag{1}$$

In Avellaneda et al. (2002, 2003) we asked the question how to integrate volatility skew information for index components more explicitly into (1) and proposed a method for relating the implied volatility skew of the index to the implied volatility skew of the components. Practical motivation, much related to Marco's activity at the time as head of the options research team at Gargoyle Strategic Investments, comes from dispersion trading: the strategy of selling (buying) index options, while buying (selling) options on the index components. The topic stayed close to Marco's heart and dispersion trading remained a topic in his NYU classes for years to come.

The basic idea of these works was the use of short-time large deviations from diffusion processes, pioneered by Varadhan (1967). This topic also stayed close to the heart of the first author of this note, as witnessed by De Marco et al. (2013), Deuschel et al. (2014), Friz et al. (2015), Bayer et al. (2015), De Marco and Friz (2018), Friz et al. (2021b), Friz et al. (2021a).

The starting point of Avellaneda et al. (2002, 2003) is the familiar relation[1]

$$\sigma_{I,\mathrm{loc}}^2(t, I_t) = \mathbb{E}\left[\sigma_{I,\mathrm{stoch}}^2(t,\omega)\middle| I_t\right] \tag{2}$$

together with assumed local volatility dynamics for the components of the index, that is

$$\sigma_i(t,\omega) = \sigma_i(t, S^i), \quad i = 1, \dots, N.$$

Setting $\tilde{\sigma}_i(x) = \sigma_i\left(0, S_0^i e^{x_i}\right)$ and also $\tilde{\sigma}_{I,\mathrm{loc}}(\bar{x}) = \sigma_{I,\mathrm{loc}}(0, I_0 e^{\bar{x}})$ it holds in the short-time limit that

$$\tilde{\sigma}_{I,\mathrm{loc}}^2(\bar{x}) = \sum_{i,j=1}^{N} \rho_{ij}\tilde{\sigma}_i(x_i^*)\tilde{\sigma}_j(x_j^*)p_i(\mathbf{x}^*)p_j(\mathbf{x}^*).$$

where $\mathbf{x}^* \in \Gamma_{\bar{x}} = \{\mathbf{x} : \sum_{i=1}^N w_i S_0^i e^{x_i} = e^{\bar{x}}\}$ minimizes the distance to the origin $\mathbf{x} = 0$, relative to the associated Riemannian metric, cf. Varadhan (1967). It is generically true (and here assumed – but see, e.g., Bayer et al. (2015)) that $\mathbf{x}^*$ is unique. We also set

$$p_i(\mathbf{x}) = \frac{w_i S_0^i e^{x_i}}{\sum_{j=1}^N w_j S_0^j e^{x_j}},$$

[1] In what follows we assume basic familiarity with stochastic, local and implied volatility, as found, for example, in Gatheral (2006); Bergomi (2016). Formula (2) goes back to Gyöngy (1986), Dupire (1994); Derman and Kani (1994), also revisited in (Brunick & Shreve, 2013; Bentata & Cont, 2015).

which represents the percentage of the stock i in the index with $S^i = S_0^i e^{x_i}$. Furthermore $\mathbf{x}^*$ solves the non-linear system

$$\int_0^{x_i^*} \frac{du}{\tilde{\sigma}_i(u)} = \lambda \sum_{j=1}^{N} \rho_{ij} p_j(\mathbf{x}^*) \tilde{\sigma}_j(x_j^*), \qquad \forall i = 1, \dots, N$$

$$e^{\bar{x}} = \sum_{i=1}^{N} w_i S^i(0) e^{x_i^*}.$$

Here λ corresponds to the Langrange multiplier of the price constraint $\Gamma_{\bar{x}}$.

Using approximate relations between local and implied volatility, notably the $1/2$-rule, valid in the short-time and ATM regime (see also Gatheral (2006); Gatheral et al. (2012) and references therein), this led us to

$$\tilde{\sigma}_{I,\text{loc}}(\bar{x}) = \sqrt{\sum_{i,j=1}^{N} \rho_{ij} p_i(\mathbf{x}^*) p_j(\mathbf{x}^*) \left(2\sigma_i^{\text{impl}}(x_i^*) - \sigma_i^{\text{impl}}(0)\right)\left(2\sigma_j^{\text{impl}}(x_j^*) - \sigma_j^{\text{impl}}(0)\right)};$$

together with a first order approximation for the most likely configuration $\mathbf{x}^*$ from (Avellaneda et al., 2002, Eq.(15))

$$x_i^* = \frac{\bar{x}}{\sigma_I^2(0)} \sum_{j=1}^{N} \rho_{ij}, w_j \sigma_i(0) \sigma_j(0), \quad i = 1, .., N. \tag{3}$$

Keep in mind, that $\tilde{\sigma}_{I,\text{loc}}$ corresponds to the local volatility at time $t = 0$, hence equality holds above, despite using approximations.

With the harmonic average formula that expresses $\sigma_I^{\text{impl}}(\bar{x})$ in terms of $\tilde{\sigma}_{I,\text{loc}}(\bar{x})$, this essentially concludes the task of refining (1) in a tractable way that allows to integrate volatility skew information.

These (index) results were revisited and extended by various authors, including (Henry-Labordère, 2008, Sec.7.2) and (Guyon and Henry-Labordère, 2014, Sec.12.9), notably towards local correlation models. The purpose of the present note is to revisit Avellaneda et al. (2002) in a way that makes it clear that one can do, in fact, without the Markovian structure that seemed rather crucial in Avellaneda et al. (2002, 2003); Henry-Labordère (2008); Bayer et al. (2015) and related works. While in a diffusion setting, short-time can be considered as special case of small noise, cf. Osajima (2015); Deuschel et al. (2014)), this is not so in a rough volatility setting and we should emphasize that we work in a small noise setting here, rather than the short-time regime of Forde and Zhang (2017). Last not least, following Avellaneda et al. (2002, 2003) and to illustrate our approach we have kept the constant correlation structure, leaving any extension to stochastic and local correlations to future work.

2 | INDEX OPTIONS UNDER ROUGH VOLATILITY

2.1 | Rough volatility index dynamics

We consider the model case where components follow rough volatility dynamics. For this let $\left(W^1, \dots, W^N, \overline{W}^1, \dots, \overline{W}^M\right)$ be independent Brownian motions and consider for $i = 1, \dots, N$ a model of the form

$$dS_t^i/S_t^i = f_i(\widehat{W}_t^i)\,dB_t^i, \tag{4}$$
$$B^i = c_i W^i + \bar{c}_i \overline{W}^i, \qquad c_i^2 + \bar{c}_i^2 = 1,$$
$$\widehat{W}_t^i = \int_0^t K^{H_i}(t,s)\,dW_s^i, \quad K^H(t,s) = C(H)|t-s|^{H-1/2}.$$

Here $\widehat{W}^i$ is a Riemann-Liouville fractional Brownian motion with Hurst parameter $H_i \in (0, 1/2]$. The constant $C(H)$ is usually chosen such that $\widehat{W}$ has unit variance. Furthermore, the c_i quantify the correlation between B^i and W^i, that is, between the respective driving factors of the underlying and the stochastic volatility process. In general, one would need to specify the full correlation structure of

$$\left(B^1, \dots, B^N; W^1, \dots, W^N\right).$$

To keep things simple, we assume $c_i = -1$, which is not an unreasonable assumption at all in equity. Here the sign of c_i does not matter, as one could redefine f_i accordingly. We are led to a path-dependent one factor stochastic volatility model,

$$dS_t^i/S_t^i = f_i(\widehat{W}_t^i)\,dW_t^i,$$

somewhat similar in spirit to Hobson and Rogers (1998) and Guyon and Henry-Labordère (2014). As before we set

$$d\langle W^i, W^j\rangle_t/dt = d\langle B^i, B^j\rangle_t/dt = \rho_{ij}.$$

2.2 | Small noise LDP for the index under rough volatility

We introduce the small noise problem

$$dS_t^{i,\varepsilon}/S_t^{i,\varepsilon} = f_i\left(\varepsilon\widehat{W}_t^{\,i}\right) d(\varepsilon W_t^i)$$
$$I_t^\varepsilon := \sum_{i=1}^N w_i S_t^{i,\varepsilon} \tag{5}$$

