




ROUGH MULTI-FACTOR VOLATILITY FOR SPX AND VIX OPTIONS

ANTOINE JACQUIER ,** Imperial College London and the Alan Turing Institute
AITOR MUGURUZA,*** Imperial College London and Kaiju Capital Management
ALEXANDRE PANNIER,**** Université Paris Cité, Laboratoire de Probabilités, Statistique et Modélisation

Abstract

We provide explicit small-time formulae for the at-the-money implied volatility, skew, and curvature in a large class of models, including rough volatility models and their multi-factor versions. Our general setup encompasses both European options on a stock and VIX options, thereby providing new insights on their joint calibration. The tools used are essentially based on Malliavin calculus for Gaussian processes. We develop a detailed theoretical and numerical analysis of the two-factor rough Bergomi model and provide insights on the interplay between the different parameters for joint SPX–VIX smile calibration.

Keywords: Rough volatility; multi-factor; asymptotics; VIX; Malliavin calculus

2010 Mathematics Subject Classification: Primary 60G15

Secondary 60G22; 60H07; 91G20

1. Introduction

Exposure to the uncertain dynamics of volatility is a desirable feature of most trading strategies and has naturally generated wide interest in volatility derivatives. From a theoretical viewpoint, an adequate financial model should reproduce the volatility dynamics accurately and consistently with those of the asset price; any discrepancy may otherwise lead to arbitrage opportunity. Despite extensive research, implied volatility surfaces from options on the VIX and the S&P 500 index still display discrepancies, betraying the lack of a proper modelling framework. This issue is well known as the *SPX–VIX joint calibration problem* and has motivated a number of creative modelling innovations in the past fifteen years. Reconciling both implied volatilities requires additional factors to enrich the variance curve dynamics, as argued by Bergomi [13, 14], who proposed the multi-factor model

$$\frac{d\xi_t(T)}{\xi_t(T)} = \sum_{i=1}^N c_i e^{-\kappa_i(T-t)} dW_t^i, \quad 0 \leq t \leq T, \quad (1)$$

Received 5 April 2022; accepted 18 July 2024.

* Postal address: Department of Mathematics, Imperial College London, South Kensington, London SW7 2AZ, United Kingdom.

** Email address: a.jacquier@imperial.ac.uk

*** Email address: aitor.muguruza-gonzalez15@imperial.ac.uk

**** Postal address: 15 rue Hélène Brion, 75013 Paris, France. Email address: pannier@lpsm.paris

© The Author(s), 2024. Published by Cambridge University Press on behalf of Applied Probability Trust.

for the forward variance, with W^1, \dots, W^N correlated Brownian motions and coefficients $c_1, \kappa_1, \dots, c_N, \kappa_N > 0$. Gatheral [26] recognised the importance of the additional factor to disentangle different aspects of the implied volatility and to allow humps in the variance curve, and introduced a mean-reverting version—the double CEV model—where the instantaneous mean of the variance follows a CEV model itself. Although promising, these attempts fell short of reproducing jointly the short-time behaviour of the SPX and VIX implied volatilities. A variety of new models were suggested to tackle this issue, both with continuous paths [7, 24, 29] and with jumps [6, 18, 19, 40, 43, 47], incorporating novel ideas and increased complexity such as regime-switching volatility dynamics. Model-free bounds were also obtained in [21, 31, 32, 46], shedding light on the links between VIX and SPX and the difficulty of capturing them both simultaneously.

Getting rid of the restraining Markovian assumption that burdens classical stochastic volatility models has permitted the emergence of rough volatility models, which consistently agree with stylised facts under both the historical and the pricing measures [4, 8, 9, 11, 25, 27]. A large portion of the toolbox developed for Markovian diffusion models is not available any longer, and asymptotic methods [30, 33, 36, 37]—and more recently path-dependent PDE methods [10, 16, 38, 44, 45, 50]—thus play a prominent role in understanding the theoretical properties and numerical aspects of these models. Since the fit of the spot implied volatility skew is extremely accurate under this class of models [27], it seems reasonable to expect good results when calibrating VIX options. Moreover, the newly established hedging formula by Viens and Zhang [50] shows that a rough volatility market is complete if it also contains a proxy of the volatility of the asset; this acts as an additional motivation for our work. Still, [35] showed that the rough Bergomi model is too close to lognormal to jointly calibrate both markets. Its younger sister [34] added a stochastic volatility-of-volatility component, generating a smile sandwiched between the bid–ask prices when calibrating VIX, but the joint calibration is not provided. By incorporating a Zumbach effect, the quadratic rough Heston model [28] achieves good results for the joint calibration at one given date. Further numerical methods were developed in [15, 17, 49]. However, the lack of analytical tractability of rough volatility models is holding back the progress of theoretical results on the VIX, with the notable exception of large deviations results from [23, 41] and the small-time asymptotics of [2].

In the latter, \mathcal{F}_T -measurable random variables (with volatility derivatives in mind) are written in the form of exponential martingales thanks to the Clark–Ocone formula, allowing the application of established asymptotic methods from [4]. An expression for the short-time limit at-the-money (ATM) implied volatility skew is derived, yielding an analytical criterion that a model should satisfy to reproduce the correct short-time behaviour. The proposed mixed rough Bergomi model does meet the requirement of positive skew of the VIX implied volatility, backing its implementation with theoretical evidence. And indeed, the fits are rather satisfying. This model is built by replacing the exponential kernels of the Bergomi model (1) ($t \mapsto e^{-\kappa t}$) with fractional kernels of the type $t \mapsto t^{H-\frac{1}{2}}$ with $H \in (0, \frac{1}{2})$, but is limited to a single factor, i.e. $W^1 = W^2$. As a result, numerical computations under this model induce a linear smile, or equivalently a null curvature, unfortunately inconsistent with market observations. To remedy this, we incorporate Bergomi's and Gatheral's insights on multi-factor models (integrated by [20, 41] into rough volatility models) and extend [2] to the multi-factor case; we also compute the short-time ATM implied volatility curvature, deriving a second criterion for a more accurate model choice. In summary, the present paper goes beyond [2] in three ways:

- We consider multi-factor models, far more efficient for VIX calibration, which complicate the analysis.
- We compute the second derivative of the implied volatility to discriminate better between models; this turns out to be considerably more technical than the skew.
- We provide detailed proofs of all of our results at three levels: abstract model, generic rough volatility model for the VIX, and two-factor rough Bergomi model, checking carefully that all the assumptions are satisfied, proving technical lemmas applicable to our setting, and exhibiting definite formulas at all three levels of generality.

We gather in Section 2 our abstract framework and assumptions. The main results, which provide the short-time limits of the implied volatility level, skew, and curvature, are contained in Section 3. Our framework covers a wide range of underlying assets, including VIX (Section 4) and stock options (Section 5); see in particular in Propositions 1 and 4. We provide further a detailed analysis of the two-factor rough Bergomi model (1). Closed-form expressions that depend explicitly on the parameters of the model are provided in Proposition 3 for the VIX and Corollary 1 for the stock. These expressions give insight into the interplay between the different parameters, and make the calibration task easier by allowing us to fit some stylised facts before performing numerical computations. For instance, different combinations of parameters can yield positive or negative curvature. All the proofs are gathered in Section 6, starting with useful lemmas and then following the order of the sections.

Notation. For an integer $N \in \mathbb{N}$ and a vector $\mathbf{x} \in \mathbb{R}^N$, we define $|\mathbf{x}| := \sum_{i=1}^N x_i$ and $\|\mathbf{x}\|^2 := \sum_{i=1}^N x_i^2$. We fix a finite time horizon $T > 0$ and let $\mathbb{T} := [0, T]$. For all $p \geq 1$, L^p stands for the space $L^p(\Omega)$ for some reference sample space Ω . As we consider rough volatility models, the Hurst parameter $H \in (0, \frac{1}{2})$ is a fundamental quantity and we shall write $H_+ := H + \frac{1}{2}$ and $H_- := H - \frac{1}{2}$.

2. Framework

We consider a square-integrable strictly positive process $(A_t)_{t \in \mathbb{T}}$, adapted to the natural filtration $(\mathcal{F}_t)_{t \in \mathbb{T}}$ of an N -dimensional Brownian motion $\mathbf{W} = (W^1, \dots, W^N)$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We further introduce the true $(\mathcal{F}_t)_{t \in \mathbb{T}}$ -martingale conditional expectation process

$$M_t := \mathbb{E}_t[A_T] := \mathbb{E}[A_T | \mathcal{F}_t], \quad \text{for all } t \in \mathbb{T}.$$

The set $\mathbb{D}^{1,2}$ will denote the domain of the Malliavin derivative operator D with respect to the Brownian motion \mathbf{W} , while D^i indicates the Malliavin derivative operator with respect to W^i . It is well known that $\mathbb{D}^{1,2}$ is a dense subset of $L^2(\Omega)$ and that D is a closed and unbounded operator from $L^2(\Omega)$ into $L^2(\mathbb{T} \times \Omega)$. Analogously we define the sets of Malliavin differentiable processes $\mathbb{L}^{n,2} := L^2(\mathbb{T}; \mathbb{D}^{n,2})$. We refer to [42] for more details on Malliavin calculus. Assuming $A_T \in \mathbb{D}^{1,2}$, the Clark–Ocone formula [42, Theorem 1.3.14] reads, for each $t \in \mathbb{T}$,

$$M_t = \mathbb{E}[M_t] + (\mathbf{m} \bullet \mathbf{W})_t := \mathbb{E}[M_t] + \sum_{i=1}^N \int_0^t m_s^i dW_s^i, \tag{2}$$

where each component of \mathbf{m} is $m_s^i := \mathbb{E}[D_s^i A_T | \mathcal{F}_s]$. Since M is a martingale, we may rewrite (2) as

$$M_t = M_0 + (M\phi \bullet \mathbf{W})_t, \tag{3}$$

where $\phi_s := \mathbf{m}_s/M_s$ is defined whenever $M_s \neq 0$ almost surely. If $\phi = (\phi^1, \dots, \phi^N)$ belongs to $\mathbb{L}^{n,2}$, then the following processes are well defined for all $t < T$:

$$\begin{aligned}
 Y_t &:= \int_t^T \|\phi_r\|^2 dr, & u_t &:= \sqrt{Y_t}, & u_t &:= \frac{u_t}{\sqrt{T-t}}; \\
 \Theta_t^i &:= \left(\int_t^T D_t^i \|\phi_r\|^2 dr \right) \phi_t^i, & \text{and} & & |\Theta| &:= \sum_{i=1}^n \Theta^i.
 \end{aligned} \tag{4}$$

Note that all the processes depend implicitly on T , which will be crucial when we study the short-time limit as T tends to zero.

2.1. Level, skew, and curvature

Since M is a strictly positive martingale process, we can use it as an underlying to introduce options. A standard practice is to work with its logarithm $\mathfrak{M} := \log(M)$, so that $\mathfrak{M}_T = \log \mathbb{E}_T[A_T] = \log(A_T)$ and $\mathfrak{M}_0 = \log \mathbb{E}[A_T]$. Under no-arbitrage arguments, the price Π_t at time t of a European call option with maturity T and log-strike $k \geq 0$ is equal to

$$\Pi_t(k) := \mathbb{E}_t \left[\left(M_T - e^k \right)^+ \right] = \mathbb{E}_t \left[\left(A_T - e^k \right)^+ \right], \tag{5}$$

and the ATM value is denoted by $\Pi_t := \Pi_t(\mathfrak{M}_0) = \mathbb{E}_t[(A_T - M_t)^+]$. We adapt the usual definitions of ATM implied volatility level, skew, and curvature to the case where the underlying is a general process (later specified for the VIX and the S&P). Denote by $\text{BS}(t, x, k, \sigma)$ the Black–Scholes price of a European call option at time $t \in \mathbb{T}$, with maturity T , log-stock x , log-strike k , and volatility σ . Its closed-form expression reads

$$\text{BS}(t, x, k, \sigma) = \begin{cases} e^x \mathcal{N}(d_+(x, k, \sigma)) - e^k \mathcal{N}(d_-(x, k, \sigma)), & \text{if } \sigma \sqrt{T-t} > 0, \\ (e^x - e^k)^+, & \text{if } \sigma \sqrt{T-t} = 0, \end{cases} \tag{6}$$

with $d_{\pm}(x, k, \sigma) := \frac{x-k}{\sigma \sqrt{T-t}} \pm \frac{\sigma \sqrt{T-t}}{2}$, where \mathcal{N} denotes the Gaussian cumulative distribution function.

Definition 1.

- For any $k \in \mathbb{R}$, the implied volatility $\mathcal{I}_T(k)$ is the unique non-negative solution to $\Pi_0(k) = \text{BS}(0, \mathfrak{M}_0, k, \mathcal{I}_T(k))$; we omit the k -dependence when considering it ATM ($k = \mathfrak{M}_0$).
- The ATM implied skew \mathcal{S} and curvature \mathcal{C} at time zero are defined as

$$\mathcal{S}_T := |\partial_k \mathcal{I}_T(k)|_{k=\mathfrak{M}_0} \quad \text{and} \quad \mathcal{C}_T := \left| \partial_k^2 \mathcal{I}_T(k) \right|_{k=\mathfrak{M}_0}.$$

2.2. Examples

The framework (3) encompasses a large class of models, including stochastic volatility models ubiquitous in quantitative finance. Consider a stock price process $(S_t)_{t \in \mathbb{T}}$ satisfying

$$\frac{dS_t}{S_t} = \sqrt{v_t} dB_t = \sqrt{v_t} \sum_{i=1}^N \rho_i dW_t^i,$$

where v is a stochastic process adapted to $(\mathcal{F}_t)_{t \in \mathbb{T}}$, $\rho := (\rho_1, \dots, \rho_N) \in [-1, 1]^N$ with $\rho \rho^\top = 1$.

2.2.1. *Asset price.* For $N = 2$, the model (3) corresponds to a one-dimensional stochastic volatility model under the identification $A = M = S$, $\phi^1 = \rho_1\sqrt{v}$, and $\phi^2 = \rho_2\sqrt{v}$, and v is a process driven by W^1 . Our analysis generalises [4, Equation (2.1)] to the multi-factor case (in the continuous-path case). We refer to Section 5 for the details in the multi-factor setting and the analysis of the implied volatility.

2.2.2. *VIX.* The VIX is defined as $VIX_T = \sqrt{\frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}_T[v_t] dt}$, where Δ is one month. The representation (2) yields that the underlying is the VIX future

$$M_t^{VIX} := \mathbb{E}_t[VIX_T] = \mathbb{E}[VIX_T] + (\mathbf{m} \bullet \mathbf{W})_t, \quad \text{with} \quad m_s^i = \frac{1}{2\Delta} \mathbb{E}_s \left[\frac{1}{VIX_T} \int_T^{T+\Delta} D_s^i v_r dr \right].$$

2.2.3. *Asian options.* For Asian options, the process of interest is $\mathcal{A}_T := \frac{1}{T} \int_0^T S_t dt$. Using (2) we find

$$M_t^A := \mathbb{E}_t[\mathcal{A}_T] = \mathbb{E}[\mathcal{A}_T] + (\mathbf{m} \bullet \mathbf{W})_t, \quad \text{with} \quad m_s^i = \frac{1}{T} \int_s^T \mathbb{E}_s[D_s^i S_r] dr.$$

2.2.4. *Multi-factor rough Bergomi.* Rough volatility models can be written as $v_t = f(\mathbf{W}_t^H)$, where \mathbf{W}^H is an N -dimensional fractional Brownian motion with correlated components and $f: \mathbb{R}^N \rightarrow \mathbb{R}$. For instance, in the two-factor rough Bergomi model,

$$v_t = v_0 \left(\chi \exp \left\{ v W_t^{1,H} - \frac{v^2}{2} \mathbb{E} \left[\left(W_t^{1,H} \right)^2 \right] \right\} + (1 - \chi) \exp \left\{ \eta W_t^{2,H} - \frac{\eta^2}{2} \mathbb{E} \left[\left(W_t^{2,H} \right)^2 \right] \right\} \right),$$

with $\chi \in (0, 1)$, $v, \eta, v_0 > 0$. In Example 2.2.2 we set $A = VIX$ and hence $N = 2$, but in the asset price case we set $A = S$ and therefore $N = 3$ even though the variance depends on only two factors.

2.3. General assumptions

We introduce the following broad assumptions, which are key to our entire analysis; in Section 4 we provide sufficient conditions to simplify them in the VIX case:

- (H1) $A \in \mathbb{L}^{4,p}$.
- (H2) $\frac{1}{M_t} \in L^p$, for all $p > 1$, and all $t \in \mathbb{T}$.
- (H3) The term $\mathbb{E}_t \left[\int_t^T \frac{|\Theta_s|}{u_s^2} ds \right]$ is well defined for all $t \in \mathbb{T}$.
- (H4) The term $\frac{1}{\sqrt{T}} \mathbb{E} \left[\int_0^T \frac{|\Theta_s|}{u_s^2} ds \right]$ tends to zero as T tends to zero.
- (H5) There exists $p \geq 1$ such that $\sup_{T \in [0,1]} u_0^p < \infty$ almost surely and, for all random variables $Z \in L^p$ and all $i \in \llbracket 1, N \rrbracket$, the following terms are well defined and tend to zero as T tends to zero:

$$\int_0^T \mathbb{E} \left[Z \left(\mathbb{E}_s \left[\frac{1}{u_0} \int_0^T D_s^i \|\phi_r\|^2 dr \right] \right)^2 \right] ds.$$

There exists $\lambda \in (-\frac{1}{2}, 0]$ such that the following hold:

- (H6 $^\lambda$) The following expressions converge to zero as T tends to zero:

$$\frac{1}{T^{\frac{1}{2}+\lambda}} \mathbb{E} \left[\int_0^T \frac{|\Theta_s| \int_s^T |\Theta_r| dr}{u_s^6} ds \right] \quad \text{and} \quad \frac{1}{T^{\frac{1}{2}+\lambda}} \mathbb{E} \left[\int_0^T \frac{1}{u_s^4} \sum_{k=1}^N \left\{ \phi_s^k D_s^k \left(\int_s^T |\Theta_r| dr \right) \right\} ds \right].$$

(H₇^λ) The random variable $\mathfrak{K}_T := \frac{\int_0^T |\Theta_s| ds}{T^{\frac{1}{2}+\lambda} u_0^3}$ is such that $\mathbb{E}[u_0^2 \mathfrak{K}_T]$ tends to zero and $\mathbb{E}[\mathfrak{K}_T]$

has a finite limit as T tends to zero.

There exists (H₈^γ) such that the following hold:

$\gamma \in (-1, 0]$ The following expressions converge to zero as T tends to zero:

$$\begin{aligned} & \frac{1}{T^{\frac{1}{2}+\gamma}} \mathbb{E} \left[\int_0^T u_s^{-10} |\Theta_s| \left(\int_s^T |\Theta_r| \left(\int_r^T |\Theta_y| dy \right) dr \right) ds \right], \\ & \frac{1}{T^{\frac{1}{2}+\gamma}} \mathbb{E} \left[\int_0^T u_s^{-8} \sum_{j=1}^N \left(\phi_s^j D_s^j \left(\int_s^T |\Theta_r| \left(\int_r^T |\Theta_y| dy \right) dr \right) ds \right) ds \right], \\ & \frac{1}{T^{\frac{1}{2}+\gamma}} \mathbb{E} \left[\int_0^T u_s^{-8} |\Theta_s| \int_s^T \sum_{j=1}^N \left\{ \phi_r^j D_r^j \left(\int_r^T |\Theta_y| dy \right) \right\} dr ds \right], \\ & \frac{1}{T^{\frac{1}{2}+\gamma}} \mathbb{E} \left[\int_0^T u_s^{-6} \sum_{k=1}^N \left\{ \phi_s^k D_s^k \left(\int_s^T \sum_{j=1}^N \left\{ \phi_r^j D_r^j \left(\int_r^T |\Theta_y| dy \right) \right\} dr \right) \right\} ds \right]. \end{aligned}$$

(H₉^γ) The random variables

$$\begin{aligned} \mathfrak{H}_T^1 & := \frac{1}{T^{\frac{1}{2}+\gamma} u_0^7} \int_0^T |\Theta_s| \left(\int_s^T |\Theta_r| dr \right) ds \\ \text{and } \mathfrak{H}_T^2 & := \frac{1}{T^{\frac{1}{2}+\gamma} u_0^5} \int_0^T \sum_{j=1}^N \left\{ \phi_s^j D_s^j \left(\int_s^T |\Theta_r| dr \right) \right\} ds \end{aligned}$$

are such that $\mathbb{E}[(u_0^6 + u_0^4 + u_0^2) \mathfrak{H}_T^1 + (u_0^4 + u_0^2) \mathfrak{H}_T^2]$ tends to zero and both $\mathbb{E}[\mathfrak{H}_T^1]$ and $\mathbb{E}[\mathfrak{H}_T^2]$ have a finite limit as T tends to zero.

Remark 1.

- (H₁) requires A to be four times Malliavin differentiable. This is necessary to prove the curvature formula using the Clark–Ocone formula (2) and using the anticipative Itô formula three times.
- When the underlying is the stock price (as in Section 2.2.1), it satisfies Equation (3) where ϕ corresponds to its volatility \sqrt{v} . One can then directly make assumptions on the variance process, as in [3–5]. We make this explicit in Proposition 4 for example. In the case of the VIX (Section 4.1) we refrain from doing the same, since ϕ is much more intricate. Nevertheless, sufficient conditions are given by (C̄).

3. Main results

We gather here our main asymptotic results for the general framework above, with the proofs postponed to Section 6.2 to ease the flow. The first theorem states that the small-time limit of the implied volatility is equal to the limit of the forward volatility. This is well known for Markovian stochastic volatility models [5, 12] and in a one-factor setting [2]. To streamline

the call to the assumptions, we shall group them using mixed subscript notation; for example (H_{123}) corresponds to $(H_1)-(H_2)-(H_3)$, and we further write (\bar{H}^λ) to mean $(H_{12345})-(H_{67}^\lambda)$ and $(\bar{H}^{\lambda\gamma})$ as short for $(H_{12345})-(H_{67}^\lambda)-(H_{89}^\gamma)$.

Theorem 1. *If (H_{12345}) hold, then*

$$\lim_{T \downarrow 0} (\mathcal{I}_T - \mathbb{E}[u_0]) = 0.$$

Note that we did not assume the limit of $\mathbb{E}[u_0]$ to be finite. The proof, in Section 6.2.1, builds on arguments from [5, Proposition 3.1]. We then turn our attention to the ATM skew, defined in 1. This short-time asymptotic is reminiscent of [4, Proposition 6.2] and [2, Theorem 8].

