Modern Stochastics: Theory and Applications 11 (2) (2024) 169–194
https://doi.org/10.15559/24-VMSTA246

# Power law in Sandwiched Volterra Volatility model

**Giulia Di Nunno**[a,b], **Anton Yurchenko-Tytarenko**[a,*]

[a]*Department of Mathematics, University of Oslo, Postboks 1053 Blindern, 0316 Oslo, Norway*
[b]*Department of Business and Management Science, NHH Norwegian School of Economics, Bergen, Norway*

giulian@math.uio.no (G. Di Nunno), antony@math.uio.no (A. Yurchenko-Tytarenko)

Received: 2 November 2023, Revised: 11 January 2024, Accepted: 12 January 2024,
Published online: 23 January 2024

**Abstract** The paper presents an analytical proof demonstrating that the Sandwiched Volterra Volatility (SVV) model is able to reproduce the power-law behavior of the at-the-money implied volatility skew, provided the correct choice of the Volterra kernel. To obtain this result, the second-order Malliavin differentiability of the volatility process is assessed and the conditions that lead to explosive behavior in the Malliavin derivative are investigated. As a supplementary result, a general Malliavin product rule is proved.

**Keywords** SVV model, stochastic volatility, sandwiched process, Gaussian Volterra noise, Malliavin calculus

**2010 MSC** 91G30, 60H10, 60H35, 60G22

## 1 Introduction

One of the well-established benchmarks for evaluating option pricing models is comparing the model-generated Black–Scholes implied volatility surface $(\tau,\kappa)\mapsto\widehat{\sigma}(\tau,\kappa)$ with the empirically observed one $(\tau,\kappa)\mapsto\widehat{\sigma}_{\mathrm{emp}}(\tau,\kappa)$. In this context, $\tau$ represents the *time to maturity* and $\kappa := \log\frac{K}{e^{r\tau}S_0}$ is the *log-moneyness* with $K$ denoting the strike, $S_0$ the current price of an underlying asset and $r$ being the instantaneous interest rate. In particular, for any fixed $\tau$, the values of $\widehat{\sigma}_{\mathrm{emp}}(\tau,\kappa)$ plotted against $\kappa$ are

*Corresponding author.

known to produce convex "*smiley*" patterns with negative slopes at-the-money (i.e. when $\kappa \approx 0$). Furthermore, as reported in, e.g., [8, 16, 20] or [12, Subsection 2.2], the smile at-the-money becomes progressively steeper as $\tau \to 0$ with a rule-of-thumb behavior

$$\left| \frac{\widehat{\sigma}_{\text{emp}}(\tau,\kappa) - \widehat{\sigma}_{\text{emp}}(\tau,\kappa')}{\kappa - \kappa'} \right| \propto \tau^{-\frac{1}{2}+H}, \quad \kappa, \kappa' \approx 0, \; H \in \left(0, \frac{1}{2}\right). \tag{1}$$

The phenomenon (1) is known as the *power law* of the at-the-money implied volatility skew, and if one wants to replicate it, one may look for a model with

$$\left| \frac{\partial \widehat{\sigma}}{\partial \kappa}(\tau,\kappa) \right|_{\kappa=0} = O\left(\tau^{-\frac{1}{2}+H}\right), \quad \tau \to 0. \tag{2}$$

However, it turns out that the property (2) is not easy to obtain: for example, as discussed in [1, Section 7.1] or [23, Remark 11.3.21], classical Brownian diffusion stochastic volatility models fail to produce implied volatilities with power law (2). In the literature, (2) is usually replicated by introducing a volatility process with a very low Hölder regularity within the *rough* volatility framework popularized by Gatheral, Jaisson and Rosenbaum in their landmark paper [20]. The efficiency of this approach can be explained as follows.

- On the one hand, a theoretical result of Fukasawa [17] suggests that the volatility process cannot be Hölder continuous of a high order in continuous nonarbitrage models exhibiting the property (2). In other words, the roughness of volatility is, in some sense, a necessary condition to reproduce (2) (at least in the fully continuous setting).
- On the other hand, as proved in the seminal 2007 paper [1] by Alòs, León and Vives, the short-term explosion (2) of the implied volatility skew can be deduced from the explosion of the Malliavin derivative of volatility. In particular, the latter characteristic is exhibited by *fractional Brownian motion with* $H < 1/2$, a common driver in the rough volatility literature.

However, despite the ability to reproduce the power law (2), rough volatility models are not perfect. In particular,

– in the specific context of a fractional Brownian motion, roughness contradicts the observations [6, 14, 15, 24, 29] of long memory on the market;

– in addition, volatility processes with long memory seem to be better in replicating the shape of implied volatility for longer maturities [7, 18, 19];

– furthermore, there is no guaranteed procedure of transition between physical and pricing measures: it is not always clear whether the volatility process $\sigma = \{\sigma(t),\ t \in [0,T]\}$ hits zero and therefore the integral $\int_0^t \frac{1}{\sigma^2(s)} ds$ that is typically present in martingale densities (see, e.g., [5]) may be poorly defined;

– just like many classical Brownian stochastic volatility models (see, e.g., [2]), they may suffer from moment explosions in price, which results in complications with the pricing of some assets, quadratic hedging, and numerical methods.

For more details on rough volatility, we refer the reader to the recent review [12, Subsection 3.3.2] or the regularly updated literature list on the subject [28].

Recently, a series of papers [9–11] introduced the *Sandwiched Volterra Volatility (SVV) model* which accounts for all the problems mentioned above. More precisely, the volatility process $Y = \{Y(t),\ t \in [0, T]\}$ is assumed to follow the stochastic differential equation

$$Y(t) = y_0 + \int_0^t b\big(s, Y(s)\big) ds + Z(t)$$

driven by a general Hölder continuous Gaussian Volterra process

$$Z(t) = \int_0^t \mathcal{K}(t, s) dB(s).$$

The special part of the equation above is the drift $b$. It is assumed that there are two continuous functions $0 < \varphi < \psi$ such that for some $\varepsilon > 0$

$$\begin{aligned} b(t, y) &\geq \frac{C}{(y - \varphi(t))^\gamma}, && y \in \big(\varphi(t), \varphi(t) + \varepsilon\big), \\ b(t, y) &\leq -\frac{C}{(\psi(t) - y)^\gamma}, && y \in \big(\psi(t) - \varepsilon, \psi(t)\big). \end{aligned}$$

Such an explosive nature of the drift resembling the one in SDEs for Bessel processes (see, e.g., [27, Chapter XI]) or singular SDEs of [21] ensures that, with probability 1,

$$0 < \varphi(t) < Y(t) < \psi(t),$$

which immediately solves the moment explosion problem (see, e.g., [9, Theorem 2.6]) and allows for a transparent transition between physical and pricing measures [9, Subsection 2.2]. In addition, the flexibility in the choice of the kernel $\mathcal{K}$ *should* allow to replicate both the long memory and the power law behavior (2).

The main goal of this paper is to give the theoretical justification to the latter claim: we prove that, with the correct choice of the Volterra kernel $\mathcal{K}$, the SVV model indeed reproduces (2). In order to do that, we employ the fundamental result [1, Theorem 6.3] by Alòs, León and Vives mentioned above and check that the Malliavin derivative $DY(t)$ indeed exhibits explosive behavior. The difficulty of this approach is as follows. While the first-order Malliavin differentiability of $Y(t)$ is established in [9, Section 3] with

$$D_s Y(t) = \mathcal{K}(t, s) + \int_s^t \mathcal{K}(u, s) b'_y\big(u, Y(u)\big) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big) dv\right\} du,$$

[1, Theorem 6.3] actually demands the existence of the second-order Malliavin derivative. In principle, it is intuitively clear how this derivative should look like:

$$
\begin{aligned}
D_r D_s Y(t) &= D_r \int_s^t \mathcal{K}(u,s) b'_y\big(u, Y(u)\big) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\}du \\
&= \int_s^t \mathcal{K}(u,s) D_r\left[b'_y\big(u, Y(u)\big) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\}\right]du \\
&= \int_s^t \mathcal{K}(u,s) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\} D_r\big[b'_y\big(u, Y(u)\big)\big]du \\
&\quad + \int_s^t \mathcal{K}(u,s) b'_y\big(u, Y(u)\big) D_r\left[\exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\}\right]du \\
&= \int_s^t \mathcal{K}(u,s) b''_{yy}\big(u, Y(u)\big) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\} D_r\big[Y(u)\big]du \\
&\quad + \int_s^t \mathcal{K}(u,s) b'_y\big(u, Y(u)\big) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\} \times \\
&\qquad \times \int_u^t b''_{yy}\big(v, Y(v)\big) D_r\big[Y(v)\big]dvdu. 
\end{aligned}
\tag{3}
$$

However, justifying the computations in (3) is far from straightforward. For example, the functions $y \mapsto b'_y(t,y)$ and $y \mapsto b''_{yy}(t,y)$ demonstrate explosive behavior as $y \to \varphi(t)+$ and $y \to \psi(t)-$ for any $t \in [0,T]$. This makes it impossible to use the classical Malliavin chain rules such as [25, Proposition 1.2.3] requiring boundedness of the derivative or [25, Proposition 1.2.4] demanding the Lipschitz condition. In order to overcome this issue, we have to use some special properties of the volatility process established in [11] and tailor a version of the Malliavin chain rule specifically for our needs.

The paper is organized as follows. In Section 2, we provide some necessary details about the sandwiched volatility process $Y$. In Section 3, we prove second-order Malliavin differentiability of $Y(t)$. Finally, in Section 4, we use [1, Theorem 6.3] to determine conditions on the kernel under which the SVV model reproduces (2). In Appendix A, we gather some necessary facts from Malliavin calculus, list some of the notation and, in addition, we prove a general Malliavin product rule to fit our purposes and that we were not able to find in the literature.

## 2 Preliminaries on sandwiched processes

In this section, we gather all the necessary details about the main object of our study: the class of *sandwiched processes driven by Hölder-continuous Gaussian Volterra noises*.