Assume (w.l.o.g.) that $S_0^{i,\varepsilon} \equiv 1$ and $\sum_{i=1}^N w_i = 1$, so that $I_0^\varepsilon = 1$. Introduce the index log-price process

$$J^\varepsilon = \log(I^\varepsilon), \quad \text{such that } I^\varepsilon = e^{J^\varepsilon}.$$

Write $\mathcal{H}^N$ for the absolutely continuous paths from $[0,1] \to \mathbb{R}^N$, started at zero, with L^2-derivative. Writing $\langle .,. \rangle$ for the L^2-inner product of both $\mathbb{R}^N$ and $\mathbb{R}$ valued paths, we have the usual Cameron-Martin inner product

$$\langle h, h\rangle_{\mathcal{H}^N} = \langle \dot{h}, \dot{h}\rangle = \sum_{i=1}^{N} \langle \dot{h}^i, \dot{h}^i\rangle.$$

For invertible ρ, we also define

$$\langle h, h\rangle_\rho := \langle \dot{h}, \rho^{-1}\dot{h}\rangle = \sum_{i,j=1}^{N} \langle \dot{h}^i, (\rho^{-1})_{ij}\dot{h}^j\rangle.$$

If W denotes a N-dimensional Brownian motion with covariance matrix ρ, then εW, viewed as $C([0,1],\mathbb{R}^N)$-valued random variable, satisfies a LDP with good rate function $\langle h, h\rangle_\rho/2$ whenever $h \in \mathcal{H}^N$, and $+\infty$ otherwise. One can also treat non-invertible ρ, at the price of working with degenerate inner products, so that only a proper subspace $\mathcal{H}_\rho \subset \mathcal{H}$ has a finite rate function.

Theorem 2.1. *Assume f_i is smooth for $i = 1, \dots, N$. Assume ρ is invertible. Then $J_1^\varepsilon = \log(I_1^\varepsilon)$ satisfies a LDP with speed ε^2 and good rate function*

$$\Lambda(x) = \inf_{h\in\mathcal{H}^N} \left\{ \frac{1}{2}\langle h, h\rangle_\rho \,:\, \phi(h) = x \right\} = \frac{1}{2}\langle h^x, h^x\rangle_\rho,$$

called energy function, *where*

$$\phi(h) = \log\left(\sum_{i=1}^{N} w_i \exp\big(\phi_i(h^i)\big)\right), \qquad \phi_i(h^i) = \int_0^1 f_i(\hat{h}^i)\, dh^i.$$

The infimum is attained at some not necessarily unique $h^x \in \mathcal{H}^N$.

Proof. If we had $J_1^\varepsilon = \phi(\varepsilon W^1, \dots, \varepsilon W^N)$ for a continuous map ϕ, this would be a plain consequence of Schilder's theorem (LDP for Brownian motion) and the contraction principle. There are many ways to show that the results still hold. A standard method is by means of the so-called extended contraction principle, as in Forde and Zhang (2017), see also Jacquier and Pannier (2022); Gulisashvili (2022). An alternative and arguably quite elegant argument was put forward in Bayer et al. (2020), see also Fukasawa and Takano (2022). Namely J_1^ε *is* the continuous image of $(\varepsilon W^1, \dots, \varepsilon W^N)$ *plus* certain iterated (Itô) integrals, in the spirit of rough paths. For this enhanced noise, Schilder type large deviations are known and the result follows. □

Remark 2.2. Note that $\mathcal{H}^N \ni h = (h^1, \dots, h^N) \mapsto \phi(h) \in \mathbb{R}$ is smooth, to the extend the f_i permit, and maps $0 \in \mathcal{H}^N$ to $\phi(0) = 0 \in \mathbb{R}$. The function $\hat{h}^i$ is given by $\hat{h}^i_t = \int_0^t K^{H_i}(t,s)\, dh^i_s$.

2.3 | Expansion of abstract energy function

Motivated by numerous papers on large deviations for stochastic and rough volatility we make the following definition.

Definition 2.3. (i) Call C^k*-reasonable* any map $\phi : \mathcal{H}^N \to \mathbb{R}$ which is weakly continuous and C^k in Fréchet sense with $k \geq 0$, further to

$$\phi(0) = 0 \in \mathbb{R}, \quad D\phi(0) \neq 0 \in \mathcal{H}^N.$$

(ii) Call $\Lambda : \mathbb{R} \to [0, \infty]$ a C^k*-good energy function* if, for some C^k-reasonable ϕ, it is a good rate function of the form

$$\Lambda(x) = \inf_{h \in \mathcal{H}^N} \left\{ \frac{1}{2} \langle h, h \rangle_\rho : \phi(h) = x \right\}.$$

Note $\Lambda(0) = 0$, with infimum trivially attained by $h^0 = 0 \in \mathcal{H}^N$.

Proposition 2.4. *Any C^1-good energy function is continuous and increasing (resp. decreasing) on $\mathbb{R}^+$ (resp. $\mathbb{R}^-$).*

Proof. Follows from Proposition 6.1 and Proposition 6.6. □

Remark 2.5. In the special case of the rough Bergomi model, monoticity of the rate functions is shown in (Gulisashvili, 2018, Lemma 15), with proof attributed to C. Bayer.

The following "abstract" theorem gives an expansion of the rate function $\Lambda(x)$ and only involves the Fréchet derivatives

$$\phi_0' := D\phi(0) \in \mathcal{H}^N, \qquad \phi_0'' := D^2\phi(0) \in (\mathcal{H}^N \times \mathcal{H}^N)^\star.$$

As usual $(\mathcal{H}^N \times \mathcal{H}^N)^\star$ denotes the topological dual space of $\mathcal{H}^N \times \mathcal{H}^N$.

Theorem 2.6. *Assume Λ is a C^3-good energy function. Then, as $x \to 0$,*

$$\Lambda(x) = \frac{1}{2}\langle h^x, h^x \rangle_\rho = \left(\frac{x^2}{2} - \frac{x^3}{2\sigma_0^4} \phi_0''\left(\rho\phi_0', \rho\phi_0'\right) \right) \Big/ \sigma_0^2 + o(x^3)$$

where

$$\sigma_0^2 := \langle \rho\phi_0', \rho\phi_0' \rangle_\rho = \langle \phi_0', \rho\phi_0' \rangle_{\mathcal{H}^N}.$$

2.4 | Consequences for implied volatility in the small noise regime

The following result has nothing to do with indices, and only assumes that the asset price process I_t^ε comes with a parameter $\varepsilon > 0$, so that $\log(I_1^\varepsilon)$, the log-price at time 1, satisfies a LDP with a good rate function $\Lambda = \Lambda(x)$. Λ is assumed to be continuous and such that $\Lambda(x) = 0$ iff $x = 0$; cf. Assumption (A1) in Friz et al. (2021a).

The following theorem can be seen as variation of the "BBF formula", Berestycki et al. (2004) (short time) and also appears in Osajima (2015) (small noise). Remark that in short-time asymptotics for diffusion models, also discussed in Deuschel et al. (2014), the energy function Λ has the interpretation as geodesic point-subspace distance. Continuity and monotonicity of Λ is then clear. In absence of this structure, authors including Forde and Zhang (2017) and Gulisashvili (2022) express their large deviations in terms of $\Lambda^\star(x) = \inf_{y>x} \Lambda(y)$ whenever $x \geq 0$ and $\Lambda^\star(x) = \inf_{y<x} \Lambda(y)$ for $x < 0$. As will be pointed out in the proof, this is not necessary in our setting, despite dealing with a somewhat generic non-Markovian small noise situation.

