On the integrability of the supremum of stochastic volatility models and other martingales

Stefan Gerhold
TU Wien
sgerhold@fam.tuwien.ac.at

Julian Pachschwöll
Universität Wien
julian.pachschwoell@univie.ac.at

Johannes Ruf
CUHK Shenzhen & LSE
j.ruf@lse.ac.uk

March 19, 2026

Abstract

We propose a method to bound the expectation of the supremum of the price process in stochastic volatility models. It can be applied, for example, to the rough Bergomi model, avoiding the need to discuss finiteness of higher moments. Our motivation stems from the theory of American option pricing, as an integrable supremum implies the existence of an optimal stopping time for any linearly bounded payoff. Moreover, we survey the literature on martingales with non-integrable supremum, and give a new construction that yields uniformly integrable martingales with this property.

MSC 2020: 60G44, 60G40, 60G46, 91G20

1 Introduction

This paper deals mainly with the question whether a nonnegative martingale $S = (S_t)_{0 \le t \le T}$, where $T \in (0, \infty]$ is deterministic, satisfies

$$\mathbb{E}\left[\sup_{t \in [0,T]} S_t\right] < \infty. \tag{1.1}$$

From the viewpoint of mathematical finance, this property is of interest in the theory of American option pricing. If S is the price process of the underlying and the expectation is under the chosen pricing measure, condition (1.1) implies the existence of an optimal exercise time for all American payoffs of at most linear growth [27, Theorem D.12]. A standard example is the straddle [6], with payoff $|S_t - K|$. In the European case, its price decomposes into a call price plus a put price, where both put and call have strike K. Under American exercise rights, such a decomposition does not hold. Thus, (1.1) serves as theoretical basis of any convergence analysis regarding approximate optimal exercise strategies, as it guarantees the existence of the limiting object.

If $T < \infty$ and $\mathbb{E}[S_T^p] < \infty$ for some $p > 1$, which may depend on T, then (1.1) follows from Doob's L^p inequality. A convenient way of verifying finiteness of moments $\mathbb{E}[S_T^p]$ is to study the domain of the characteristic function, and so the existence of higher moments is well understood for affine models [28] and affine Volterra models [16, 29]. Further examples of stochastic volatility models with known characteristic function can be found in [20, Chapter 4]. The 3/2 model admits a particularly simple statement in this regard, which has not been made explicit in the literature: If the asset and variance processes are negatively correlated, any moment $\mathbb{E}[S_T^p]$ with exponent $p \in [1, u_+)$ is finite, irrespective of the value of T. We refer to [17] for the definition of $u_+ > 1$ in terms of the model parameters, and for refined tail asymptotics of the marginal density. For models that do not feature an explicit characteristic function, a tool that could be used to establish (1.1) is [34, Theorem 4.1], which gives sufficient conditions for the finiteness of the exponential Orlicz norm of the maximum of a continuous local martingale.

For rough volatility models outside the affine Volterra class, it seems that there are only results on the *non*-existence of moments so far [15, 21], which motivates studying (1.1) directly. We introduce a class of generic rough stochastic volatility models for which we give a sufficient condition for (1.1) based on Doob's L^1 inequality and a change of measure. We demonstrate the applicability of our approach by treating several specific instances of our model class.

We give a full characterization of (1.1) for the multi-factor rough Bergomi model [31], which contains the classical rough Bergomi model as a special case. We further prove that (1.1) holds generically for models satisfying certain linear growth conditions, effectively recovering the integrability of the supremum for volatilities given by affine Volterra equations without requiring existence of higher order moments. Finally, we slightly extend results in [1] by showing integrability of the supremum for multi-factor signature volatility models if and only if the price process is a true martingale.

We note in passing that the stronger condition that the supremum of the expectation in (1.1) over all equivalent martingale measures be finite is of interest as well. Under this assumption, *all* arbitrage free price processes of a linearly bounded American payoff can be represented by Snell envelopes with respect to some equivalent martingale measure. This follows from [24, Theorem 13.2.9], by specializing from game options to American options. The statement is not found in many textbooks, while the European counterpart of the theorem, viz. that all arbitrage free price processes arise from equivalent martingale measures, is standard material in mathematical finance courses – see [14, Theorem 5.29] for this (the European) statement in discrete time, which avoids technicalities.

A-priori, it is not obvious whether there actually are martingales that do *not* satisfy (1.1). We discuss this question in Section 4. We survey the literature and show how known results easily yield explicit examples of such martingales. Moreover, we present a new construction that gives plenty of examples, by appropriately stopping a strict local martingale.