Fix some $T \in (0,\infty)$ and consider a kernel $\mathcal{K} : [0,T]^2 \to \mathbb{R}$ satisfying the following assumptions.

**Assumption 1.** The kernel $\mathcal{K}$ is of Volterra type, i.e. $\mathcal{K}(t,s) = 0$ whenever $t \le s$, and

(K1) $\mathcal{K}$ is square-integrable, i.e.

$$\int_0^T \int_0^T \mathcal{K}^2(t,s)dsdt < \infty,$$

(K2) there exists $H \in (0,1)$ such that for all $\lambda \in (0,H)$ and $0 \le t_1 \le t_2 \le T$

$$\int_0^T \big(\mathcal{K}(t_2,s) - \mathcal{K}(t_1,s)\big)^2 ds \le C_\lambda |t_2 - t_1|^{2\lambda},$$

where $C_\lambda > 0$ is some constant depending on $\lambda$.

**Remark 1.** Note that items (K1) and (K2) of Assumption 1 jointly imply that

$$\sup_{t\in[0,T]} \int_0^T \mathcal{K}^2(t,s)ds < \infty. \tag{4}$$

Let $B = \{B(t),\ t \in [0,T]\}$ be a standard Brownian motion. Assumption 1 allows to define a *Gaussian Volterra process*

$$Z(t) := \int_0^t \mathcal{K}(t,s)dB(s), \quad t \in [0,T], \tag{5}$$

and, moreover, Assumption 1(K2) together with [3, Theorem 1 and Corollary 4] implies that $Z$ has a modification with Hölder continuous trajectories of any order $\lambda \in (0,H)$. In what follows, we always use this modification of $Z$: in other words, with probability 1, for any $\lambda \in (0,H)$ there exists a random variable $\Lambda = \Lambda(\lambda) > 0$ such that for all $0 \le t_1 \le t_2 \le T$

$$\big|Z(t_2) - Z(t_1)\big| \le \Lambda |t_2 - t_1|^\lambda. \tag{6}$$

Furthermore, as stated in [3, Theorem 1], the random variable $\Lambda$ from (6) can be chosen such that

$$\mathbb{E}\big[\Lambda^r\big] < \infty \quad \text{for all } r \in \mathbb{R}. \tag{7}$$

In what follows, we assume that (7) always holds.

Next, denote

$$\begin{aligned} \mathcal{D} &:= \big\{(t,y) \in [0,T] \times \mathbb{R} \mid \varphi(t) < y < \psi(t)\big\}, \\ \overline{\mathcal{D}} &:= \big\{(t,y) \in [0,T] \times \mathbb{R} \mid \varphi(t) \le y \le \psi(t)\big\}. \end{aligned} \tag{8}$$

Take $H \in (0,1)$ from Assumption 1(K2), consider two $H$-Hölder continuous functions $\varphi, \psi\colon [0,T] \to \mathbb{R}$ such that

$$0 < \varphi(t) < \psi(t) \quad \text{for all } t \in [0,T],$$

and define a function $b\colon \mathcal{D} \to \mathbb{R}$ as

$$b(t,y) := \frac{\theta_1(t)}{(y - \varphi(t))^{\gamma_1}} - \frac{\theta_2(t)}{(\psi(t) - y)^{\gamma_2}} + a(t,y), \tag{9}$$

where the coefficients in (9) satisfy the following assumption.

**Assumption 2.** The constants $\gamma_1, \gamma_2 > 0$ and functions $\theta_1, \theta_2, a$ are such that

(B1) $\gamma_1 > \frac{1}{H} - 1$, $\gamma_2 > \frac{1}{H} - 1$ with $H \in (0, 1)$ being from Assumption 1(K2);

(B2) the functions $\theta_1, \theta_2$: $[0, T] \to \mathbb{R}$ are strictly positive and continuous;

(B3) the function $a$: $[0, T] \times \mathbb{R} \to \mathbb{R}$ is locally Lipschitz in $y$ uniformly in $t$, i.e. for any $N > 0$ there exists a constant $C_N > 0$ that does not depend on $t$ such that

$$\left|a(t, y_2) - a(t, y_1)\right| \le C_N |y_2 - y_1|, \quad t \in [0, T],\ y_1, y_2 \in [-N, N];$$

(B4) $a$: $[0, T] \times \mathbb{R} \to \mathbb{R}$ is two times differentiable w.r.t. the spatial variable $y$ with $a, a'_y, a''_{yy}$ all being continuous on $[0, T] \times \mathbb{R}$.

**Remark 2.** Note that $b'_y$ is bounded from above on $\mathcal{D}$: indeed,

$$\begin{aligned} b'_y(t, y) &= -\frac{\gamma_1\theta_1(t)}{(y - \varphi(t))^{\gamma_1+1}} - \frac{\gamma_2\theta_2(t)}{(\psi(t) - y)^{\gamma_2+1}} + a'_y(t, y) \\ &< \max_{(t,y)\in\overline{\mathcal{D}}} a'_y(t, y) < \infty. \end{aligned}$$

Finally, fix $\varphi(0) < y_0 < \psi(0)$ and consider a stochastic differential equation of the form

$$Y(t) = y_0 + \int_0^t b\big(s, Y(s)\big)ds + Z(t), \quad t \in [0, T]. \tag{10}$$

By [11, Theorem 4.1], under Assumptions 1 and 2, the SDE (10) has a unique strong solution $Y = \{Y(t),\ t \in [0, T]\}$. Moreover, with probability 1,

$$\varphi(t) < Y(t) < \psi(t) \quad \text{for all } t \in [0, T]. \tag{11}$$

**Remark 3.** Motivated by the property (11), we will call the solution $Y$ of (10) a *sandwiched* process.

In what follows, we will need to analyze the behavior of the stochastic processes $|b(t, Y(t))|$, $|b'_y(t, Y(t))|$ and $|b''(t, Y(t))|$, $t \in [0, T]$. In this regard, the property (11) alone is not sufficient: the process $Y$ can, in principle, approach the bounds $\varphi$ and $\psi$ which results in an explosive growth of the processes mentioned above. Luckily, [11, Theorem 4.2] provides a refinement of (11) allowing for a more precise control of $Y$ near $\varphi$ and $\psi$. We give a slightly reformulated version of this result below.

**Theorem 1.** *Let Assumptions 1 and 2 hold and $\lambda \in (0, H)$, $\Lambda = \Lambda(\lambda) > 0$ be from (6). Then there exist deterministic constants $C_Y = C_Y(\lambda) > 0$ and $\beta = \beta(\lambda) > 0$ such that*

$$\varphi(t) + \frac{C_Y}{(1 + \Lambda)^{\beta}} \le Y(t) \le \psi(t) - \frac{C_Y}{(1 + \Lambda)^{\beta}} \quad \textit{for all } t \in [0, T].$$

*In particular, since $\Lambda$ can be chosen to have moments of all orders, for all $r \ge 0$*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \frac{1}{(Y(t) - \varphi(t))^r}\right] < \infty, \quad \mathbb{E}\left[\sup_{t\in[0,T]} \frac{1}{(\psi(t) - Y(t))^r}\right] < \infty.$$

We finalize this section by citing the first-order Malliavin differentiability result for the sandwiched process (10) proved in [9, Section 3].

**Theorem 2.** *Let Assumptions 1 and 2 hold and $Y$ be the sandwiched process given by (10). Then, for any $t \in [0,T]$, $Y(t) \in \mathbb{D}^{1,2}$ and, with probability 1, for a.a. $s \in [0,T]$*

$$D_s Y(t) = \mathcal{K}(t,s) + \int_s^t \mathcal{K}(u,s) b'_y\big(u, Y(u)\big) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big) dv\right\} du. \tag{12}$$

**Remark 4.** The result above actually holds for more general drifts than the one given in (9). The same is also, in principle, true for the results of the subsequent sections. Namely, it would be sufficient to assume that there exist deterministic constants $c > 0$, $r > 0$, $\gamma > \frac{1}{H} - 1$ and $0 < y_* < \max_{t\in[0,T]} |\psi(t) - \varphi(t)|$ such that

- $b\colon \mathcal{D} \to \mathbb{R}$ is continuous on $\mathcal{D}$ and has continuous partial derivatives $b'_y, b''_{yy}$;
- for any $0 < \varepsilon < \frac{1}{2}\max_{t\in[0,T]} |\psi(t) - \varphi(t)|$,

  $$\big|b(t,y_2) - b(t,y_1)\big| \le \frac{c}{\varepsilon^r}|y_2 - y_1|, \quad t \in [0,T],\ \varphi(t) + \varepsilon \le y_1 \le y_2 \le \psi(t) - \varepsilon;$$

- $b$ has an explosive growth to $\infty$ near $\varphi$ and explosive decay to $-\infty$ near $\psi$ of order $\gamma > \frac{1}{H} - 1$, i.e.

  $$\begin{aligned} b(t,y) &\ge \frac{c}{(y - \varphi(t))^\gamma}, && y \in \big(\varphi(t), \varphi(t) + y_*\big),\\ b(t,y) &\le -\frac{c}{(\psi(t) - y)^\gamma}, && y \in \big(\psi(t) - y_*, \psi(t)\big); \end{aligned}$$

- for all $(t,y) \in \mathcal{D}$, the partial derivatives $b'_y$ and $b''_{yy}$ satisfy

  $$-C\left(1 + \frac{c}{(y - \varphi(t))^r} + \frac{c}{(\psi(t) - y)^r}\right) < b'_y(t,y) < C$$

  and

  $$|b''_{yy}| \le C\left(1 + \frac{c}{(y - \varphi(t))^r} + \frac{c}{(\psi(t) - y)^r}\right).$$

However, since (9) is the most natural choice satisfying these assumptions, we stick to this shape for notational convenience.

## 3 Second-order Malliavin differentiability

Let Assumptions 1 and 2 hold and $Y = \{Y(t),\ t \in [0,T]\}$ be the sandwiched process defined by (10) with the drift (9).