Theorem 2.7 (Implied volatility). *Under the above assumption it follows that*

$$(\sigma^\varepsilon_{impl})^2(x,1) \sim_{\varepsilon\downarrow 0} \frac{x^2}{2\Lambda(x)}.$$

Corollary 2.8 (Spot variance and skew). *Let Λ have the local expansion of Theorem 2.6. Then*

$$(\sigma^\varepsilon_{impl})^2(x,1) \sim_{\varepsilon\downarrow 0} \frac{x^2}{2\Lambda(x)} \sim_{x\downarrow 0} \sigma_0^2 + xS_0 + o(x)$$

with spot variance implied variance skew given by, respectively,

$$\begin{aligned} \sigma_0^2 &= \langle \phi_0', \rho\phi_0' \rangle_{\mathcal{H}^N}, \\ S_0 &= \phi_0''(\rho\phi_0', \rho\phi_0')/\sigma_0^2. \end{aligned}$$

Proof of Theorem 2.7. We content ourselves with a sketch of the argument. The LDP together with Proposition 2.4 give directly OTM binary call price aymptotics,

$$\mathbb{P}[X_1^\varepsilon > x] \approx \exp\left\{-\frac{\Lambda(x)}{\varepsilon^2}\right\}.$$

Matching exponents with OTM Bachelier prices and their Gaussian tail behavior,

$$\mathbb{P}[\sigma\varepsilon W_1 > x] \approx \exp\left\{-\frac{x^2}{2\sigma^2\varepsilon^2}\right\}.$$

For the effective normal implied volatility, as $\varepsilon \downarrow 0$, we see that

$$(\sigma^\varepsilon_{\text{norm}})^2(x,1) \sim \frac{x^2}{2\Lambda(x)}.$$

The same asymptotics is valid for $\sigma^\varepsilon_{\text{impl}}$, the Black-Scholes implied volatility. This follows from the fact that large asymptotics for binary and classical (OTM) option are identical. The only caveat here is a 1+ moment assumption to treat call options, cf. Friz et al. (2021a). □

Proof of Corollary 2.8. We only need to give a simple expansion for small x. Due to Theorem 2.6:

$$\frac{x^2}{2\Lambda(x)} = \sigma_0^2\left(\frac{1}{1 - x\sigma_0^{-4}\phi_0''(\rho\phi_0', \phi_0') + o(x)}\right)$$
$$= \sigma_0^2\left(1 + x\sigma_0^{-4}\phi_0''(\rho\phi_0', \phi_0') + o(x)\right).$$

□

2.5 | Index spot variance and skew

We now re-introduce, step-by-step, the structure of interest to us. We start with a general index result, that applies, for instance to the index with local volatility components considered in Avellaneda et al. (2002). In this case $\phi_i(h^i) = y_1^\varepsilon$ is the time-1 solution map of the ODE

$$dy^i = f_i(y^i)\,dh^i,$$

with initial value $y^i = 0$. For Lipschitz f_i this solution map is well-posed. This result also applies to rough volatility, with

$$\phi_i(h^i) = \int_0^1 f_i(\hat{h}^i)\,dh^i.$$

Note that in this case $\langle\phi_{i0}', k^i\rangle_{\mathcal{H}} = \int_0^1 f_i(0)dk^i$ implying that $\phi_{i0}' \equiv D\phi_i(0) \in \mathcal{H}$ has, as element of the Cameron-Martin space, constant velocity $\sigma_i = f_i(0)$. A similar statement holds in the local volatility example. This motivates our condition (6) below. The following result is a consequence of the general Corollary 2.8, injecting the additional information of weights and correlations.

Proposition 2.9 (Index energy). *Assume that* $\sum_{i=1}^N w_i = 1$ *and let* $\phi_i : \mathcal{H} \to \mathbb{R}$ *be* C^3*-reasonable*[2] *with* $i = 1, \dots, N$. *Let*

$$\exp\left(\phi(h^1, \dots, h^N)\right) := \sum_{i=1}^N w_i \exp\left(\phi_i(h^i)\right).$$

Set $\phi_{i0}' = D\phi_i(0) \in \mathcal{H}$ *as well as* $\phi_{i0}'' = D^2\phi_i(0) \in (\mathcal{H}\times\mathcal{H})^\star$ *and assume*

$$\phi_{i0}' = \sigma_i Id \in \mathcal{H}, \qquad Id : t \mapsto t. \tag{6}$$

Then

$$(\sigma_{impl}^\varepsilon)^2(x, 1) \sim_{x\downarrow 0} \sigma_I^2 + xS_I + o(x),$$

[2] In the sense of Definition 2.3 with $N = 1$.

with

$$\sigma_I^2 = \sum_{i,j=1}^{N} w_i w_j \rho_{ij} \sigma_i \sigma_j,$$

$$\mathcal{S}_I = -\sigma_I^2 + \left(\sum_{i=1}^{N} w_i \Sigma_i^2 \left(\phi_{i0}''(Id, Id) + \sigma_i^2 \right) \right) \Big/ \sigma_I^2, \qquad \Sigma_i := \sum_{j=1}^{N} w_j \rho_{ij} \sigma_j. \tag{7}$$

Remark 2.10. (i) Assumption (6) is satisfied in large classes of examples. However, as the proof shows one can do without but the formulae are a bit less pretty. (ii) The expression for spot variance σ_I^2 is consistent with the classical formula for Index Options that we gave in (1), cf. Avellaneda et al. (2002) or Guyon and Henry-Labordère (2014). The expression for $\mathcal{S}_I$ can be seen an answer to the problem, first tackled in Avellaneda et al. (2002), of how to integrate volatility skew information into such a classical formula.

Remark 2.11 (Most-likely configuration). Consider, up to first order, $\bar{x} = \phi(\bar{h}) := \sum_{i=1}^{N} w_i \phi_i(\bar{h}^i)$ where $\bar{h} = (\bar{h}^1, \dots, \bar{h}^N)$ is the minimizer for the constraint $\bar{x} = \phi(\bar{h})$. We then know $\bar{h} = \bar{x}a$, as well as $\phi(\bar{h}) := \bar{x}\langle \phi_0', a\rangle_{H^N}$ with

$$a^i := \frac{(\rho\phi_0')^i}{\langle \rho\phi_0', \rho\phi_0'\rangle_\rho} = \frac{(\rho\phi_0')^i}{\sigma_I^2(0)} \Rightarrow \bar{h}^i = \frac{\bar{x}}{\sigma_I^2(0)} (\rho\phi_0')^i \quad i = 1, \dots, N.$$

We thus have that $\bar{x}^i = \phi_i(\bar{h}^i)$ equals, to first order and in agreement with (3),

$$\langle \phi_{i0}', \bar{h}^i \rangle_H = \frac{\bar{x}}{\sigma_I^2(0)} \langle \phi_{i0}', (\rho\phi_0')^i \rangle_H = \frac{\bar{x}}{\sigma_I^2(0)} \sum_{j=1}^{N} \rho_{ij} w_j \langle \phi_{i0}', \phi_{j0}' \rangle_H$$

$$= \frac{\bar{x}}{\sigma_I^2(0)} \sum_{j=1}^{N} \rho_{ij} w_j \sigma_i(0) \sigma_j(0),$$

where we used $\phi_0' = \sum_{j=1}^{N} w_j \phi_{j0}' \Rightarrow (\rho\phi_0')^i = \sum_{j=1}^{N} \rho_{ij} w_j \phi_{j0}'$.