In terms of notation, for a given probability space, the class $L \log L$ consists of the random variables X with $\mathbb{E}[|X \log |X||] < \infty$. Throughout this paper, $d, m \in \mathbb{N}$ denote two integers. We also consider the lower triangle

$$\Delta = \{(t, s) \in [0, \infty)^2 : t > s\}.$$

2 Integrability of the supremum in stochastic volatility models

The following lemma states a version of Doob's L^1-inequality [5, 18].

Lemma 2.1. Let $(X_t)_{t \geq 0}$ be a nonnegative right-continuous submartingale and let $T > 0$. Then we have

$$\mathbb{E}\left[\sup_{t \in [0,T]} X_t\right] \leq \frac{e}{e-1}\Big(\mathbb{E}\left[X_T \log(X_T)\right] + \mathbb{E}\left[X_0(1 - \log(X_0))\right]\Big).$$

If the submartingale is additionally uniformly integrable, the extended process $(X_t)_{t\in[0,\infty]}$ is a submartingale as well by Doob's L^1-convergence theorem. In this case, the inequality also holds for $T = \infty$.

Doob's original statement has $(\log(\cdot))^+$ instead of log, which is slightly less convenient for our application. While some sources present the inequality in discrete time, it is clear that we can easily pass to continuous time, by means of the monotone convergence theorem. We note in passing that an application of the inequality to pathwise hedging is presented in [4, Section 5].

We start with a rather generic modeling setup for the spot price process S that includes various rough and non-rough stochastic volatility models as special cases. Our main theoretical results on the integrability of the supremum are given below in Theorem 2.5 and Corollary 2.6. We show in the next section that various commonly used models are in scope of our framework.

Definition 2.2 (Progressive functional)**.** For $t \geq 0$ define the coordinate projection $\pi_t : C([0,\infty),\mathbb{R}^m) \to \mathbb{R}^m$ by $\pi_t(w) = w(t)$. Define the sigma algebra $\mathcal{G}_t = \sigma(\pi_s : s \in [0,t])$ and the filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$. Then a function $g : [0,\infty) \times C([0,\infty),\mathbb{R}^m) \to \mathbb{R}^d$ is a progressive functional if for each $t \geq 0$, the restriction

$$g : [0,t] \times C([0,\infty),\mathbb{R}^m) \to \mathbb{R}^d$$

is $\mathcal{B}|_{[0,t]} \otimes \mathcal{G}_t$-measurable.

A progressive functional depends at each time $t \geq 0$ only on path information up until time t. In particular, for an adapted and continuous $\mathbb{R}^m$-valued process X and a progressive functional $g : [0,\infty) \times C([0,\infty),\mathbb{R}^m) \to \mathbb{R}^d$, the map $(\omega, t) \to f(t, X(\omega))$ is progressive in the usual sense.

Definition 2.3 (Generic stochastic volatility model)**.** Set $E = [0,\infty) \times C([0,\infty),\mathbb{R}^m)$ and consider progressive functionals

$$\psi : E \to \mathbb{R}, \quad \mu : E \to \mathbb{R}^m, \quad \sigma^S : E \to \mathbb{R}^{1\times d}, \quad \sigma^Y : E \to \mathbb{R}^{m\times d}$$

with $\|\sigma^S\|_2 = 1$, and some measurable kernel $K : \Delta \to [0,\infty)$. We consider a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, where $\mathscr{F}_0$ is $\mathbb{P}$-trivial, satisfying the usual conditions and supporting an $\mathbb{R}^d$-valued Brownian motion W. Then a *generic stochastic volatility model* is a continuous and adapted $\mathbb{R}^{1+m}$-valued process (S, Y) on that space such that, for any $t \geq 0$,

$$\int_0^t \psi(s,Y)^2 \mathrm{d}s + \int_0^t K(t,s)|\mu(s,Y)|\mathrm{d}s + \int_0^t K^2(t,s)\mathrm{tr}(\sigma^Y(s,Y)\sigma^Y(s,Y)^\top)\mathrm{d}s < \infty$$

and

$$S_t = S_0 \exp\left(\int_0^t \psi(s,Y)\sigma^S(s,Y)\mathrm{d}W_s - \frac{1}{2}\int_0^t \psi(s,Y)^2\mathrm{d}s \right), \tag{SV1}$$

$$Y_t = Y_0 + \int_0^t K(t,s)\mu(s,Y)\mathrm{d}s + \int_0^t K(t,s)\sigma^Y(s,Y)\,\mathrm{d}W_s, \tag{SV2}$$

for $S_0 > 0$ and $Y_0 \in \mathbb{R}^m$.