Theorem 2. *If there exists $\lambda \in (-\frac{1}{2}, 0]$ such that (\bar{H}^λ) are satisfied, then*

$$\lim_{T \downarrow 0} \frac{\mathcal{S}_T}{T^\lambda} = \frac{1}{2} \lim_{T \downarrow 0} \mathbb{E} \left[\frac{1}{T^{\frac{1}{2}+\lambda}} \frac{\int_0^T |\Theta_s| ds}{u_0^3} \right]. \tag{7}$$

Note that (7) still holds without (H_7^λ) , but in that case both sides are infinite. In the rough-volatility setting of Section 2.2.1 with $v_t = f(\mathbf{W}_t^H)$, λ corresponds to $H - \frac{1}{2}$ so that (7) matches the slope of the observed ATM skew of SPX implied volatility. We prove this theorem in Section 6.2.2. We also provide the short-term curvature, in the following theorem, which is proved in Section 6.2.3.

Theorem 3. *If there exist $\lambda \in (-\frac{1}{2}, 0]$ and $\gamma \in (-1, \lambda]$ ensuring $(\bar{H}^{\lambda\gamma})$, then*

$$\begin{aligned} \lim_{T \downarrow 0} \frac{\mathcal{C}_T}{T^\gamma} = \lim_{T \downarrow 0} \frac{1}{T^\gamma} \left\{ \mathcal{S}_T - \frac{15}{2\sqrt{T}} \mathbb{E} \left[\frac{1}{u_0^7} \int_0^T |\Theta_r| \left(\int_r^T |\Theta_y| dy \right) dr \right] \right. \\ \left. + \frac{3}{2\sqrt{T}} \mathbb{E} \left[\frac{1}{u_0^5} \int_0^T \sum_{j=1}^N \left\{ \phi_s^j D_s^j \left(\int_s^T |\Theta_y| dy \right) \right\} ds \right] \right\}. \tag{8} \end{aligned}$$

The limit still holds without (H_7^γ) , but in that case the second and third terms are infinite.

Note that (H_7^λ) with $\lambda \geq \gamma$ guarantees that $T^{-\gamma} \mathcal{S}_T$ converges. By Theorem 2,

$$\lim_{T \downarrow 0} \frac{\mathcal{S}_T}{T^\gamma} = \begin{cases} 0, & \text{if } \lambda > \gamma, \\ \frac{1}{2} \lim_{T \downarrow 0} \mathbb{E} \left[\frac{1}{T^{\frac{1}{2}+\lambda}} \frac{\int_0^T |\Theta_s| ds}{u_0^3} \right], & \text{if } \lambda = \gamma, \\ +\infty, & \text{if } \lambda < \gamma. \end{cases}$$

4. Asymptotic results in the VIX case

As advertised, our framework includes the VIX case where

$$A_T = \text{VIX}_T = \sqrt{\frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}_T[v_r] dr},$$

for $v_r \in \mathbb{D}^{3,2}$ for all $r \in [0, T + \Delta]$, and we provide simple sufficient conditions for $(\bar{H}^{\lambda\gamma})$ to hold.

4.1. A generic volatility model

Consider the following four conditions, which we gather under the notation (\bar{C}) . There exist $H \in (0, \frac{1}{2})$ and $X \in L^p$ for all $p > 1$ such that the following hold:

- (C₁) For all $t \geq 0$, $\frac{1}{M_t} \leq X$ almost surely.
- (C₂) For all $i, j, k \in \llbracket 1, N \rrbracket$ and $t \leq s \leq y \leq T \leq r$, we have, almost surely
 - $v_r \leq X$,
 - $D_y^i v_r \leq X(r - y)^{H-}$,
 - $D_s^j D_y^i v_r \leq X(r - s)^{H-} (r - y)^{H-}$,
 - $D_t^k D_s^j D_y^i v_r \leq X(r - t)^{H-} (r - s)^{H-} (r - y)^{H-}$.

(C₃) For all $p > 1$, $\mathbb{E}[u_s^{-p}]$ is uniformly bounded in s and T , with $s \leq T$.

(C₄) For all $i, j, k \in \llbracket 1, N \rrbracket$ and $r \geq 0$, the mappings $y \mapsto D_y^i v_r$, $s \mapsto D_s^j D_y^i v_r$, and $t \mapsto D_t^k D_s^j D_y^i v_r$ are almost surely continuous in a neighbourhood of zero.

Recall the notation H_- and H_+ from the introduction. We compute the level, skew, and curvature of the VIX implied volatility in a model which satisfies the sufficient conditions. Let us define the following limits:

$$J_i := \int_0^\Delta \mathbb{E}[D_0^i v_r] dr, \quad G_{ij} := \int_0^\Delta \mathbb{E}[D_0^j D_0^i v_r] dr, \quad \text{for all } i, j \in \llbracket 1, N \rrbracket. \quad (9)$$

Proposition 1. *Under (\bar{C}) , the following limits hold:*

$$\begin{aligned} \lim_{T \downarrow 0} \mathcal{I}_T &= \frac{\|J\|}{2\Delta \text{VIX}_0^2}, & \text{if } H \in \left(0, \frac{1}{2}\right), \\ \lim_{T \downarrow 0} \mathcal{S}_T &= \frac{\sum_{i,j=1}^N J_i J_j \left(G_{ij} - \frac{J_i J_j}{\Delta \text{VIX}_0^2}\right)}{2\|J\|^3}, & \text{if } H \in \left(0, \frac{1}{2}\right), \\ \lim_{T \downarrow 0} \frac{\mathcal{C}_T}{T^{3H-\frac{1}{2}}} &= \frac{2\Delta \text{VIX}_0^2}{3\|J\|^5} \sum_{i,j,k=1}^N J_i J_j J_k \lim_{T \downarrow 0} \frac{\int_T^{T+\Delta} \mathbb{E}\left[D_0^k D_0^j D_0^i v_r\right] dr}{T^{3H-\frac{1}{2}}}, & \text{if } H \in \left(0, \frac{1}{6}\right). \end{aligned}$$

Remark 2. Our results stand under the fairly general set of assumptions (\bar{C}) . If v is a reasonably well-behaved function of an N -dimensional Gaussian Volterra process $(W^{1,H}, \dots, W^{N,H})$, then these should be relatively easy to check, as Proposition 2 suggests. For other rough stochastic volatility models, such as the rough Heston model [22, 39], it might be harder to verify the assumptions. Indeed, the latter is not even known to be Malliavin differentiable to this day, and thus does not lie within the scope of the present study.

We split the proof into two steps, collected in Section 6.3. First we show that (C₁), (C₂), (C₃) are sufficient to apply our main theorems, as they imply $(\bar{H}^{\lambda\gamma})$ for any $\lambda \in (-\frac{1}{2}, 0]$ and $\gamma \in (-1, 3H - \frac{1}{2}]$. Thanks to (C₄) we can also compute the limits—after a rigorous statement of convergence results—starting with \mathcal{I}_T and the skew with $\lambda = 0$. Restricting H to $(0, 1/6)$, which is the most relevant regime for rough volatility models, we can set $\gamma = 3H - \frac{1}{2} < \lambda$ and

compute the short-time curvature, with only the second term in (H_0^γ) contributing to the limit. The curvature limit in Proposition 1 is finite by the last item of (C_2) .

Remark 3. In the regime $H \in [1/6, 1/2)$, the rescaling becomes $\gamma = 0$, and many more terms that would just vanish when $H < 1/6$ now make a non-trivial contribution in the limit. Informally (that is, without a proof), the limit reads

$$\begin{aligned} \lim_{T \downarrow 0} \mathcal{C}_T &= \lim_{T \downarrow 0} \mathcal{S}_T - \frac{15\Delta \text{VIX}_0^2}{\|J\|^7} \left(\sum_{i,j=1}^N J_i J_j \left(G_{ij} - \frac{J_i J_j}{\Delta \text{VIX}_0^2} \right) \right)^2 \\ &+ \frac{12\Delta \text{VIX}_0^2}{\|J\|^5} \sum_{i,j,k=1}^N J_i \left(G_{jk} - \frac{J_j J_k}{\Delta \text{VIX}_0^2} \right) \left(G_{ik} - \frac{J_i J_k}{\Delta \text{VIX}_0^2} \right) \\ &+ \frac{1}{\|J\|^5} \sum_{i,j,k=1}^N \left\{ \frac{9(J_i J_j J_k)^2}{2\Delta \text{VIX}_0^2} - 6J_i J_j J_k (G_{ij} J_k + G_{ik} J_j + G_{jk} J_i) \right\} \\ &+ \frac{2\Delta \text{VIX}_0^2}{3\|J\|^5} \sum_{i,j,k=1}^N J_i J_j J_k \int_0^\Delta \mathbb{E} \left[D_0^k D_0^j D_0^i v_r \right] dr. \end{aligned}$$

4.2. The two-factor rough Bergomi

We consider the two-factor exponential model

$$v_t = v_0 \left[\chi \mathcal{E} \left(v W_t^{1,H} \right) + \bar{\chi} \mathcal{E} \left(\eta \left(\rho W_t^{1,H} + \bar{\rho} W_t^{2,H} \right) \right) \right] = : v_0 (\chi \mathcal{E}_t^1 + \bar{\chi} \mathcal{E}_t^2), \tag{10}$$

where $H \in (0, \frac{1}{2}]$, $W_t^{i,H} = \int_0^t (t-s)^{H-1} dW_s^i$, W^1, W^2 are independent Brownian motions, the Wick exponential is defined as $\mathcal{E}(X) := \exp\{X - \frac{1}{2}\mathbb{E}[X^2]\}$ for any random variable X , and $\chi \in [0, 1]$, $\bar{\chi} := 1 - \chi$, $v_0, v, \eta > 0$, $\rho \in [-1, 1]$, $\bar{\rho} = \sqrt{1 - \rho^2}$. This model is an extension of the Bergomi model [14], where the exponential kernel is replaced by a fractional one and an extension of the rough Bergomi model [9] to the two-factor case. It combines Bergomi’s insights on the need for several factors with the benefits of rough volatility. As proved in Section 6.4.1, it satisfies our conditions.;

Proposition 2. *If $\rho \in (-\sqrt{2}/2, 1]$, the model (10) satisfies (\bar{C}) .*

The restriction of the range of ρ is equivalent to $\rho + \bar{\rho} > 0$, a necessary requirement in the proof. Proposition 1 therefore applies and we obtain the following limits, as proved in Section 6.4.2.

Proposition 3. *Let $\psi(\rho, v, \eta, \chi) := \sqrt{(\chi v + \bar{\chi} \eta \rho)^2 + \bar{\chi}^2 \eta^2 \rho^2}$. If $H \in (0, \frac{1}{6})$ and $\rho \in (-\frac{\sqrt{2}}{2}, 1]$, then*

$$\begin{aligned} \lim_{T \downarrow 0} \mathcal{I}_T &= \frac{\Delta^{H-}}{2H+1} \psi(\rho, v, \eta, \chi), \\ \lim_{T \downarrow 0} \mathcal{S}_T &= \frac{H_+ \Delta^{H-}}{2\psi(\rho, v, \eta, \chi)^3} \left\{ (\chi v + \bar{\chi} \eta \rho)^2 \left[\frac{\chi v^2 + \bar{\chi} \eta^2 \rho^2}{2H} - \left(\frac{\chi v + \bar{\chi} \eta \rho}{H_+} \right)^2 \right] \right. \\ &\quad \left. + 2(\chi v + \bar{\chi} \eta \rho) \bar{\chi}^2 \eta^2 \rho^2 \left[\frac{\eta \rho}{2H} - \frac{v + \eta \rho}{H_+^2} \right] + \bar{\chi}^3 \eta^4 \rho^4 \left(\frac{1}{2H} - \frac{1}{H_+^2} \right) \right\}, \end{aligned}$$

$$\lim_{T \downarrow 0} \frac{C_T}{T^{3H-\frac{1}{2}}} = \frac{128\Delta^{-2H}H_+^2}{3\psi(\rho, \nu, \eta, \chi)^5(1-6H)} \left\{ (\chi\nu + \bar{\chi}\eta\rho)^3(\chi\nu^3 + \bar{\chi}\eta^3\rho^3) \right. \\ \left. + 3(\chi\nu + \bar{\chi}\eta\rho)^2\bar{\chi}^2\eta^4\rho^2\rho^2 + 3(\chi\nu + \bar{\chi}\eta\rho)\bar{\chi}^3\eta^5\rho^4\rho + \bar{\chi}^4\eta^6\rho^6 \right\}.$$

The limits depend explicitly on the parameters of the model $(H, \chi, \nu, \eta, \rho)$ and can be used to gain insight on their impact over the quantities of interest.

Remark 4.

- In the case $\rho = 1$ (and hence $\bar{\rho} = 0$), the above limits simplify to

$$\lim_{T \downarrow 0} \mathcal{I}_T = \frac{\Delta^{H-}}{2H_+} (\chi\nu + \bar{\chi}\eta),$$

$$\lim_{T \downarrow 0} \mathcal{S}_T = \frac{1}{2} \frac{H_+\Delta^{H-}}{\chi\nu + \bar{\chi}\eta} \left[\frac{\chi\nu^2 + \bar{\chi}\eta^2}{2H} - \left(\frac{\chi\nu + \bar{\chi}\eta}{H_+} \right)^2 \right],$$

$$\lim_{T \downarrow 0} \frac{C_T}{T^{3H-\frac{1}{2}}} = \frac{128\Delta^{-2H}H_+^2}{3-18H} \frac{\chi\nu^3 + \bar{\chi}\eta^3}{(\chi\nu + \bar{\chi}\eta)^2}.$$

- If we set $\rho = 0$ (and hence $\bar{\rho} = 1$), we obtain

$$\lim_{T \downarrow 0} \mathcal{I}_T = \frac{\Delta^{H-}}{2H_+} \sqrt{\chi^2\nu^2 + \bar{\chi}^2\eta^2},$$

$$\lim_{T \downarrow 0} \mathcal{S}_T = \frac{H_+\Delta^{H-}}{2(\chi^2\nu^2 + \bar{\chi}^2\eta^2)^{3/2}} \left\{ \chi^3\nu^4 \left[\frac{1}{2H} - \frac{\chi}{H_+^2} \right] - \frac{2\chi\nu^2\bar{\chi}^2\eta^2}{H_+^2} + \bar{\chi}^3\eta^4 \left(\frac{1}{2H} - \frac{1}{H_+^2} \right) \right\},$$

$$\lim_{T \downarrow 0} \frac{C_T}{T^{3H-\frac{1}{2}}} = \frac{128\Delta^{-2H}H_+^2}{3-18H} \frac{\chi^4\nu^6 + \bar{\chi}^4\eta^6}{(\chi^2\nu^2 + \bar{\chi}^2\eta^2)^{5/2}}.$$

- When $\rho = -1$ (not covered per se by the proposition), the above limits simplify to

$$\lim_{T \downarrow 0} \mathcal{I}_T = \frac{\Delta^{H-}}{2H+1} |\chi\nu - \bar{\chi}\eta|,$$

$$\lim_{T \downarrow 0} \mathcal{S}_T = \frac{H_+\Delta^{H-}}{2|\chi\nu - \bar{\chi}\eta|} \left[\frac{\chi\nu^2 + \bar{\chi}\eta^2}{2H} - \left(\frac{\chi\nu - \bar{\chi}\eta}{H_+} \right)^2 \right],$$

$$\lim_{T \downarrow 0} \frac{C_T}{T^{3H-\frac{1}{2}}} = \frac{128\Delta^{-2H}H_+^2}{3-18H} \frac{\chi\nu^3 - \bar{\chi}\eta^3}{(\chi\nu - \bar{\chi}\eta)^2} \operatorname{sgn}(\chi\nu - \bar{\chi}\eta).$$

Some tedious yet straightforward manipulations allow us to obtain some information about the sign of the limiting curvature.

Lemma 1. *For any $\eta, \nu > 0, \chi \in [0, 1]$, there exists $\rho_{\chi, \nu}^* < 0$ such that $\lim_{T \downarrow 0} \frac{C_T}{T^{3H-\frac{1}{2}}}$ is strictly positive when $\rho > \rho_{\chi, \nu}^*$ and strictly negative when $\rho < \rho_{\chi, \nu}^*$. When*

$$(\chi\nu - \bar{\chi}\eta) (\chi\nu^3 - \bar{\chi}\eta^3) > 0,$$

we have $\rho_{\chi, v, \eta}^* < -1$, and hence the limiting curvature is strictly positive for all $\rho \in [-1, 1]$.

Proof. The expression we are interested in, given in Proposition 3, and ignoring the obviously strictly positive multiplicative factor, reads

$$\begin{aligned} \Phi(\rho) &= [\chi v + \bar{\chi} \eta \rho]^3 (\chi v^3 + \bar{\chi} \eta^3 \rho^3) + 3\{\chi v + \bar{\chi} \eta \rho\}^2 \bar{\chi}^2 \eta^4 \rho^2 \rho^2 \\ &\quad + 3(\chi v + \bar{\chi} \eta \rho) \bar{\chi}^3 \eta^5 \rho^4 \rho + \bar{\chi}^4 \eta^6 \rho^6 \\ &= \left[\chi^3 v^3 + 3\chi^2 v^2 \bar{\chi} \eta \rho + 3\chi v \bar{\chi}^2 \eta^2 \rho^2 + \bar{\chi}^3 \eta^3 \rho^3 \right] (\chi v^3 + \bar{\chi} \eta^3 \rho^3) \\ &\quad + 3 \left\{ \chi^2 v^2 + 2\chi v \bar{\chi} \eta \rho + \bar{\chi}^2 \eta^2 \rho^2 \right\} \bar{\chi}^2 \eta^4 \rho^2 \rho^2 + 3(\chi v + \bar{\chi} \eta \rho) \bar{\chi}^3 \eta^5 \rho^4 \rho + \bar{\chi}^4 \eta^6 \rho^6 \\ &= \left[\chi^3 v^3 + 3\chi^2 v^2 \bar{\chi} \eta \rho + 3\chi v \bar{\chi}^2 \eta^2 \rho^2 + \bar{\chi}^3 \eta^3 \rho^3 \right] (\chi v^3 + \bar{\chi} \eta^3 \rho^3) \\ &\quad + 3 \left\{ \chi^2 v^2 + 2\chi v \bar{\chi} \eta \rho + \bar{\chi}^2 \eta^2 \rho^2 \right\} \bar{\chi}^2 \eta^4 (1 - \rho^2) \rho^2 \\ &\quad + 3(\chi v + \bar{\chi} \eta \rho) \bar{\chi}^3 \eta^5 (1 - 2\rho^2 + \rho^4) \rho + \bar{\chi}^4 \eta^6 (1 - 3\rho^2 + 3\rho^4 - \rho^6) \\ &=: \sum_{i=0}^6 \alpha_i \rho^i, \end{aligned}$$

where

$$\begin{aligned} \alpha_0 &= \chi^4 v^6 + \bar{\chi}^4 \eta^6 &&= \chi^4 v^6 + \bar{\chi}^4 \eta^6, \\ \alpha_1 &= 3\chi^3 v^5 \bar{\chi} \eta + 3\chi v \bar{\chi}^3 \eta^5 &&= 3(\chi^2 v^4 + \bar{\chi}^2 \eta^4) \chi v \eta \bar{\chi}, \\ \alpha_2 &= 3\chi^2 \bar{\chi}^2 \eta^2 v^4 + 3\chi^2 v^2 \bar{\chi}^2 \eta^4 + 3\bar{\chi}^4 \eta^6 \\ &\quad - 3\bar{\chi}^4 \eta^6 = 3\chi^2 \bar{\chi}^2 \eta^2 v^4 + 3\chi^2 v^2 \bar{\chi}^2 \eta^4 &&= 3(v^2 + \eta^2) \chi^2 \bar{\chi}^2 \eta^2 v^2, \\ \alpha_3 &= \chi^3 v^3 \bar{\chi} \eta^3 + \bar{\chi}^3 \eta^3 \chi v^3 + 6\chi v \bar{\chi}^3 \eta^5 \\ &\quad - 6\chi v \bar{\chi}^3 \eta^5 = (\chi^2 + \bar{\chi}^2) \chi \bar{\chi} v^3 \eta^3 &&= (\chi^2 + \bar{\chi}^2) \chi \bar{\chi} v^3 \eta^3, \\ \alpha_4 &= 3\chi^2 v^2 \bar{\chi}^2 \eta^4 - 3\chi^2 v^2 \bar{\chi}^2 \eta^4 + 3\bar{\chi}^4 \eta^6 - 6\bar{\chi}^4 \eta^6 + 3\bar{\chi}^4 \eta^6 &&= 0, \\ \alpha_5 &= 3\bar{\chi}^3 \eta^5 \chi v - 6\bar{\chi}^3 \eta^5 \chi v + 3\chi v \bar{\chi}^3 \eta^5 &&= 0, \\ \alpha_6 &= \bar{\chi}^4 \eta^6 - 3\bar{\chi}^4 \eta^6 + 3\bar{\chi}^4 \eta^6 - \bar{\chi}^4 \eta^6 &&= 0. \end{aligned}$$

These surprising simplifications show that Φ is in fact a polynomial in ρ of order three, with a strictly positive leading coefficient α_3 , so that Φ'' is linear and increasing in ρ and is such that

$$\Phi''(\rho) = 0 \quad \text{if and only if} \quad \rho = -\frac{(\eta^2 + v^2) \bar{\chi} \chi}{v \eta (\chi^2 + \bar{\chi}^2)} =: \rho_{\bullet}(\chi, v, \eta).$$