Proof of Proposition 2.9. Let $k = (k^1, \dots, k^N) \in \mathcal{H}^N$. We apply the function $\exp\big(\phi(h)\big) = \sum_{i=1}^{N} w_i \exp\big(\phi_i(h^i)\big)$ with $h = \varepsilon k$. The l.h.s. of $\exp(\phi(\varepsilon k))$ then expands to

$$1 + \phi(\varepsilon k) + \phi^2(\varepsilon k)/2 + o(\varepsilon^2) = 1 + \varepsilon \langle \phi_0', k \rangle_{\mathcal{H}^N} + \frac{\varepsilon^2}{2} \left(\phi_0''(k, k) + \langle \phi_0', k \rangle_{\mathcal{H}^N}^2 \right) + o(\varepsilon^2).$$

The r.h.s. expands to the weighted sum of the same expression with $\phi(\varepsilon k)$ replaced by $\phi_i(\varepsilon k^i)$, namely

$$1 + \varepsilon \sum_{i=1}^{N} w_i \langle \phi_{i0}', k^i \rangle_{\mathcal{H}} + \frac{\varepsilon^2}{2} \sum_{i=1}^{N} w_i \left(\phi_{i0}''(k^i, k^i) + \langle \phi_{i0}', k^i \rangle_{\mathcal{H}}^2 \right) + o\big(\varepsilon^2\big).$$

Power matching gives

$$\langle \phi_0', k\rangle_{\mathcal{H}^N} = \sum_{i=1}^N w_i \langle \phi_{i0}', k^i\rangle_{\mathcal{H}},$$

implying that $\phi_0' = \left(w_1\phi_{10}', \dots, w_N\phi_{N0}'\right) \in \mathcal{H}^N$. For the second order

$$\phi_0''(k,k) = -\langle \phi_0', k\rangle_{\mathcal{H}^N}^2 + \sum_{i=1}^N w_i\left(\phi_{i0}''(k^i,k^i) + \langle \phi_{i0}', k^i\rangle_{\mathcal{H}}^2\right).$$

We note that

$$\left(\rho\phi_0'\right)^i = \sum_{j=1}^N \rho_{ij} w_j \phi_{j0}'.$$

We enter this expression into $\mathcal{S}_I = \phi_0''(\rho\phi_0', \rho\phi_0')/\sigma_I^2$, see Corollary 2.8. This gives

$$\mathcal{S}_I = (\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3)/\sigma_I^2,$$

with

$$\mathcal{S}_1 = -\left(\sum_{i,j=1}^N w_i w_j \rho_{ij}\langle \phi_{i0}', \phi_{j0}'\rangle_{\mathcal{H}}\right)^2 = -\sigma_I^4$$

and

$$\begin{aligned}\mathcal{S}_2 &= \sum_{i=1}^N w_i \phi_{i0}''(k^i,k^i) = \sum_{i,j,\ell=1}^N w_i w_j w_\ell \rho_{ij}\rho_{i\ell}\phi_{i0}''(\phi_{j0}', \phi_{\ell 0}') \\ &= \sum_{i=1}^N w_i \phi_{i0}''(\mathrm{Id},\mathrm{Id})\left(\sum_{j=1}^N w_j\rho_{ij}\sigma_j\right)^2.\end{aligned}$$

where the last equality holds under the assumption of (6). Here Id denotes the scalar Cameron–Martin path $t \mapsto t$ with velocity 1. At last,

$$\mathcal{S}_3 = \sum_{i=1}^N w_i\left(\sum_{j=1}^N \rho_{ij} w_j \langle \phi_{i0}', \phi_{j0}'\rangle_{\mathcal{H}}\right)^2 = \sum_{i=1}^N w_i\sigma_i^2\left(\sum_{j=1}^N w_j\rho_{ij}\sigma_j\right)^2,$$

where the last equality again holds if we assume (6).

Set $\Sigma_i = \sum_{j=1}^N w_j\rho_{ij}\sigma_j$ and note, always under the assumption of (6),

$$\mathcal{S}_2 = \sum_{i=1}^N w_i\phi_{i0}''(\mathrm{Id},\mathrm{Id})\Sigma_i^2, \qquad \mathcal{S}_3 = \sum_{i=1}^N w_i\sigma_i^2\Sigma_i^2.$$

In summary, we conclude by writing

$$\mathcal{S}_I = \frac{\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3}{\sigma_I^2} = -\sigma_I^2 + \left(\sum_{i=1}^N w_i \Sigma_i^2 \left(\phi_{i0}''(\mathrm{Id}, \mathrm{Id}) + \sigma_i^2 \right) \right) \Big/ \sigma_I^2.$$

□

Remark 2.12. Consider the case $N = 1$. In this case $w_1 = 1, \Sigma_1 = \sigma_1$ and $\sigma_I = \sigma_1$. Hence $\mathcal{S}_I$ reduces to $\phi_{10}''(\mathrm{Id}, \mathrm{Id})$ in agreement with the skew expression of Corollary 2.8, applied with $\rho = \pm 1$.

3 | RETURN TO ROUGH VOLATILITY

We now return to the model defined by (5), that is, dynamics of the form

$$dS_t^{i,\varepsilon} / S_t^{i,\varepsilon} = f_i \left(\varepsilon \widehat{W}_t^i \right) d(\varepsilon W_t^i)$$

$$I_t^\varepsilon = \sum_{i=1}^N w_i S_t^{i,\varepsilon}$$

Given $H \in (0, 1/2]$ we choose the kernel $K^H(t,s) = \sqrt{2H}(t-s)^{H-1/2}$, such that $\widehat{W}_1 = \int_0^1 K^H(1,s) dW_s$ has unit variance, cf. Bayerq et al. (2019). With this kernel we define the operator $\mathcal{K}^H : \mathcal{H} \to \mathcal{H}$ such that $\left(\mathcal{K}^H h \right)(t) = \int_0^t h(s) K^H(t,s)\, ds$. Note that a short calculation shows

$$\langle \mathcal{K}^H 1, 1 \rangle = \frac{\sqrt{2H}}{\left(H + \frac{1}{2} \right)\left(H + \frac{3}{2} \right)}.$$

3.1 | Single asset, one-factor rough volatility dynamics

As as warm-up, we consider the case of $N = 1$ asset, with trivial correlation "matrix" 1. By Corollary 2.8,

$$\sigma_0{}^2 = \langle \phi_0', \phi_0' \rangle_{\mathcal{H}}, \mathcal{S} = \phi_0''(\phi_0', \phi_0') / \sigma_0^2.$$

We can therefore find all relevant terms by expanding $\phi(\varepsilon k)$ to order $o(\varepsilon^2)$:

$$\begin{aligned} \phi(\varepsilon k) &= \int_0^1 f(\varepsilon \hat{k}) d(\varepsilon k) \\ &\approx \varepsilon \left(\int_0^1 f(0) dk \right) + \varepsilon^2 \int_0^1 f'(0) \hat{k} dk = \varepsilon f_0 \langle 1, \dot{k} \rangle + \frac{\varepsilon^2}{2} 2 f_0' \langle \mathcal{K}^H \dot{k}, \dot{k} \rangle. \end{aligned}$$

From this we read off $\langle \phi_0', k\rangle_{\mathcal{H}} = f_0\langle 1, \dot{k}\rangle$ as well as $\phi_0''(k,k) = 2f_0'\langle \mathcal{K}^H \dot{k}, \dot{k}\rangle$. In particular ϕ_0' has constant velocity

$$\dot{\phi}_0' \equiv f_0.$$

If $f \in C^2$ and $f_0 \neq 0$, then ϕ is also C^k-reasonable, see Definition 2.3. By Corollary 2.8 we see

$$\sigma_0^2 = \langle \phi_0', \phi_0'\rangle_{\mathcal{H}} = \langle \dot{\phi}_0', \dot{\phi}_0'\rangle = f_0^2,$$
$$\mathcal{S}_0 = \phi_0''(\phi_0', \phi_0')/\sigma_I^2 = 2f_0'\langle \mathcal{K}^H 1, 1\rangle.$$

By the chain-rule,

$$\mathcal{S}_0 := \partial_x \left(\sigma_{\mathrm{impl}}^0\right)^2 (x,1)|_{x=0} = 2f_0 \partial_x \sigma_{\mathrm{impl}}^0 (x,1)|_{x=0}.$$

Hence in the small noise regime we have the following ATM implied volatility skew:

$$\partial_x \sigma_{\mathrm{impl}}^0 (x,1)|_{x=0} = \frac{\mathcal{S}_0}{2f_0} = \frac{f_0'}{f_0}\langle \mathcal{K}^H 1, 1\rangle.$$

Remark 3.1. This result is consistent with the skew formula in Bayerq et al. (2019), see equation (8) below.

$$\frac{\sigma_{\mathrm{impl}}\left(yt^{1/2-H+\beta}, t\right) - \sigma_{\mathrm{impl}}\left(zt^{1/2-H+\beta}, t\right)}{(y-z)t^{1/2-H+\beta}} \sim \rho \frac{\sigma_0'}{\sigma_0}\langle \mathcal{K}^H 1, 1\rangle t^{H-1/2}. \tag{8}$$

Since we deal with small noise rather than short time, there is no extra $t^{H-1/2}$ blowup factor here!