Remark 2.4. We remark that Definition 2.3 is quite generic. Suppose that strong existence holds for (SV2). This allows us to start on any filtered probability space satisfying the usual conditions and supporting a Brownian motion, and obtain a continuous process Y satisfying (SV2). On the other hand, if weak existence holds for (SV2), then the filtered probability space and the Brownian motion are part of the solution. Note that uniqueness of the solution of (SV2) is not relevant.

The following results provide tools for establishing (1.1) for various stochastic volatility models, presented in Section 3. We start with a quite general theorem as our basic result. It bounds (1.1) in terms of the expected integrated squared variance under a new measure $\mathbb{P}'$.

Theorem 2.5. Assume the setting of Definition 2.3; in particular, we are given continuous processes S and Y satisfying (SV1) and (SV2), respectively. Let τ be any stopping time such that S^τ is a uniformly integrable martingale. Then there exists a probability measure $\mathbb{P}'$ on $(\Omega, \mathscr{F})$ and an $\mathbb{R}^d$-valued $\mathbb{P}'$-Brownian motion W' such that Y satisfies

$$\begin{aligned} Y_t = Y_0 &+ \int_0^t K(t,s)\mu(s,Y)\mathrm{d}s \\ &+ \int_0^{t\wedge\tau} K(t,s)\sigma^Y(s,Y)\sigma^S(s,Y)^\top\psi(s,Y)\mathrm{d}s \\ &+ \int_0^t K(t,s)\sigma^Y(s,Y)\,\mathrm{d}W'_s, \quad t \geq 0. \end{aligned} \tag{2.1}$$

Moreover, we have the estimate

$$\mathbb{E}\left[\sup_{t\in[0,\tau]} S_t\right] \leq S_0 \frac{e}{e-1}\left(1 + \mathbb{E}_{\mathbb{P}'}\left[\int_0^\tau \psi(s,Y)^2\mathrm{d}s\right]\right). \tag{2.2}$$

Proof. Assume w.l.o.g. that $S_0 = 1$ and define the (stopped) share measures $\mathrm{d}\mathbb{P}' = S^\tau_\infty \mathrm{d}\mathbb{P}$. Applying Girsanov's theorem yields that

$$W' = W - \int_0^{\cdot\wedge\tau} \sigma^S(s,Y)^\top\psi(s,Y)\mathrm{d}s$$

is an $\mathbb{R}^d$-valued Brownian motion under $\mathbb{P}'$, and hence (2.1). For the second assertion we assume w.l.o.g. that

$$\mathbb{E}_{\mathbb{P}'}\bigg[\int_0^\tau \psi(s,Y)^2 \mathrm{d}s\bigg] < \infty,$$

since (2.2) would otherwise be trivially satisfied. Thus, the random variable $\int_0^\tau \psi(s,Y)\sigma^S(s,Y)\mathrm{d}W'_s$ has zero expectation under $\mathbb{P}'$. Note that a simple rearrangement shows for $t \geq 0$

$$\int_0^t \psi(s,Y)\sigma^S(s,Y)\mathrm{d}W_s = \int_0^t \psi(s,Y)\sigma^S(s,Y)\mathrm{d}W'_s + \int_0^{t\wedge\tau} \psi(s,Y)^2\mathrm{d}s.$$

Combining this with an application of Lemma 2.1 to the uniformly integrable martingale S^τ yields

$$\begin{aligned}
\mathbb{E}\bigg[\sup_{t\in[0,\tau]} S_t\bigg] &\leq c\bigg(1 + \mathbb{E}\bigg[S_\tau\bigg(\int_0^\tau \psi(s,Y)\sigma^S(s,Y)\mathrm{d}W_s - \frac{1}{2}\int_0^\tau \psi(s,Y)^2\mathrm{d}s\bigg)\bigg]\bigg) \\
&= c\bigg(1 + \mathbb{E}_{\mathbb{P}'}\bigg[\int_0^\tau \psi(s,Y)\sigma^S(s,Y)\mathrm{d}W'_s + \frac{1}{2}\int_0^\tau \psi(s,Y)^2\mathrm{d}s\bigg]\bigg) \\
&= c\bigg(1 + \mathbb{E}_{\mathbb{P}'}\bigg[\frac{1}{2}\int_0^\tau \psi(s,Y)^2\mathrm{d}s\bigg]\bigg),
\end{aligned}$$

where $c = e/(e-1)$. This concludes the proof. □

Now we combine Theorem 2.5 with a straightforward localization argument.