**Notation.** Here and in the sequel, $C$ will denote any positive deterministic constant the exact value of which is not relevant. Note that $C$ may change from line to line (or even within one line).

The main goal of this section is to establish the second-order Malliavin differentiability of the sandwiched process (10) and compute the corresponding derivative explicitly. As mentioned above, the main difficulty lies in controlling the behavior of $b(t, Y(t))$, $b'_y(t, Y(t))$ and $b''_{yy}(t, Y(t))$ whenever $Y(t)$ approaces the bounds. Luckily, Theorem 1 gives all the necessary tools to do that as summarized in the following proposition.

**Proposition 1.** *There exists a random variable $\xi > 0$ such that*

- *for any $p \geq 1$, $\mathbb{E}[\xi^p] < \infty$;*
- *for any $t \in [0, T]$,*

$$\big|b\big(t, Y(t)\big)\big| + \big|b'_y\big(t, Y(t)\big)\big| + \big|b''_{yy}\big(t, Y(t)\big)\big| < \xi.$$

*In particular, for any $p \geq 1$,*

$$\mathbb{E}\Big[\sup_{t\in[0,T]}\Big(\big|b\big(t, Y(t)\big)\big|^p + \big|b'_y\big(t, Y(t)\big)\big|^p + \big|b''_{yy}\big(t, Y(t)\big)\big|^p\Big)\Big] < \infty.$$

**Proof.** Fix $\lambda \in (0, H)$ and take the corresponding $\Lambda > 0$ from (6) and $C_Y, \beta > 0$ being from Theorem 1. Then

$$\begin{aligned}
\big|b\big(t, Y(t)\big)\big| &= \frac{|\theta_1(t)|}{(Y(t) - \varphi(t))^{\gamma_1}} + \frac{|\theta_2(t)|}{(\psi(t) - Y(t))^{\gamma_2}} + \big|a\big(t, Y(t)\big)\big| \\
&\leq \frac{\sup_{t\in[0,T]} |\theta_1(t)|(1 + \Lambda)^{\beta\gamma_1}}{C_Y^{\gamma_1}} \\
&\quad + \frac{\sup_{t\in[0,T]} |\theta_2(t)|(1 + \Lambda)^{\beta\gamma_2}}{C_Y^{\gamma_2}} \\
&\quad + \sup_{(t,y)\in\mathcal{D}} \big|a(t, y)\big| \\
&:= \xi_0, \\
\big|b'_y\big(t, Y(t)\big)\big| &= \frac{\gamma_1|\theta_1(t)|}{(Y(t) - \varphi(t))^{\gamma_1+1}} + \frac{\gamma_2|\theta_2(t)|}{(\psi(t) - Y(t))^{\gamma_2+1}} + \big|a'_y\big(t, Y(t)\big)\big| \\
&\leq \frac{\gamma_1 \sup_{t\in[0,T]} |\theta_1(t)|(1 + \Lambda)^{\beta(\gamma_1+1)}}{C_Y^{\gamma_1+1}} \\
&\quad + \frac{\gamma_2 \sup_{t\in[0,T]} |\theta_2(t)|(1 + \Lambda)^{\beta(\gamma_2+1)}}{C_Y^{\gamma_2+1}} \\
&\quad + \sup_{(t,y)\in\mathcal{D}} \big|a'_y(t, y)\big| \\
&:= \xi_1, \\
\big|b''_{yy}\big(t, Y(t)\big)\big| &= \frac{\gamma_1(\gamma_1 + 1)|\theta_1(t)|}{(Y(t) - \varphi(t))^{\gamma_1+2}} + \frac{\gamma_2(\gamma_2 + 1)|\theta_2(t)|}{(\psi(t) - Y(t))^{\gamma_2+2}} + \big|a''_{yy}\big(t, Y(t)\big)\big| \\
&\leq \frac{\gamma_1(\gamma_1 + 1) \sup_{t\in[0,T]} |\theta_1(t)|(1 + \Lambda)^{\beta(\gamma_1+2)}}{C_Y^{\gamma_1+2}}
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{\gamma_2(\gamma_2+1)\sup_{t\in[0,T]}|\theta_2(t)|(1+\Lambda)^{\beta(\gamma_2+2)}}{C_Y^{\gamma_2+2}} \\
&+ \sup_{(t,y)\in\mathcal{D}} \left|a''_{yy}(t,y)\right| \\
:= \xi_2.
\end{aligned}
$$

Note that $\xi_0$, $\xi_1$ and $\xi_2$ have moments of all orders by the properties of $\Lambda$, see (7), and hence, putting

$$\xi := \xi_0 + \xi_1 + \xi_2,$$

we obtain the required result. $\square$

As noted in Theorem 2, $Y(t) \in \mathbb{D}^{1,2}$ for each $t \geq 0$. In fact, Proposition 1 together with the shape (12) of the derivative allows to establish a more general result.

**Proposition 2.** *For any $t \in [0,T]$ and $p > 1$, $Y(t) \in \mathbb{D}^{1,p}$.*

**Proof.** Note that, by (11), $\mathbb{E}[|Y(t)|^p] < \infty$ for any $p > 1$, so, by Lemma 1 from the Appendix, it is sufficient to prove that

$$\mathbb{E}\left[\left(\int_0^T \left(D_s Y(t)\right)^2 ds\right)^{\frac{p}{2}}\right] < \infty$$

for any $p > 1$. Note that, by Remark 2,

$$\exp\left\{\int_s^t b'_y\left(v, Y(v)\right) dv\right\} < \exp\{cT\},$$

where

$$c := \max_{(t,y)\in\overline{\mathcal{D}}} a'_y(t,y),$$

and, by Proposition 1, there exists a random variable $\xi$ having all moments such that

$$\sup_{s\in[0,T]} \left|b'_y\left(s, Y(s)\right)\right| \leq \xi.$$

Hence

$$
\begin{aligned}
&\left|D_s Y(t)\right| \\
&\quad \leq \left|\mathcal{K}(t,s)\right| + \int_s^t \left|\mathcal{K}(u,s)\right|\left|b'_y\left(u, Y(u)\right)\right| \exp\left\{\int_u^t b'_y\left(v, Y(v)\right) dv\right\} du \\
&\quad \leq \left|\mathcal{K}(t,s)\right| + \xi \exp\{cT\} \int_s^t \left|\mathcal{K}(u,s)\right| du. \qquad (13)
\end{aligned}
$$

By Assumption 1 and Remark 1,

$$\left(\int_0^T \mathcal{K}^2(t,s) ds\right)^{\frac{p}{2}} < \infty,$$

therefore

$$
\begin{aligned}
&\mathbb{E}\left[\left(\int_0^T \big(D_s Y(t)\big)^2 ds\right)^{\frac{p}{2}}\right] \\
&\quad \le C\left(\int_0^T \mathcal{K}^2(t,s)ds\right)^{\frac{p}{2}} \\
&\qquad + C\mathbb{E}\left[\left(\int_0^T \int_0^t \mathcal{K}^2(u,s)\big(b'_y(u,Y(u))\big)^2 \exp\left\{2\int_u^t b'_y\big(v,Y(v)\big)dv\right\}duds\right)^{\frac{p}{2}}\right] \\
&\quad \le C\left(\int_0^T \mathcal{K}^2(t,s)ds\right)^{\frac{p}{2}} + C\mathbb{E}\big[\xi^p\big]\exp\{pcT\}\left(\int_0^T \int_0^t \mathcal{K}^2(u,s)duds\right)^{\frac{p}{2}} \\
&\quad < \infty,
\end{aligned}
\tag{14}
$$

which ends the proof. □

Our next goal is to establish the Malliavin chain rule for the random variables $b'_y(t, Y(t))$ and $\exp\{\int_u^t b'_y(v, Y(v))dv\}$.

**Proposition 3.** *For any* $0 \le u \le t \le T$ *and* $p > 1$*,*

*1)* $b'_y(t, Y(t)) \in \mathbb{D}^{1,p}$ *with*

$$
D_s\big[b'_y\big(t, Y(t)\big)\big] = b''_{yy}\big(t, Y(t)\big)D_s Y(t),
\tag{15}
$$

*2)* $\exp\{\int_u^t b'_y(v, Y(v))dv\} \in \mathbb{D}^{1,p}$ *with*

$$
\begin{aligned}
&D_s\left[\exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\}\right] \\
&\qquad = \exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\}\int_u^t b''_{yy}\big(v, Y(v)\big)D_s Y(v)dv.
\end{aligned}
\tag{16}
$$

**Proof.** 1) We shall start from proving that $b'_y(t, Y(t)) \in \mathbb{D}^{1,p}$. Note that $b'_y$ is not a bounded function itself and it does not have bounded derivatives – hence the classical chain rule from [25, Section 1.2] cannot be applied here in a straightforward manner. In order to overcome this issue, we will use the approach in the spirit of [26, Lemma A.1] or [9, Proposition 3.4]. For the reader's convenience, we divide the proof into steps.

**Step 0.** First of all, observe that $b'(t, Y(t)) \in L^2(\Omega)$ as a direct consequence of Proposition 1. Also, for any $p > 1$,

$$
\mathbb{E}\left[\left(\int_0^T \big(b''_{yy}\big(t, Y(t)\big)D_s Y(t)\big)^2 ds\right)^{\frac{p}{2}}\right] < \infty.
$$

Indeed, again by Proposition 1 together with the proof of Proposition 2, we have

$$
\mathbb{E}\left[\left(\int_0^T \big(b''_{yy}\big(t, Y(t)\big)D_s Y(t)\big)^2 ds\right)^{\frac{p}{2}}\right]
$$

$$\leq \mathbb{E}\left[\xi^p\left(\int_0^T \big(D_sY(t)\big)^2 ds\right)^{\frac{p}{2}}\right]$$
$$\leq \left(\mathbb{E}\left[\xi^{2p}\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\left(\int_0^T \big(D_sY(t)\big)^2 ds\right)^p\right]\right)^{\frac{1}{2}}$$
$$< \infty.$$

Therefore, by Lemma 1, it is sufficient to prove that $b'_y(t, Y(t)) \in \mathbb{D}^{1,2}$ with (15) being the corresponding Malliavin derivative.