3.2 | Index with one-factor rough volatility components

Consider now $N \in \mathbb{N}$ assets, with (non-degenerate) correlation matrix ρ. Using notation from Proposition 2.9, we can recycle the single asset computations. For $i = 1, \dots, N$,

$$\sigma_{i0} = f_{i0} := f_i(0), \qquad \phi_{i0}'' = 2f_{i0}'\langle \mathcal{K}^{H_i} 1, 1\rangle.$$

The proof of Proposition 2.9 shows that $\phi_0' \neq 0$ if $f_{i0} \neq 0$ for some i, which together with C^2-regularity of f implies that ϕ is C^3-reasonable, see Definition 2.3. Application of the second part of Proposition 2.9 gives index spot variance and skew

$$\begin{aligned} \sigma_I^2 &= \sum_{i,j=1}^N w_i w_j \rho_{ij} f_{i0} f_{j0}, \\ \mathcal{S}_I &= -\sigma_I^2 + \left(\sum_{i=1}^N w_i \Sigma_i^2 \left(2f_{i0}'\langle \mathcal{K}^{H_i} 1, 1\rangle + f_{i0}^2\right)\right) \Big/ \sigma_I^2, \end{aligned} \tag{9}$$

where $\Sigma_i = \sum_{j=1}^N w_j \rho_{ij} f_{j0}$.

Example 3.2 (Index with One-factor rough Bergomi components). For $i = 1, \dots, N$ consider component dynamics of "rough Bergomi" type, following the terminology of Bayer et al. (2016), that is,

$$dS_t^i / S_t^i = \sigma_i e^{\eta_i \widehat{W}_t^i} \left(c_i \, dW^i + \bar{c}_i \, d\overline{W}^i \right),$$

with vvol $\eta_i > 0$ and spot volatility σ_i. We consider the "fully correlated" case, $c_i = -1$, hence $\bar{c}_i = 0$. In law, this is the same as

$$dS_t^i / S_t^i = \left(\sigma_i e^{-\eta_i \widehat{W}_t^i} \right) dW_t^i \equiv f_i \left(\widehat{W}_t^i \right) dW_t^i.$$

In the notation of this section we have $f_{i0} := f_i(0) = \sigma_i$ and $f'_{i0} = -\eta_i \sigma_i$. Thus the spot volatility and skew are given by

$$\sigma_{i0} = \sigma_i, \qquad \frac{\mathcal{S}_{i0}}{2\sigma_i} = -\eta_i \langle K^{H_i} 1, 1 \rangle = -\eta_i \frac{\sqrt{2H_i}}{\left(H_i + \frac{1}{2} \right) \left(H_i + \frac{3}{2} \right)}.$$

Concerning the index $I = \sum_{i=1}^N w_i S^i$, we have the usual spot volatility

$$\sigma_I = \sqrt{\sum_{i,j=1}^N w_i w_j \rho_{ij} \sigma_i \sigma_j}.$$

For the implied variance skew we leave it to the reader to substitute $f_{i0} = \sigma_i$ and $f'_{i0} = -\eta_i \sigma_i$ into the formula in Equation (9). Of course, $\mathcal{S}_I / (2\sigma_I)$ then gives the implied volatility skew (at unit time, in the small noise limit).

4 | PROOF OF THEOREM 2.6

To emphasize the generality of the argument, we write $\left(H, \langle ., . \rangle_H \right)$ instead of $\mathcal{H}^N$ with the Cameron-Martin inner product. The abstract minimization then concerns $\langle h, \rho^{-1} h \rangle_H / 2$ for some invertible (linear) operator $\rho : H \to H$ subject to a constraint $\phi(h) = x$, where $\phi : H \to \mathbb{R}$ is weakly continuous and thrice Fréchet differentiable. The optimization problem can be written as

$$\Lambda(x) = \inf \left\{ \frac{1}{2} \langle h, \rho^{-1} h \rangle_H \, : \, \phi(h) = x \right\}.$$

Proof of Theorem 2.6. Define $\psi : \mathbb{R} \times \mathbb{R} \times H \to \mathbb{R} \times H$ via

$$\psi(x, \lambda, h) = \left(\phi(h) - x, \rho^{-1} h - \lambda D\phi(h) \right).$$

We want h^x, λ^x s.t. $\psi(x, \lambda^x, h^x) = (0, 0)$, which corresponds to the first order condition of the minimization problem. Regularity of $x \mapsto h^x$ implies regularity of $x \mapsto \Lambda(x)$ in which case we know $\Lambda' = \lambda$. By the implicit function theorem

$$\begin{pmatrix} \lambda'(x) \\ h'(x) \end{pmatrix} = -\underbrace{\begin{pmatrix} \partial_\lambda \psi_1 & \partial_\lambda \psi_2 \\ \partial_h \psi_1 & \partial_h \psi_2 \end{pmatrix}^{-1}}_{=:J^{-1}} \begin{pmatrix} \partial_x \psi_1 \\ \partial_x \psi_2 \end{pmatrix}$$

A simple calculation tells us, evaluated at h^x,

$$J = \begin{pmatrix} 0 & -D\phi(h^x) \\ D\phi(h^x) & \rho^{-1} - \lambda^x\Big(\dots\Big). \end{pmatrix}$$

Here the bracket term has no contribution because $\lambda^x = 0$ for $x = 0$. Note that

$$\begin{pmatrix} \partial_x \psi_1 \\ \partial_x \psi_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \in \mathbb{R} \times H.$$

Therefore, we only care about the first column of J. Note that for a block matrix it follows that

$$\begin{pmatrix} 0 & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} -\Big(BD^{-1}C\Big)^{-1} & \Big(BD^{-1}C\Big)^{-1} BD^{-1} \\ D^{-1}C\Big(BD^{-1}C\Big)^{-1} & \dots \end{pmatrix}^{-1}$$

Using the block form of J we see that

$$J^{-1} = \begin{pmatrix} \Big(D\phi(h^x)\,\rho\, D\phi(h^x)\Big)^{-1} & \Big(D\phi(h^x)\,\rho\, D\phi(h^x)\Big)^{-1} \rho D\phi(h^x) \\ -\Big(D\phi(h^x)\,\rho\, D\phi(h^x)\Big)^{-1} \rho D\phi(h^x) & \dots \end{pmatrix}$$

implying that

$$\begin{pmatrix} \lambda'(0) \\ h'(0) \end{pmatrix} = \begin{pmatrix} \frac{1}{\langle \phi_0', \rho\phi_0' \rangle_H} \\ -\frac{1}{\langle \phi_0', \rho\phi_0' \rangle_H} \rho\phi_0' \end{pmatrix}. \tag{10}$$

For the second derivative we start with a short calculation. Let $g(x) = (\lambda^x, h^x)$ such that $\psi(x, g(x)) = 0$. Then

$$\begin{aligned} 0 = \partial_x^2 \psi(x, g(x)) &= \partial_x \partial_1 \psi(x, g(x)) + \partial_x \Big(\partial_2 \psi(x, g(x)) g'(x)\Big) \\ &= \partial_1^2 \psi(x, g(x)) + 2\partial_1 \partial_2 \psi(x, g(x)) g'(x) + \partial_2^2 \psi(x, g(x)) \partial_x g(x) \partial_x g(x) \\ &\quad + \partial_2 \psi(x, g(x)) \partial_x^2 g(x). \end{aligned}$$

Note that

$$\partial_1^2\psi(0,g(0)) = 0 = 2\partial_1\partial_2\psi(0,g(0))g'(0),$$

and $\partial_2\psi(x,g(x)) = J$. We therefore only need to calculate the $\mathbb{R}^2 \times \left(\mathbb{R}\times H\right)^* \times \left(\mathbb{R}\times H\right)^*$-tensor $\partial_2^2\psi(0,g(0))$, which we do component wise.

For this we see $\partial_\lambda^2\psi_1 = \partial_\lambda\partial_h\psi_1 = \partial_\lambda^2\psi_2 = 0$. Also $\partial_h^2\psi_2 = 0$ at $x = 0$. At last $\partial_h^2\psi_1 = D^2\phi(h^x)(.,.)$ and $\partial_\lambda\partial_h\psi_2 = -D^2\phi(h^x)(.,.)$.