Corollary 2.6. Assume the setup of Definition 2.3 and let $(\tau_n)_{n\in\mathbb{N}}$ be a localizing sequence for the local martingale S. Then, for any $T > 0$,

$$\mathbb{E}\bigg[\sup_{t\in[0,T]} S_t\bigg] \leq S_0 \frac{e}{e-1}\bigg(1 + \lim_{n\uparrow\infty} \mathbb{E}_{\mathbb{P}'_n}\bigg[\int_0^{\tau_n\wedge T} \psi(s,Y)^2\mathrm{d}s\bigg]\bigg), \tag{2.3}$$

where $\mathrm{d}\mathbb{P}'_n = S_T^{\tau_n}/S_0\mathrm{d}\mathbb{P}$. For each $n \in \mathbb{N}$, the process Y satisfies (2.1) with $\tau = \tau_n \wedge T$ and $W' = W'_n$.

Proof. This is an immediate consequence of Theorem 2.5, since $S^{\tau_n\wedge T}$ is a uniformly integrable martingale for each $n \in \mathbb{N}$. An application of monotone convergence yields (2.3). □

Theorem 2.5 and Corollary 2.6 give two ways to show (1.1) for a given model. If we know a priori that S is a martingale, then the deterministic stopping time T turns S^T into a uniformly integrable martingale. Thus,

Theorem 2.5 gives a concrete upper bound we can check; for an example of this approach, see the proof of Theorem 3.8. On the other hand, if we do not know whether S is a martingale, then we first choose a localizing sequence, e.g., given by $\tau_n = \inf\{t \geq 0 : |Y_t| \geq n\}$ for each $n \in \mathbb{N}$, and then try to bound the right hand side of (2.3) uniformly in n. If this works, then S is a martingale with integrable supremum. Theorem 3.2 provides an example for this approach. See also [36] for a related test of the martingale property, using an appropriate version of Föllmer's measure.

3 Applications

We apply the results of Section 2, in particular Theorem 2.5 and Corollary 2.6, to the validity of (1.1) for three (classes of) stochastic volatility models. First we treat the mixed multi-factor rough Bergomi model from [31], which contains the classical rough Bergomi model as a special case. We continue with a class of models whose coefficients satisfy certain growth assumptions, which are in particular satisfied by the rough Heston model. Lastly, we show integrability of the supremum for multi-factor signature volatility models driven by multiple Brownian motions, as treated in [1].

3.1 Rough Bergomi model and multi-factor extensions

We introduce the multi-factor extension of the rough Bergomi model from [31]. This definition extends the classical rough Bergomi model from [7] by incorporating multiple driving Brownian motions. To this end, we first define the fractional Riemann-Liouville kernel

$$K_{\alpha,\eta}(t,s) = \eta\sqrt{2\alpha-1}(t-s)^{\alpha-1}, \quad 0 \leq s < t, \tag{3.1}$$

with $\alpha > 1/2$ and $\eta > 0$. Further, for $m \in \mathbb{N}$ consider a $d = (m+1)$-dimensional Brownian motion $W = (W^1, \dots, W^d)$ and define the vector-valued Riemann-Liouville process Y by

$$Y_t = (Y_t^1, \dots, Y_t^m) = \left(\int_0^t K_{\alpha,\eta}(t,s)\mathrm{d}W_s^1, \dots, \int_0^t K_{\alpha,\eta}(t,s)\mathrm{d}W_s^m \right), \quad t \geq 0.$$