**Step 1.** Let $\phi \in C^1(\mathbb{R})$ be a compactly supported function such that $\phi(x) = x$ whenever $|x| \leq 1$ and $|\phi(x)| \leq |x|$ for all $|x| > 1$. Fix $t \in [0, T]$ and, for $m \geq 1$, put

$$f_m(y) := m\phi\left(\frac{b'_y(t, y)}{m}\right).$$

Observe that

$$f'_m(y) = b''_{yy}(t, y)\phi'\left(\frac{b'_y(t, y)}{m}\right)$$

is bounded. Indeed, let $0 < \varepsilon_m < \psi(t) - \varphi(t)$ be such that

$$-\frac{\gamma_1\theta_1(t)}{\varepsilon_m^{\gamma_1+1}} + \max_{\varphi(t)\leq x\leq\psi(t)} a'_y(t, x) < m \inf \operatorname{supp} \phi$$

and

$$-\frac{\gamma_2\theta_2(t)}{\varepsilon_m^{\gamma_2+1}} + \max_{\varphi(t)\leq x\leq\psi(t)} a'_y(t, x) < m \inf \operatorname{supp} \phi.$$

Then,

- if $y \in (\varphi(t), \varphi(t) + \varepsilon_m)$, then
$$\begin{aligned} b'_y(t, y) &= -\frac{\gamma_1\theta_1(t)}{(y - \varphi(t))^{\gamma_1+1}} - \frac{\gamma_2\theta_2(t)}{(\psi(t) - y)^{\gamma_2+1}} + a'_y(t, y) \\ &\leq -\frac{\gamma_1\theta_1(t)}{\varepsilon_m^{\gamma_1+1}} + \max_{\varphi(t)\leq x\leq\psi(t)} a'_y(t, x) \\ &< m \inf \operatorname{supp} \phi, \end{aligned}$$
  so $\frac{b'_y(t,y)}{m} \notin \operatorname{supp} \phi$, $f_m(y) = 0$ and $f'_m(y) = 0$;
- if $y \in (\psi(t) - \varepsilon_m, \psi(t))$, then, similarly,
$$\begin{aligned} b'_y(t, y) &= -\frac{\gamma_1\theta_1(t)}{(y - \varphi(t))^{\gamma_1+1}} - \frac{\gamma_2\theta_2(t)}{(\psi(t) - y)^{\gamma_2+1}} + a'_y(t, y) \\ &\leq -\frac{\gamma_2\theta_2(t)}{\varepsilon_m^{\gamma_2+1}} + \max_{\varphi(t)\leq x\leq\psi(t)} a'_y(t, x) \\ &< m \inf \operatorname{supp} \phi, \end{aligned}$$
  so $\frac{b'_y(t,y)}{m} \notin \operatorname{supp} \phi$, $f_m(y) = 0$ and $f'_m(y) = 0$;

- on the compact set $[\varphi(t) + \varepsilon_m, \psi(t) - \varepsilon_m]$, both $f_m$ and its derivative $f_m'$ are continuous and hence bounded.

Therefore, the function $f_m$ satisfies the conditions of the classical Malliavin chain rule [25, Proposition 1.2.3], so $f_m(Y(t)) \in \mathbb{D}^{1,2}$ and, with probability 1 for a.a. $s \in [0, T]$,

$$D_s f_m\big(Y(t)\big) = b''_{yy}\big(t, Y(t)\big)\phi'\left(\frac{b'_y(t, Y(t))}{m}\right) D_s Y(t).$$

Now it remains to prove that

$$f_m\big(Y(t)\big) \to b'\big(t, Y(t)\big)$$

in $L^2(\Omega)$ and

$$Df_m\big(Y(t)\big) \to b''_{yy}\big(t, Y(t)\big) DY(t)$$

in $L^2(\Omega \times [0, T])$ as $m \to \infty$; then the result will follow immediately from the closedness of the Malliavin derivative operator $D$.

**Step 2:** $f_m(Y(t)) \to b'(t, Y(t))$ in $L^2(\Omega)$ as $m \to \infty$. By the definitions of $f_m$ and $\phi$, $f_m(Y(t)) \to b'(t, Y(t))$ a.s. as $m \to \infty$. Moreover, with probability 1, $|f_m(Y(t))| \le |b'_y(t, Y(t))| \in L^2(\Omega)$ and hence the required convergence follows from the dominated convergence theorem.

**Step 3:** $Df_m(Y(t)) \to b''_{yy}(t, Y(t))DY(t)$ in $L^2(\Omega \times [0, T])$ as $m \to \infty$. By the definitions of $f_m$ and $\phi$, with probability 1,

$$\left(b''_{yy}\big(t, Y(t)\big)\phi'\left(\frac{b'_y(t, Y(t))}{m}\right)\right)^2 \int_0^T \big(D_s Y(t)\big)^2 ds$$
$$\to \big(b''_{yy}\big(t, Y(t)\big)\big)^2 \int_0^T \big(D_s Y(t)\big)^2 ds$$

as $m \to \infty$. Moreover, since $\phi$ has compact support, $\max_{y\in\mathbb{R}}(\phi'(y))^2 < \infty$, so we can write

$$\int_0^T \big(D_s f_m\big(Y(t)\big)\big)^2 ds = \left(b''_{yy}\big(t, Y(t)\big)\phi'\left(\frac{b'_y(t, Y(t))}{m}\right)\right)^2 \int_0^T \big(D_s Y(t)\big)^2 ds$$
$$\le \max_{y\in\mathbb{R}}\big(\phi'(y)\big)^2 \big(b''_{yy}\big(t, Y(t)\big)\big)^2 \int_0^T \big(D_s Y(t)\big)^2 ds \in L^2(\Omega).$$

Therefore, by the dominated convergence theorem,

$$\mathbb{E}\left[\int_0^T \big(D_s f_m\big(Y(t)\big) - b''_{yy}\big(t, Y(t)\big)D_s Y(t)\big)^2 ds\right] \to 0, \quad m \to \infty,$$

which proves the first claim of the Proposition.

2) Let us proceed with the second claim and verify that

$$\exp\left\{\int_u^t b'_y\big(v, Y(v)\big) dv\right\} \in \mathbb{D}^{1,p}$$

with (16) being the corresponding Malliavin derivative. Note that, since $b'_y$ is bounded from above, $\exp\{\int_u^t b'_y(v, Y(v))dv\}$ is also bounded from above and hence is an element of $L^p(\Omega)$ for any $p > 1$. Moreover, by Proposition 1, boundedness of $\exp\{\int_u^t b'_y(v, Y(v))dv\}$ and (13), we can write

$$\begin{aligned}
&\mathbb{E}\Bigg[\Bigg(\int_0^T \Bigg(\exp\Bigg\{\int_u^t b'_y\big(v, Y(v)\big)dv\Bigg\}\int_u^t b''_{yy}\big(v, Y(v)\big)D_s Y(v)dv\Bigg)^2 ds\Bigg)^{\frac{p}{2}}\Bigg] \\
&\quad \le C\mathbb{E}\Bigg[\xi^p\Bigg(\int_0^T \int_u^t \big(D_s Y(v)\big)^2 dvds\Bigg)^{\frac{p}{2}}\Bigg] \\
&\quad \le C\mathbb{E}\Bigg[\xi^p\Bigg(\int_0^T \int_u^t \mathcal{K}^2(v, s)dvds\Bigg)^{\frac{p}{2}}\Bigg] \\
&\qquad + C\exp\{pcT\}\mathbb{E}\big[\xi^{2p}\big]\Bigg(\int_0^T \int_u^t \int_s^v \mathcal{K}^2(u, s)dudvds\Bigg)^{\frac{p}{2}} \\
&\quad < \infty,
\end{aligned}$$

and hence it is sufficient to prove that $\exp\{\int_u^t b'_y(v, Y(v))dv\} \in \mathbb{D}^{1,2}$.

Since the Malliavin derivative operator $D$ is closed and the expression $\int_u^t b''_{yy}(v, Y(v))D_s Y(v)dv$ is well-defined by Proposition 1, Step 1 of the current proof and Hille's theorem [22, Theorem 1.2.4] guarantee that

$$\int_u^t b'_y\big(v, Y(v)\big)Y(v)dv \in \mathbb{D}^{1,2}$$

and

$$D_s \int_u^t b'_y\big(v, Y(v)\big)Y(v)dv = \int_u^t b''_{yy}\big(v, Y(v)\big)D_s Y(v)dv.$$

Finally, the function $x \mapsto e^x$ satisfies the conditions of the chain rule from [9, Proposition 3.4] and hence $\exp\{\int_u^t b'_y(v, Y(v))dv\} \in \mathbb{D}^{1,2}$ and (16) holds. $\square$

Proposition 3 and Lemma 2 together allow us to deduce the following corollary.

**Corollary 1.** *For any* $0 \le s < t \le T$ *and* $p > 1$,

$$b'_y\big(s, Y(s)\big)\exp\Bigg\{\int_s^t b'_y\big(v, Y(v)\big)dv\Bigg\} \in \mathbb{D}^{1,p}$$

*and*

$$\begin{aligned}
&D_u\Bigg[b'_y\big(s, Y(s)\big)\exp\Bigg\{\int_s^t b'_y\big(v, Y(v)\big)dv\Bigg\}\Bigg] \\
&\quad = b''_{yy}\big(s, Y(s)\big)\exp\Bigg\{\int_s^t b'_y\big(v, Y(v)\big)dv\Bigg\}D_u Y(s) \\
&\qquad + b'_y\big(s, Y(s)\big)\exp\Bigg\{\int_s^t b'_y\big(v, Y(v)\big)dv\Bigg\}\int_s^t b''_{yy}\big(v, Y(v)\big)D_u Y(v)dv. \qquad (17)
\end{aligned}$$

**Proof.** For fixed $0 \le s < t \le T$, denote

$$X_1 := b'_y\big(s, Y(s)\big), \quad X_2 := \exp\left\{\int_s^t b'_y\big(v, Y(v)\big)dv\right\}.$$

By Proposition 3 and Lemma 2 from the Appendix, it is sufficient to check that for all $p \ge 2$

(i) the product $X_1 X_2 \in L^p(\Omega)$,

(ii) $\mathbb{E}[(\int_0^T (X_2 D_u X_1)^2)^{\frac{p}{2}}] < \infty$ and

(iii) $\mathbb{E}[(\int_0^T (X_1 D_u X_2)^2)^{\frac{p}{2}}] < \infty$.