Evaluating the tensor $\partial_2^2\psi(0,g(0))$ at $\partial_x g(x)$ as well as $\partial_x g(x)$ and using that $\partial_2\psi(x,g(x)) = J$ we see

$$\begin{pmatrix}\lambda''(x)\\ h''(x)\end{pmatrix} = -J^{-1}\begin{pmatrix} D^2\phi(h^x)(h'(x),h'(x))\\ -2D^2\phi(h^x)(h'(x),.)\lambda'(x)\end{pmatrix}$$

Finally, by substituting back we get by (10) that

$$\lambda''(0) = \frac{-3}{\langle\phi_0',\rho\phi_0'\rangle_H}\phi_0''\Big(\frac{1}{\langle\phi_0',\rho\phi_0'\rangle_H}\rho\phi_0', \frac{1}{\langle\phi_0',\rho\phi_0'\rangle_H}\rho\phi_0'\Big).$$

□

5 | EXTENSION TO STOCHASTIC VOLATILITY

We return to the setting where each component of the index has rough volatility dynamics, as specified in (4), with 2 Brownian factors. An index with N-components thus involves a total of $2N$ Brownians, which we assume given as $2N$-dimensional Brownian motion W with non-singular correlation matrix

$$\mathbb{E}(W_1 \otimes W_1) = \rho \in \mathbb{R}^{2N\times 2N}.$$

As before, in a small noise regime, the precise form of the model (4) is not so important – what matters are the rate functions. To this end, we have

Proposition 5.1 (Correlated Index energy). *Assume that $\sum_{i=1}^N w_i = 1$ and let $\phi_i : \mathcal{H} \to \mathbb{R}$ be C^k-reasonable according to Definition 2.3. Let*

$$\exp\big(\phi(h^1,\dots,h^{2N})\big) := \sum_{i=1}^N w_i \exp\big(\phi_i(h^i,h^{N+i})\big).$$

Then ϕ is also C^k-reasonable. Set $\phi_{i0}' = D\phi_i(0) \in \mathcal{H}^2$ as well as $\phi_{i0}'' = D^2\phi_i(0) \in (\mathcal{H}^2\times\mathcal{H}^2)^\star$ and assume

$$\phi_{i0}' = \begin{pmatrix}\sigma_i\\ 0\end{pmatrix} Id \in \mathcal{H}^2, \qquad Id : t \mapsto t, \tag{11}$$

Then

$$(\sigma^{\varepsilon}_{impl})^2(x,1) \sim_{x\downarrow 0} \sigma_I^2 + x\mathcal{S}_I + o(x),$$

with

$$\sigma_I^2 = \sum_{i,j=1}^{N} w_i w_j \rho_{ij}\sigma_i\sigma_j,$$

$$\mathcal{S}_I = -\sigma_I^2 + \sum_{i=1}^{N} w_i \left(\left(\sum_{j,\ell=1}^{N} w_j w_\ell \sigma_j \sigma_\ell P_{i\ell} \cdot \Phi_i P_{ij} \right) + \left(\sigma_i \Sigma_i\right)^2 \right) \Big/ \sigma_I^2,$$

$$\Sigma_i := \sum_{j=1}^{N} w_j \rho_{ij}\sigma_j, \qquad P_{i\ell} = \begin{pmatrix} \rho_{i\ell} \\ \rho_{(i+N)\ell} \end{pmatrix},$$

$$\Phi_i = \begin{pmatrix} \phi''_{i0}\big((Id,0);(Id,0)\big) & \phi''_{i0}\big((Id,0);(0,Id)\big) \\ \phi''_{i0}\big((Id,0);(0,Id)\big) & \phi''_{i0}\big((0,Id);(0,Id)\big) \end{pmatrix}.$$

Proof. We only sketch the proof, because all calculations are similar to the one in the proof of Proposition 2.9. Similar to before, by power matching we get

$$\langle \phi'_0, k\rangle_{\mathcal{H}^N} = \sum_{i=1}^{N} w_i \langle \phi'_{i0}, (k^i, k^{N+i})\rangle_{\mathcal{H}^2},$$

implying that

$$\phi'_0 = \left(w_1 (\phi'_{10})^1, \dots, w_N (\phi'_{N0})^1; w_1 (\phi'_{10})^2, \dots, w_N (\phi'_{N0})^2 \right) \in \mathcal{H}^{2N}.$$

For the second order similar calculations show

$$\phi''_0(k,k) = -\langle \phi'_0, k\rangle^2_{\mathcal{H}^{2N}} + \sum_{i=1}^{N} w_i \left(\phi''_{i0}\left(k^i, k^{N+i}; k^i, k^{N+i}\right) + \langle \phi'_{i0}, (k^i, k^{N+i})\rangle^2_{\mathcal{H}^2} \right).$$

We note that

$$\left(\rho\phi'_0\right)^i = \sum_{j=1}^{2N} \rho_{ij} \left(\phi'_0\right)^j = \sum_{j=1}^{N} \rho_{ij} w_j \left(\phi'_{j0}\right)^1 + \sum_{j=1}^{N} \rho_{i(j+N)} w_j \left(\phi'_{j0}\right)^2.$$

Under the assumption of (11) we thus see

$$\sigma_I^2 = \langle \rho\phi'_0, \phi'_0\rangle_{\mathcal{H}^N} = \sum_{i,j=1}^{N} w_i w_j \rho_{ij}\sigma_i\sigma_j.$$

We split up $\mathcal{S}_I = \phi''_0(\rho\phi'_0, \rho\phi'_0)/\sigma_I^2$ into

$$\mathcal{S}_I = (\mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3)/\sigma_I^2.$$

Under assumption (11) a calculation shows that

$$\mathcal{S}_1 = -\left(\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \rho_{ij}\sigma_i\sigma_j\right)^2 = -\sigma_I^4$$

Doing similar calculations as in Proposition 2.9, under the assumption of (11) it follows that

$$\mathcal{S}_2 = \sum_{i=1}^{N} w_i \left(\sum_{j,\ell=1}^{N} w_j w_\ell \sigma_j \sigma_\ell P_{i\ell}\Phi_i P_{ij}\right).$$

as well as

$$\mathcal{S}_3 = \sum_{i=1}^{N} w_i \left(\sigma_i \sum_{j=1}^{N} w_j \rho_{ij}\sigma_j\right)^2 = \sum_{i=1}^{N} w_i \left(\sigma_i \Sigma_i\right)^2.$$

The formula for $\mathcal{S}_I$ then follows from summing up. □

Remark 5.2. As before, the proof shows that one can do without assumption (11) at the expense of less appealing formulae. Yet, this assumption is satisfied in the examples we have in mind. Indeed, let us show that equation (11) is satisfied in the rough volatility case, that is, when ϕ_i is given by[3]

$$\phi_i(h^1, h^2) = \int_0^1 f_i(\hat{h}^2)\, dh^1.$$

In that case

$$D\phi_i(h^1, h^2)(g^1, g^2) = \int_0^1 f_i(\hat{h}^2)\, dg^1 + \int_0^1 f_i'(\hat{h}^2)\hat{g}^2\, dh^1.$$

This implies that ϕ_{i0}' satisfies Assumption (11) with $\sigma_i = f_i(0)$.

6 | TECHNICAL RESULTS

In the main text we have used the classical Cameron–Martin Hilbert space. Taking a general separable Hilbert space H instead, we here consider, in this Hilbert generality,

$$\Lambda(x) = \inf\left\{\frac{1}{2}\langle h, h\rangle_H \,:\, \phi(h) = x\right\} \in [0, \infty].$$

Call *x-admissible* any $h \in H$ with $\phi(h) = x$. Set $D_\Lambda = \{x \in \mathbb{R} : \Lambda(x) < \infty\}$.

6.1 | Monotonicity of energy

We now discuss monotonicity of Λ.

[3] No need for a correlation parameter here, contained in Hilbert structure of $\mathcal{H}^2 \ni (h^1, h^2)$.