Definition 3.1 (Mixed multi-factor rough Bergomi model [31])**.** Let $N \in \mathbb{N}$. The mixed multi-factor rough Bergomi model is given in terms of the stock

price and its instantaneous variance process for $t \geq 0$ by

$$S_t = S_0 \exp\left(\int_0^t \psi(s,Y)\sigma^S \mathrm{d}W_s - \frac{1}{2}\int_0^t \psi(s,Y)^2 \mathrm{d}s\right),$$
$$\psi(t,Y) = \left(\xi_0(t)\sum_{i=1}^N \gamma_i \exp\left(\frac{\nu_i^\top L_i}{\eta} Y_t - \frac{\nu_i^\top \Sigma_i \nu_i}{2} t^{2\alpha-1}\right)\right)^{\frac{1}{2}},$$

where $S_0 > 0$, $\xi_0 : [0,\infty) \to (0,\infty)$ is some continuous and bounded function representing the initial forward variance curve, and σ^S is a deterministic d-dimensional vector with $\|\sigma^S\|_2 = 1$. Furthermore, $(\gamma_1, \ldots, \gamma_N) \in (0,1]^N$ is such that $\sum_{i=1}^N \gamma_i = 1$, and the vector $\nu_i = (\nu_i^1, \ldots, \nu_i^m)^\top$ satisfies $0 < \nu_i^1 < \cdots < \nu_i^m$ for all $i \in \{1, \ldots, N\}$. The matrix $L_i \in \mathbb{R}^{m\times m}$ is lower triangular and such that $\Sigma_i = L_i L_i^\top$ is positive definite for each $i \in \{1, \ldots, N\}$.

The next theorem fully characterizes the martingale property and integrability of the supremum (1.1) for the mixed multi-factor rough Bergomi model. The result regarding the martingale property follows the same spirit as [15]. The characterization of (1.1) is novel. Due to its significance in the literature we give a detailed remark at the end of this section on how the standard rough Bergomi model fits this definition and how the main theorem of this section applies in this situation.

Theorem 3.2. Consider the mixed multi-factor rough Bergomi model from Definition 3.1, and denote by $\sigma^S_{(1:m)}$ and $\sigma^S_{(m+1)}$ the split of σ^S into its first m and last components, respectively. Then the stock price S is a martingale if and only if $\sigma^S_{(1:m)} L_i^\top \nu_i \leq 0$ for all $i = 1, \ldots, N$. In this case, (1.1) holds for all $T > 0$.

Proof. Clearly, the mixed multi-factor rough Bergomi model fits Definition 2.3. We first show that S is a martingale and (1.1) holds under the stated assumption. To this end, let $(\tau_n)_{n\in\mathbb{N}}$ be a localizing sequence such that S^{τ_n} is a bounded martingale. Fix now $n \in \mathbb{N}$. By Corollary 2.6, there exists a probability measure $\mathbb{P}'_n$ and a $\mathbb{P}'_n$-Brownian motion W'_n such that

$$Y_t = \int_0^{t\wedge\tau_n\wedge T} K_{\alpha,\eta}(t,s)\psi(s,Y)(\sigma^S_{(1:m)})^\top \mathrm{d}s + \int_0^t K_{\alpha,\eta}(t,s)\mathrm{d}W'_{n,1:m,s}, \quad t \geq 0.$$

For convenience, we write M_n for the Gaussian process defined by $M_{n,t} = \int_0^t K_{\alpha,\eta}(t,s)\mathrm{d}W'_{n,1:m,s}$. Using this representation of Y and the assumption $\sigma^S_{(1:m)} L_i^\top \nu_i \leq 0$ for all $i = 1, \ldots, N$, for any $t \in [0,T]$ one has

$$\psi(t,Y)^2 \leq C_T \sum_{i=1}^N \exp\left(\frac{\nu_i^\top L_i}{\eta} M_{n,t} - \frac{\nu_i^\top \Sigma_i \nu_i}{2} t^{2\alpha-1}\right),$$

where $C_T > 0$ is some appropriate constant. Thus one gets

$$\mathbb{E}_{\mathbb{P}'_n}\bigg[\int_0^{\tau_n\wedge T}\psi(s,Y)^2\mathrm{d}s\bigg] \leq C_T\sum_{i=1}^N\int_0^T\mathbb{E}_{\mathbb{P}'_n}\bigg[\exp\Big(\frac{\nu_i^\top L_i}{\eta}M_{n,s}-\frac{\nu_i^\top\Sigma_i\nu_i}{2}s^{2\alpha-1}\Big)\bigg]\mathrm{d}s = C_TNT.$$

Hence S is a martingale with integrable supremum by Corollary 2.6. The converse implication is more involved and deferred to Appendix B. □

Remark 3.3 (Standard rough Bergomi model [7])**.** In case of the standard rough Bergomi model [7], Definition 3.1 and Theorem 3.2 simplify significantly. Indeed, choosing $m = N = \gamma_1 = L_1 = 1$ and $\nu_1^1 = \eta$ gives the instantaneous variance

$$\xi_0(t)\exp\bigg(\int_0^t K_{\alpha,\eta}(t,s)\mathrm{d}W_s^1-\frac{\eta^2}{2}t^{2\alpha-1}\bigg).$$