Now,

$$\begin{aligned} &\Phi'(\rho_{\bullet}(\chi, v, \eta)) \\ &= 3\alpha_3 \rho^*(\eta, v \chi)^2 + 2\alpha_2 \rho^*(\eta, v \chi) + \alpha_1 \end{aligned}$$

$$\begin{aligned}
 &= 3 \left(\chi^2 + \bar{\chi}^2 \right) \chi \bar{\chi} v^3 \eta^3 \rho^*(\eta, v \chi)^2 + 6 \left(v^2 + \eta^2 \right) \chi^2 \bar{\chi}^2 \eta^2 v^2 \rho^*(\eta, v \chi) \\
 &\quad + 3 \left(\chi^2 v^4 + \bar{\chi}^2 \eta^4 \right) \chi v \eta \bar{\chi} \\
 &= 3 \left(\chi^2 + \bar{\chi}^2 \right) \chi \bar{\chi} v^3 \eta^3 \left[\frac{(\eta^2 + v^2) \bar{\chi} \chi}{v \eta (\chi^2 + \bar{\chi}^2)} \right]^2 - 6 \left(v^2 + \eta^2 \right) \chi^2 \bar{\chi}^2 \eta^2 v^2 \frac{(\eta^2 + v^2) \bar{\chi} \chi}{v \eta (\chi^2 + \bar{\chi}^2)} \\
 &\quad + 3 \left(\chi^2 v^4 + \bar{\chi}^2 \eta^4 \right) \chi v \eta \bar{\chi} \\
 &= 3 \chi^3 \bar{\chi}^3 v \eta \frac{(\eta^2 + v^2)^2}{\chi^2 + \bar{\chi}^2} - 6 \chi^3 \bar{\chi}^3 \eta v \frac{(\eta^2 + v^2)^2}{\chi^2 + \bar{\chi}^2} + 3 \left(\chi^2 v^4 + \bar{\chi}^2 \eta^4 \right) \chi v \eta \bar{\chi} \\
 &= \frac{1}{\chi^2 + \bar{\chi}^2} \left(-3 \chi^3 \bar{\chi}^3 v \eta (\eta^2 + v^2)^2 + 3 (\chi^2 + \bar{\chi}^2) \left(\chi^2 v^4 + \bar{\chi}^2 \eta^4 \right) \chi v \eta \bar{\chi} \right) \\
 &= \frac{3 v \eta \chi \bar{\chi}}{\chi^2 + \bar{\chi}^2} \left((\chi^2 + \bar{\chi}^2) \left(\chi^2 v^4 + \bar{\chi}^2 \eta^4 \right) - \chi^2 \bar{\chi}^2 (\eta^2 + v^2)^2 \right) \\
 &= \frac{3 v \eta \chi \bar{\chi}}{\chi^2 + \bar{\chi}^2} \left(\chi^4 v^4 + \chi^2 \bar{\chi}^2 \eta^4 + \chi^2 \bar{\chi}^2 v^4 + \bar{\chi}^4 \eta^4 - \chi^2 \bar{\chi}^2 \left(\eta^4 + 2 \eta^2 v^2 + v^4 \right) \right) \\
 &= \frac{3 v \eta \chi \bar{\chi}}{\chi^2 + \bar{\chi}^2} \left(\chi^4 v^4 + \bar{\chi}^4 \eta^4 - 2 \chi^2 \bar{\chi}^2 \eta^2 v^2 \right) = \frac{3 v \eta \chi \bar{\chi}}{\chi^2 + \bar{\chi}^2} (\chi v + \bar{\chi} \eta)^2 (\chi v - \bar{\chi} \eta)^2 > 0.
 \end{aligned}$$

Since Φ' is an upward parabola with strictly positive minimum, it is always strictly positive; hence Φ is a strictly increasing function (of ρ), and the lemma follows. Let $\rho_{\chi, v, \eta}^*$ denote the unique solution to $\Phi(\rho_{\chi, v, \eta}^*) = 0$; there is an explicit closed-form expression for this solution, but its exact representation is messy and not particularly informative. We can, however, provide an upper bound. Indeed,

$$\begin{aligned}
 \Phi(-1) &= \alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 \\
 &= \chi^4 v^6 + \bar{\chi}^4 \eta^6 - 3 \left(\chi^2 v^4 + \bar{\chi}^2 \eta^4 \right) \chi v \eta \bar{\chi} + 3 \left(v^2 + \eta^2 \right) \chi^2 \bar{\chi}^2 \eta^2 v^2 \\
 &\quad - \left(\chi^2 + \bar{\chi}^2 \right) \chi \bar{\chi} v^3 \eta^3 \\
 &= \left(\chi^3 v^3 - 3 \chi^2 v^2 \eta \bar{\chi} + 3 \chi \bar{\chi}^2 \eta^2 v - \bar{\chi}^3 \eta^3 \right) \chi v^3 \\
 &\quad + \left(\bar{\chi}^3 \eta^3 - 3 \bar{\chi}^2 \eta^2 \chi v + 3 \chi^2 \bar{\chi} \eta v^2 - \chi^3 v^3 \right) \bar{\chi} \eta^3 \\
 &= (\chi v - \bar{\chi} \eta)^3 \left(\chi v^3 - \bar{\chi} \eta^3 \right).
 \end{aligned}$$

As soon as $\Phi(-1) > 0$, clearly $\rho_{\chi, v, \eta}^* < -1$, so the limiting curvature is always strictly positive. The sign of $\Phi(-1)$ is given by that of $(\chi v - \bar{\chi} \eta) (\chi v^3 - \bar{\chi} \eta^3)$, which is an upward parabola in χ . □

5. The stock smile under multi-factor models

We use the setting of Section 2.2.1 to apply our results to an asset price of the form

$$S_t = S_0 + \int_0^t S_r \sqrt{v_r} dB_r,$$

where B is correlated with the other N Brownian motions as $B = \sum_{i=1}^N \rho_i W^i$ with $\sum_{i=1}^N \rho_i^2 = 1$, $\rho_i \in [-1, 1]$ for all $i \in \llbracket 1, N \rrbracket$. The volatility is a function of $(N - 1)$ Brownian motions, such that the stock price features one additional and independent source of randomness. To fit this model into (3) we set $A = S$ and identify ϕ^i with $\rho_i \sqrt{v}$. We modify the notation slightly to differentiate from the VIX framework: the implied volatility is denoted by $\widehat{\mathcal{I}}_T$ and the skew by $\widehat{\mathcal{S}}_T$. We do not consider the curvature in this setting, for lack of an explicit formula. The proof of this proposition and the following corollary are postponed to Section 6.5.

Proposition 4. Assume that there exist $H \in (0, \frac{1}{2})$ and a random variable X such that, for all $0 \leq s \leq y$, $j \in \llbracket 1, N \rrbracket$, and $p \geq 1$, $X \in L^p$, the following hold:

- (i) $v_s \leq X$;
- (ii) $D_s^j v_y \leq X(y - s)^{H-}$;
- (iii) $\sup_{s \leq T} \mathbb{E}[u_s^{-p}] < \infty$;
- (iv) $\limsup_{T \downarrow 0} \mathbb{E}[(\sqrt{v_T/v_0} - 1)^2] = 0$.

Then the short-time limits of the implied volatility and skew are

$$\lim_{T \downarrow 0} \widehat{\mathcal{I}}_T = \sqrt{v_0} \quad \text{and} \quad \lim_{T \downarrow 0} \frac{\widehat{\mathcal{S}}_T}{T^{H-}} = \frac{1}{2v_0} \sum_{j=1}^N \rho_j \lim_{T \downarrow 0} \frac{\int_0^T \int_s^T \mathbb{E} [D_s^j v_y] dy ds}{T^{H+3/2}}.$$

Remark 5.

- The second limit is finite because of the condition (ii).
- The one-dimensional version ($N = 2$) agrees with [4, Theorem 6.3] up to the sign because the authors of [4] derive with respect to the spot x and not to the log-strike k .

In the two-factor rough Bergomi model (10) we can compute the short-time skew more explicitly. Recall from Example 2.2.4 that, for all $t \geq 0$, it means setting $N = 3$ and defining

$$\begin{cases} S_t = S_0 + \int_0^t S_r \sqrt{v_r} dB_r, \\ v_t = v_0 [\chi \mathcal{E} (v W^{1,H})_t + \bar{\chi} \mathcal{E} (\eta (\rho W^{1,H} + \bar{\rho} W^{2,H}))_t], \end{cases}$$

where $W_t^{i,H} = \int_0^t (t - s)^{H-} dW_s^i$, for $i = 1, 2$ and $B = \sum_{i=1}^3 \rho_i W^i$, with W^1, W^2, W^3 being independent Brownian motions. Hence W^3 influences only the asset price, not the variance.

Corollary 1. In the two-factor rough Bergomi model we have the short-time skew limit

$$\lim_{T \downarrow 0} \frac{\widehat{\mathcal{S}}_T}{T^{H-}} = \frac{\rho_1 \chi v + \eta \bar{\chi} (\rho_1 \rho + \rho_2 \bar{\rho})}{2H_+(1 + H_+)}. \tag{11}$$

5.1. Tips for joint calibration in the two-factor rough Bergomi model

Assuming we can observe the short-time limits of the spot ATM implied volatility, it grants us v_0 for free, while the slope of its skew gives us H by (11). Next, we simplify the expressions from Proposition 3 in the case $\chi = \frac{1}{2}$. Denote by \mathcal{I}_0, S_0 , and \mathcal{C}_0 the three limits of Proposition 3,

and let $H_{\pm} := H \pm \frac{1}{2}$, $\alpha := \eta\rho$, $\beta := \eta\bar{\rho}$. Introduce further the normalised parameters

$$\tilde{\alpha} := \frac{\alpha}{\nu}, \quad \tilde{\beta} := \frac{\beta}{\nu},$$

so that, defining $\tilde{\psi}(\tilde{\alpha}, \tilde{\beta}) := \sqrt{(1 + \tilde{\alpha})^2 + \tilde{\beta}^2}$, we have, after simplifications,

$$\begin{aligned} \mathcal{I}_0 &= \frac{\nu\Delta H_-}{4H_+} \sqrt{(1 + \tilde{\alpha})^2 + \tilde{\beta}^2} &&= : \nu C_I \tilde{\psi}(\tilde{\alpha}, \tilde{\beta}), \\ \mathcal{S}_0 &= \frac{\nu H_+ \Delta H_-}{2} \frac{(1 + \tilde{\alpha})^2 \left[\frac{1 + \tilde{\alpha}^2}{2H} - \left(\frac{1 + \tilde{\alpha}}{H_+} \right)^2 \right] + 2(1 + \tilde{\alpha})\tilde{\beta}^2 \left[\frac{\tilde{\alpha}}{2H} - \frac{1 + \tilde{\alpha}}{H_+} \right] + \tilde{\alpha}^4 \left[\frac{1}{2H} - \frac{1}{H_+^2} \right]}{(1 + \tilde{\alpha})^2 + \tilde{\beta}^2)^{3/2}} &&= : \nu C_S \frac{\Phi_S(\tilde{\alpha}, \tilde{\beta})}{\tilde{\psi}(\tilde{\alpha}, \tilde{\beta})^3}, \\ \mathcal{C}_0 &= \frac{128\nu H_+^2}{3\Delta^{2H}} \frac{\left\{ (1 + \tilde{\alpha})^3(1 + \tilde{\alpha}^3) + 3(1 + \tilde{\alpha})^2\tilde{\alpha}^2\tilde{\beta}^2 + 3(1 + \tilde{\alpha})\tilde{\alpha}\tilde{\beta}^4 + \tilde{\beta}^6 \right\}}{(1 + \tilde{\alpha})^2 + \tilde{\beta}^2)^{5/2} (1 - 6H)} &&= : \nu C_C \frac{\Phi_C(\tilde{\alpha}, \tilde{\beta})}{\tilde{\psi}(\tilde{\alpha}, \tilde{\beta})^5}, \end{aligned}$$

where the constants C_I , C_S , C_C only depend on Δ and H . Provided we can observe an approximation of these three limits, we can numerically solve for $\nu, \tilde{\alpha}, \tilde{\beta}$ in a system with three equations. Alternatively, since all three quantities have the factor ν , any quotient of two of them is a function of only $\tilde{\alpha}, \tilde{\beta}$, which we can plot and match to observed data. Both methods allow us to deduce $\nu, \tilde{\alpha}, \tilde{\beta}$, in turn yielding η and ρ . Finally, we are left with ρ_1 and ρ_2 to play with so that the right-hand-side of (11) matches the market observations.

Remark 6. We are not here—as in fact in most papers related to asymptotics—advocating the use of these formulae for actual direct option pricing, since they are asymptotics. In particular, this raises several calibration issues (shared with most results on the topic): (i) very short-maturity options on the VIX are hardly available, and the computation of the curvature, in particular, is a matter of personal choice (the result will change drastically depending on the number of data points around the ATM), which is left to the trader; (ii) such asymptotic formulae serve to provide some intuition about the roles of the model parameters, in particular on which one helps for each part of the smile. One key message of our result, for example, is that the model is able to disentangle (over short time horizons) the role of H and that of ν, η, ρ , and to a certain extent the role of ν and that of η, ρ . Compared to simpler models (one-factor (rough) Bergomi, classical Heston), we have more parameters here, and our results should be combined with more asymptotics (for smile wings and large expiries) to be fully meaningful. Unfortunately, these are not fully available yet, and we would rather leave a full-scale numerical calibration scheme to future endeavours.

6. Proofs

6.1. Useful results

We start by adapting to the multivariate case a well-known decomposition formula and then prove a lemma which will be used extensively in the rest of the proofs. Both proofs build on the multidimensional anticipative Itô formula [42, Theorem 3.2.4].

Proposition 5. (Price decomposition.) *Under (H123), the following decomposition formula holds, for all $t \in \mathbb{T}$, for the price (5), with u_t defined in (4) and $G := (\partial_x^2 - \partial_x)\text{BS}$:*

$$\Pi_t(k) = \mathbb{E}_t [\text{BS}(t, \mathfrak{M}_t, k, u_t)] + \frac{1}{2} \mathbb{E}_t \left[\int_t^T \partial_x G(s, \mathfrak{M}_s, k, u_s) | \Theta_s | ds \right].$$

Proof. Define $\widehat{\text{BS}}(t, x, k, \sigma^2(T-t)) := \text{BS}(t, x, k, \sigma)$ and write for simplicity $\widehat{\text{BS}}_T := \widehat{\text{BS}}(t, \mathfrak{M}_t, k, Y_t) = \text{BS}(t, \mathfrak{M}_t, k, u_t)$, where we recall that $Y_t = u_t^2(T-t)$. Note that $\Pi_T = \widehat{\text{BS}}_T$; hence $\Pi_t = \mathbb{E}_t[\widehat{\text{BS}}_T]$ by no-arbitrage arguments. Thanks to (H₁) and (H₂), we can then apply a multidimensional anticipative Itô's formula [42, Theorem 3.2.4] with respect to (t, \mathfrak{M}, Y) :

$$\begin{aligned} \text{BS}(T, \mathfrak{M}_T, k, u_T) &= \widehat{\text{BS}}_T = \widehat{\text{BS}}_t + \int_t^T \partial_s \widehat{\text{BS}}_s ds + \int_t^T \partial_x \widehat{\text{BS}}_s \left(d(\phi \bullet \mathbf{W})_s - \frac{1}{2} \|\phi_s\|^2 ds \right) \\ &\quad - \int_t^T \partial_y \widehat{\text{BS}}_s \|\phi_s\|^2 ds + \frac{1}{2} \int_t^T \partial_x^2 \widehat{\text{BS}}_s \|\phi_s\|^2 ds + \int_t^T \partial_{xy} \widehat{\text{BS}}_s |\Theta_s| ds, \end{aligned}$$

with Θ as in (4). The derivatives of the Black–Scholes price read as follows (for simplicity, we omit the argument):

$$\partial_s \widehat{\text{BS}}_s = \partial_s \text{BS}_s + \frac{u_s \partial_u \text{BS}}{2(T-s)}, \quad \partial_y \widehat{\text{BS}}_s = \frac{\partial_u \text{BS}}{2u_s(T-s)}, \quad \text{and} \quad G = \frac{\partial_u \text{BS}}{u_s(T-s)}.$$

Putting everything together, using the gamma–vega–delta relation

$$\frac{\partial_\sigma \text{BS}(t, x, k, \sigma)}{\sigma(T-t)} = (\partial_x^2 - \partial_x) \text{BS}(t, x, k, \sigma), \tag{12}$$

and applying conditional expectation, we obtain

$$\Pi_t = \mathbb{E}_t[\text{BS}(t, \mathfrak{M}_t, k, u_t)] + \mathbb{E}_t \left[\int_t^T \mathcal{L}_{\text{BS}}(s, u_s) ds \right] + \mathbb{E}_t \left[\int_t^T \frac{\partial_{xu} \text{BS}(s, \mathfrak{M}_s, k, u_s)}{2u_s(T-s)} |\Theta_s| ds \right], \tag{13}$$

where $\mathcal{L}_{\text{BS}}(s, u_s) := \frac{1}{2} [u_s^2 (\partial_x^2 - \partial_x) + \partial_s] \text{BS}(s, \mathfrak{M}_s, k, u_s)$ is the Black–Scholes operator applied to the Black–Scholes function. Since $\mathcal{L}_{\text{BS}}(s, u_s) = 0$ by construction and

$$\partial_x G(s, x, k, \sigma) = \frac{e^x \mathcal{N}'(d_+(x, k, \sigma))}{\sigma \sqrt{T-s}} \left(1 - \frac{d_+(x, k, \sigma)}{\sigma \sqrt{T-s}} \right),$$

the last term in (13) is well defined by (H₃) and the proposition follows.

Lemma 2. For all $t \in \mathbb{T}$, let $J_t := \int_t^T a_s ds$, for some adapted process $a \in \mathbb{L}^{1,2}$, and let $\mathcal{L} := \sum_{i=1}^n c_i \partial_x^i$ be a linear combination of partial derivatives, with weights $c_i \in \mathbb{R}$. Then, writing for clarity $\text{BS}_t := \text{BS}(t, \mathfrak{M}_t, \mathfrak{M}_0, u_t)$, we have

$$\begin{aligned} \mathbb{E} \left[\int_0^T \mathcal{L} \text{BS}_s a_s ds \right] &= \mathbb{E} \left[\mathcal{L} \text{BS}_0 J_0 + \int_0^T \left(\partial_x^3 - \partial_x^2 \right) \mathcal{L} \text{BS}_s |\Theta_s| J_s ds \right. \\ &\quad \left. + \int_0^T \partial_x \mathcal{L} \text{BS}_s \sum_{k=1}^N \left(\phi_s^k D_s^k J_s \right) ds \right]. \end{aligned} \tag{14}$$

Remark 7. We will use this lemma freely below, with the justification that the condition $a \in \mathbb{L}^{1,2}$ is always satisfied thanks to (H₁). □

Proof. As in the proof of Proposition 5, we define $\widehat{\text{BS}}(t, x, k, \sigma^2(T-t)) := \text{BS}(t, x, k, \sigma)$ and write for simplicity $\widehat{\text{BS}}_T := \widehat{\text{BS}}(t, \mathfrak{M}_t, \mathfrak{M}_0, Y_t) = \text{BS}(t, \mathfrak{M}_t, \mathfrak{M}_0, u_t)$. Define

$\widehat{P}(t, x, k, y, j) := \widehat{\mathcal{LBS}}(t, x, k, u)j$ and denote $\widehat{P}(t, \mathfrak{M}_s, \mathfrak{M}_0, u_s, J_s)$ by \widehat{P}_t for simplicity. We then apply the multidimensional anticipative Itô's formula [42, Theorem 3.2.4] with respect to (t, \mathfrak{M}, Y, J) :

$$\begin{aligned} \widehat{P}_T &= \widehat{P}_0 + \int_0^T \partial_s \widehat{P}_s ds + \int_0^T \partial_x \widehat{P}_s \left(d(\boldsymbol{\phi} \bullet \mathbf{W})_s - \frac{1}{2} \|\boldsymbol{\phi}_s\|^2 ds \right) - \int_t^T \partial_y \widehat{P}_s \|\boldsymbol{\phi}_s\|^2 ds \\ &\quad + \frac{1}{2} \int_t^T \partial_x^2 \widehat{P}_s \|\boldsymbol{\phi}_s\|^2 ds + \int_t^T \partial_{xy} \widehat{P}_s | \boldsymbol{\Theta}_s | ds + \int_0^T \partial_j \widehat{P}_s dJ_s + \int_0^T \partial_{xj} \widehat{P}_s \sum_{k=1}^N \left(\phi_s^k D_s^k J_s \right) ds. \end{aligned}$$

One first notices that $\widehat{P}_0 = \widehat{\mathcal{LBS}}_0 J_0$ and $\widehat{P}_T = 0$. Moreover we observe that $\int_0^T \partial_j \widehat{P}_s dJ_s = - \int_0^T \widehat{\mathcal{LBS}}_s a_s ds$, which corresponds to the left-hand-side of (14), and

$$\int_0^T \partial_{xj} \widehat{P}_s \sum_{k=1}^N \left(\phi_s^k D_s^k J_s \right) ds = \int_0^T \partial_x \widehat{\mathcal{LBS}}_s \sum_{k=1}^N \left(\phi_s^k D_s^k J_s \right) ds.$$

Since \mathcal{L} is a linear operator, the partial derivatives in s, x , and u cancel as in the proof of Proposition 5. That means we are left with

$$\begin{aligned} \int_0^T \widehat{\mathcal{LBS}}_s a_s ds &= \widehat{\mathcal{LBS}}_0 J_0 + \int_0^T \left(\partial_x^3 - \partial_x^2 \right) \widehat{\mathcal{LBS}}_s | \boldsymbol{\Theta}_s | J_s ds + \int_0^T \partial_x \widehat{P}_s d(\boldsymbol{\phi} \bullet \mathbf{W})_s \\ &\quad + \int_0^T \partial_x \widehat{\mathcal{LBS}}_s \sum_{k=1}^N \left(\phi_s^k D_s^k J_s \right) ds. \end{aligned}$$

Since $\partial_x^n \widehat{\mathcal{BS}}(s, x, u) = \partial_x^n \widehat{\mathcal{BS}}(s, x, u^2(T - s))$ for any $n \in \mathbb{N}$, summing everything and taking expectations imply the claim. \square

We adapt and clarify [4, Lemma 4.1] to obtain a convenient bound for the partial derivatives of G . For notational simplicity, since σ and $T - t$ are fixed, we write $\varsigma := \sigma \sqrt{T - t}$ and $\mathfrak{G}(x, k, \varsigma) := G(t, x, k, \sigma)$.

Proposition 6. *For any $n \in \mathbb{N}$ and $p \in \mathbb{R}$, there exists $C_{n,p} > 0$ independent of x and ς such that, for all $\varsigma > 0$ and $x \in \mathbb{R} \setminus \left\{ 0, \frac{\varsigma^2}{2} \right\}$,*

$$\partial_x^n \mathfrak{G}(x, k, \varsigma) \leq \frac{C_{n,p} e^k}{\varsigma^{p+1}}. \tag{15}$$

If $x = 0$, then for any $n \in \mathbb{N}$ the bound (15) holds with $p = n$.