Proposition 6.1. *Assume $\phi : H \to \mathbb{R}$ is continuous and $\phi(0) = 0$. Let $0 < x < y$. Then $0 = \Lambda(0) \leqslant \Lambda(x) \leqslant \Lambda(y)$.*

Proof. If $\Lambda(y) = +\infty$ there is nothing to show, else have, for every $\varepsilon > 0$, some y-admissible $h^y \in H$: $\frac{1}{2}\langle h^y, h^y\rangle_H < \Lambda(y) + \varepsilon$. The real function $[0, 1] \ni c \mapsto \phi(ch^y) \in \mathbb{R}$ is continuous, by continuity of ϕ, with end-points $\phi(0) = 0$ and $\phi(h^y) = y$, respectively. By the intermediate value theorem, for every $x \in (0, y)$, there is $c_0 \in (0, 1)$ such that $\phi(c_0 h^y) = x$ and so

$$\Lambda(x) \leqslant \frac{1}{2}\langle c_0 h^y, c_0 h^y\rangle_H < \frac{1}{2}\langle h^y, h^y\rangle_H < \Lambda(y) + \varepsilon.$$

Conclude by taking $\varepsilon \downarrow 0$. □

6.2 | Existence of a minimizer

Proposition 6.2. *Assume $\phi : H \to \mathbb{R}$ is weakly continuous (hence continuous).*

(i) Let $x \in \phi(H)$. Then there exists an x-admissible h^x s.t. $\Lambda(x) = \frac{1}{2}\langle h^x, h^x\rangle_H$.
(ii) The map Λ is LSC.

Proof. We only consider $x > 0$. (i) Fix $x \in \phi(H)$, so that $\Lambda(x) < \infty$ and pick x-admissible $h^n \in H$ so that $\frac{1}{2}\langle h^n, h^n\rangle_H \downarrow \Lambda(x)$. By weak compactness, there exists $h \in H$ and a subsequence (n_k) such that $h^{n_k} \to h$ weakly in H. Hence $x = \phi(h^n) \to \phi(h)$ which shows that h is x-admissible. If follows that

$$\Lambda(x) \leqslant \frac{1}{2}\|h\|_H^2 \leqslant \liminf_{k\to\infty} \frac{1}{2}\|h^{n_k}\|_H^2 = \Lambda(x).$$

(ii) Consider $x_n \to x$. We need to see $\Lambda(x) \leq \liminf_n \Lambda(x_n)$. We may assume that $\liminf_n \Lambda(x_n) < \infty$ and that all $x_n \in D_\Lambda$ as otherwise it is inconsequential to remove all n with $\Lambda(x_n) = \infty$. Consider $(h^{x_n})_n$, then there is a subsequence (n_k) such that $(h^{x_{n_k}})$ is bounded in H and therefore weakly compact. By weak continuity any weak limit point h is x-admissible, implying that $\Lambda(x) \leq \frac{1}{2}\|h\|_H^2$. We conclude with lower semi-continuity of the norm in H under weak convergence. □

We give a general criterion for weak continuity.

6.3 | Rough path type continuity implies weak continuity

Consider a Banach space $\mathbf{W} = W \oplus \tilde{W}$ which is a direct sum of two Banach spaces $W, \tilde{W}$. Write $\pi : \mathbf{W} \to W$ for the canonical projection on the first component. Let H be a Hilbert space, with compact embedding $\iota : H \hookrightarrow W$, and further consider a lift, that is a map

$$\mathcal{L} : H \to \mathbf{W}$$

so that $\pi \circ \mathcal{L} = \iota$. Assume that the function

$$I(\mathbf{w}) = \begin{cases} \frac{1}{2}\langle h, h\rangle_H & \text{when } \mathbf{w} = \mathcal{L}(h), \\ +\infty & \text{otherwise.} \end{cases}$$

is a good rate function on $\mathbf{W}$, that is, LSC with compact level sets. This situation is typical for small noise large deviations of Gaussian rough paths.

Theorem 6.3. *Assume I is a good rate function, then $\mathcal{L}$ is weakly continuous. As a consequence, any map*

$$\phi : H \to \mathbb{R}, \quad \phi = \bar{\phi} \circ \mathcal{L}, \quad \bar{\phi} \in C(\mathbf{W}, \mathbb{R})$$

is also weakly continuous.

Remark 6.4. (i) We do not (want to) assume that $\phi = \hat{\phi} \circ \iota$ for $\hat{\phi} \in C(\mathrm{W}, \mathbb{R})$

(ii) Consider

$$\Lambda(x) = \inf \left\{ I(\mathbf{w}) : \bar{\phi}(\mathbf{w}) = x \right\} \in [0, \infty].$$

The contraction principle then tells us that Λ is LSC, has compact levels sets, and also gives the existence of h^x. Our presentation highlights the role of weak continuity, which may (or may not) be checked with rough path type continuity arguments.

Proof. Consider $h^n \to h$ weakly in H. Such a sequence is necessarily bounded, as a consequence of the uniform boundedness principle. By goodness of the rate function, $\mathcal{L}(h^{n_k}) = \mathbf{w}^k \to \mathbf{w}$ in $\mathbf{W}$, for some $\mathbf{w} = (w, \tilde{w}) \in \mathbf{W}$ and some subsequence (n_k). LSC implies

$$I(\mathbf{w}) \leqslant \liminf_{k \to \infty} I(\mathbf{w}^k) \leqslant \sup_k \frac{1}{2}\langle h^k, h^k\rangle_H < \infty.$$

It follows that $\mathbf{w} = \mathcal{L}(w)$ with $w \in H$. To identify w, note $h^{n_k} \to w$ in W and since $\mathrm{W}^\star \subset H^\star \cong H$, we see $\langle \theta, h^{n_k}\rangle_H \to \langle \theta, w\rangle_H$ for $\theta \in \mathrm{W}$. But by weak convergence we also have $\langle \theta, h^{n_k}\rangle_{\mathrm{W}} \to \langle \theta, h\rangle_{\mathrm{W}} = \langle \theta, h\rangle_H$. This implies $w = h$. We have shown that $h^n \to h$ weakly implies $\mathcal{L}(h^n) \to \mathcal{L}(h)$ in $\mathbf{W}$ along a subsequence. By a standard argument this also shows convergence (without subsequence). □

6.4 | Continuity of rate function

We start with an explicit example where one has discontinuities.

Example 6.5. Assume $\phi(h) = F(\langle g, h\rangle_H)$ for some continuous F with $F(0) = 0$ and a fixed unit vector $g \in H$. By scaling, this applies to any non-zero $g \in H$ and the case $g = 0$ is trivial anyway. Such ϕ is obviously weakly continuous on H. Assume $0 \leqslant x \in \phi(H)$ and F strictly increasing, then $\phi(h) = x$ iff $\langle g, h\rangle_H = F^{-1}(x)$. Obviously the minimal h is colinear to g and explicitly

$h = F^{-1}(x)\, g$. Then

$$\Lambda(x) = \frac{1}{2}(F^{-1}(x))^2,$$

which is in fact continuous since F^{-1} is. If F is only assumed to be increasing (in sense of non-decreasingness), we write F^{-} for the generalized-inverse of F which is defined on the interval $F(\mathbb{R})$ by

$$F^{-}(y) = \inf\{t \in \mathbb{R} : F(t) = y\}, \quad y \in F(\mathbb{R}).$$

Such F^{-} is also increasing and left-continuous, hence LSC. Flat parts of F precisely correspond to jumps in F^{-}. As above, $h^x = F^{-}(x)g$ and hence

$$\Lambda(x) = \frac{1}{2}\left(F^{-}(x)\right)^2.$$

This function need not be continuous because F^{-} may have jumps.

If F is not increasing, this form of the rate function *in general* fails and one can only say

$$\Lambda(x) = \min\left\{\frac{1}{2}|y|^2 : F(y) = x\right\}.$$

Proposition 6.6. *Assume $x \in D_\Lambda$ admits an x-admissible minimizer $h^x \in H$. Assume $\phi : H \to \mathbb{R}$ has no local maximum at h^x. Then Λ is continuous at x. This holds in particular, if ϕ is C^1-reasonable, see Definition 2.3.*

Proof. We already know that Λ is monotone and LSC, implying that Λ is left-continuous. Thus, the only possible disconuity can be a jump at some point $x \in D_\Lambda$.