All conditions (i)–(iii) can be checked in a straightforward manner using Proposition 1 and the arguments similar to the proof of Proposition 2. □

We are now ready to formulate the main result of this section.

**Theorem 3.** *For any $t \in [0, T]$ and $p \ge 2$,*

*1) $Y(t) \in \mathbb{D}^{2,p}$,*

*2) with probability 1 and for a.a. $r, s \in [0, T]$,*

$$D_r D_s Y(t) = \int_s^t \mathcal{K}(u, s) F_1(t, u)\left(\int_u^t b''_{yy}\big(v, Y(v)\big) D_r Y(v) dv\right) du$$
$$+ \int_s^t \mathcal{K}(u, s) F_2(t, u) D_r Y(u) du, \tag{18}$$

*where*

$$F_1(t, u) := b'_y\big(u, Y(u)\big) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\},$$
$$F_2(t, u) := b''_{yy}\big(u, Y(u)\big) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\}.$$

**Proof.** Our goal is to prove that $Y(t) \in \mathbb{D}^{2,p}$ and

$$D_r D_s Y(t) = \int_s^t \mathcal{K}(u, s) D_r\left[b'_y\big(u, Y(u)\big) \exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\}\right] du$$
$$= \int_s^t \mathcal{K}(u, s) D_r\big[F_1(t, u)\big] du,$$

since, in such case, (18) follows immediately from Corollary 1. Recall that

$$D_s Y(t) = \mathcal{K}(t, s) + \int_s^t \mathcal{K}(u, s) F_1(t, u) du.$$

Clearly, for any $0 \le r, s < t \le T$,

$$D_r \mathcal{K}(t, s) = 0,$$

so, by closedness of $D$ and Hille's theorem [22, Theorem 1.2.4], it is enough to show that

(i) for a.a. $0 \le s \le u < t \le T$, $\mathcal{K}(u,s)F_1(t,u) \in \mathbb{D}^{1,p}$ and

(ii) for a.a. $0 \le s < t \le T$,

$$\begin{aligned}
&\int_0^T \left(\mathbb{E}\left[\left(\int_0^T \left(D_r\left[\mathcal{K}(u,s)F_1(t,u)\right]\right)^2 dr\right)^{\frac{p}{2}}\right]\right)^{\frac{1}{p}} du \\
&\quad = \int_0^T \mathcal{K}(u,s)\left(\mathbb{E}\left[\left(\int_0^T \left(D_r\left[F_1(t,u)\right]\right)^2 dr\right)^{\frac{p}{2}}\right]\right)^{\frac{1}{p}} du \\
&\quad < \infty.
\end{aligned}$$

Item (i) above follows immediately from Corollary 1. As for item (ii), observe that, by Proposition 1, (13) as well as the boundedness of $\exp\{\int_u^t b'_y(v, Y(v))dv\}$, we have

$$\begin{aligned}
\left(D_r\left[F_1(t,u)\right]\right)^2 &\le C\Bigg(\left(b''_{yy}\left(u, Y(u)\right)\right)^2 \exp\left\{2\int_u^t b'_y\left(v, Y(v)\right)dv\right\}\left(D_r Y(u)\right)^2 \\
&\qquad + \left(b'_y\left(u, Y(u)\right)\right)^2 \exp\left\{2\int_u^t b'_y\left(v, Y(v)\right)dv\right\}\times \\
&\qquad \times \int_u^t \left(b''_{yy}\left(v, Y(v)\right)D_r Y(v)\right)^2 dv\Bigg) \\
&\le C\left(\xi^2\left(D_r Y(u)\right)^2 + \xi^4\int_u^t \left(D_r Y(v)\right)^2 dv\right) \\
&\le C\xi^2\left(\mathcal{K}^2(u,r) + \int_r^u \mathcal{K}^2(z,r)dz\right) \\
&\qquad + C\xi^4\left(\int_u^t \mathcal{K}^2(v,r)dv + \int_u^t\int_r^v \mathcal{K}^2(z,r)dzdv\right).
\end{aligned}$$

Hence, for any $p \ge 2$, Remark 1 implies

$$\begin{aligned}
\int_0^T \left(D_r\left[F_1(t,u)\right]\right)^2 dr &\le C\xi^2\left(\int_0^T \mathcal{K}^2(u,r)dr + \int_0^T\int_r^u \mathcal{K}^2(z,r)dzdr\right) \\
&\qquad + C\xi^4\int_0^T\int_u^t \mathcal{K}^2(v,r)dvdr \\
&\qquad + C\xi^4\int_0^T\int_u^t\int_r^v \mathcal{K}^2(z,r)dzdvdr \\
&\le C\left(\xi^2 + \xi^4\right),
\end{aligned}$$

so, since $\xi$ has moments of all orders, (ii) holds, which finalizes the proof. □

Finally, denote $\mathbb{L}^{2,p} := L^p([0,T]; \mathbb{D}^{2,p})$. We complete the section with the following result.

**Corollary 2.** *For any* $p \ge 2$, $Y \in \mathbb{L}^{2,p}$.

**Proof.** By the definition of the $\|\cdot\|_{2,p}$-norm in (32) from Appendix A, it is sufficient to check that

$$\int_0^T \mathbb{E}\big[|Y(t)|^p\big] < \infty, \tag{19}$$

$$\int_0^T \mathbb{E}\left[\left(\int_0^T \big(D_s Y(t)\big)^2 ds\right)^{\frac{p}{2}}\right] dt < \infty \tag{20}$$

and

$$\int_0^T \mathbb{E}\left[\left(\int_0^T \int_0^T \big(D_r D_s Y(t)\big)^2 ds dr\right)^{\frac{p}{2}}\right] dt < \infty. \tag{21}$$

By (11), (19) holds automatically. Next, (20) can be easily deduced from (14). Finally, using Proposition (1) and the boundedness of $\exp\{\int_u^t b'_y(v, Y(v))dv\}$, it is easy to prove a bound similar to (14) for

$$\mathbb{E}\left[\left(\int_0^T \int_0^T \big(D_r D_s Y(t)\big)^2 ds dr\right)^{\frac{p}{2}}\right],$$

which implies (21). By this, the proof is complete. □

## 4 Power law in SVV model

Having the second-order Malliavin differentiability in place, we now possess all the necessary tools to analyze the behavior of implied volatility skew of a model with the sandwiched process (10) as stochastic volatility. Namely, we consider a (risk-free) market model with the price process $S = \{S(t),\ t \in [0, T]\}$ of the form

$$\begin{aligned} S(t) &= e^{X(t)}, \\ X(t) &= x_0 + rt - \frac{1}{2}\int_0^t Y^2(s)ds + \int_0^t Y(s)\big(\rho dB_1(s) + \sqrt{1-\rho^2} dB_2(s)\big), \\ Y(t) &= y_0 + \int_0^t b\big(s, Y(s)\big)ds + \int_0^t \mathcal{K}(t,s)dB_1(s), \end{aligned} \tag{22}$$

where $B_1$, $B_2$ are two independent Brownian motions, $X = \{X(t), t \in [0, T]\}$ denotes the (risk-free) log-price of an asset starting from some level $x_0 \in \mathbb{R}$, $r$ is a constant instantaneous interest rate, and $\rho \in (-1, 1)$ is a correlation coefficient that accounts for the leverage effect. As previously, the drift $b$ and the Volterra kernel $\mathcal{K}$ satisfy Assumptions 1 and 2.

**Remark 5.** The model (22) was initially introduced in [9] and, given the nature of the volatility process, is called the *Sandwiched Volterra Volatility* (SVV) model.

The goal of this section is to establish conditions under which (22) reproduces the power law (2) of the short-term at-the-money implied volatility. Namely, we have the following result.

**Theorem 4.** *Let Assumptions 1 and 2 hold with $H \in (\frac{1}{6}, \frac{1}{2})$. Assume that the Volterra kernel $\mathcal{K}$ is such that, for any $0 \le s < t \le T$,*

$$\left|\mathcal{K}(t,s)\right| \le C|t-s|^{-\frac{1}{2}+H}$$

*for some constant $C > 0$, and*

$$\frac{1}{\tau^{\frac{3}{2}+H}} \int_0^\tau \int_s^\tau \mathcal{K}(t,s)dtds \to K_Y, \quad \tau \to 0+, \tag{23}$$

*for some finite constant $K_Y$. Then, with probability 1, the SVV implied volatility $\widehat{\sigma}$ exhibits the property*

$$\lim_{\tau \to 0} \tau^{\frac{1}{2}-H} \frac{\partial \widehat{\sigma}}{\partial \kappa}(\tau,\kappa)\bigg|_{\kappa=0} = \frac{\rho}{y_0} K_Y.$$

*In particular, if $\rho K_Y \neq 0$, the SVV model (22) reproduces the power law (2) of the at-the-money implied volatility skew.*

**Remark 6.** The behavior of empirically observed implied volatilities (see, e.g., [12]) shows that realistic market models should produce $\widehat{\sigma}$ with

$$\frac{\partial \widehat{\sigma}}{\partial \kappa}(\tau,\kappa)\bigg|_{\kappa=0} < 0. \tag{24}$$

In the SVV setting (22), Theorem 4 guarantees that (24) holds for all small enough $\tau$ provided that $\rho K_Y < 0$.

**Remark 7.** The condition $H > \frac{1}{6}$ in Theorem 4 is consistent with the recent empirical estimate $H \approx 0.19$ for the SPX implied volatility obtained in [8].