If $x = \frac{1}{2} \varsigma^2$, there exists a strictly positive constant C_n independent of ς such that

$$\partial_x^n \mathfrak{G} \left(\frac{\varsigma^2}{2}, k, \varsigma \right) = \begin{cases} \frac{C_n e^k}{\varsigma^{n+1}}, & \text{if } n \text{ is even,} \\ 0, & \text{if } n \text{ is odd.} \end{cases}$$

The following simplification (and extension) will be useful later.

Corollary 2. *For any $n \in \mathbb{N}$, there exists a non-negative $C_{n,k}$ independent of x and ς such that, for all $\varsigma > 0$ and $x \in \mathbb{R}$,*

$$\left| \partial_x^n \mathfrak{G}(x, k, \varsigma) \right| \leq \frac{C_{n,k}}{\varsigma^{n+1}}, \quad \left| \partial_k \partial_x^n \mathfrak{G}(x, k, \varsigma) \right| \leq \frac{C_{n,k}}{\varsigma^{n+2}}, \quad \text{and} \quad \left| \partial_k^2 \partial_x^n \mathfrak{G}(x, k, \varsigma) \right| \leq \frac{C_{n,k}}{\varsigma^{n+3}}.$$

Proof of Proposition 6. We first consider the case $k = 0$. Since

$$\mathfrak{G}(x, 0, \varsigma) := (\partial_{xx} - \partial_x)\text{BS}(t, x, 0, \varsigma) = \frac{1}{\varsigma\sqrt{2\pi}} \exp\left\{x - \frac{1}{2}d_+(x, \varsigma)^2\right\},$$

where $d_+(x, \varsigma) := d_+(x, 0, \sigma) = \frac{x}{\varsigma} + \frac{\varsigma}{2}$, direct computation (proof by recursion) yields, for any $n \in \mathbb{N}$,

$$\partial_x^n \mathfrak{G}(x, 0, \varsigma) = \exp\left\{-\frac{(\varsigma^2 - 2x)^2}{8\varsigma^2}\right\} \sum_{j=0}^n \alpha_j \frac{P_j(x)}{\varsigma^{2j+1}}, \tag{16}$$

where, for each j , P_j is a polynomial of degree j independent of ς .

Since $d_+(\frac{\varsigma^2}{2}, \varsigma) = \partial_x d_+(\frac{\varsigma^2}{2}, \varsigma) = 0$, $\partial_x^2 d_+(\frac{\varsigma^2}{2}, \varsigma) = -\frac{1}{\varsigma^2}$, the induction simplifies to

$$\partial_x^n \mathfrak{G}(x, 0, \varsigma)\Big|_{x=\frac{\varsigma^2}{2}} = \begin{cases} C_n & \text{if } n \text{ is even,} \\ \varsigma^{n+1} & \text{if } n \text{ is odd,} \end{cases}$$

for some constant $C_n > 0$ independent of ς , proving the third statement in the proposition.

Similarly, if $x = 0$, simplifications occur which yield, for any $n \in \mathbb{N}$,

$$\partial_x^n \mathfrak{G}(x, 0, \varsigma)\Big|_{x=0} = \exp\left\{-\frac{\varsigma^2}{8}\right\} \sum_{j=0}^n \frac{\alpha_j}{\varsigma^{j+1}} = \frac{1}{\varsigma^{n+1}} \exp\left\{-\frac{\varsigma^2}{8}\right\} \sum_{j=0}^n \alpha_j \varsigma^{n-j},$$

and the second statement in the proposition follows.

Finally, in the general case $x \in \mathbb{R} \setminus \left\{0, \frac{\varsigma^2}{2}\right\}$, we can rewrite (16) for any $p \in \mathbb{R}$ as

$$\begin{aligned} \partial_x^n \mathfrak{G}(x, 0, \varsigma) &= \frac{1}{\varsigma^{p+1}} \exp\left\{-\frac{(\varsigma^2 - 2x)^2}{8\varsigma^2}\right\} \sum_{j=0}^n \alpha_k P_j(x) \varsigma^{p-2j} \\ &=: \frac{1}{\varsigma^{p+1}} \exp\left\{-\frac{(\varsigma^2 - 2x)^2}{8\varsigma^2}\right\} H_{n,p}(x, \varsigma). \end{aligned}$$

For each $n \in \mathbb{N}$, $p \in \mathbb{N}$, $H_{n,p}$ is a two-dimensional function consisting only of powers of ς^2 and x^2/ς^2 . Since the exponential factor contains these very same terms, there exists a strictly positive constant $C_{n,p}$, independent of x and ς , such that

$$\exp\left\{-\frac{(\varsigma^2 - 2x)^2}{8\varsigma^2}\right\} H_{n,p}(x, \varsigma) \leq C_{n,p},$$

proving the proposition in the case $k = 0$.

The case $k \in \mathbb{R}$ follows directly from the observation that $\mathfrak{G}(x, 0, \varsigma) = \mathfrak{G}(x - k, 0, \varsigma)e^k$. Finally, since $\partial_k d_+(x, k, \sigma) = -\partial_x d_+(x, k, \sigma)$ and $\partial_k^2 d_+(x, k, \sigma) = -\partial_x^2 d_+(x, k, \sigma)$, the same simplifications occur if we take a partial derivative with respect to k instead of x .

6.2. Proofs of the main results

6.2.1. *Proof of Theorem 1: level.* To prove this result, we draw insights from the proofs of [2, Theorem 8] and [5, Proposition 3.1]. By definition

$$\mathcal{I}_T = \text{BS}^{\leftarrow}(0, \mathfrak{M}_0, \mathfrak{M}_0, \Pi_0) = : \overleftarrow{\text{BS}}(\Pi_0),$$

and we write $\overleftarrow{\text{BS}}(x) := \text{BS}(0, x, x, u_0)$. Using Proposition 5 at time 0, we see that $\Pi_0 = \Gamma_T$, where

$$\Gamma_t := \mathbb{E} \left[\overleftarrow{\text{BS}}(\mathfrak{M}_0) + \frac{1}{2} \int_0^t \partial_x G(s, \mathfrak{M}_s, \mathfrak{M}_0, u_s) | \Theta_s | ds \right], \quad \text{for } t \in \mathbb{T},$$

which is a deterministic path. The fundamental theorem of integration reads

$$\begin{aligned} \mathcal{I}_T = \overleftarrow{\text{BS}}(\Gamma_T) &= \overleftarrow{\text{BS}}(\Gamma_0) + \int_0^T \partial_t \overleftarrow{\text{BS}}(\Gamma_t) dt = \overleftarrow{\text{BS}}(\Gamma_0) + \int_0^T \overleftarrow{\text{BS}}'(\Gamma_t) \partial_t \Gamma_t dt \\ &= \overleftarrow{\text{BS}}(\Gamma_0) + \frac{1}{2} \int_0^T \overleftarrow{\text{BS}}'(\Gamma_t) \mathbb{E}[|\Theta_t| \partial_x G_t] dt, \end{aligned}$$

where $G_t := G(t, \mathfrak{M}_t, \mathfrak{M}_0, u_t)$. We can deal with the integral by computing $\overleftarrow{\text{BS}}'$ and $\partial_x G$ explicitly:

$$\begin{aligned} \overleftarrow{\text{BS}}'(\Gamma_t) &= \left(e^{\mathfrak{M}_0 \mathcal{N}'} \left(d_+ \left(\mathfrak{M}_t, \mathfrak{M}_0, \overleftarrow{\text{BS}}(\Gamma_t) \right) \sqrt{T-t} \right) \right)^{-1}, \\ \partial_x G(s, x, k, \sigma) &= \frac{e^x \mathcal{N}'(d_+(x, k, \sigma))}{\sigma \sqrt{T-s}} \left(1 - \frac{d_+(x, k, \sigma)}{\sigma \sqrt{T-s}} \right). \end{aligned}$$

Since $\Gamma: \mathbb{R}_+ \rightarrow \mathbb{R}$ and $\overleftarrow{\text{BS}}: \mathbb{R} \rightarrow \mathbb{R}$ are continuous, the following is uniformly bounded for all $T \leq 1$:

$$\frac{\mathcal{N}'(d_+(\mathfrak{M}_s, \mathfrak{M}_0, u_0))}{\mathcal{N}'(d_+(\mathfrak{M}_s, \mathfrak{M}_0, \overleftarrow{\text{BS}}(\Gamma_s)))} = \exp \left\{ \frac{1}{8} \left((T-s) \overleftarrow{\text{BS}}(\Gamma_s)^2 - u_0^2 \right) \right\}.$$

Therefore, by (H4) we obtain

$$\lim_{T \downarrow 0} \mathbb{E} \left[\int_0^T \overleftarrow{\text{BS}}'(\Gamma_t) | \Theta_t | \partial_x G_t dt \right] = \lim_{T \downarrow 0} \mathbb{E} \left[\int_0^T \frac{\mathcal{N}'(d_+(\mathfrak{M}_t, \mathfrak{M}_0, u_t))}{\mathcal{N}'(d_+(\mathfrak{M}_t, \mathfrak{M}_0, \overleftarrow{\text{BS}}(\Gamma_t)))} \frac{|\Theta_t|}{2u_t^2 \sqrt{T}} dt \right] = 0.$$

Since $\Gamma_0 = \mathbb{E}[\overleftarrow{\text{BS}}(\mathfrak{M}_0)]$ and $u_0 = \overleftarrow{\text{BS}}(\overleftarrow{\text{BS}}(\mathfrak{M}_0))$, we have

$$\overleftarrow{\text{BS}}(\Gamma_0) = \overleftarrow{\text{BS}}(\mathbb{E}[\overleftarrow{\text{BS}}(\mathfrak{M}_0)]) - \mathbb{E}[u_0 - u_0] = \mathbb{E}[\overleftarrow{\text{BS}}(\mathbb{E}[\overleftarrow{\text{BS}}(\mathfrak{M}_0)])] - \overleftarrow{\text{BS}}(\overleftarrow{\text{BS}}(\mathfrak{M}_0)) + \mathbb{E}[u_0]. \tag{17}$$

The Clark–Ocone formula yields

$$\overleftarrow{\text{BS}}(\mathfrak{M}_0) = \mathbb{E}[\overleftarrow{\text{BS}}(\mathfrak{M}_0)] + \sum_{i=1}^N \int_0^T \mathbb{E}_s [D_s^i \overleftarrow{\text{BS}}(\mathfrak{M}_0)] dW_s^i,$$

and by the gamma–vega–delta relation (12) we have

$$\partial_\sigma \text{BS}(0, x, x, \sigma) = \exp \left\{ x - \frac{\sigma^2}{8} T \right\} \sqrt{\frac{T}{2\pi}}, \tag{18}$$

which in turn implies

$$U_s^i := \mathbb{E}_s \left[D_s^i \widetilde{\text{BS}}(\mathfrak{M}_0) \right] = \mathbb{E}_s \left[\frac{\partial_\sigma \widetilde{\text{BS}}(\mathfrak{M}_0)}{2u_0\sqrt{T}} \int_0^T D_s^i \|\phi_r\|^2 dr \right] = \mathbb{E}_s \left[\frac{e^{\mathfrak{M}_0 - \frac{1}{8}u_0^2}}{2\sqrt{2\pi}u_0} \int_0^T D_s^i \|\phi_r\|^2 dr \right]. \tag{19}$$

Define $\Lambda_r := \mathbb{E}_r \left[\widetilde{\text{BS}}(\mathfrak{M}_0) \right]$, so that the difference we are interested in from (17) reads, after we apply the standard Itô’s formula,

$$\begin{aligned} \overleftarrow{\text{BS}}(\Gamma_0) - \mathbb{E}[u_0] &= \mathbb{E} \left[\overleftarrow{\text{BS}}(\Lambda_0) - \overleftarrow{\text{BS}}(\Lambda_T) \right] \\ &= - \sum_{i=1}^N \mathbb{E} \left[\int_0^T \overleftarrow{\text{BS}}'(\Lambda_s) U_s^i dW_s^i + \frac{1}{2} \int_0^T \overleftarrow{\text{BS}}''(\Lambda_s) (U_s^i)^2 ds \right]. \end{aligned} \tag{20}$$

The stochastic integral above has zero expectation by the same argument as used for [5, Proposition 3.1]. Moreover, (H₅) states that u_0 is dominated almost surely by $Z \in L^p$, and therefore so are Λ and

$$\overleftarrow{\text{BS}}''(\Lambda_s) = \frac{\overleftarrow{\text{BS}}(\Lambda_s)}{4e^{2\mathfrak{M}_s} \mathcal{N}' \left(d_+(\mathfrak{M}_s, \mathfrak{M}_0, \overleftarrow{\text{BS}}(\Lambda_s)) \right)^2},$$

by continuity. Plugging in the expression for U^i from (19), we apply (H₅) to conclude that the second integral of (20) tends to zero.

6.2.2. *Proof of Theorem 2: skew.* This proof follows from arguments similar to those of [4, Proposition 5.1]. We recall that $\Pi_0(k) = \text{BS}(0, \mathfrak{M}_0, k, \mathcal{I}_T(k))$. On the one hand, by the chain rule we have

$$\partial_k \Pi_0(k) = \partial_k \text{BS}(0, \mathfrak{M}_0, k, \mathcal{I}_T(k)) + \partial_\sigma \text{BS}(0, \mathfrak{M}_0, k, \mathcal{I}_T(k)) \partial_k \mathcal{I}_T(k). \tag{21}$$

On the other hand, the decomposition obtained in Proposition 5 yields

$$\partial_k \Pi_0(k) = \mathbb{E} \left[\partial_k \text{BS}(0, \mathfrak{M}_T, k, u_0) \right] + \mathbb{E} \left[\int_0^T \frac{1}{2} \partial_{xk} G(s, \mathfrak{M}_s, \mathfrak{M}_0, u_s) | \Theta_s | ds \right]. \tag{22}$$

Equating (21) and (22) gives

$$\begin{aligned} \partial_k \mathcal{I}_T(k) &= \frac{\mathbb{E} \left[\partial_k \text{BS}(0, \mathfrak{M}_T, k, u_0) \right] - \partial_k \text{BS}(0, \mathfrak{M}_0, k, \mathcal{I}_T(k))}{\partial_\sigma \text{BS}(0, \mathfrak{M}_0, k, \mathcal{I}_T(k))} \\ &\quad + \frac{\mathbb{E} \left[\int_0^T \partial_{xk} G(s, \mathfrak{M}_s, \mathfrak{M}_0, u_s) | \Theta_s | ds \right]}{2\partial_\sigma \text{BS}(0, \mathfrak{M}_0, k, \mathcal{I}_T(k))}, \end{aligned} \tag{23}$$

which in particular also holds for $k = \mathfrak{M}_0$. Performing simple algebraic manipulations and using the derivatives of the Black–Scholes function ATM as in [4, Proposition 5.1], we find the following (remember we drop the k -dependence in \mathcal{I}_T when ATM):

$$\mathbb{E}[\partial_k \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, u_0)] - \partial_k \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T) = \frac{1}{2} \mathbb{E} \left[\int_0^T \frac{1}{2} \partial_x G(s, \mathfrak{M}_s, \mathfrak{M}_0, u_s) |\Theta_s| ds \right].$$

By (23), this in turn yields

$$\partial_k \mathcal{I}_T = \frac{\mathbb{E} \left[\int_0^T L(s, \mathfrak{M}_s, \mathfrak{M}_0, u_s) |\Theta_s| ds \right]}{\partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)}, \tag{24}$$

where $L := (\frac{1}{2} + \partial_k) \frac{1}{2} \partial_x G$. We write $L_s := L(s, \mathfrak{M}_s, \mathfrak{M}_0, u_s)$ for simplicity and apply Lemma 2 to $L_s \int_s^T |\Theta_r| dr$, which yields

$$\begin{aligned} \mathbb{E} \left[\int_0^T L_s |\Theta_s| ds \right] &= \mathbb{E} \left[L_0 \int_0^T |\Theta_s| ds \right] + \mathbb{E} \left[\int_0^T (\partial_x^3 - \partial_x^2) L_s |\Theta_s| \left(\int_s^T |\Theta_r| dr \right) ds \right] \\ &\quad + \mathbb{E} \left[\int_0^T \partial_x L_s \sum_{j=1}^N \left\{ \phi_s^j D_s^j \left(\int_s^T |\Theta_r| dr \right) \right\} ds \right] = :R_1 + R_2 + R_3. \end{aligned}$$

We combine (18) with the bound $\partial_k \partial_x^n G(t, x, k, \sigma) \leq C(\sigma \sqrt{T-t})^{-n-2}$ from Corollary 2 to obtain

$$\begin{aligned} \frac{R_2}{\partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)} &\leq \frac{C}{\sqrt{T}} \mathbb{E} \left[\int_0^T \frac{|\Theta_s|}{u_s^6} \left(\int_s^T |\Theta_r| dr \right) ds \right], \\ \frac{R_3}{\partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)} &\leq \frac{C}{\sqrt{T}} \mathbb{E} \left[\int_0^T \frac{1}{u_s^4} \sum_{j=1}^N \left\{ \phi_s^j D_s^j \left(\int_s^T |\Theta_r| dr \right) \right\} ds \right], \end{aligned}$$

and both converge to zero by (H_6^λ) . We are left with R_1 . From Section 6.6, we have

$$L(0, x, x, u) = \frac{\exp\{x - \frac{u^2}{8} T\}}{u \sqrt{2\pi T}} \left(\frac{1}{4} + \frac{1}{2u^2 T} \right),$$

and therefore by (18),

$$\frac{L_0}{\partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)} = \left(\frac{1}{4} + \frac{1}{2u_0^2 T} \right) \frac{1}{u_0 T} \exp \left\{ -\frac{T}{8} (u_0^2 - \mathcal{I}_T^2) \right\}.$$

This yields

$$\frac{R_1}{T^\lambda \partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)} = \mathbb{E} \left[\left(\frac{u_0^2}{2} + 1 \right) \exp \left\{ -\frac{T}{8} (u_0^2 - \mathcal{I}_T^2) \right\} \mathfrak{K}_T \right],$$

where $\mathfrak{K}_T := \frac{\int_0^T |\Theta_s| ds}{2T^{\frac{1}{2} + \lambda} u_0^3}$. Furthermore,

$$\sup_{\omega \in \Omega} \left| \exp \left\{ -\frac{T}{8} (u_0^2 - \mathcal{I}_T^2) \right\} - 1 \right| = \sup_{\omega \in \Omega} \left| \exp \left\{ \frac{1}{8} (T\mathcal{I}_T^2 - u_0^2) \right\} - 1 \right|$$

not only is finite but converges to zero as T goes to zero. Hence,

$$\lim_{T \downarrow 0} \mathbb{E} \left[\left(\frac{u_0^2}{2} + 1 \right) \exp \left\{ -\frac{T}{8} (u_0^2 - \mathcal{I}_T^2) \right\} \mathfrak{R}_T \right] = \lim_{T \downarrow 0} \mathbb{E} \left[\left(\frac{u_0^2}{2} + 1 \right) \mathfrak{R}_T \right].$$

We can finally conclude by (H_7^λ) that

$$\lim_{T \downarrow 0} \frac{R_1}{T^\lambda \partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)} = \lim_{T \downarrow 0} \mathbb{E}[\mathfrak{R}_T],$$

which has a finite limit.

6.2.3. *Proof of Theorem 3: curvature. Step 1.* Let us start by simply taking a second derivative with respect to k (we write $\text{BS}(\mathfrak{M}_0, \mathcal{I}_T(k))$ as short for $\text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T(k))$):

$$\begin{aligned} \partial_k \left(\partial_\sigma \text{BS}(\mathfrak{M}_0, \mathcal{I}_T(k)) \partial_k \mathcal{I}_T(k) \right) &= \partial_\sigma \text{BS}(\mathfrak{M}_0, \mathcal{I}_T(k)) \partial_k^2 \mathcal{I}_T(k) \\ &\quad + \left[\partial_{k\sigma} \text{BS}(\mathfrak{M}_0, \mathcal{I}_T(k)) + \partial_\sigma^2 \text{BS}(\mathfrak{M}_0, \mathcal{I}_T(k)) \partial_k \mathcal{I}_T(k) \right] \partial_k \mathcal{I}_T(k). \end{aligned}$$

Taking the derivative with respect to k in (24) and equating with the above formula yields

$$\begin{aligned} \partial_k^2 \mathcal{I}_T(k) \partial_\sigma \text{BS}(\mathfrak{M}_0, \mathcal{I}_T) &= -\partial_\sigma^2 \text{BS}(\mathfrak{M}_0, \mathcal{I}_T(k)) \partial_k \mathcal{I}_T(k)^2 - \partial_{k\sigma} \text{BS}(\mathfrak{M}_0, \mathcal{I}_T(k)) \partial_k \mathcal{I}_T(k) \\ &\quad + \mathbb{E} \left[\int_0^T \partial_k L(s, \mathfrak{M}_s, \mathfrak{M}_0, u_s) | \Theta_s | ds \right] = : T_1 + T_2 + T_3. \end{aligned}$$

A similar expression is presented in [3], and we notice that T_1 and T_2 in the expression above, after being multiplied by $T^{-\lambda}$, are identical to those from [3, Equation (25)] and can therefore be dealt with in the same way. Step 1 shows that $T^{-\lambda} T_1$ tends to zero as $T \downarrow 0$, and Step 2 yields $T_2 = -\frac{1}{2} \partial_k \mathcal{I}_T(k)$.