Setting $\sigma^S \equiv (\rho,\ \sqrt{1-\rho^2})$ for $\rho\in[-1,1]$ yields a correlation ρ between the Brownian motion driving the stock S and Y, i.e., $\sigma^S W = \rho W^1+\sqrt{1-\rho^2}W^2$. In this case, the condition in Theorem 3.2 reduces to $\rho\leq 0$, which is thus a necessary and sufficient condition for the integrability of the supremum in the rough Bergomi model. Regarding the characterization of the martingale property, this exactly recovers the results in [15].

Remark 3.4 (General Gaussian Volterra models)**.** The assumption that the variance is an exponential of Y is not essential to show that S is a martingale with integrable supremum. Many papers, such as [9], consider models with more general linking functions. In the uni-variate case, suppose we replace the linking function $(t,y)\mapsto\exp(y-\eta^2/2t^{2\alpha-1})$ in the rough Bergomi model by a more general function $\zeta(t,y)$. As long as ζ is non-negative and increasing in the y-argument, and the first component of σ^S is non-positive, it suffices to check that

$$\int_0^T\mathbb{E}\bigg[\zeta\Big(t,\int_0^t K_{\alpha,\eta}(t,s)dW_s\Big)\bigg]\mathrm{d}t<\infty.$$

3.2 Models with coefficients of linear growth

We apply the results of Section 2 to stochastic volatility models where the coefficient functions of Definition (2.3) satisfy certain growth conditions.

Assumption 3.5 (Linear growth)**.** Let (S, Y) be a generic stochastic volatility model in the sense of Definition 2.3. We assume that there exists an increasing function $D : [0, \infty) \to [0, \infty)$ such that

$$|\mu(t,y)| + |\psi(t,y)| + |\sigma^Y(t,y)| + |\sigma^Y(t,y)\sigma^S(t,y)^\top \psi(t,y)| \leq D_t\left(1 + \sup_{s\leq t}|y_s|\right)$$

holds for any $t \geq 0$ and $y \in C([0,\infty), \mathbb{R}^m)$ and one of the following two conditions is satisfied:

(A) The kernel K satisfies $K(t,s) = 1$ for all $(t,s) \in \Delta$.

(B) Assume that the functions $\mu, \sigma^S, \sigma^Y, \psi$ depend at time $t \geq 0$ on Y only through the value Y_t and in the growth assumptions from above $\sup_{s\leq t}|y_s|$ is replaced by $|y_t|$. Additionally assume there exists $\overline{K} \in L^2_{\text{loc}}([0,\infty), \mathbb{R})$ such that $K(t,s) = \overline{K}(t-s)$, i.e., K depends on s, t only through the difference $t - s$.

Case (A) retains path dependency, but does not allow for singular kernels. Case (B) allows for path dependency only through the kernel K, which is important for applying the results of [3]. Using this assumption on generic stochastic volatility models, we can always conclude that S is a true martingale with integrable supremum by means of Corollary 2.6.

Theorem 3.6. Let (S, Y) define a generic stochastic volatility model satisfying Assumption 3.5. Then we have

$$\mathbb{E}\left[\sup_{t\in[0,T]} S_t\right] < \infty$$

for any $T > 0$. In particular, S is a true martingale with integrable supremum.

Proof. Fix $T > 0$ and define the stopping times $\tau_n = \inf\{t \geq 0 : |Y_t| \geq n\}$ for each $n \in \mathbb{N}$. By Novikov's condition this yields a localizing sequence for S. Define next the stopped share measures $\mathrm{d}\mathbb{P}'_n = S_T^{\tau_n}/S_0 \mathrm{d}\mathbb{P}$ for each $n \in \mathbb{N}$. Thanks to Corollary 2.6, finiteness of the expression in (2.3) implies the assertion. By the linear growth assumption, it suffices to show, for case (A), that

$$\lim_{n\uparrow\infty} \mathbb{E}_{\mathbb{P}'_n}\left[\sup_{t\leq T}|Y_{t\wedge\tau_n}|^2\right] < \infty,$$

and, for case (B), that

$$\lim_{n\uparrow\infty}\int_0^T \mathbb{E}_{\mathbb{P}'_n}[|Y_t|^2\mathbf{1}_{\{t\leq\tau_n\}}]\mathrm{d}s<\infty.$$