To prove Theorem 4, we will apply the fundamental result [1, Theorem 6.3] which connects the shape of the skew with the Malliavin derivative of the volatility.

**Remark 8.** In the recent literature (see, e.g., [4, 8, 12, 20]), it is typical to characterize the implied volatility skew in terms of $\frac{\partial \widehat{\sigma}}{\partial \kappa}$ with $\kappa = \log \frac{K}{e^{r\tau + x_0}}$ being the log-moneyness. In [1], a slightly different parametrization $\widehat{\sigma}_{\text{log-price}}(\tau, x_0)$ is considered with

$$\widehat{\sigma}_{\text{log-price}}(\tau,x) = \widehat{\sigma}\left(\tau, \log\frac{K}{e^{r\tau}} - x\right).$$

With this parametrization,

$$\frac{\partial \widehat{\sigma}_{\text{log-price}}(\tau,x)}{\partial x} = -\frac{\partial \widehat{\sigma}(\tau, \log\frac{K}{e^{r\tau}} - x)}{\partial \kappa}$$

and the power law (2) is equivalent to

$$\left|\frac{\partial \widehat{\sigma}_{\text{log-price}}}{\partial x}(\tau,x)\right|_{x=\log\frac{K}{e^{r\tau}}} = O\left(\tau^{-\frac{1}{2}+H}\right), \quad \tau \to 0.$$

With Remark 8 in mind, let us provide a slightly adjusted version of [1, Theorem 6.3].

**Theorem 5.** *Consider a risk-free log-price*

$$X(t) = x_0 + rt - \frac{1}{2}\int_0^t \sigma^2(s)ds + \int_0^t \sigma(s)\big(\rho dB_1(s) + \sqrt{1-\rho^2}dB_2(s)\big), \quad (25)$$

*where $B_1$, $B_2$ are two independent Brownian motions, $x_0 \in \mathbb{R}$ is a deterministic initial value, $r$ is an instantaneous interest rate, $\rho \in (-1, 1)$ is a correlation coefficient and $\sigma = \{\sigma(t),\ t \in [0, T]\}$ is a square-integrable stochastic process with right-continuous trajectories adapted to the filtration $\mathcal{F} = \{\mathcal{F}_t,\ t \in [0, T]\}$ generated by $B_1$.*

*Assume that*

*(H1)* $\sigma \in \mathbb{L}^{2,4}$ *with respect to $B_1$;*

*(H2) there exists a constant $\varphi_* > 0$ such that, with probability 1, $\sigma(t) > \varphi_*$ for all $t \in [0, T]$;*

*(H3) there exists a constant $H \in (0, \frac{1}{2})$ such that, with probability 1, for any $0 < s < t < T$,*

$$\mathbb{E}\big[\big(D_s\sigma(t)\big)^2\big] \le \frac{C}{(t-s)^{1-2H}}, \quad (26)$$

$$\mathbb{E}\big[\big(D_r D_s\sigma(t)\big)^2\big] \le C\left(\frac{t-r}{t-s}\right)^{1-2H}, \quad (27)$$

*where $C > 0$ is some constant;*

*(H4) $\sigma$ has a.s. right-continuous trajectories;*

*(H5)* $\sup_{r,s,t\in[0,\tau]} \mathbb{E}[(\sigma(s)\sigma(t) - \sigma^2(r))^2] \to 0$ *when $\tau \to 0+$.*

*Finally, assume that there exists a constant $K_\sigma > 0$ such that, with probability 1,*

$$\frac{1}{\tau^{\frac{3}{2}+H}}\int_0^\tau\int_s^\tau \mathbb{E}\big[D_s\sigma(t)\big]dtds - K_\sigma \to 0, \quad \tau \to 0+. \quad (28)$$

*Then, with probability 1,*

$$\lim_{\tau\to 0} \tau^{\frac{1}{2}-H}\frac{\partial\widehat{\sigma}_{\textit{log-price}}}{\partial x}(\tau, x)\bigg|_{x=\log\frac{K}{e^{r\tau}}} = -\frac{\rho}{\sigma(0)}K_\sigma.$$

**Remark 9.** The original formulation of [1, Theorem 6.3] is slightly more general than Theorem 5 above in the sense that

1) in [1, Theorem 6.3], the log-price $X$ is allowed to have jumps;

2) the result in [1] is formulated for the *future* implied volatility surfaces $\widehat{\sigma}_{\text{log-price}}(t_0, \tau, X(t_0))$, $t_0 \ge 0$.

Since we are interested in the continuous model (22), we removed the jump component in (25) and, for the simplicity of notation, we put $t_0 = 0$.

Observe that the SVV model (22) automatically satisfies a number of assumptions of Theorem 5:

- assumption (H2) with $\varphi^* := \min_{t\in[0,T]} \varphi(t) > 0$;
- assumption (H4) since $Y$ is continuous a.s.;
- assumption (H1) by the results of Section 3 above.

Therefore, it remains to check (H3), (H5), and (28). Naturally, given the shape of the Malliavin derivative (12), both (H3) and (28) require additional assumptions on the kernel, so let us start with (H5).

**Proposition 4.** *Let Assumptions 1 and 2 hold. Then with probability 1,*

$$\sup_{r,s,t\in[0,\tau]} \mathbb{E}\big[\big(Y(s)Y(t) - Y^2(r)\big)^2\big] \to 0, \quad \tau \to 0.$$

**Proof.** By [10, Lemma 3.6], there exists a positive random variable $\Upsilon = \Upsilon_T$ such that, for all $t_1, t_2 \in [0, T]$,

$$\big|Y(t_1) - Y(t_2)\big| \le \Upsilon |t_1 - t_2|^\lambda$$

and, for any $r > 0$,

$$\mathbb{E}\big[\Upsilon^r\big] < \infty.$$

Therefore, given that $\max_{t\in[0,T]} Y(t) < \max_{t\in[0,T]} \psi(t)$ by (11),

$$\begin{aligned}
&\mathbb{E}\big[\big(Y(s)Y(t) - Y^2(r)\big)^2\big] \\
&\quad = \mathbb{E}\Big[\Big(Y(s)\big(Y(t) - Y(r)\big) + Y(r)\big(Y(s) - Y(r)\big)\Big)^2\Big] \\
&\quad \le 2\mathbb{E}\big[(Y^2(s)\big(Y(t) - Y(r)\big)^2\big] + 2\mathbb{E}\big[Y^2(r)\big(Y(s) - Y(r)\big)^2\big] \\
&\quad \le 2|t-r|^{2\lambda} \max_{s\in[0,T]} \psi^2(s)\mathbb{E}\big[\Upsilon^2\big] + 2|s-r|^{2\lambda} \max_{s\in[0,T]} \psi^2(s)\mathbb{E}\big[\Upsilon^2\big]
\end{aligned}$$

and hence, with probability 1,

$$\sup_{r,s,t\in[0,\tau]} \mathbb{E}\big[\big(Y(s)Y(t) - Y^2(r)\big)^2\big] \le 4\tau^{2\lambda} \max_{s\in[0,T]} \psi^2(s)\mathbb{E}\big[\Upsilon^2\big] \to 0$$

as $\tau \to 0+$. □

Our next step is to handle (28).

**Proposition 5.** *Let Assumptions 1 and 2 hold and the Volterra kernel $\mathcal{K}$ satisfy (23) for some finite constant $K_Y$. Then, with probability 1,*

$$\frac{1}{\tau^{\frac{3}{2}+H}} \int_0^\tau \int_s^\tau \mathbb{E}\big[D_s Y(t)\big] dt ds - K_Y \to 0, \quad \tau \to 0+.$$

**Proof.** Recall that

$$F_1(t,u) := b'_y\big(u, Y(u)\big)\exp\left\{\int_u^t b'_y\big(v, Y(v)\big)dv\right\}$$

and that, by Proposition 1,

$$\big|F_1(t,u)\big| \le e^{cT}\xi, \tag{29}$$

where $c := \max_{(t,y)\in\overline{\mathcal{D}}} a'_y(t,y)$. Then we can write

$$\begin{aligned}
&\frac{1}{\tau^{\frac{3}{2}+H}}\int_0^\tau\int_s^\tau \mathbb{E}\big[D_s Y(t)\big]dtds\\
&\quad= \frac{1}{\tau^{\frac{3}{2}+H}}\int_0^\tau\int_s^\tau \mathcal{K}(t,s)dtds\\
&\qquad+ \frac{1}{\tau^{\frac{3}{2}+H}}\int_0^\tau\int_s^\tau\int_s^t \mathcal{K}(u,s)\mathbb{E}\big[F_1(t,u)\big]dudtds\\
&\quad= \frac{1}{\tau^{\frac{3}{2}+H}}\int_0^\tau\int_s^\tau \mathcal{K}(t,s)dtds\\
&\qquad+ \frac{1}{\tau^{\frac{3}{2}+H}}\int_0^\tau\int_s^\tau \mathcal{K}(u,s)\left(\int_u^\tau \mathbb{E}\big[F_1(t,u)\big]dt\right)duds.
\end{aligned}$$

The term $\frac{1}{\tau^{\frac{3}{2}+H}}\int_0^\tau\int_s^\tau \mathcal{K}(t,s)dtds$ converges to $K_Y$ by (23). As for the second term, note that, with probability 1, for any $u \in [0,\tau]$,

$$\int_u^\tau \big|\mathbb{E}\big[F_1(t,u)\big]\big|dt \le C\mathbb{E}[\xi]\tau$$

and hence, given (23), with probability 1,

$$\frac{1}{\tau^{\frac{3}{2}+H}}\int_0^\tau\int_s^\tau \mathcal{K}(u,s)\left(\int_u^\tau \mathbb{E}\big[F_1(t,u)\big]dt\right)duds \to 0, \quad \tau \to 0+,$$

which ends the proof. □

Finally, let us deal with (H3).