Let $x \in D_\Lambda, x > 0$ and $\varepsilon > 0$ be arbirtray. By Assumption there is some x-admissible $h^x \in H$ minimizing $\Lambda(x)$. Choose δ so small that

$$\left|\frac{1}{2}\langle h, h\rangle_H - \frac{1}{2}\langle h^x, h^x\rangle_H\right| < \frac{\varepsilon}{2}$$

for all $h \in U_\delta(h^x)$. By assumption h^x is not a local maxima therefore there is some $\tilde{h} \in U_\delta(h^x)$ such that $x = \phi(h^x) < \phi(\tilde{h}) =: \tilde{y}$. But by monotonicity of Λ and by construction

$$\Lambda(x) < \Lambda(\tilde{y}) \le \frac{1}{2}\langle \tilde{h}, \tilde{h}\rangle \le \frac{1}{2}\langle h^x, h^x\rangle + \frac{\varepsilon}{2} = \Lambda(x) + \frac{\varepsilon}{2}$$

As ε was arbitrary we see that $\lim_{\lambda\to 0}\Lambda(x+\lambda) = \Lambda(x)$ implying that Λ is continuous at point x and therefore in D_Λ. □

Even if the energy function is continuous, it need not be smooth. Similar facts for (sub)Riemannian square-distance are well-known. In the context of Example 6.5 we can exhibit this directly via the function $F(y) = |y|^\alpha, \alpha > 0$ with inverse $\pm|y|^{1/\alpha}$. Then

$$\Lambda(x) = \frac{1}{2}\left(F^{-1}(x)\right)^2 = \frac{1}{2}|x|^{2/\alpha}.$$

Note that $\phi : h \mapsto F(\langle g, h\rangle)$ is weakly continuous and inherits Fréchet regularity from F. For instance, $F(y) = y^2$ makes ϕ Fréchet smooth, but Λ fails to be C^1 at $x = 0$; the problem, as seem below, is that $D\phi(0) = F'(0)g = 0$ in this example.

6.5 | Smoothness of energy function

The following is a consequence of a more general statement that can be found in the appendix of Friz et al. (2021a).

Theorem 6.7. *Assume $\phi : H \to \mathbb{R}$ is C^n-reasonable, see Definition 2.3. Then for all sufficiently small x, there exists a unique x-admissible minimizer h^x, such that $x \mapsto h^x \in H$ is C^{n-1}. Moreover, $x \mapsto \Lambda(x)$ is C^n near $x = 0$, hence*

$$\Lambda(x) = \Lambda''(0)x^2/2! + \Lambda'''(0)x^3/3! + \cdots + \Lambda^{(n)}(0)x^n/n! + o(|x|^n).$$

ACKNOWLEDGMENTS

The first author acknowledges support from the German science foundation (DFG) via the cluster of excellence MATH+, as PI of project AA4-2. The second author was supported by the German science foundation (DFG) via MATH+, as PhD student in the Berlin Mathematical School (BMS). We thank the referee and several colleagues for valuable feedback.

Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

No data source was used for this study.

ORCID

Peter K. Friz https://orcid.org/0000-0003-2571-8388
Thomas Wagenhofer https://orcid.org/0000-0002-5609-2624

REFERENCES

Avellaneda, M., Boyer-Olson, D., Busca, J., & Friz, P. (2002). Reconstruction of volatility: Pricing index options by the steepest descent approximation. *Risk*, *15*, 87–91.

Avellaneda, M., Boyer-Olson, D., Busca, J., & Friz, P. (2003). Application of large deviation methods to the pricing of index options in finance. *Comptes Rendus Mathematique*, *336*(3), 263–266.

Berestycki, H., Busca, J., & Florent, I. (2004). Computing the implied volatility in stochastic volatility models. *Communications on Pure and Applied Mathematics*, *57*(10), 1352–1373.

Bergomi, L. (2016). *Stochastic Volatility Modeling*. CRC Press.

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

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

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

Bayer, C., Friz, P., & Laurence, P. (2015). On the probability density function of baskets. In Peter K. Friz, Jim Gatheral, Archil Gulisashvili, Antoine Jacquier, Josef Teichmann, *Large deviations and asymptotic methods in finance* (pp. 449–472). Springer.

Bentata, A., & Cont, R. (2015). Forward equations for option prices in semimartingale models. *Finance and Stochastics*, *19*(3), 617–651.

Brunick, G., & Shreve, S. (2013). Mimicking an Itô process by a solution of a stochastic differential equation. *The Annals of Applied Probability*, *23*(4), 1584–1628.

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

Deuschel, J.-D., Friz, P. K., Jacquier, A., & Violante, S. (2014). Marginal density expansions for diffusions and stochastic volatility II: Applications. *Communications on Pure and Applied Mathematics*, *67*(2), 321–350.

Derman, E., & Kani, I. (1994). Riding on a smile. *Risk*, *7*(2), 32–39.

De Marco, S., & Friz, P. K. (2018). Local volatility, conditioned diffusions, and Varadhan's formula. *SIAM Journal on Financial Mathematics*, *9*(2), 835–874.

De Marco, S., Friz, P., & Gerhold, S. (2013). Rational shapes of local volatility. *Risk*, *26*(2), 70.

Dupire, B. (1994). Pricing with a smile. *Risk*, *7*(1), 18–20.

Friz, P. K., Gatheral, J., Gulisashvili, A., Jacquier, A., & Teichmann, J. (2015). *Large deviations and asymptotic methods in finance*. Series: Springer Proceedings in Mathematics & Statistics, Vol. 110, 2015.

Friz, P., Gassiat, P., & Pigato, P. (2021a). Precise asymptotics: Robust stochastic volatility models. *The Annals of Applied Probability*, *31*(2), 896–940.

Friz, P., Pigato, P., & Seibel, J. (2021b). The step stochastic volatility model. *Risk*. https://www.risk.net/cutting-edge/banking/7835841/the-step-stochastic-volatility-model

Fukasawa, M., & Takano, R. (2022). A partial rough path space for rough volatility. *arXiv preprint arXiv:2205.09958*, 2022.

Forde, M., & Zhang, H. (2017). Asymptotics for rough stochastic volatility models. *SIAM Journal on Financial Mathematics*, *8*(1), 114–145.

Gatheral, J. (2006). *The volatility surface: A practitioner's guide*. John Wiley & Sons.

Gatheral, J., Hsu, E., Laurence, P., Ouyang, C., & Wang, T.-H. (2012). Asymptotics of implied volatility in local volatility models. *Mathematical Finance*, *22*(4), 591–620.

Guyon, J., & Henry-Labordère, P. (2014). *Nonlinear option pricing*. Chapman & Hall/CRC Financial Mathematics Series. CRC Press.

Gulisashvili, A. (2018). Large deviation principle for volterra type fractional stochastic volatility models. *SIAM Journal on Financial Mathematics*, *9*(3), 1102–1136.

Gulisashvili, A. (2022). Multivariate stochastic volatility models and large deviation principles. *arXiv preprint arXiv:2203.09015*.

Gyöngy, I. (1986). Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probability Theory and Related Fields*, *71*(4), 501–516.

Henry-Labordère, P. (2008). *Analysis, geometry, and modeling in finance: Advanced methods in option pricing*. Chapman and Hall/CRC.

Hobson, D., & Rogers, L. C. G. (1998). Complete models with stochastic volatility. *Mathematical Finance*, *8*(1), 27–48.

Jacquier, A., & Pannier, A. (2022). Large and moderate deviations for stochastic Volterra systems. *Stochastic Processes and their Applications*, *149*, 142–187.

Osajima, Y. (2015). General asymptotics of Wiener functionals and application to implied volatilities. In *Large deviations and asymptotic methods in finance*. (pp. 137–173). Springer.

Varadhan, S. R. S. (1967). Diffusion processes in a small time interval. *Communications on Pure and Applied Mathematics*, *20*(4), 659–685.

How to cite this article: Friz, P. K., & Wagenhofer, T. (2023). Reconstructing volatility: Pricing of index options under rough volatility. *Mathematical Finance*, *33,* 19–40. https://doi.org/10.1111/mafi.12374