Step 2. Recall that $L = \frac{1}{2} \left(\frac{1}{2} + \partial_k \right) \partial_x G$. We need the anticipative Itô's formula (Lemma 2) twice on T_3 . Indeed, even though the bound on $\partial_x^n G$ worsens as n increases, it is more than compensated for by the additional integrations. The terms with more integrals (i.e. more regularity) tend to zero as T goes to zero, by (H_8^λ) , and we compute the others in closed form. For clarity we write $L_s = L(s, \mathfrak{M}_s, \mathfrak{M}_0, u_s)$ for all $s \geq 0$. By a first application of Lemma 2 on $\partial_k L_s \int_s^T | \Theta_s | ds$ we obtain

$$\begin{aligned} T_3 &= \mathbb{E} \left[\partial_k L_0 \int_0^T | \Theta_s | ds \right] + \mathbb{E} \left[\int_0^T (\partial_x^3 - \partial_x^2) \partial_k L_s | \Theta_s | \left(\int_s^T | \Theta_r | dr \right) ds \right] \\ &\quad + \mathbb{E} \left[\int_0^T \partial_{xk} L_s \sum_{j=1}^N \left\{ \phi_s^j D_s^j \left(\int_s^T | \Theta_r | dr \right) \right\} ds \right] = : S_1 + S_2 + S_3. \end{aligned}$$

To deal with S_2 , we apply Lemma 2 again on $(\partial_x^3 - \partial_x^2) \partial_k L_s \int_s^T | \Theta_r | \left(\int_r^T | \Theta_y | dy \right) dr = : H_s Z_s$, which yields

$$S_2 = \mathbb{E} \left[H_0 Z_0 + \int_0^T (\partial_x^3 - \partial_x^2) H_s | \Theta_s | Z_s ds + \int_0^T \partial_x H_s \sum_{j=1}^N (\phi_s^j D_s^j Z_s) ds \right] = : S_2^a + S_2^b + S_2^c.$$

We will deal with these terms in the last step. For S_3 , we apply Lemma 2 once more to

$$\partial_{xk} L_s \int_s^T \sum_{j=1}^N \left\{ \phi_r^j D_r^j \left(\int_r^T | \Theta_y | dy \right) \right\} dr = : \tilde{H}_s \tilde{Z}_s,$$

and obtain

$$S_3 = \mathbb{E} \left[\tilde{H}_0 \tilde{Z}_0 + \int_0^T (\partial_x^3 - \partial_x^2) \tilde{H}_s | \Theta_s | \tilde{Z}_s ds + \int_0^T \partial_x \tilde{H}_s \sum_{j=1}^N (\phi_s^j D_s^j \tilde{Z}_s) ds \right] = : S_3^a + S_3^b + S_3^c.$$

Step 3. We now evaluate the derivative at $k = \mathfrak{M}_0$ and drop the k -dependence. To summarise,

$$\partial_k^2 \mathcal{I}_T = \frac{T_1 + T_2 + S_1 + S_2^a + S_2^b + S_2^c + S_3^a + S_3^b + S_3^c}{\partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)},$$

where

$$\begin{aligned} S_1 &= \mathbb{E} \left[\partial_k L_0 \int_0^T | \Theta_s | ds \right], \\ S_2^a &= \mathbb{E} \left[H_0 \int_0^T | \Theta_r | \left(\int_r^T | \Theta_y | dy \right) dr \right], \\ S_3^a &= \mathbb{E} \left[\tilde{H}_0 \int_0^T \sum_{j=1}^N \left\{ \phi_r^j D_r^j \left(\int_r^T | \Theta_y | dy \right) \right\} dr \right], \\ S_2^b &= \mathbb{E} \left[\int_0^T (\partial_x^3 - \partial_x^2) H_s | \Theta_s | \left(\int_s^T | \Theta_r | \left(\int_r^T | \Theta_y | dy \right) dr \right) ds \right], \\ S_2^c &= \mathbb{E} \left[\int_0^T \partial_x H_s \sum_{j=1}^N \left(\phi_s^j D_s^j \left(\int_s^T | \Theta_r | \left(\int_r^T | \Theta_y | dy \right) dr \right) ds \right) ds \right], \\ S_3^b &= \mathbb{E} \left[\int_0^T (\partial_x^3 - \partial_x^2) \tilde{H}_s | \Theta_s | \int_s^T \sum_{j=1}^N \left\{ \phi_r^j D_r^j \left(\int_r^T | \Theta_y | dy \right) \right\} dr ds \right], \\ S_3^c &= \mathbb{E} \left[\int_0^T \partial_x \tilde{H}_s \sum_{k=1}^N \left\{ \phi_s^k D_s^k \left(\int_s^T \sum_{j=1}^N \left\{ \phi_r^j D_r^j \left(\int_r^T | \Theta_y | dy \right) \right\} dr \right) \right\} ds \right]. \end{aligned}$$

We recall once again the bound $\partial_x^n G(t, x, k, \sigma) \leq C(\sigma \sqrt{T-t})^{-n-1}$ as $T-t$ goes to zero. We observe that H and \tilde{H} consist of derivatives of G up to the sixth and the fourth order, respectively; therefore $S_2^b, S_2^c, S_3^b, S_3^c$ tend to zero by (H_6') . In order to deal with $S_1, S_2^a,$ and S_3^a , we use the explicit partial derivatives from Section 6.6 and (18); as in the proof of Theorem 2, (H_6') implies that only the higher derivatives of u_0 remain in the limit:

$$\begin{aligned}
 \lim_{T \downarrow 0} \frac{S_1}{T^\lambda \partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)} &= \lim_{T \downarrow 0} \frac{1}{T^\lambda} \frac{\mathbb{E} \left[\partial_k L_0 \int_0^T |\Theta_s| ds \right]}{\partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)} \\
 &= \lim_{T \downarrow 0} \frac{1}{T^\lambda} \mathbb{E} \left[\frac{1}{u_0 T} \left(\frac{1}{8} + \frac{1}{2u_0^2 T} \right) \int_0^T |\Theta_s| ds \right] \\
 &= \lim_{T \downarrow 0} \mathbb{E} \left[\frac{1}{2u_0^3 T^{\frac{1}{2} + \lambda}} \int_0^T |\Theta_s| ds \right] = \lim_{T \downarrow 0} \frac{S_T}{T^\lambda}, \\
 \lim_{T \downarrow 0} \frac{S_2^a}{T^\lambda \partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)} &= \lim_{T \downarrow 0} \frac{1}{T^\lambda} \mathbb{E} \left[\frac{Z_0}{u_0 \sqrt{T}} \left(-\frac{15}{2u_0^6} - \frac{3}{2u_0^4} - \frac{5}{32u_0^2} - \frac{1}{64} \right) \right] \\
 &= \lim_{T \downarrow 0} \mathbb{E} \left[\frac{-15}{2u_0^7 T^{\frac{1}{2} + \lambda}} \int_0^T |\Theta_r| \left(\int_r^T |\Theta_y| dy \right) dr \right], \\
 \lim_{T \downarrow 0} \frac{S_3^a}{T^\lambda \partial_\sigma \text{BS}(0, \mathfrak{M}_0, \mathfrak{M}_0, \mathcal{I}_T)} &= \lim_{T \downarrow 0} \frac{1}{T^\lambda} \mathbb{E} \left[\frac{\tilde{Z}_0}{u_0 \sqrt{T}} \left(\frac{3}{2u_0^4} + \frac{3}{8u_0^2} + \frac{1}{16} \right) \right] \\
 &= \lim_{T \downarrow 0} \frac{3}{2} \mathbb{E} \left[\frac{1}{u_0^5 T^{\frac{1}{2} + \lambda}} \int_0^T \sum_{j=1}^N \left\{ \phi_r^j D_r^j \left(\int_r^T |\Theta_y| dy \right) \right\} dr \right].
 \end{aligned}$$

Hence, to conclude, the claim follows from

$$\lim_{T \downarrow 0} \frac{C_T}{T^\lambda} = \lim_{T \downarrow 0} \frac{T_2 + S_1 + S_2^a + S_3^a}{T^\lambda \partial_\sigma \text{BS}(\mathfrak{M}_0, \mathcal{I}_T)}.$$

6.3. Proof of Proposition 1: VIX asymptotics

In this section, we will repeatedly interchange the Malliavin derivative and conditional expectation, which is justified by [42, Proposition 1.2.8].

Proposition 7. *In the case where $A = \text{VIX}$, the conditions in (\bar{C}) imply the assumptions $(\bar{H}^{\lambda, \gamma})$ for any $\lambda \in (-\frac{1}{2}, 0]$ and $\gamma \in (-1, 3H - \frac{1}{2}]$.*

Proof. We write $a \lesssim b$ when there exists $X \in L^p$ such that $a \leq Xb$ almost surely, and $a \approx b$ if $a \lesssim b$ and $b \lesssim a$. The assumption (H_1) is given by the first item of (C_2) , and (H_2) corresponds to (C_1) . Since $1/M$ is dominated, so is $1/\text{VIX}$. We then have, for $i = 1, 2$ and by Cauchy–Schwarz,

$$m_y^i = \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} D_y^i v_r dr}{2\Delta \text{VIX}_T} \right] \lesssim \int_T^{T+\Delta} (r - y)^{H_-} dr = \frac{(T + \Delta - y)^{H_+} - (T - y)^{H_+}}{H_+}.$$

If $H < \frac{1}{2}$, then the incremental function $x \mapsto (x + \Delta)^{H_+} - x^{H_+}$ is decreasing by concavity. For $j = 1, 2$ and $t \leq s$, this implies by domination of $1/M$ that $\phi^i \approx m^i$ is also dominated and

$$\begin{aligned}
 D_s^j \phi_y^i &= \frac{D_s^j m_y^i}{M_y} - \frac{m_y^i D_s^j M_y}{M_y^2} \\
 &\lesssim \int_T^{T+\Delta} (r-y)^{H-} (r-s)^{H-} dr + \int_T^{T+\Delta} (r-y)^{H-} dr \int_T^{T+\Delta} (r-s)^{H-} dr \\
 &\leq \frac{\Delta^{2H}}{2H} + \frac{\Delta^{H+1}}{H_+^2}.
 \end{aligned}
 \tag{25}$$

Combining these two estimates, we obtain

$$\Theta_s^j = 2\phi_s^j \int_s^T \left(\sum_{i=1}^N \phi_y^i D_s^j \phi_y^i \right) dy \lesssim T - s.$$

It is clear by now that indices and sums do not influence the estimates, so we informally drop them for more clarity and continue with the higher derivatives:

$$D_t \Theta_s = D_t \phi_s \int_s^T \phi_r D_s \phi_r dr + \phi_s \int_s^T D_t \phi_r D_s \phi_r dr + \phi_s \int_s^T \phi_r D_t D_s \phi_r dr,$$

where the first and second terms behave like $T - s$. For $t \leq s \leq y \leq T$, we deduce from (25) that $D_t D_s \phi_y$ consists of five terms, of which four behave like $(T - s)$, and only one features three derivatives:

$$\begin{aligned}
 D_t D_s m_y &\lesssim \int_T^{T+\Delta} (r-s)^{H-} (r-t)^{H-} (r-y)^{H-} dr \leq \int_T^{T+\Delta} (r-y)^{3H-\frac{3}{2}} dy \\
 &\approx (T + \Delta - y)^{3H-\frac{1}{2}} - (T - y)^{3H-\frac{1}{2}}.
 \end{aligned}$$

If $H \geq \frac{1}{6}$, then concavity implies $D_t \Theta_s \lesssim (T - s)$. Otherwise, if $H < \frac{1}{6}$, then

$$D_t \Theta_s \lesssim (T - s) + \left[(T + \Delta - s)^{3H+\frac{1}{2}} - \Delta^{3H+\frac{1}{2}} \right] + (T - s)^{3H+\frac{1}{2}} \leq (T - s) + 2(T - s)^{3H+\frac{1}{2}}.$$

In the second derivative of Θ , the first and second terms behave like $(T - s)$ and $D_t \Theta_s \lesssim (T - s) + (T - s)^{(3H+\frac{1}{2}) \wedge 1}$ respectively; hence we focus on $\int_s^T D_w D_t D_s \phi_y dy$, where the new term is

$$D_w D_t D_s m_y \approx \int_T^{T+\Delta} (r-w)^{H-} (r-s)^{H-} (r-t)^{H-} (r-y)^{H-} dr \lesssim (T + \Delta - y)^{4H-1} - (T - y)^{4H-1}.$$

If $H \geq \frac{1}{4}$, then $D_t \Theta_s \lesssim (T - s)$ by concavity. Otherwise, when $H < \frac{1}{4}$,

$$\begin{aligned}
 D_w D_t \Theta_s &\lesssim (T - s) + (T - s)^{(3H+\frac{1}{2}) \wedge 1} + \left[(T + \Delta - s)^{4H} - \Delta^{4H} \right] + (T - s)^{4H} \\
 &\leq (T - s) + (T - s)^{(3H+\frac{1}{2}) \wedge 1} + 2(T - s)^{4H},
 \end{aligned}$$

where the last inequality holds by yet again the same concavity argument. □

This yields a rule for checking that the quantities in our assumptions indeed converge. We summarise the above estimates in the case $H \leq \frac{1}{2}$: there exists $Z \in L^p$ such that for $s \leq T$ and T small enough,

$$\Theta_s \leq Z(T - s), \quad D\Theta_s \leq Z(T - s)^{(3H + \frac{1}{2}) \wedge 1}, \quad D D\Theta_s \leq Z(T - s)^{(4H) \wedge 1}$$

hold almost surely. Thanks to the Cauchy–Schwarz inequality we can disentangle the numerators (integrals and derivatives of Θ) and denominators (powers of u) of the assumptions, which are both uniformly bounded in L^p . We can easily deduce that (H_3) , (H_4) , (H_5) , (H_6^λ) , (H_8^γ) are satisfied (convergence to zero). In (H_7^λ) , $\mathbb{E}[\mathfrak{R}_T]$ behaves as $T^{-\lambda}$, so it converges for any $\lambda \in (-\frac{1}{2}, 0]$, and the uniform L^2 bound is satisfied thanks to (C_3) . Moreover, in the limit the first term in (H_9^γ) behaves as $T^{-\gamma}$ and the second behaves as $T^{3H - \frac{1}{2} - \gamma}$; therefore both assumptions are satisfied for any $\lambda \in (-\frac{1}{2}, 0]$ and $\gamma \in (-1, 3H - \frac{1}{2}]$. Similarly, (C_3) ensures the uniform L^2 bounds.

6.3.1. *Convergence lemmas.* We require some preliminaries before we dive into the computations. We present three versions of integral convergence tailored to our purposes, which are essential for computing the limits in Theorems 1, 2, and 3. The conditions they require hold thanks to the continuity of (C_4) . Recall the local Taylor theorem: if a function $g(\cdot)$ is continuous on $[0, \delta]$ for some $\delta > 0$, then there exists a continuous function $\varepsilon(\cdot)$ on $[0, \delta]$ with $\lim_{x \downarrow 0} \varepsilon(x) = 0$ such that $g(x) = g(0) + \varepsilon(x)$ for any $x \in [0, \delta]$.

Lemma 3. *If $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ is such that $f(T, \cdot)$ is continuous on $[0, \delta_0]$ for some $\delta_0 > 0$ and $\lim_{T \downarrow 0} f(T, 0) = f(0, 0)$, then*

$$\lim_{T \downarrow 0} \frac{1}{T} \int_0^T f(T, y) \, dy = f(0, 0). \tag{26}$$

Proof. For $T < \delta_0$, we can write

$$\frac{1}{T} \int_0^T f(T, y) \, dy = \frac{1}{T} \int_0^T [f(T, 0) + \varepsilon_0(y)] \, dy = f(T, 0) + \frac{1}{T} \int_0^T \varepsilon_0(y) \, dy,$$

where the function ε_0 is continuous on $[0, \delta_0]$ and converges to zero at the origin. Hence, for any $\eta_0 > 0$, there exists $\tilde{\delta}_0 > 0$ such that, for any $y \leq \tilde{\delta}_0$, $|\varepsilon_0(y)| < \eta_0$. For all $T < \tilde{\delta}_0 \wedge \delta_0$,

$$\left| \frac{1}{T} \int_0^T \varepsilon_0(y) \, dy \right| \leq \eta_0.$$

Since η_0 can be taken as small as desired, the fact that $\lim_{T \downarrow 0} f(T, 0) = f(0, 0)$ concludes the proof. □

Lemma 4. *Let $f : \mathbb{R}_+^3 \rightarrow \mathbb{R}$ be such that, for each $y \leq T$, $f(T, y, \cdot)$ is continuous on $[0, \delta_0]$ with $\delta_0 > 0$, $f(T, \cdot, 0)$ is continuous on $[0, \delta_1]$ with $\delta_1 > 0$, and $\lim_{T \downarrow 0} f(T, 0, 0) = f(0, 0, 0)$. Then*

$$\lim_{T \downarrow 0} \frac{1}{T^2} \int_0^T \int_0^y f(T, y, s) \, ds \, dy = \frac{f(0, 0, 0)}{2}. \tag{27}$$

Proof. For $T < \delta_0 \wedge \delta_1$, we can write

$$\begin{aligned} \frac{1}{T^2} \int_0^T \left\{ \int_0^y f(T, y, s) ds \right\} dy &= \frac{1}{T^2} \int_0^T \left\{ \int_0^y [f(T, y, 0) + \varepsilon_0(s)] ds \right\} dy \\ &= \frac{1}{T^2} \int_0^T \left\{ f(T, y, 0)y + \int_0^y \varepsilon_0(s) ds \right\} dy \\ &= \frac{1}{T^2} \int_0^T \left\{ (f(T, 0, 0) + \varepsilon_1(y))y + \int_0^y \varepsilon_0(s) ds \right\} dy \\ &= \frac{f(T, 0, 0)}{2} + \frac{1}{T^2} \int_0^T \left\{ \varepsilon_1(y)y + \int_0^y \varepsilon_0(s) ds \right\} dy, \end{aligned}$$

where $\varepsilon_1(\cdot)$ is continuous on $[0, \delta_1]$ and $\varepsilon_0(\cdot)$ is continuous on $[0, \delta_0]$, and both are null at the origin. For any $\eta_1 > 0$, there exists $\tilde{\delta}_1 > 0$ such that, for any $y \in [0, \tilde{\delta}_1]$, we have $|\varepsilon_1(y)| < \eta_1$. Therefore, for the first integral, we have, for $T < \tilde{\delta}_1 \wedge \delta_0 \wedge \delta_1$,

$$\left| \frac{1}{T^2} \int_0^T \varepsilon_1(y)y dy \right| \leq \frac{1}{T^2} \int_0^T |\varepsilon_1(y)| y dy \leq \frac{1}{T^2} \int_0^T \eta_1 y dy \leq \frac{\eta_1}{2}.$$

Likewise, since $\varepsilon_0(\cdot)$ tends to zero at the origin, for any $\eta_0 > 0$ there exists $\tilde{\delta}_0 > 0$ such that, for any $y \in [0, \tilde{\delta}_0]$, we have $|\varepsilon_0(y)| < \eta_0$. Therefore, for the second integral, we have, for $T < \tilde{\delta}_0 \wedge \delta_0 \wedge \delta_1$,

$$\left| \frac{1}{T^2} \int_0^T \int_0^y \varepsilon_0(s) ds dy \right| \leq \frac{1}{T^2} \int_0^T \int_0^y |\varepsilon_0(s)| ds dy \leq \frac{1}{T^2} \int_0^T \int_0^y \eta_0 ds dy \leq \frac{\eta_0}{2}.$$

Since η_1 and η_0 can be taken as small as desired, taking the limit of $f(T, 0, 0)$ as T goes to zero concludes the proof. \square

Lemma 5. Let $f : \mathbb{R}_+^4 \rightarrow \mathbb{R}$ be such that, for all $0 \leq s \leq y \leq T$, the functions $f(T, y, s, \cdot)$, $f(T, y, \cdot, 0)$, $f(T, \cdot, 0, 0)$ are continuous on $[0, \delta_0]$, $[0, \delta_1]$, $[0, \delta_2]$, respectively, for some $\delta_0, \delta_1, \delta_2 > 0$, and $\lim_{T \downarrow 0} f(T, 0, 0, 0) = f(0, 0, 0, 0)$. Then the following limit holds:

$$\lim_{T \downarrow 0} \frac{1}{T^3} \int_0^T \int_0^y \int_0^s f(T, y, s, t) dt ds dy = \frac{f(0, 0, 0, 0)}{6}. \quad (28)$$

Proof. For $T < \delta_0 \wedge \delta_1 \wedge \delta_2$, we can write

$$\begin{aligned} \frac{1}{T^3} \int_0^T \left\{ \int_0^y \left(\int_0^s f(T, y, s, t) dt \right) ds \right\} dy \\ = \frac{1}{T^3} \int_0^T \left\{ \int_0^y \left(\int_0^s [f(T, y, s, 0) + \varepsilon_0(t)] dt \right) ds \right\} dy \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{T^3} \int_0^T \left\{ \int_0^y \left(f(T, y, s, 0)s + \int_0^s \varepsilon_0(t)dt \right) ds \right\} dy \\
 &= \frac{1}{T^3} \int_0^T \left\{ \int_0^y \left([f(T, y, 0, 0) + \varepsilon_1(s)]s + \int_0^s \varepsilon_0(t)dt \right) ds \right\} dy \\
 &= \frac{1}{T^3} \int_0^T \left\{ f(T, y, 0, 0)\frac{y^2}{2} + \int_0^y \left(\varepsilon_1(s)s + \int_0^s \varepsilon_0(t)dt \right) ds \right\} dy \\
 &= \frac{1}{T^3} \int_0^T \left\{ [f(T, 0, 0, 0) + \varepsilon_2(y)]\frac{y^2}{2} + \int_0^y \left(\varepsilon_1(s)s + \int_0^s \varepsilon_0(t)dt \right) ds \right\} dy \\
 &= \frac{f(T, 0, 0, 0)}{6} + \frac{1}{T^3} \int_0^T \left\{ \varepsilon_2(y)\frac{y^2}{2} + \int_0^y \left(\varepsilon_1(s)s + \int_0^s \varepsilon_0(t)dt \right) ds \right\} dy,
 \end{aligned}$$

where the function ε_2 is continuous on $[0, \delta_2]$, the function ε_1 is continuous on $[0, \delta_1]$, and the function ε_0 is continuous on $[0, \delta_0]$, all converging to zero at the origin. By the same argument as in the previous proof, for any $\eta_0, \eta_1, \eta_2 > 0$, there exists $\tilde{\delta} > 0$ such that for all $T \leq \tilde{\delta}$, we have $|\varepsilon_0(T)| \leq \eta_0, |\varepsilon_1(T)| \leq \eta_1$, and $|\varepsilon_2(T)| \leq \eta_2$. This implies

$$\left| \frac{1}{T^3} \int_0^T \left\{ \varepsilon_2(y)\frac{y^2}{2} + \int_0^y \left(\varepsilon_1(s)s + \int_0^s \varepsilon_0(t)dt \right) ds \right\} dy \right| \leq \frac{\eta_2 + \eta_1 + \eta_0}{6}.$$

Since η_2, η_1 and η_0 can be taken as small as desired, taking the limit of $f(T,0,0,0)$ as T goes to zero concludes the proof. □

To apply these lemmas, we will use a modified version of the martingale convergence theorem, which holds in our setting thanks to domination provided by (C₁) and (C₂) and the continuity of (C₄).