**Proposition 6.** *Let Assumptions 1 and 2 hold with $H \in (\frac{1}{6}, \frac{1}{2})$ and the Volterra kernel $\mathcal{K}$ be such that for any $0 \le s < t \le T$*

$$\big|\mathcal{K}(t,s)\big| \le C|t-s|^{-\frac{1}{2}+H} \tag{30}$$

*for some constant $C > 0$. Then the hypothesis (H3) from Theorem 5 holds for the sandwiched volatility process $\sigma = Y$.*

**Proof.** Fix $0 < r, s < t$. Then, taking into account (29), with probability 1,

$$\begin{aligned}\big|D_s Y(t)\big| &\le \big|\mathcal{K}(t,s)\big| + \int_s^t \big|\mathcal{K}(u,s)\big|\big|F_1(t,u)\big|du\\ &\le C\bigg(|t-s|^{-\frac{1}{2}+H} + \xi\int_s^t |u-s|^{-\frac{1}{2}+H}du\bigg)\\ &\le C(1+T\xi)|t-s|^{-\frac{1}{2}+H}\\ &=: \zeta\,|t-s|^{-\frac{1}{2}+H},\end{aligned} \tag{31}$$

which immediately implies (26). Next, by Proposition 1,

$$\big|b''_{yy}\big(v, Y(v)\big)\big| \le \xi$$

for any $v \in [0, T]$ and, for any $0 \le u \le t \le T$,

$$\begin{aligned}\big|F_2(t,u)\big| &= \bigg|b''_{yy}\big(u, Y(u)\big)\exp\bigg\{\int_u^t b'_y\big(v, Y(v)\big)dv\bigg\}\bigg|\\ &\le e^{cT}\xi\end{aligned}$$

with $c := \max_{(t,y)\in\overline{\mathcal{D}}} a'_y(t,y)$, so we can write

$$\begin{aligned}&\big|D_r D_s Y(t)\big|\\ &\quad\le \int_s^t \big|\mathcal{K}(u,s)\big|\big|F_1(t,u)\big|\bigg(\int_u^t \big|b''_{yy}\big(v, Y(v)\big)\big|\big|D_r Y(v)\big|dv\bigg)du\\ &\qquad+ \int_s^t \big|\mathcal{K}(u,s)\big|\big|F_2(t,u)\big|\big|D_r Y(u)\big|du\\ &\quad\le C\bigg(\xi^2\int_s^t \big|\mathcal{K}(u,s)\big|\bigg(\int_u^t \big|D_r Y(v)\big|dv\bigg)du + \xi\int_s^t \big|\mathcal{K}(u,s)\big|\big|D_r Y(u)\big|du\bigg)\\ &\quad= C\bigg(\xi^2\int_s^t \big|\mathcal{K}(u,s)\big|\bigg(\int_{u\vee r}^t \big|D_r Y(v)\big|dv\bigg)du + \xi\int_{r\vee s}^t \big|\mathcal{K}(u,s)\big|\big|D_r Y(u)\big|du\bigg).\end{aligned}$$

Taking into account (30) and (31),

$$\begin{aligned}\big|D_r D_s Y(t)\big| &\le C\bigg(\xi^2\zeta\int_s^t |u-s|^{-\frac{1}{2}+H}\bigg(\int_{u\vee r}^t |v-r|^{-\frac{1}{2}+H}dv\bigg)du\\ &\qquad+ \xi\zeta\int_{r\vee s}^t |u-s|^{-\frac{1}{2}+H}|u-r|^{-\frac{1}{2}+H}du\bigg)\\ &\le C\bigg(\xi^2\zeta\int_s^t |u-s|^{-\frac{1}{2}+H}|t-r|^{\frac{1}{2}+H}du\\ &\qquad+ \xi\zeta\int_{r\vee s}^t |u-s|^{-\frac{1}{2}+H}|u-r|^{-\frac{1}{2}+H}du\bigg).\end{aligned}$$

Note that

$$\int_s^t |u-s|^{-\frac{1}{2}+H}|t-r|^{\frac{1}{2}+H}du \le C|t-r|^{\frac{1}{2}+H}|t-s|^{\frac{1}{2}+H}$$

$$\le C\left(\frac{t-r}{t-s}\right)^{\frac{1}{2}-H}.$$

As for the integral $\int_{r\vee s}^t |u-s|^{-\frac{1}{2}+H}|u-r|^{-\frac{1}{2}+H}du$, we have two cases:

- if $0 < r \le s < t$, we can write

$$\int_s^t |u-s|^{-\frac{1}{2}+H}|u-r|^{-\frac{1}{2}+H}du \le \int_s^t |u-s|^{-1+2H}du$$

$$\le C(t-s)^{2H} \le C\left(\frac{t-r}{t-s}\right)^{\frac{1}{2}-H};$$

- similarly, if $0 < s < r < t$ and given that $H > \frac{1}{6}$, we have

$$\int_r^t |u-s|^{-\frac{1}{2}+H}|u-r|^{-\frac{1}{2}+H}du \le \int_r^t |u-r|^{-1+2H}du$$

$$\le C(t-r)^{2H} \le C\left(\frac{t-r}{t-s}\right)^{\frac{1}{2}-H}.$$

In any case,

$$\left|D_r D_s Y(t)\right| \le C\xi\zeta(\xi+1)\left(\frac{t-r}{t-s}\right)^{\frac{1}{2}-H},$$

where $\xi$ and $\zeta$ are random variables having all moments, and hence (27) holds. □

Having in mind all of the results above, we are ready to prove the main result of this section, namely Theorem 4.

**Proof of Theorem 4.** The results above show that the SVV model satisfies conditions (H1)–(H5) of Theorem 5. Therefore, taking into account the reparametrization described in Remark 8, Theorem 4 follows immediately from Theorem 5. □

**Example 1.** Let $\frac{1}{6} < H_0 < H_1 < \cdots < H_n < 1$ be such that $H_0 < \frac{1}{2}$ and $\alpha_k > 0$, $k = 0, \dots, n$. Then the kernel

$$\mathcal{K}(t,s) = \left(\sum_{k=0}^n \alpha_k (t-s)^{H_k - \frac{1}{2}}\right)\mathbb{1}_{s<t}$$

satisfies the assumptions of Theorem 4, so the corresponding SVV model generates power law (2) with $H = H_0$ provided that $\rho < 0$ in (22).

## A Selected results from the Malliavin calculus

### *A.1 The Malliavin derivative and the space $\mathbb{D}^{k,p}$*

Hereafter, we summarize the essentials of the Malliavin derivative with respect to the classical Brownian motion. For more details, we refer the reader to the classical books [25] or [13].

Denote by $C_p^{(\infty)}(\mathbb{R}^n)$ the space of all infinitely differentiable functions with the derivatives of at most polynomial growth. Let $B = \{B(t),\ t \in [0,T]\}$ be a standard Brownian motion. For any $h \in L^2([0,T])$, denote

$$B(h) := \int_0^T h(t)dB(t).$$

**Definition 1.** The random variables $X$ of the form

$$X = f\big(B(h_1), \dots, B(h_n)\big),$$

where $n \geq 1$, $f \in C_p^{(\infty)}(\mathbb{R}^n)$ and $h_1, \dots, h_n \in L^2([0,T])$ are called smooth. The set of all smooth random variables is denoted by $\mathcal{S}$.

**Definition 2.** Let $X \in \mathcal{S}$. The Malliavin derivative of $X$ (with respect to $B$) is the $L^2([0,T])$-valued random variable of the form

$$DX := \sum_{k=1}^n \frac{\partial f}{\partial x_k}\big(B(h_1), \dots, B(h_n)\big)h_k.$$

By [25, Proposition 1.2.1], the operator $D$ is closable from $L^p(\Omega)$ to $L^p(\Omega \times [0,T])$ for any $p \geq 1$, and we use the same notation $D$ for the closure. The domain of this closure $D$ in $L^p(\Omega)$, i.e. the closure of the class $\mathcal{S}$ with respect to the norm

$$\|X\|_{1,p} := \left(\mathbb{E}\big[|X|^p\big] + \mathbb{E}\left[\left(\int_0^T (D_s X)^2 ds\right)^{\frac{p}{2}}\right]\right)^{\frac{1}{p}},$$

is traditionally denoted by $\mathbb{D}^{1,p}$. This definition can be iterated as described in [25, p. 27] to introduce the iterated derivative $D^k X$ as a random variable with values in $(L^2([0,T]))^{\otimes k} \sim L^2([0,T]^k)$. One can also define $\mathbb{D}^{k,p}$ as the completion of $\mathcal{S}$ with respect to the seminorm

$$\|X\|_{k,p} := \left(\mathbb{E}\big[|X|^p\big] + \sum_{j=1}^k \mathbb{E}\big[\|D^j X\|^p_{L^2([0,T]^k)}\big]\right)^{\frac{1}{p}}. \tag{32}$$

Throughout the paper, we often use the following lemma which is essentially a simplified version of [25, Proposition 1.5.5].

**Lemma 1.** *Let $p > 1$ and $X \in \mathbb{D}^{1,2}$ be such that*

$$\mathbb{E}\big[|X|^p\big] < \infty$$

*and*

$$\mathbb{E}\left[\left(\int_0^T (D_s X)^2 ds\right)^{\frac{p}{2}}\right] < \infty.$$

*Then $X \in \mathbb{D}^{1,p}$.*

### A.2 Generalized Malliavin product rule

Finally, let us prove a generalized version of the product rule from [25, Exercise 1.2.12] or [13, Theorem 3.4].

**Lemma 2.** *Let $X_1, X_2 \in \mathbb{D}^{1,2}$ be such that*

*(i)* $X_1 X_2 \in L^2(\Omega)$;

*(ii)* $X_2 DX_1, X_1 DX_2 \in L^2(\Omega \times [0, T])$.