In particular, without loss of generality, we may assume that $Y_0=0$. To make headway, note that for each $n\in\mathbb{N}$, thanks to Corollary 2.6, we have the representation

$$\begin{aligned}Y_t=\int_0^{t\wedge\tau_n}K(t,s)\bar{\mu}(s,Y)\mathrm{d}s+\int_{t\wedge\tau_n}^t K(t,s)\mu(s,Y)\mathrm{d}s\\ +\int_0^t K(t,s)\sigma^Y(s,Y)\,\mathrm{d}W'_{n,s}\end{aligned}\tag{3.2}$$

for $t\in[0,T]$ and $\mathbb{P}'_n$-Brownian motions W'_n, where $\bar{\mu}=\mu+\sigma^Y(\sigma^S)^\top\psi$ is of linear growth by assumption.

First assume that case (A) of Assumption 3.5 is in force and recall that $K\equiv 1$ in this case. Fix $n\in\mathbb{N}$ for the moment. Then the BDG inequality, together with Jensen's inequality, yields the existence of some $C_T>0$, independent of n, such that

$$\mathbb{E}_{\mathbb{P}'_n}\Big[\sup_{s\leq T}|Y_{s\wedge\tau_n}|^2\Big]\leq C_T\bigg(1+\int_0^T\mathbb{E}_{\mathbb{P}'_n}\big[|\bar{\mu}(t,Y)|^2+|\sigma^Y(t,Y)|^2\big]\mathrm{d}t\bigg).$$

Using the linear growth assumptions on $\bar{\mu}$ and σ^Y yields the existence of another constant C'_T, independent of n, such that

$$\mathbb{E}_{\mathbb{P}'_n}\Big[\sup_{t\leq T}|Y_{t\wedge\tau_n}|^2\Big]\leq C'_T\bigg(1+\int_0^T\mathbb{E}_{\mathbb{P}'_n}\Big[\sup_{s\leq t}|Y_{s\wedge\tau_n}|^2\Big]\mathrm{d}t\bigg).$$

Now Grönwall's inequality implies

$$\mathbb{E}_{\mathbb{P}'_n}\Big[\sup_{t\leq T}|Y_{t\wedge\tau_n}|^2\Big]\leq C'_Te^{C'_TT},$$

yielding case (A).

Assume now that case (B) is in force. We follow the argument from [3, Proof of Lemma 3.1] with slight alterations. Due to the assumption that all coefficient functions depend at time s on the path Y only through Y_s, we abuse notation and view them as function of (s,Y_s). We fix again $n\in\mathbb{N}$. We first claim that for $t\in[0,T]$,

$$\begin{aligned}|Y_t|^2\mathbf{1}_{\{t<\tau_n\}}\leq\bigg|\int_0^t K(t,s)\bar{\mu}(s,Y_s\mathbf{1}_{\{s<\tau_n\}})\mathrm{d}s\\ +\int_0^t K(t,s)\sigma^Y(s,Y_s\mathbf{1}_{\{s<\tau_n\}})\mathrm{d}W'_{n,s}\bigg|^2.\end{aligned}\tag{3.3}$$

Indeed, on the event $\{t \geq \tau_n\}$, the left hand side is zero and the right hand side is nonnegative. Moreover, on the event $\{t < \tau_n\}$, we have, thanks to Itô isometry,

$$\int_0^t K(t,s)\sigma^Y(s,Y_s)\mathrm{d}W'_{n,s} = \int_0^t K(t,s)\sigma^Y(s,Y_s\mathbf{1}_{\{s<\tau_n\}})\mathrm{d}W'_{n,s}$$

for all $t \in \mathbb{Q} \cap [0,T]$, hence by continuity also for all $t \in [0,T]$. A similar statement holds for the finite-variation integral on the right-hand side of (3.3). Recalling (3.2) thus yields (3.3).