Lemma 6. *Let $(X_t)_{t \geq 0}$ be almost surely continuous in a neighbourhood of zero, with $\sup_{t \leq 1} |X_t| \leq Z \in L^1$. Then the conditional expectation process $(\mathbb{E}_t[X_t])_{t \geq 0}$ is also almost surely continuous in a neighbourhood of zero. In particular,*

$$\lim_{t \downarrow 0} \mathbb{E}_t[X_t] = \mathbb{E}[X_0].$$

Remark 8. The process $(X_t)_{t \geq 0}$ is not necessarily adapted.

Proof. All the limits are taken in the almost sure sense. Let $\delta > 0$ be such that X is continuous on $[0, \delta]$, and fix $t < \delta$. We set a sequence $\{t_n\}_{n \in \mathbb{N}}$ on $[0, \delta]$ which converges to t as n goes infinity. Assume first that $\{t_n\}_{n \in \mathbb{N}}$ is a monotone sequence. Since \mathcal{F}_{t_n} tends monotonically to \mathcal{F}_t and X is dominated, the classical martingale convergence theorem (MCT) asserts that $\lim_{n \uparrow \infty} \mathbb{E}_{t_n}[X_t] = \mathbb{E}_t[X_t]$. For fixed $n \in \mathbb{N}$ and any $q \geq |t_n - t|$,

$$|X_{t_n} - X_t| \leq \sup_{|p-t| \leq q} |X_p - X_t|. \tag{29}$$

Let us fix $\varepsilon > 0$. By the MCT, there exists $n_0 \in \mathbb{N}$ such that, if $n \geq n_0$, then

$$\left| \mathbb{E}_{t_n} \left[\sup_{|p-t| \leq q} |X_p - X_t| \right] - \mathbb{E}_t \left[\sup_{|p-t| \leq q} |X_p - X_t| \right] \right| < \varepsilon,$$

and by dominated convergence there exists $\delta' > 0$ with $\mathbb{E}_t \left[\sup_{|p-t| \leq \delta'} |X_p - X_t| \right] < \varepsilon$. There exists $n_1 \in \mathbb{N}$ such that $|t_n - t| \leq \delta'$ for all $n \geq n_1$; thus if $n \geq n_0 \vee n_1$, then (29) yields

$\mathbb{E}_{t_n}[|X_{t_n} - X_t|] < 2\varepsilon$ and

$$\lim_{n \uparrow \infty} \mathbb{E}_{t_n}[X_{t_n}] = \mathbb{E}_t[X_t]. \tag{30}$$

Now we consider the general case where $\{t_n\}_{n \in \mathbb{N}}$ is not monotone. From every subsequence of $\{t_n\}_{n \in \mathbb{N}}$, one can extract a further subsequence which is monotone. Let us call this subsequence $\{t_{n_k}\}_{k \in \mathbb{N}}$. Therefore, (30) holds with t_{n_k} instead of t_n . Since every subsequence of $(\mathbb{E}_{t_n}[X_{t_n}])_{n \in \mathbb{N}}$ has a further subsequence that converges to the same limit, the original sequence also converges to this limit. \square

For convenience, we use the following definition.

Definition 2. Let $k, n \in \mathbb{N}$ with $k \leq n$. For a function $f : \mathbb{R}_+^n \rightarrow \mathbb{R}$, we define

$$\lim_{0 \leq x_1 \leq x_2 \leq \dots \leq x_k \downarrow 0} f(x_1, \dots, x_n) := \lim_{x_k \downarrow 0} \dots \lim_{x_2 \downarrow 0} \lim_{x_1 \downarrow 0} f(x_1, \dots, x_n).$$

Notice that the right-hand sides of (26), (27), and (28) correspond to

$$\lim_{y \leq T \downarrow 0} f(T, y), \quad \frac{1}{2} \lim_{s \leq y \leq T \downarrow 0} f(T, y, s), \quad \text{and} \quad \frac{1}{6} \lim_{t \leq s \leq y \leq T \downarrow 0} f(T, y, s, t),$$

respectively.

6.3.2. *Proof of Proposition 1.* Let us recall some important quantities:

$$\begin{aligned} M_y &= \mathbb{E}_y[\text{VIX}_T] = \mathbb{E}_y \left[\sqrt{\frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}_T v_r \, dr} \right], \\ m_y^i &= \mathbb{E}_y[\text{D}_y^i M_y] = \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} \text{D}_y^i \mathbb{E}_T v_r \, dr}{2\Delta \text{VIX}_T} \right] = \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} \text{D}_y^i v_r \, dr}{2\Delta \text{VIX}_T} \right], \\ \phi_y^i &= \frac{m_y^i}{M_y} = \frac{\mathbb{E}_y \left[\left(\int_T^{T+\Delta} \text{D}_y^i v_r \, dr \right) / (2\Delta \text{VIX}_T) \right]}{\mathbb{E}_y[\text{VIX}_T]}. \end{aligned} \tag{31}$$

We also recall that J_i and G_{ij} , $i, j \in \llbracket 1, N \rrbracket$, were defined in (9). In this proof we will define $f(0) := \lim_{x \downarrow 0} f(x)$, for every $f : \mathbb{R}_+ \rightarrow \mathbb{R}$, as soon as the limit exists and even if f is not actually continuous around zero. In this way we make it continuous, which allows us to apply the convergence lemmas.

Level. By (C₁) and the MCT, $\lim_{y \downarrow 0} \mathbb{E}_y[\text{VIX}_T] = \mathbb{E}[\text{VIX}_T]$ and $(M_y)_{y \geq 0}$ is continuous around zero, almost surely. By (C₄) and the dominated convergence theorem (DCT), we have $\lim_{y \downarrow 0} \int_T^{T+\Delta} \text{D}_y^i v_r \, dr = \int_T^{T+\Delta} \text{D}_0^i v_r \, dr$ and $(\int_T^{T+\Delta} \text{D}_y^i v_r \, dr)_{y \geq 0}$ is continuous around zero, almost surely. Let $i \in \llbracket 1, N \rrbracket$; from (C₁) and (C₂) we also obtain that almost surely

$$\frac{1}{\text{VIX}_T} \int_T^{T+\Delta} \text{D}_y^i v_r \, dr \leq X^2 \left\{ (T + \Delta - y)^{H_+} - (T - y)^{H_+} \right\},$$

for some $X \in L^2$. Therefore it is dominated, and by Lemma 6, almost surely m_y^i is continuous at zero and

$$\lim_{y \downarrow 0} m_y^i = \mathbb{E} \left[\frac{\int_T^{T+\Delta} \text{D}_0^i v_r \, dr}{2\Delta \text{VIX}_T} \right].$$

Since $M_y > 0$ for all $y \leq T$, ϕ^i is also continuous at zero and $\lim_{y \leq T \downarrow 0} \phi_y^i = J_i / (2\Delta \text{VIX}_0^2)$. By virtue of Theorem 1 and Lemma 3, we obtain

$$\lim_{T \downarrow 0} \mathcal{I}_T = \lim_{T \downarrow 0} \mathbb{E}[u_0] = \lim_{y \leq T \downarrow 0} \|\phi_y\| = \frac{\|J\|}{2\Delta \text{VIX}_0^2}.$$

Skew. To obtain the skew limit we need to compute a few Malliavin derivatives. For all $i, j \in \llbracket 1, N \rrbracket$,

$$\begin{aligned} D_s^j m_y^i &= \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} D_s^j D_y^i v_r dr \text{VIX}_T - \int_T^{T+\Delta} D_y^i v_r dr D_s^j \text{VIX}_T}{2\Delta \text{VIX}_T^2} \right] \\ &= \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} D_s^j D_y^i v_r dr}{2\Delta \text{VIX}_T} - \frac{\int_T^{T+\Delta} D_y^i v_r dr \int_T^{T+\Delta} D_s^j v_r dr}{4\Delta^2 \text{VIX}_T^3} \right], \end{aligned}$$

which yields

$$\begin{aligned} D_s^j \phi_y^i &= \frac{D_s^j m_y^i}{M_y} - \frac{m_y^i D_s^j M_y}{M_y^2} \\ &= \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} D_s^j D_y^i v_r dr}{2\Delta \text{VIX}_T M_y} - \frac{\int_T^{T+\Delta} D_y^i v_r dr \int_T^{T+\Delta} D_s^j v_r dr}{4\Delta^2 \text{VIX}_T^3 M_y} - \frac{m_y^i \int_T^{T+\Delta} D_s^j v_r dr}{2\Delta \text{VIX}_T M_y^2} \right] \\ &=: \mathbb{E}_y \left[A_T^{ij}(y, s) + B_T^{ij}(y, s) + C_T^{ij}(y, s) \right]. \end{aligned}$$

Based on (C₁), (C₂), and (C₄), for each $T \geq 0$, A_T^{ij} , B_T^{ij} , and C_T^{ij} are dominated and almost surely continuous in both arguments. For each $s \geq 0$, Lemma 6 and the DCT yield, almost surely, that $(D_s^j \phi_y^i)_{y \geq 0}$ and $(D_s^j \phi_0^i)_{s \geq 0}$ are continuous around zero. In particular,

$$\lim_{s \downarrow 0} \mathbb{E}_y [A_T^{ij}(y, s) + B_T^{ij}(y, s) + C_T^{ij}(y, s)] = \mathbb{E} [A_T^{ij}(y, 0) + B_T^{ij}(y, 0) + C_T^{ij}(y, 0)],$$

$$\lim_{y \downarrow 0} \mathbb{E}_y [A_T^{ij}(y, 0) + B_T^{ij}(y, 0) + C_T^{ij}(y, 0)] = \mathbb{E} [A_T^{ij}(0, 0) + B_T^{ij}(0, 0) + C_T^{ij}(0, 0)].$$

By the DCT again this yields

$$\lim_{T \downarrow 0} \mathbb{E} [A_T^{ij}(0, 0)] = \frac{G_{ij}}{2\Delta \text{VIX}_0^2} \quad \text{and} \quad \lim_{T \downarrow 0} \mathbb{E} [B_T^{ij}(0, 0)] = \lim_{T \downarrow 0} \mathbb{E} [C_T^{ij}(0, 0)] = -\frac{J_i J_j}{4\Delta^2 \text{VIX}_0^4}.$$

Therefore $\phi_s^j D_s^j (\phi_y^i)^2$ satisfies the continuity requirements of $f(T, y, s)$ in Lemma 4. We combine this lemma with the limits above to see that, almost surely,

$$\begin{aligned} \lim_{T \downarrow 0} \frac{1}{T^2} \int_0^T \phi_s^j \int_s^T D_s^j (\phi_y^i)^2 dy ds &= \lim_{T \downarrow 0} \frac{1}{T^2} \int_0^T \int_0^y \phi_s^j D_s^j (\phi_y^i)^2 ds dy = \frac{1}{2} \lim_{s \leq y \leq T \downarrow 0} \phi_s^j D_s^j (\phi_y^i)^2 \\ &= \frac{J_j}{4\Delta \text{VIX}_0^2} \left[\frac{J_i G_{ij}}{2\Delta^2 \text{VIX}_0^4} - \frac{J_i^2 J_j}{2\Delta^3 \text{VIX}_0^6} \right]. \end{aligned}$$

We also recall that $\lim_{T \downarrow 0} u_0 = \frac{\|J\|}{2\Delta\text{VIX}_0^2}$ almost surely; hence, with (C₂) and (C₃), the DCT implies

$$\lim_{T \downarrow 0} \mathcal{S}_T = \sum_{i,j=1}^N \lim_{T \downarrow 0} \frac{1}{2} \mathbb{E} \left[\frac{\int_0^T \phi_s^j \int_s^T D_s^j(\phi_y^i)^2 dy ds}{u_0^3 T^2} \right] = \frac{1}{2\|J\|^3} \sum_{i,j=1}^N J_i J_j \left(G_{ij} - \frac{J_i J_j}{\Delta\text{VIX}_0^2} \right). \tag{32}$$

Curvature. We now turn our attention to the curvature. By the same arguments as above we have

$$\lim_{T \downarrow 0} \mathbb{E} \left[\frac{\left(\sum_{i,j=1}^N \int_0^T \phi_s^j \int_s^T D_s^j(\phi_y^i)^2 dy ds \right)^2}{u_0^4 T^4} \right] = \frac{2\Delta\text{VIX}_0^2}{\|J\|^7} \left(\sum_{i,j=1}^N J_i J_j \left(G_{ij} - \frac{J_i J_j}{\Delta\text{VIX}_0^2} \right) \right)^2.$$

For the last term of (8) we need to go one step further and compute more Malliavin derivatives, since

$$\begin{aligned} D_t^k \Theta_s^j &= \sum_{i=1}^N \left(D_t^k \phi_s^j \int_s^T D_s^j(\phi_y^i)^2 dy + 2\phi_s^j \int_s^T D_t^k \phi_y^i D_s^j \phi_y^i dy + 2\phi_s^j \int_s^T \phi_y^i D_t^k D_s^j \phi_y^i dy \right) \\ &=: \sum_{i=1}^N \int_s^T \Upsilon^{ijk}(t, s, y, T) dy. \end{aligned}$$

Thus we zoom in on the last term of the display above:

$$\begin{aligned} D_t^k D_s^j \phi_y^i &= \frac{D_t^k D_s^j m_y^i M_y - D_s^j m_y^i D_t^k M_y}{M_y^2} - \frac{D_t^k m_y^i D_s^j M_y + m_y^i D_t^k D_s^j M_y}{M_y^2} + \frac{2m_y^i D_s^j M_y D_t^k M_y}{M_y^3} \\ &=: \sum_{n=1}^5 Q_n^{ijk}(t, s, y, T), \end{aligned}$$

We zoom in again on $Q_1^{ijk}(t, s, y, T)$:

$$\begin{aligned} D_t^k D_s^j m_y^i &= D_t^k D_s^j D_y^i M_y = D_t^k \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} D_s^j D_y^i v_r dr}{2\Delta\text{VIX}_T} - \frac{\int_T^{T+\Delta} D_y^i v_r dr \int_T^{T+\Delta} D_s^j v_r dr}{4\Delta^2\text{VIX}_T^3} \right] \\ &=: \mathbb{E}_y \left[\alpha_T^{ijk} + \beta_T^{ijk} \right]. \end{aligned}$$

Some additional computations lead to

$$\begin{aligned} \alpha_T^{ijk} &= \frac{\text{VIX}_T \int_T^{T+\Delta} D_t^k D_s^j D_y^i v_r dr - D_t^k \text{VIX}_T \int_T^{T+\Delta} D_s^j D_y^i v_r dr}{2\Delta\text{VIX}_T^2} \\ &= \frac{\int_T^{T+\Delta} D_t^k D_s^j D_y^i v_r dr}{2\Delta\text{VIX}_T} - \frac{\int_T^{T+\Delta} D_t^k v_r dr \int_T^{T+\Delta} D_s^j D_y^i v_r dr}{4\Delta^2\text{VIX}_T^3}, \end{aligned}$$

$$\begin{aligned} \beta_T^{ijk} &= -\frac{\int_T^{T+\Delta} D_t^k D_y^j v_r dr \int_T^{T+\Delta} D_s^j v_r dr + \int_T^{T+\Delta} D_y^i v_r dr \int_T^{T+\Delta} D_t^k D_s^j v_r dr}{4\Delta^2 \text{VIX}_T^3} \\ &\quad + \frac{D_t^k \text{VIX}_T^3 \int_T^{T+\Delta} D_y^i v_r dr \int_T^{T+\Delta} D_s^j v_r dr}{4\Delta^2 \text{VIX}_T^6} \\ &= -\frac{\int_T^{T+\Delta} D_t^k D_y^j v_r dr \int_T^{T+\Delta} D_s^j v_r dr + \int_T^{T+\Delta} D_y^i v_r dr \int_T^{T+\Delta} D_t^k D_s^j v_r dr}{4\Delta^2 \text{VIX}_T^3} \\ &\quad + \frac{3 \int_T^{T+\Delta} D_t^k v_r dr \int_T^{T+\Delta} D_y^i v_r dr \int_T^{T+\Delta} D_s^j v_r dr}{8\Delta^3 \text{VIX}_T^5}. \end{aligned}$$

We notice, crucially, that we have already justified the continuity of ϕ and $D\phi$ around zero in the proofs of level and skew, respectively. Furthermore, by Lemma 6, the first two terms in Υ^{ijk} as well as Q_2, Q_3, Q_4, Q_5 all converge to some finite limit as $t \leq s \leq y \downarrow 0$ and are continuous around zero, almost surely. Similarly, β_T and the second term in α_T are almost surely continuous around zero, and their conditional expectation converges almost surely to some finite limit as $t \leq s \leq y \downarrow 0$ by the DCT and Lemma 6. Taking the limit as T goes to zero afterwards, we see that all of the aforementioned terms tend to a finite limit. On the other hand, by (C4), the DCT, and Lemma 6 we know that the conditional expectation of the first term in α_T is almost surely continuous around zero, and its limit is

$$\lim_{t \leq s \leq y \downarrow 0} \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} D_t^k D_s^j D_y^i v_r dr}{2\Delta \text{VIX}_T} \right] = \mathbb{E} \left[\frac{\int_T^{T+\Delta} D_0^k D_0^j D_0^i v_r dr}{2\Delta \text{VIX}_T} \right].$$

Since $\gamma < 0$, only this term contributes in the limit:

$$\begin{aligned} \lim_{t \leq s \leq y \leq T \downarrow 0} \frac{\phi_t^k \Upsilon^{ijk}(t, s, y, T)}{T^\gamma} &= \lim_{t \leq s \leq y \leq T \downarrow 0} 2\phi_t^k \phi_s^j \phi_y^i \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} D_t^k D_s^j D_y^i v_r dr}{2T^\gamma \Delta \text{VIX}_T M_y} \right] \\ &= \frac{J_i J_j J_k}{8\Delta^4 \text{VIX}_0^8} \lim_{T \downarrow 0} \frac{\mathbb{E} \left[\int_T^{T+\Delta} D_0^k D_0^j D_0^i v_r dr \right]}{T^\gamma}, \end{aligned}$$

where we applied the DCT at the end. Moreover, we know by (C2) that this limit is finite for $\gamma = 3H - \frac{1}{2}$; hence the conditions of Lemma 5 are satisfied. We also recall that $\lim_{T \downarrow 0} u_0 = \frac{\|J\|}{2\Delta \text{VIX}_0^2}$ almost surely; hence Lemma 5 yields the almost sure limit

$$\begin{aligned} &\lim_{T \downarrow 0} \frac{1}{u_0^5 T^{3+\gamma}} \int_0^T \sum_{k=1}^N \left\{ \phi_t^k D_t^k \left(\int_t^T |\Theta_s| ds \right) \right\} dt \\ &= \sum_{i,j,k=1}^N \lim_{T \downarrow 0} \frac{1}{u_0^5 T^3} \int_0^T \int_0^y \int_0^s \frac{\phi_t^k \Upsilon^{ijk}(t, s, y, T)}{T^{3H-\frac{1}{2}}} dy ds dt \\ &= \frac{2\Delta \text{VIX}_0^2}{3\|J\|^5} \sum_{i,j,k=1}^N J_i J_j J_k \lim_{T \downarrow 0} \frac{\mathbb{E} \left[\int_T^{T+\Delta} D_0^k D_0^j D_0^i v_r dr \right]}{T^{3H-\frac{1}{2}}}. \end{aligned}$$

The first two terms in (8) tend to zero since $\gamma < 0$; hence Theorem 3 and the DCT yield the final result:

$$\lim_{T \downarrow 0} \frac{C_T}{T^{3H-\frac{1}{2}}} = \frac{2\Delta \text{VIX}_0^2}{3\|J\|^5} \sum_{i,j,k=1}^N J_i J_j J_k \lim_{T \downarrow 0} \frac{\int_T^{T+\Delta} \mathbb{E} \left[D_0^k D_0^j D_0^i v_r \right] dr}{T^{3H-\frac{1}{2}}}.$$

6.4. Proofs in the two-factor rough Bergomi model

6.4.1. *Proof of Proposition 2.* We start with a useful lemma for Gaussian processes.

Lemma 7. *If B is a Gaussian process with $\|B\|_T := \sup_{t \leq T} |B_t|$, then $\mathbb{E}[e^{p\|B\|_T}]$ is finite for all $p \in \mathbb{R}$.*

Proof. The Borell–TIS inequality asserts that $\mathbb{E}[\|B\|_T] < \infty$ and

$$\mathbb{P}(\|B\|_T - \mathbb{E}\|B\|_T > x) \leq \exp \left\{ -\frac{x^2}{2\sigma_T^2} \right\},$$

where $\sigma_T^2 := \sup_{t \leq T} \mathbb{E}[B_t^2]$; see [1, Theorem 2.1.1]. We then follow the proof of [1, Theorem 2.1.2]:

$$\mathbb{E} \left[e^{p\|B\|_T} \right] = \int_0^\infty \mathbb{P} \left(e^{p\|B\|_T} > x \right) dx \leq e^p + \mathbb{E}[\|B\|_T] + \int_{e^p \vee \mathbb{E}[\|B\|_T]}^\infty \mathbb{P} \left(\|B\|_T > \frac{\log(x)}{p} \right) dx.$$

The Borell–TIS inequality in particular reads as follows:

$$\mathbb{P} \left(\|B\|_T > \log(x^{1/p}) \right) \leq \exp \left\{ -\frac{(\log(x^{1/p}) - \mathbb{E}[\|B\|_T])^2}{2\sigma_T^2} \right\}, \quad \text{for all } u > \mathbb{E}[\|B\|_T].$$

After a change of variable this yields

$$\int_{e^p \vee \mathbb{E}[\|B\|_T]}^\infty \mathbb{P} \left(\|B\|_T > \frac{\log(x)}{p} \right) dx \leq \int_{\frac{\log(e^p \vee \mathbb{E}[\|B\|_T])}{p}}^\infty \exp \left\{ -\frac{(x - \mathbb{E}[\|B\|_T])^2}{2\sigma_T^2} \right\} p e^{px} dx,$$

which is finite as desired. □

By the above lemma, $\|v\|_T \in L^p$, so that we can compute its Malliavin derivatives

$$D_y^1 v_r = v_0(r-y)^{H-} \left(\chi v \mathcal{E}_r^1 + \bar{\chi} \eta \rho \mathcal{E}_r^2 \right) \quad \text{and} \quad D_y^2 v_r = v_0 \bar{\chi} \eta \bar{\rho} (r-y)^{H-} \mathcal{E}_r^2. \tag{33}$$

Without explicitly computing further derivatives, one notices that (C4) holds and that there exist $C > 0$ and a random variable $X = C \|\mathcal{E}_r^1 + \mathcal{E}_r^2\|_T \in L^p$ for all $p > 1$ such that $D_y^i v_r \leq X(r-y)^{H-}$, $D_s^j D_y^i v_r \leq X(r-s)^{H-} (r-y)^{H-}$, and $D_t^k D_s^j D_y^i v_r \leq X(r-t)^{H-} (r-s)^{H-} (r-y)^{H-}$, implying (C2). The following lemma yields (C1).