*Then $X_1 X_2 \in \mathbb{D}^{1,2}$ and*

$$D[X_1 X_2] = X_2 DX_1 + X_1 DX_2.$$

*If, in addition,*

$$\mathbb{E}\big[|X_1 X_2|^p\big] < \infty, \quad \mathbb{E}\bigg[\bigg(\int_0^T (X_2 D_u X_1 + X_1 D_u X_2)^2 du\bigg)^{\frac{p}{2}}\bigg] < \infty$$

*for some $p \geq 2$, then $X_1 X_2 \in \mathbb{D}^{1,p}$.*

**Proof.** Let $\phi \in C^\infty(\mathbb{R})$ be a compactly supported function such that $\phi(x) = x$ whenever $|x| \leq 1$ and $|\phi(x)| \leq |x|$ for all $|x| > 1$. For $m \geq 1$, put

$$f_m(x_1, x_2) := m^2 \phi\left(\frac{x_1}{m}\right)\phi\left(\frac{x_2}{m}\right)$$

and observe that both partial derivatives

$$\frac{\partial f_m}{\partial x_1}(x_1, x_2) = m\phi'\left(\frac{x_1}{m}\right)\phi\left(\frac{x_2}{m}\right), \quad \frac{\partial f_m}{\partial x_2}(x_1, x_2) = m\phi\left(\frac{x_1}{m}\right)\phi'\left(\frac{x_2}{m}\right)$$

are bounded. Therefore, by the classical chain rule [25, Proposition 1.2.3],

$$Df_m(X_1, X_2) = m\left(\phi'\left(\frac{X_1}{m}\right)\phi\left(\frac{X_2}{m}\right)DX_1 + \phi\left(\frac{X_1}{m}\right)\phi'\left(\frac{X_2}{m}\right)DX_2\right).$$

Now it is sufficient to prove that

$$f_m(X_1, X_2) \to X_1 X_2 \tag{33}$$

in $L^2(\Omega)$ and

$$m\left(\phi'\left(\frac{X_1}{m}\right)\phi\left(\frac{X_2}{m}\right)DX_1 + \phi\left(\frac{X_1}{m}\right)\phi'\left(\frac{X_2}{m}\right)DX_2\right) \to X_2 DX_1 + X_1 DX_2 \tag{34}$$

in $L^2(\Omega \times [0, T])$ as $m \to \infty$.

Observe that $|f_m(X_1, X_2)| \to X_1 X_2$ a.s. as $m \to \infty$ and

$$\big|f_m(X_1, X_2)\big| \leq X_1 X_2 \in L^2(\Omega),$$

so (33) holds by the dominated convergence theorem. Next, since $\phi'$ is bounded, we have that, with probability 1,

$$\begin{aligned} m\left|\phi'\left(\frac{X_1}{m}\right)\phi\left(\frac{X_2}{m}\right)DX_1 + \phi\left(\frac{X_1}{m}\right)\phi'\left(\frac{X_2}{m}\right)DX_2\right| \\ \le \max_{x\in\mathbb{R}}\left|\phi'(x)\right|\left(|X_2DX_1| + |X_1DX_2|\right) \in L^2\left(\Omega\times[0,T]\right). \end{aligned}$$

Therefore, since $m\phi'(\frac{X_1}{m})\phi(\frac{X_2}{m}) \to X_2$ a.s. and $m\phi(\frac{X_1}{m})\phi'(\frac{X_2}{m}) \to X_1$ a.s. as $m \to \infty$, (34) holds by the dominated convergence, which ends the proof of the first claim.

The second claim immediately follows from Lemma 1. □

## Funding

The present research is carried out within the frame and support of the ToppForsk project nr. 274410 of the Research Council of Norway with the title STORM: Stochastics for Time-Space Risk Models.

## References

[1] Alòs, E., León, J.A., Vives, J.: On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. Finance Stoch. **11**(4), 571–589 (2007). MR2335834. https://doi.org/10.1007/s00780-007-0049-1

[2] Andersen, L.B.G., Piterbarg, V.V.: Moment explosions in stochastic volatility models. Finance Stoch. **11**(1), 29–50 (2006). MR2284011. https://doi.org/10.1007/s00780-006-0011-7

[3] Azmoodeh, E., Sottinen, T., Viitasaari, L., Yazigi, A.: Necessary and sufficient conditions for Hölder continuity of Gaussian processes. Stat. Probab. Lett. **94**, 230–235 (2014). https://doi.org/10.1016/j.spl.2014.07.030. MR3257384

[4] Bayer, C., Friz, P., Gatheral, J.: Pricing under rough volatility. Quant. Finance **16**(6), 887–904 (2016). MR3494612. https://doi.org/10.1080/14697688.2015.1099717

[5] Biagini, F., Guasoni, P., Pratelli, M.: Mean-variance hedging for stochastic volatility models. Math. Finance **10**(2), 109–123 (2000). MR1802593. https://doi.org/10.1111/1467-9965.00084

[6] Bollerslev, T., Mikkelsen, H.O.: Modeling and pricing long memory in stock market volatility. J. Econom. **73**(1), 151–184 (1996). https://doi.org/10.1016/0304-4076(95)01736-4

[7] Comte, F., Renault, E.: Long memory in continuous-time stochastic volatility models. Math. Finance **8**(4), 291–323 (1998). MR1645101. https://doi.org/10.1111/1467-9965.00057

[8] Delemotte, J., De Marco, S., Segonne, F.: Yet another analysis of the SP500 at-the-money skew: Crossover of different power-law behaviours. SSRN Electron. J. (2023). https://doi.org/10.2139/ssrn.4428407

[9] Di Nunno, G., Mishura, Y., Yurchenko-Tytarenko, A.: Option pricing in Volterra sandwiched volatility model. arXiv:2209.10688 (2022). https://doi.org/10.48550/ARXIV.2209.10688

[10] Di Nunno, G., Mishura, Y., Yurchenko-Tytarenko, A.: Drift-implicit Euler scheme for sandwiched processes driven by Hölder noises. Numer. Algorithms **93**(2), 459–491 (2023). MR4586188. https://doi.org/10.1007/s11075-022-01424-6

[11] Di Nunno, G., Mishura, Y., Yurchenko-Tytarenko, A.: Sandwiched SDEs with unbounded drift driven by Hölder noises. Adv. Appl. Probab. **55**(3), 927–964 (2023). MR4624032. https://doi.org/10.1017/apr.2022.56

[12] Di Nunno, G., Kubilius, K., Mishura, Y., Yurchenko-Tytarenko, A.: From constant to rough: A survey of continuous volatility modeling. Mathematics **11**(19), 4201 (2023). https://doi.org/10.3390/math11194201

[13] Di Nunno, G., Øksendal, B., Proske, F.: Malliavin Calculus for Lévy Processes with Applications to Finance, 1st edn. Springer, Berlin, Heidelberg (2009). MR2460554

[14] Ding, Z., Granger, C.W.J.: Modeling volatility persistence of speculative returns: A new approach. J. Econom. **73**(1), 185–215 (1996). MR1410004. https://doi.org/10.1016/0304-4076(95)01737-2

[15] Ding, Z., Granger, C.W.J., Engle, R.F.: A long memory property of stock market returns and a new model. J. Empir. Finance **1**(1), 83–106 (1993). https://doi.org/10.1016/0927-5398(93)90006-d

[16] Fouque, J.-P., Papanicolaou, G., Sircar, R., Sølna, K.: Maturity cycles in implied volatility. Finance Stoch. **8**(4) (2004). MR2212113. https://doi.org/10.1007/s00780-004-0126-7

[17] Fukasawa, M.: Volatility has to be rough. Quant. Finance **21**(1), 1–8 (2021). MR4188876. https://doi.org/10.1080/14697688.2020.1825781

[18] Funahashi, H., Kijima, M.: Does the Hurst index matter for option prices under fractional volatility? Ann. Finance **13**(1), 55–74 (2017). MR3623799. https://doi.org/10.1007/s10436-016-0289-1

[19] Funahashi, H., Kijima, M.: A solution to the time-scale fractional puzzle in the implied volatility. Fractal Fract. **1**(1), 14 (2017). https://doi.org/10.3390/fractalfract1010014

[20] Gatheral, J., Jaisson, T., Rosenbaum, M.: Volatility is rough. Quant. Finance **18**(6), 933–949 (2018). MR3805308. https://doi.org/10.1080/14697688.2017.1393551

[21] Hu, Y., Nualart, D., Song, X.: A singular stochastic differential equation driven by fractional Brownian motion. Stat. Probab. Lett. **78**(14), 2075–2085 (2008). MR2458016. https://doi.org/10.1016/j.spl.2008.01.080

[22] Hytönen, T., van Neerven, J., Veraar, M., Weis, L.: Analysis in Banach Spaces. Volume I: Martingales and Littlewood-Paley theory, 1st edn. Springer, Cham, Switzerland (2016). MR3617205

[23] Lee, R.W.: Implied volatility: Statics, dynamics, and probabilistic interpretation. In: Recent Advances in Applied Probability, pp. 241–268. Kluwer Academic Publishers, Boston (2006). MR2102956. https://doi.org/10.1007/0-387-23394-6_11

[24] Lobato, I.N., Velasco, C.: Long memory in stock-market trading volume. J. Bus. Econ. Stat. **18**(4), 410–427 (2000). MR1802045. https://doi.org/10.2307/1392223

[25] Nualart, D.: The Malliavin Calculus and Related Topics. Springer, Berlin/Heidelberg (2006). MR2200233

[26] Ocone, D.L., Karatzas, I.: A generalized Clark representation formula, with application to optimal portfolios. Stoch. Stoch. Rep. **34**(3–4), 187–220 (1991). MR1124835. https://doi.org/10.1080/17442509108833682

[27] Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. Springer, Berlin, Heidelberg (1999). MR1725357. https://doi.org/10.1007/978-3-662-06400-9

[28] Rough Volatility Literature. https://sites.google.com/site/roughvol/home/rough-volatility-literature?authuser=0. Accessed: 2023-10-30.

[29] Willinger, W., Taqqu, M.S., Teverovsky, V.: Stock market prices and long-range dependence. Finance Stoch. **3**(1), 1–13 (1999). https://doi.org/10.1007/s007800050049