Lemma 8. *In the two-factor rough Bergomi model (10) with $0 \leq T_1 < T_2$,*

$$\mathbb{E} \left[\sup_{y \leq T_1} \left(\mathbb{E}_y \left[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} v_r dr \right] \right)^{-p} \right]$$

is finite for all $p > 1$. In particular, $1/M$ is dominated in L^p .

Proof. We first apply an exp–log identity, then Jensen’s inequality (using the concavity of the logarithm function), to obtain

$$\begin{aligned} \mathbb{E}_y \left[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} v_r dr \right]^{-p} &= \exp \left\{ -p \log \left(\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} v_r dr \right) \right\} \\ &\leq \exp \left\{ -\frac{p}{T_2 - T_1} \int_{T_1}^{T_2} \log (\mathbb{E}_y[v_r]) dr \right\}. \end{aligned}$$

We further bound $\log \mathbb{E}_y[v_r]$ as follows, using the concavity of the logarithm and (10):

$$-\log \mathbb{E}_y[v_r] \leq -\frac{1}{2} \left\{ \log \left(2\chi v_0 \mathbb{E}_y[\mathcal{E}_r^1] \right) + \log \left(2\bar{\chi} v_0 \mathbb{E}_y[\mathcal{E}_r^2] \right) \right\}, \tag{34}$$

which we now compute as

$$\begin{aligned} \mathbb{E}_y[\mathcal{E}_r^1] &= \exp \left\{ -\frac{v^2 r^{2H}}{4H} + v \int_0^y (r-s)^{H-} dW_s^1 \right\} \mathbb{E}_y \left[\exp \left\{ v \int_y^r (r-s)^{H-} dW_s^1 \right\} \right] \\ &= \exp \left\{ \frac{v^2}{4H} \left[(r-y)^{2H} - r^{2H} \right] + v \int_0^y (r-s)^{H-} dW_s^1 \right\}, \\ \mathbb{E}_y[\mathcal{E}_r^2] &= \exp \left\{ \frac{\eta^2}{4H} \left[(r-y)^{2H} - r^{2H} \right] + \eta \int_0^y (r-s)^{H-} d(\rho W_s^1 + \bar{\rho} W_s^2) \right\}. \end{aligned} \tag{35}$$

Let us deal with the first term of (34), as the second one is analogous. We have

$$\int_{T_1}^{T_2} \left[(r-y)^{2H} - r^{2H} \right] dr = \frac{(T_2-u)^{2H_+} - (T_1-y)^{2H_+} - T_2^{2H_+} + T_1^{2H_+}}{2H_+},$$

which is clearly bounded below for all $0 \leq u \leq T_1$. Moreover, by Fubini’s theorem,

$$\int_{T_1}^{T_2} \int_0^y (r-t)^{H-} dW_t^1 dr = \int_0^y \int_{T_1}^{T_2} (r-t)^{H-} dr dW_t^1 = \int_0^y \frac{(T_2-t)^{H_+} - (T_1-t)^{H_+}}{H_+} dW_t^1 =: \bar{B}_t$$

is a Gaussian process. Since $\exp\{\cdot\}$ is increasing, $\sup_{t \in [0, T]} \exp\{\bar{B}_t\} = \exp\{\sup_{t \in [0, T]} \bar{B}_t\}$; thus

$$\mathbb{E} \left[\sup_{y \leq T_1} \exp \left(-\frac{p}{T_2 - T_1} \int_{T_1}^{T_2} \int_0^y (r-s)^{H-} dW_s^1 dr \right) \right] \leq \mathbb{E} \left[\exp \left(\frac{p}{T_2 - T_1} \|\bar{B}\|_T \right) \right] < \infty,$$

by Lemma 7, which concludes the proof. □

Combining (33) and (35), we obtain $\mathbb{E}_y[D_y^i v_r]$, $i = 1, 2$. The following lemma proves that (C3) is satisfied.

Lemma 9. For any $p > 1$, $\mathbb{E}[u_s^{-p}]$ is uniformly bounded in s and T , with $s \leq T$.

Proof. Since $v, \eta, \rho + \bar{\rho} > 0$, we have $D_y^1 v_r + D_y^2 v_r > 0$ almost surely for all $y \leq r$. Moreover, VIX and $1/\text{VIX}$ are dominated by some $X \in L^p$ for all $p > 1$, so, almost surely and independently of the sign of the numerator, we obtain

$$m_y^i = \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} D_y^i v_r dr}{2\Delta \text{VIX}_T} \right] \geq \mathbb{E}_y \left[\frac{\int_T^{T+\Delta} D_y^i v_r dr}{2\Delta X} \right].$$

Therefore, using that $1/M$ is dominated by X and Jensen’s inequality, we get

$$\begin{aligned} \frac{1}{u_s^2} &= \frac{T-s}{\int_s^T \sum_{i=1}^N (\phi_y^i)^2 dy} \leq \frac{X^2(T-s)}{\int_s^T \sum_{i=1}^N (m_y^i)^2 dy} \leq \frac{X^2 N(T-s)}{\int_s^T \left(\sum_{i=1}^N m_y^i\right)^2 dy} \leq X^2 N \left(\frac{T-s}{\int_s^T \sum_{i=1}^N m_y^i dy}\right)^2 \\ &\leq 4X^2 N \left(\frac{\int_s^T \int_T^{T+\Delta} \sum_{i=1}^N \mathbb{E}_y \left[D_y^i v_r / X\right] dr dy}{\Delta(T-s)}\right)^{-2}. \end{aligned} \tag{36}$$

Hence we turn our attention to

$$\begin{aligned} &\mathbb{E} \left[\left(\frac{1}{\Delta(T-s)} \int_s^T \int_T^{T+\Delta} \mathbb{E}_y \left[\frac{D_y^1 v_r + D_y^2 v_r}{X} \right] dr dy \right)^{-p} \right] \\ &= \mathbb{E} \left[\exp \left\{ -p \log \left(\frac{1}{\Delta(T-s)} \int_s^T \int_T^{T+\Delta} \mathbb{E}_y \left[\frac{D_y^1 v_r + D_y^2 v_r}{X} \right] dr dy \right) \right\} \right] \\ &\leq \mathbb{E} \left[\exp \left\{ -\frac{p}{\Delta(T-s)} \int_s^T \int_T^{T+\Delta} \mathbb{E}_y \left[\log \left(D_y^1 v_r + D_y^2 v_r \right) - \log(X) \right] dr dy \right\} \right] \\ &\leq \left(\mathbb{E} \left[\exp \left\{ -\frac{2p}{\Delta(T-s)} \int_s^T \int_T^{T+\Delta} \mathbb{E}_y \left[\log \left(D_y^1 v_r + D_y^2 v_r \right) \right] dr dy \right\} \right] \right)^{\frac{1}{2}} \sqrt{\mathbb{E} [X^{2p}]}, \end{aligned} \tag{37}$$

using Jensen’s inequality, the Cauchy–Schwarz inequality, and the fact that $e^{p\mathbb{E}_y[\log(X)]} \leq \mathbb{E}_y[X^p]$. Convexity and (33) imply

$$-\log \left(D_y^1 v_r + D_y^2 v_r \right) \leq -\frac{1}{2} \left\{ \log \left(2v_0 \chi v(r-y)^{H-} \mathcal{E}_r^1 \right) + \log \left(2v_0 \bar{\chi} \eta(\rho + \bar{\rho})(r-y)^{H-} \mathcal{E}_r^2 \right) \right\}.$$

We focus on the first term; the other can be treated identically. From (35) we have

$$\mathbb{E}_y \left[\log \left(2v_0 \chi v(r-y)^{H-} \mathcal{E}_r^1 \right) \right] = \log \left(2v_0 \chi v(r-y)^{H-} \right) - \frac{v^2 r^{2H}}{4H} + v \int_0^y (r-t)^{H-} dW_t^1. \tag{38}$$

Let us start with

$$\begin{aligned} &\int_s^T \int_T^{T+\Delta} \log \left(2v_0 \chi v(r-y)^{H-} \right) dr dy \\ &= 2(T-s)\Delta v_0 \chi v + H_- \int_s^T \int_T^{T+\Delta} \log(r-y) dr dy \\ &= 2(T-s)\Delta v_0 \chi v + H_- \\ &\quad \int_s^T \left[(T+\Delta-y) \log(T+\Delta-y) - (T+\Delta-y) - (T-y) \log(T-y) + (T-y) \right] dy \\ &= 2(T-s)\Delta v_0 \chi v + H_- \left\{ -\Delta(T-s) - \int_\Delta^{T+\Delta-s} x \log(x) dx + \int_0^{T-s} x \log(x) dx \right\} \end{aligned}$$

$$\begin{aligned}
 &= 2(T-s)\Delta v_0 \chi v + H_- \left\{ -\Delta(T-s) + \left(\frac{(T-s)^2 \log(T-s)}{2} - \frac{(T-s)^2}{4} \right) \right. \\
 &\quad \left. - \left(\frac{(T+\Delta-s)^2 \log(T+\Delta-s)}{2} - \frac{(T+\Delta-s)^2}{4} \frac{\Delta^2 \log(\Delta)}{2} + \frac{\Delta^2}{4} \right) \right\} \\
 &= 2(T-s)\Delta v_0 \chi v + (T-s)H_- \left\{ -\Delta + \left(\frac{(T-s) \log(T-s)}{2} - \frac{(T-s)}{4} \right) + \frac{2\Delta + (T-s)}{4} \right. \\
 &\quad \left. - \Delta \log(T+\Delta-s) - \frac{(T-s) \log(T+\Delta-s)}{2} \right\} \\
 &\quad - \frac{\Delta^2}{2} (\log(T+\Delta-s) - \log(\Delta)).
 \end{aligned}$$

By Taylor’s theorem, $\log(T+\Delta-s) - \log(\Delta) = \frac{T-s}{\Delta} + \varepsilon(T-s)$, where $\varepsilon: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is such that $\varepsilon(x)/x$ tends to zero at the origin. We conclude that

$$-\frac{p}{2\Delta(T-s)} \int_s^T \int_T^{T+\Delta} \log(2v_0 \chi v(r-y)^{H_-}) \, dr dy$$

is uniformly bounded. Now we study the second term of (38):

$$-\int_s^T \int_T^{T+\Delta} r^{2H} \, dr dy = (T-s) \frac{T^{2H_+} - (T+\Delta)^{2H_+}}{2H_+}.$$

Therefore the following is uniformly bounded:

$$\frac{p}{2\Delta(T-s)} \int_s^T \int_T^{T+\Delta} \frac{v^2 r^{2H}}{4H} \, dr dy.$$

For the last term, by the stochastic Fubini theorem [48, Theorem 65], we get

$$\begin{aligned}
 \int_s^T \int_T^{T+\Delta} \int_0^y (r-t)^{H_-} \, dW_t^1 \, dr dy &= \int_s^T \int_0^y \int_T^{T+\Delta} (r-t)^{H_-} \, dr dW_t^1 \, dy \\
 &= \int_0^T \int_{s \vee t}^T \frac{(T+\Delta-t)^{H_+} - (T-t)^{H_+}}{H_+} \, dy dW_t^1.
 \end{aligned}$$

Standard Gaussian computations then yield

$$\begin{aligned}
 &\mathbb{E} \left[\exp \left\{ -\frac{p}{4\Delta(T-s)} \int_s^T \int_T^{T+\Delta} v \int_0^y (r-t)^{H_-} \, dW_t^1 \, dr dy \right\} \right] \\
 &= \exp \left\{ \frac{1}{2} \left(\frac{pv}{4\Delta(T-s)} \right)^2 \int_0^T \left(\int_{s \vee t}^T \frac{(T+\Delta-t)^{H_+} - (T-t)^{H_+}}{H_+} \, dy \right)^2 \, dt \right\}. \quad (39)
 \end{aligned}$$

The incremental function $x \mapsto (x+\Delta)^{H_+} - x^{H_+}$ is decreasing by concavity; hence $(T+\Delta-t)^{H_+} - (T-t)^{H_+} \leq \Delta^{H_+}$, and we obtain

$$\int_0^T (T-s \vee t)^2 \, dt = \int_0^s (T-s)^2 \, dt + \int_s^T (T-t)^2 \, dt = s(T-s)^2 + \frac{(T-s)^3}{3},$$

which implies that (39) is uniformly bounded. We have thus shown that (37) is uniformly bounded in s, T . \square

Coming back to (36), by the Cauchy–Schwarz inequality we have

$$\mathbb{E}[u_s^{-p}]^2 \leq 2^p \mathbb{E}[X^p] \mathbb{E} \left[\left(\frac{T-s}{\int_s^T (m_y^1 + m_y^2) dy} \right)^{2p} \right],$$

which is uniformly bounded for all $s \leq T$, and this concludes the proof.

6.4.2. *Proof of Proposition 3. Level.* We start with the derivatives

$$D_s^1 v_t = v_0 \left[\chi v (t-s)^{H-} \mathcal{E}_t^1 + \bar{\chi} \eta \rho (t-s)^{H-} \mathcal{E}_t^2 \right] \quad \text{and} \quad D_s^2 v_t = v_0 \bar{\chi} \eta \bar{\rho} (t-s)^{H-} \mathcal{E}_t^2,$$

and recall, from the definitions in (9),

$$J_1 = \int_0^\Delta v_0 \mathbb{E} \left[\chi v r^{H-} \mathcal{E}_r^1 + \bar{\chi} \eta \rho r^{H-} \mathcal{E}_r^2 \right] dr = v_0 (\chi v + \bar{\chi} \eta \rho) \frac{\Delta^{H_+}}{H_+} \quad \text{and} \quad J_2 = v_0 \bar{\chi} \eta \bar{\rho} \frac{\Delta^{H_+}}{H_+}.$$

We also note that $\mathbb{E}[\mathcal{E}_t^i] = 1$. This yields the norm

$$\|J\| := \left(J_1^2 + J_2^2 \right)^{\frac{1}{2}} = \frac{v_0 \Delta^{H_+}}{H_+} \sqrt{(\chi v + \bar{\chi} \eta \rho)^2 + \bar{\chi}^2 \eta^2 \bar{\rho}^2} = \frac{v_0 \Delta^{H_+}}{H_+} \psi(\rho, v, \eta, \chi),$$

with the function ψ defined in the proposition, which grants us the first limit by Proposition 1. To simplify the notation below, we introduce $\mathfrak{w} := \chi v + \bar{\chi} \eta \rho$.

Skew. We compute the further derivatives

$$D_0^1 D_0^1 v_t = v_0 \left(\chi v^2 t^{2H-1} \mathcal{E}_t^1 + \bar{\chi} \eta^2 \rho^2 t^{2H-1} \mathcal{E}_t^2 \right),$$

$$D_0^1 D_0^2 v_t = v_0 \bar{\chi} \eta^2 \rho \bar{\rho} t^{2H-1} \mathcal{E}_t^2,$$

$$D_0^2 D_0^2 v_t = v_0 \bar{\chi} \eta^2 \bar{\rho}^2 t^{2H-1} \mathcal{E}_t^2.$$

Similarly to J , we recall that $G_{ij} = \int_0^\Delta \mathbb{E}[D_0^i D_0^j v_r] dr$, so that

$$G_{11} = \frac{\Delta^{2H}}{2H} v_0 (\chi v^2 + \bar{\chi} \eta^2 \rho^2), \quad G_{12} = \frac{\Delta^{2H}}{2H} v_0 \bar{\chi} \eta^2 \rho \bar{\rho}, \quad G_{22} = \frac{\Delta^{2H}}{2H} v_0 \bar{\chi} \eta^2 \bar{\rho}^2.$$

Notice that $\text{VIX}_0^2 = v_0$; thus we have

$$\begin{aligned} J_1^2 \left(G_{11} - \frac{J_1^2}{\Delta \text{VIX}_0^2} \right) &= \frac{v_0^3 \Delta^{4H+1}}{2HH_+^2} \mathfrak{w}^2 (\chi v^2 + \bar{\chi} \eta^2 \rho^2) - \frac{v_0^3 \Delta^{4H+1}}{H_+^4} \mathfrak{w}^4 \\ &= v_0^3 \frac{\Delta^{4H+1}}{H_+^2} \mathfrak{w}^2 \left[\frac{\chi v^2 + \bar{\chi} \eta^2 \rho^2}{2H} - \frac{\mathfrak{w}^2}{H_+^2} \right], \end{aligned}$$

$$\begin{aligned}
 J_1 J_2 \left(G_{12} - \frac{J_1 J_2}{\Delta \text{VIX}_0^2} \right) &= \frac{v_0^3 \Delta^{4H+1}}{2HH_+^2} \mathfrak{w} \bar{\chi}^2 \eta^3 \rho \bar{\rho}^2 - \frac{v_0^3 \Delta^{4H+1}}{H_+^4} \mathfrak{w}^2 \bar{\chi}^2 \eta^2 \bar{\rho}^2 \\
 &= v_0^3 \frac{\Delta^{4H+1}}{H_+^2} \mathfrak{w} \bar{\chi}^2 \eta^2 \bar{\rho}^2 \left[\frac{\eta \rho}{2H} - \frac{\mathfrak{w}}{H_+^2} \right], \\
 J_2^2 \left(G_{22} - \frac{J_2^2}{\Delta \text{VIX}_0^2} \right) &= v_0^3 \frac{\Delta^{4H+1}}{2HH_+^2} \bar{\chi}^3 \eta^4 \bar{\rho}^4 - v_0^3 \bar{\chi}^4 \frac{\Delta^{4H+1}}{H_+^4} \eta^4 \bar{\rho}^4 \\
 &= v_0^3 \frac{\Delta^{4H+1}}{H_+^2} \bar{\chi}^3 \eta^4 \bar{\rho}^4 \left(\frac{1}{2H} - \frac{\bar{\chi}}{H_+^2} \right).
 \end{aligned}$$

Finally, by Proposition 1 we obtain

$$\begin{aligned}
 \lim_{T \downarrow 0} S_T &= \frac{H_+ \Delta^{H-}}{2\psi(\rho, v, \eta, \chi)^3} \left\{ \mathfrak{w}^2 \left[\frac{\chi v^2 + \bar{\chi} \eta^2 \rho^2}{2H} - \left(\frac{\mathfrak{w}}{H_+} \right)^2 \right] \right. \\
 &\quad \left. + 2\mathfrak{w} \bar{\chi}^2 \eta^2 \bar{\rho}^2 \left[\frac{\eta \rho}{2H} - \frac{v + \eta \rho}{H_+^2} \right] + \bar{\chi}^3 \eta^4 \bar{\rho}^4 \left(\frac{1}{2H} - \frac{1}{H_+^2} \right) \right\}.
 \end{aligned}$$

Curvature. For the last step we go one step further:

$$\begin{aligned}
 D_0^1 D_0^1 D_0^1 v_t &= v_0 \left(\chi v^3 \mathcal{E}_t^1 + \bar{\chi} \eta^3 \rho^3 \mathcal{E}_t^2 \right) t^{3H-}, & D_0^2 D_0^1 D_0^1 v_t &= v_0 \bar{\chi} \eta^3 \rho^2 \bar{\rho} t^{3H-} \mathcal{E}_t^2, \\
 D_0^2 D_0^2 D_0^1 v_t &= v_0 \bar{\chi} \eta^3 \rho \bar{\rho}^2 t^{3H-} \mathcal{E}_t^2, & D_0^2 D_0^2 D_0^2 v_t &= v_0 \bar{\chi} \eta^3 \bar{\rho}^3 t^{3H-} \mathcal{E}_t^2.
 \end{aligned}$$

We notice that

$$\lim_{T \downarrow 0} \frac{\int_T^{T+\Delta} r^{3H-} dr}{T^{3H-\frac{1}{2}}} = \lim_{T \downarrow 0} \frac{(T + \Delta)^{3H-\frac{1}{2}} - T^{3H-\frac{1}{2}}}{T^{3H-\frac{1}{2}}(3H - \frac{1}{2})} = \frac{2}{1 - 6H}.$$

By the curvature limit in Proposition 1, we have

$$\begin{aligned}
 \lim_{T \downarrow 0} \frac{\mathcal{C}_T}{T^{3H-\frac{1}{2}}} &= \frac{2\Delta v_0}{3 \left(\frac{v_0 \Delta^{H+}}{2H_+} \right)^5 \psi(\rho, v, \eta, \chi)^5} \left\{ v_0^3 \frac{\Delta^{3H+}}{H_+^3} \mathfrak{w}^3 v_0 \frac{\chi v^3 + \bar{\chi} \eta^3 \rho^3}{\frac{1}{2} - 3H} \right. \\
 &\quad \left. + 3v_0^3 \frac{\Delta^{3H+}}{H_+^3} \mathfrak{w}^2 \bar{\chi} \eta \bar{\rho} v_0 \frac{\bar{\chi} \eta^3 \rho^2 \bar{\rho}}{\frac{1}{2} - 3H} + 3v_0^3 \frac{\Delta^{3H+}}{H_+^3} \mathfrak{w} \bar{\chi}^2 \eta^2 \bar{\rho}^2 v_0 \frac{\bar{\chi} \eta^3 \rho \bar{\rho}^2}{\frac{1}{2} - 3H} + v_0 \frac{\Delta^{3H+}}{H_+^3} \bar{\chi}^3 \eta^3 \bar{\rho}^3 v_0 \frac{\bar{\chi} \eta^3 \bar{\rho}^3}{\frac{1}{2} - 3H} \right\} \\
 &= \frac{128 \Delta^{-2H} H_+^2}{3\psi(\rho, v, \eta, \chi)^5 (1 - 6H)} \\
 &\quad \left\{ \mathfrak{w}^3 (\chi v^3 + \bar{\chi} \eta^3 \rho^3) + 3\mathfrak{w}^2 \bar{\chi}^2 \eta^4 \bar{\rho}^2 \rho^2 + 3\mathfrak{w} \bar{\chi}^3 \eta^5 \bar{\rho}^4 \rho + \bar{\chi}^4 \eta^6 \bar{\rho}^6 \right\},
 \end{aligned}$$

which yields the claim.

6.5. Proofs for the stock price

6.5.1. *Proof of Proposition 4.* Since ϕ and u^{-p} are dominated by the conditions (i) and (iii) respectively, with the same notation as in the proof of Proposition 7, we obtain by (ii), as T goes to zero,

$$D\phi_s \lesssim (T - s)^{H-}, \quad \Theta_s \lesssim (T - s)^{H+}, \quad D\Theta_s \lesssim (T - s)^{2H}, \quad DD\Theta_s \lesssim (T - s)^{3H-\frac{1}{2}}.$$

Under our three assumptions it is straightforward to see that (H_{12345}) are satisfied. Moreover, the terms in (H_6^λ) behave as $T^{2H-\lambda}$ and the one in (H_7^λ) as $T^{H-\lambda}$, which means that if we set $\lambda = H_-$ the former vanishes and the second yields a non-trivial behaviour.

Let us have a look at the short-time implied volatility. By Lemma 3 and the continuity of v we have $\lim_{T \downarrow 0} u_0 = \sqrt{\sum_{i=1}^N v_0 \rho_i^2} = \sqrt{v_0}$ almost surely; hence by Theorem 1 and the DCT,

$$\lim_{T \downarrow 0} \widehat{L}_T = \lim_{T \downarrow 0} \mathbb{E}[u_0] = \sqrt{v_0}.$$

We then turn our attention to the short-time skew. With $\lambda = H_-$, Theorem 2 and the DCT imply

$$\lim_{T \downarrow 0} \frac{\widehat{S}_T}{T^{H-}} = \sum_{i,j=1}^N \lim_{T \downarrow 0} \frac{1}{2} \mathbb{E} \left[\frac{\int_0^T \phi_s^j \int_s^T D_s^j(\phi_y^i)^2 dy ds}{u_0^3 T^{\frac{3}{2}+H}} \right] = \sum_{j=1}^N \frac{\rho_j}{2v_0^{3/2}} \mathbb{E} \left[\lim_{T \downarrow 0} \frac{\int_0^T \int_s^T \sqrt{v_s} D_s^j v_y dy ds}{T^{\frac{3}{2}+H}} \right],$$

where we used $\sum_{i=1}^N \rho_i^2 = 1$. For any $j \in \llbracket 1, N \rrbracket c$, the Cauchy–Schwarz inequality yields

$$\mathbb{E} \left[\left(\sqrt{\frac{v_s}{v_0}} - 1 \right) D_s^j v_y \right] \leq \mathbb{E} \left[\left(\sqrt{\frac{v_s}{v_0}} - 1 \right)^2 \right]^{\frac{1}{2}} \mathbb{E}[(D_s^j v_y)^2]^{\frac{1}{2}},$$

where $\mathbb{E}[(D_s^j v_y)^2]^{\frac{1}{2}} \leq C(y - s)^{H-}$ for some finite constant C by (ii). Therefore,

$$\begin{aligned} \lim_{T \downarrow 0} \left(\frac{\int_0^T \int_s^T \mathbb{E}[\sqrt{v_s} D_s^j v_y] dy ds}{\sqrt{v_0} T^{\frac{3}{2}+H}} - \frac{\int_0^T \int_s^T \mathbb{E}[D_s^j v_y] dy ds}{T^{\frac{3}{2}+H}} \right) \\ \leq C \lim_{T \downarrow 0} \left(\sup_{t \leq T} \mathbb{E} \left[\left(\sqrt{\frac{v_t}{v_0}} - 1 \right)^2 \right]^{\frac{1}{2}} \frac{\int_0^T \int_s^T (y - s)^{H-} dy ds}{T^{\frac{3}{2}+H}} \right). \end{aligned}$$

Since the fraction is equal to $((H + \frac{3}{2})H_+)^{-1}$ and $\limsup_{T \downarrow 0} \mathbb{E}[(\sqrt{v_t/v_0} - 1)^2]$ is null by (iv), we obtain

$$\lim_{T \downarrow 0} \frac{\widehat{S}_T}{T^{H-}} = \sum_{j=1}^N \frac{\rho_j}{2v_0} \mathbb{E} \left[\lim_{T \downarrow 0} \frac{\int_0^T \int_s^T D_s^j v_y dy ds}{T^{\frac{3}{2}+H}} \right].$$

6.5.2. *Proof of Corollary 1.* Since

$$\mathbb{E}[u_s^{-p}] = \mathbb{E} \left[\left(\frac{1}{T - s} \int_s^T v_r dr \right)^{-\frac{p}{2}} \right],$$

Lemmas 7 and 8 show that the assumptions (i)–(iii) of Proposition 4 hold. Moreover, v has almost-sure continuous paths; hence $\sqrt{\frac{v_t}{v_0}}$ tends to one almost surely and (iv) holds by reverse Fatou’s lemma. For $0 \leq s \leq y$, (33) implies

$$\mathbb{E}[D_s^1 v_y] = v_0(y - s)^{H-} (\chi v + \bar{\chi} \eta \rho) \quad \text{and} \quad \mathbb{E}[D_s^2 v_y] = v_0(y - s)^{H-} \bar{\chi} \eta \bar{\rho},$$

and clearly $\mathbb{E}[D_s^3 v_y] = 0$. Therefore, Proposition 4 implies

$$\lim_{T \downarrow 0} \frac{\widehat{S}_T}{T^{H-}} = \frac{\rho_1}{2v_0} \frac{v_0(\chi v + \bar{\chi} \eta \rho)}{H_+(H + \frac{3}{2})} + \frac{\rho_2}{2v_0} \frac{v_0 \bar{\chi} \eta \bar{\rho}}{H_+(H + \frac{3}{2})} = \frac{\rho_1 \chi v + \eta \bar{\chi} (\rho_1 \rho + \rho_2 \bar{\rho})}{(2H_+)(H + \frac{3}{2})}.$$

6.6. Partial derivatives of the Black–Scholes function

Recall the Black–Scholes formula (6) and assume $\varsigma := \sigma \sqrt{T - t} > 0$ fixed. Then

$$\begin{aligned} \partial_x \text{BS}(t, x, k, \sigma) &= e^x \mathcal{N}(d_+(x, k, \sigma)), \\ \partial_x^2 \text{BS}(t, x, k, \sigma) &= e^x \left\{ \mathcal{N}(d_+(x, k, \sigma)) + \frac{\mathcal{N}'(d_+(x, k, \sigma))}{\varsigma} \right\}, \end{aligned}$$

so that (we drop the dependence on t and σ in the $G(\cdot)$ notation)

$$G(x, k) := (\partial_x^2 - \partial_x) \text{BS}(t, x, k, \sigma) = \frac{e^{x - \frac{1}{2} d_+(x, k, \sigma)^2}}{\varsigma \sqrt{2\pi}} = \frac{e^{k - \frac{1}{2} d_-(x, k, \sigma)^2}}{\varsigma \sqrt{2\pi}}.$$

Now define

$$f(x, k) := x - \frac{d_+(x, k, \sigma)^2}{2} = k - \frac{d_-(x, k, \sigma)^2}{2} = \frac{x + k}{2} - \frac{(x - k)^2}{2\varsigma^2} - \frac{\varsigma^2}{8}.$$

We then have

$$\begin{aligned} \partial_x f(x, k) &= \frac{1}{2} - \frac{x - k}{\varsigma^2}, & \partial_k f(x, k) &= \frac{1}{2} + \frac{x - k}{\varsigma^2}, \\ \partial_x^2 f(x, k) &= \partial_k^2 f(x, k) = -\partial_{xk}^2 f(x, k) = -\frac{1}{\varsigma^2}. \end{aligned}$$

For the partial derivatives, noting that $\partial_x G = \frac{1}{\varsigma \sqrt{2\pi}} \partial_x f e^f$ implies the ATM formula

$$\partial_x G(x, x) = \frac{1}{2\varsigma \sqrt{2\pi}} \exp \left\{ x - \frac{\varsigma^2}{8} \right\},$$

and furthermore,

$$\partial_{xk} G = \frac{e^f}{\varsigma \sqrt{2\pi}} (\partial_{xk} f + \partial_x f \partial_k f) \quad \text{and} \quad \partial_{xk} G(x, x) = \frac{e^{f(x, x)}}{\varsigma \sqrt{2\pi}} \left(\frac{1}{\varsigma^2} + \frac{1}{4} \right).$$

We further define the partial derivatives appearing in the proof of Theorem 2, after (24):

$$\begin{aligned} L(x, k) &:= \left(\frac{1}{4} \partial_x + \frac{1}{2} \partial_{xk} \right) G(x, k) = \frac{1}{\varsigma \sqrt{2\pi}} e^{f(x, k)} \left(\frac{1}{4} + \frac{1}{4} \partial_k f(x, k) - \frac{1}{2} (\partial_k f(x, k))^2 - \frac{1}{2} \partial_{kk} f(x, k) \right), \\ L(x, x) &= \frac{e^{f(x, x)}}{\varsigma \sqrt{2\pi}} \left(\frac{1}{4} + \frac{1}{2\varsigma^2} \right). \end{aligned}$$

Using $\partial_k f = 1 - \partial_x f$ and $\partial_{xk} f = -\partial_{xx} f = -\partial_{kk} f$, we compute

$$\begin{aligned}\partial_k L &= \frac{e^f}{\zeta\sqrt{2\pi}} \left[\frac{3}{4}\partial_x f - \frac{5}{4}(\partial_x f)^2 + \frac{1}{2}(\partial_x f)^3 - \frac{5}{4}\partial_{xx} f + \frac{3}{2}\partial_x f \partial_{xx} f \right], \\ \partial_k L(x, x) &= \frac{e^{f(x,x)}}{\zeta\sqrt{2\pi}} \left(\frac{1}{8} + \frac{1}{2\zeta^2} \right).\end{aligned}$$

Finally, we need the derivatives featuring in the proof of Theorem 3. We start with

$$\begin{aligned}\tilde{H} = \partial_{xk} L &= \frac{e^f}{\zeta\sqrt{2\pi}} \\ &\left[\frac{3}{4}(\partial_x f)^2 - \frac{5}{4}(\partial_x f)^3 + \frac{1}{2}(\partial_x f)^4 + \frac{3}{4}\partial_{xx} f - \frac{15}{4}\partial_x f \partial_{xx} f + 3(\partial_x f)^2 \partial_{xx} f + \frac{3}{2}(\partial_{xx} f)^2 \right], \\ \partial_{xk} L(x, x) &= \frac{e^{f(x,x)}}{\zeta\sqrt{2\pi}} \left(\frac{1}{16} + \frac{3}{8\zeta^2} + \frac{3}{2\zeta^4} \right).\end{aligned}$$

The next partial derivative yields

$$\begin{aligned}\partial_{xxk} L &= \frac{e^f}{\zeta\sqrt{2\pi}} \left[\frac{3}{4}(\partial_x f)^3 - \frac{5}{4}(\partial_x f)^4 + \frac{1}{2}(\partial_x f)^5 + \frac{9}{4}\partial_{xx} f \partial_x f - \frac{15}{2}\partial_{xx} f (\partial_x f)^2 \right. \\ &\quad \left. - \frac{15}{4}(\partial_{xx} f)^2 + 5\partial_{xx} f (\partial_x f)^3 + \frac{15}{2}(\partial_{xx} f)^2 \partial_x f \right], \\ \partial_{xxk} L(x, x) &= \frac{e^{f(x,x)}}{\zeta\sqrt{2\pi}} \left(\frac{1}{32} + \frac{1}{8\zeta^2} \right),\end{aligned}$$

and differentiating one last time we reach

$$\begin{aligned}\partial_{xxxk} L &= \frac{e^f}{\zeta\sqrt{2\pi}} \left[\frac{3}{4}(\partial_x f)^4 - \frac{5}{4}(\partial_x f)^5 + \frac{1}{2}(\partial_x f)^6 + \frac{9}{2}\partial_{xx} f (\partial_x f)^2 - \frac{25}{2}\partial_{xx} f (\partial_x f)^3 \right. \\ &\quad \left. - \frac{75}{4}(\partial_{xx} f)^2 \partial_x f + \frac{15}{2}\partial_{xx} f (\partial_x f)^4 + \frac{45}{2}(\partial_{xx} f)^2 (\partial_x f)^2 + \frac{9}{4}(\partial_{xx} f)^2 + \frac{15}{2}(\partial_{xx} f)^3 \right], \\ \partial_{xxxk} L(x, x) &= \frac{e^{f(x,x)}}{\zeta\sqrt{2\pi}} \left(\frac{1}{64} - \frac{1}{32\zeta^2} - \frac{3}{2\zeta^4} - \frac{15}{2\zeta^6} \right).\end{aligned}$$

We conclude that

$$H(x, x) = (\partial_{xxxk} - \partial_{xxk})L(x, x) = \frac{e^{f(x,x)}}{\zeta\sqrt{2\pi}} \left(-\frac{1}{64} - \frac{5}{32\zeta^2} - \frac{3}{2\zeta^4} - \frac{15}{2\zeta^6} \right).$$

Funding information

A. P. acknowledges financial support from the EPSRC CDT in Financial Computing and Analytics, while A. P. and A. J. acknowledge support from the EPSRC EP/T032146/1 grant. Part of the project was carried out while A. P. was at Imperial.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] ADLER, R. J. AND TAYLOR, J. E. (2007). *Random Fields and Geometry*. Springer, New York.
- [2] ALÒS, E., GARCÍA-LORITE, D. AND MUGURUZA, A. (2022). On smile properties of volatility derivatives and exotic products: understanding the VIX skew. *SIAM J. Financial Math.* **13**, 32–69.
- [3] ALÒS, E. AND LEÓN, J. A. (2017). On the curvature of the smile in stochastic volatility models. *SIAM J. Financial Math.* **8**, 373–399.
- [4] ALÒS, E., LEÓN, J. A. AND VIVES, J. (2007). On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance Stoch.* **11**, 571–589.
- [5] ALÒS, E. AND SHIRAYA, K. (2019). Estimating the Hurst parameter from short term volatility swaps. *Finance Stoch.* **23**, 423–447.
- [6] BALDEAUX, J. AND BADRAN, A. (2014). Consistent modelling of VIX and equity derivatives using a 3/2 plus jumps model. *Appl. Math. Finance* **21**, 299–312.
- [7] BARLETTA, A., NICOLATO, E. AND PAGLIARANI, S. (2018). The short-time behavior of VIX implied volatilities in a multifactor stochastic volatility framework. *Math. Finance* **29**, 928–966.
- [8] BAYER, C. *et al.* (eds) (2023). *Rough Volatility*. Society for Industrial and Applied Mathematics, Philadelphia.
- [9] BAYER, C., FRIZ, P. K. AND GATHERAL, J. (2016). Pricing under rough volatility. *Quant. Finance* **16**, 887–904.
- [10] BAYER, C., QIU, J. AND YAO, Y. (2022). Pricing options under rough volatility with backward SPDEs. *SIAM J. Financial Math.* **13**, 179–212.
- [11] BENNEDSEN, M., LUNDE, A. AND PAKKANEN, M. S. (2022). Decoupling the short-and long-term behavior of stochastic volatility. *J. Financial Econometrics* **20**, 961–1006.
- [12] BERESTYCKI, H., BUSCA, J. AND FLORENT, I. (2004). Computing the implied volatility in stochastic volatility models. *Commun. Pure Appl. Math.* **57**, 1352–1373.
- [13] BERGOMI, L. (2005). Smile dynamics II. *Risk*, 67–73.
- [14] BERGOMI, L. (2008). Smile dynamics III. *Risk*, 90–96.
- [15] BONESINI, O., CALLEGARO, G. AND JACQUIER, A. (2023). Functional quantization of rough volatility and applications to the VIX. *Quant. Finance* **23**, 1769–1792.
- [16] BONESINI, O., JACQUIER, A. AND PANNIER, A. (2023). *Rough volatility, path-dependent PDEs and weak rates of convergence*. Preprint. Available at <https://arxiv.org/abs/2304.03042>.
- [17] BOURGEY, F. AND DE MARCO, S. (2022). Multilevel Monte Carlo simulation for VIX options in the rough Bergomi model. *J. Comput. Finance* **26**, 53–82.
- [18] CARR, P. AND MADAN, D. (2014). Joint modeling of VIX and SPX options at a single and common maturity with risk management applications. *IIE Trans.* **46**, 1125–1131.
- [19] CONT, R. AND KOKHOLM, T. (2013). A consistent pricing model for index options and volatility derivatives. *Math. Finance* **23**, 248–274.
- [20] DE MARCO, S. (2018). Volatility derivatives in (rough) forward variance model, Slides at Stochastic Analysis and its Applications, Oaxaca, Mai 2018 Bachelier Congress, Dublin. Available at <http://webfiles.birs.ca/cmo-workshops/2018/18w5080/files/Marco.pdf>.
- [21] DE MARCO, S. AND HENRY-LABORDÈRE, P. (2015). Linking vanillas and VIX options: a constrained martingale optimal transport problem. *SIAM J. Financial Math.* **6**, 1171–1194.
- [22] EL EUCH, O. AND ROSENBAUM, M. (2019). The characteristic function of rough Heston models. *Math. Finance* **29**, 3–38.
- [23] FORDE, M., GERHOLD, S. AND SMITH, B. (2021). *Small-time VIX smile and the stationary distribution for the rough Heston model*. Preprint. Available at <https://nms.kcl.ac.uk/martin.forde>.
- [24] FOUQUE, J.-P. AND SAPORITO, Y. F. (2018). Heston stochastic vol-of-vol model for joint calibration of VIX and S&P 500 options. *Quant. Finance* **18**, 635–654.
- [25] FUKASAWA, M. (2011). Asymptotic analysis for stochastic volatility: martingale expansion. *Finance Stoch.* **15**, 635–654.
- [26] GATHERAL, J. (2008). Slides at The Fifth World Congress of the Bachelier Finance Society London, July 18, 2008. Bachelier Congress, London. Available at <https://pdfs.semanticscholar.org/9c56/d0bc9c3f75791a623daa8474c6620b8f03e1.pdf>.
- [27] GATHERAL, J., JAISSON, T. AND ROSENBAUM, M. (2018). Volatility is rough. *Quant. Finance* **18**, 933–949.
- [28] GATHERAL, J., JAISSON, T. AND ROSENBAUM, M. (2020). The quadratic rough Heston model and the joint S&P 500/VIX smile calibration problem. *Risk*.

- [29] GOUTTE, S., ISMAIL, A. AND PHAM, G. (2017). Regime-switching stochastic volatility model: estimation and calibration to VIX options. *Appl. Math. Finance* **24**, 38–75.
- [30] GUENNON, H., JACQUIER, A., ROOME, P. AND SHI, F. (2018). Asymptotic behaviour of the fractional Heston model. *SIAM J. Financial Math.* **9**, 1017–1045.
- [31] GUYON, J. (2020). The joint S&P 500/VIX smile calibration puzzle solved. *Risk*.
- [32] GUYON, J., MENEGAUX, R. AND NUTZ, M. (2017). Bounds for VIX futures given S&P 500 smiles. *Finance Stoch.* **21**, 593–630.
- [33] HORVATH, B., JACQUIER, A. AND LACOMBE, C. (2019). Asymptotic behaviour of randomised fractional volatility models. *J. Appl. Prob.* **56**, 496–523.
- [34] HORVATH, B., JACQUIER, A. AND TANKOV, P. (2020). Volatility options in rough volatility models. *SIAM J. Financial Math.* **11**, 437–469.
- [35] JACQUIER, A., MARTINI, C. AND MUGURUZA, A. (2018). On VIX futures in the rough Bergomi model. *Quant. Finance* **18**, 45–61.
- [36] JACQUIER, A., PAKKANEN, M. AND STONE, H. (2018). Pathwise large deviations for the rough Bergomi model. *J. Appl. Prob.* **55**, 1078–1092.
- [37] JACQUIER, A. AND PANNIER, A. (2022). Large and moderate deviations for stochastic Volterra systems. *Stoch. Process. Appl.* **149**, 142–187.
- [38] JACQUIER, A. AND ZURIC, Z. (2023). *Random neural networks for rough volatility*. Preprint. Available at <https://arxiv.org/abs/2305.01035>.
- [39] JAISSON, T. AND ROSENBAUM, M. (2016). Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *Ann. Appl. Prob.* **26**, 2860–2882.
- [40] KOKHOLM, T. AND STISEN, M. (2015). Joint pricing of VIX and SPX options with stochastic volatility and jump models. *J. Risk Finance* **16**, 27–48.
- [41] LACOMBE, C., MUGURUZA, A. AND STONE, H. (2021). Asymptotics for volatility derivatives in multi-factor rough volatility models. *Math. Financial Econom.* **15**, 545–577.
- [42] NUALART, D. (2006). *The Malliavin Calculus and Related Topics*. Springer, Berlin.
- [43] PACATI, C., POMPA, G. AND RENO, R. (2018). Smiling twice: the Heston ++ model. *J. Banking Finance* **96**, 185–206.
- [44] PANNIER, A. (2023). *Path-dependent PDEs for volatility derivatives*. Preprint. Available at <https://arxiv.org/abs/2311.08289>.
- [45] PANNIER, A. AND SALVI, C. (2024). *A path-dependent PDE solver based on signature kernels*. Preprint. Available at <https://arxiv.org/abs/2403.11738>.
- [46] PAPANICOLAOU, A. (2018). Extreme-strike comparisons and structural bounds for SPX and VIX options. *SIAM J. Financial Math* **9**, 401–434.
- [47] PAPANICOLAOU, A. AND SIRCAR, R. (2014). A regime-switching Heston model for VIX and S&P 500 implied volatilities. *Quant. Finance* **14**, 1811–1827.
- [48] PROTTER, P (2005). *Stochastic Integration and Differential Equations*. Springer, Berlin.
- [49] ROSENBAUM, M. AND ZHANG, J. (2022). Deep calibration of the quadratic rough Heston model. *Risk*.
- [50] VIENS, F. AND ZHANG, J. A martingale approach for fractional Brownian motions and related path-dependent PDEs. *Ann. Appl. Prob.* **29**, 3489–3540.