Using Jensen's inequality on (3.3), there exists some $C_T > 0$, independent of $n \in \mathbb{N}$, such that, for each $t \in [0,T]$,

$$|Y_t|^2\mathbf{1}_{\{t<\tau_n\}} \leq C_T\bigg(1 + \int_0^t K(t,s)^2|\bar{\mu}(s,Y_s\mathbf{1}_{\{s<\tau_n\}})|^2\mathrm{d}s + \bigg|\int_0^t K(t,s)\sigma^Y(s,Y_s\mathbf{1}_{\{s<\tau_n\}})\mathrm{d}W'_{n,s}\bigg|^2\bigg).$$

Writing $f_n(t) = \mathbb{E}_{\mathbb{P}'_n}[|Y_t|^2\mathbf{1}_{\{t<\tau_n\}}]$ for each $t \in [0,T]$ and applying the BDG inequality to the $\mathbb{P}'_n$-local martingale $(M_r)_{r\in[0,t]}$ with

$$M_r = \int_0^r K(t,s)\sigma^Y(s,Y_s\mathbf{1}_{\{s<\tau_n\}})\mathrm{d}W'_{n,s}$$

yields, using the linear growth assumption, a constant C'_T, independent of $n \in \mathbb{N}$, such that

$$\begin{aligned} f_n(t) &\leq C_T\Big(1 + \int_0^t K(t,s)^2\mathbb{E}_{\mathbb{P}'_n}[|\bar{\mu}(s,Y_s\mathbf{1}_{\{s<\tau_n\}})|^2]\mathrm{d}s + \mathbb{E}_{\mathbb{P}'_n}\Big[\sup_{r\in[0,t]}|M_r|^2\Big]\Big) \\ &\leq C'_T\Big(1 + \int_0^T \overline{K}(s)^2\mathrm{d}s + \int_0^t K(t,s)^2 f_n(s)\mathrm{d}s\Big). \end{aligned}$$

We are now exactly in the situation $f_n \leq c' + c'K^2 * f_n$ for some $c' > 0$ on $[0,T]$. Following the same arguments as in [3, Proof of Lemma 3.1], we have $\sup_{t\in[0,T],n\in\mathbb{N}} f_n(t) \leq c$ for some $c > 0$, concluding the proof. □

Remark 3.7. The last theorem in particular applies to affine Volterra processes in the sense of [3] and hence also to the well-known rough Heston model. Note however that the existence of higher order moments in the rough Heston model is known, see [16]. In this case, Theorem 3.6 gives an alternative proof of (1.1).

3.3 Signature volatility models

Signature volatility models have recently attracted considerable attention as a flexible, nonparametric framework for modeling volatility in financial markets. Drawing on ideas from rough path theory and signature methods, they offer a way to include path-dependency in a tractable way that classical parametric models can not. Notable work includes [10], which develops a signature-based model to jointly calibrate SPX and VIX options across maturities, and [2], which uses signature based models for pricing and hedging derivatives in models based on signature of Brownian motion. In particular, for Brownian motion based signature volatility models, [1] characterizes the martingale property and provides conditions under which higher order moments exist.

Before elaborating these results we give a minimalist introduction to signature. For a more thorough introduction, see for example [2, 10, 11]. Since it suffices for our purposes, we only introduce truncated signatures and fix a truncation level $N \in \mathbb{N}$ and some dimension $d \in \mathbb{N}$ throughout this subsection. For $n \in \mathbb{N}$, we denote the n-th tensor power of $\mathbb{R}^d$ by $(\mathbb{R}^d)^{\otimes n}$ with $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$. Using this, we define the truncated tensor algebra of order N as

$$T^N(\mathbb{R}^d) = \{\mathbf{a} = (a_0, \dots, a_N) : a_n \in (\mathbb{R}^d)^{\otimes n}, n = 0, \dots, N\}.$$

For each $\mathbf{a}, \mathbf{b} \in T^N(\mathbb{R}^d)$ and $\lambda \in \mathbb{R}$ we define scalar multiplication and addition

$$\lambda \cdot \mathbf{a} = (\lambda a_0, \dots, \lambda a_N), \quad \mathbf{a} + \mathbf{b} = (a_0 + b_0, \dots, a_N + b_N)$$

and the tensor product

$$\mathbf{a} \otimes \mathbf{b} = (c_0, \dots, c_N), \quad c_n = \sum_{k=0}^{n} a_k \otimes b_{n-k}.$$

For $n \in \{0, \dots, N\}$, denote the set of words of length n over the alphabet $\{1, \dots, d\}$ by $\mathcal{W}_n$. Note that the set of words of length zero $\mathcal{W}_0$, contains only the empty word $\emptyset$. For $n \in \{0, \dots, N\}$ and a word $w = (i_1, \dots, i_n)$ we set $e_w = e_{i_1} \otimes \cdots \otimes e_{i_n}$ such that each element $\mathbf{a} \in T^N(\mathbb{R}^d)$ can be written as

$$\sum_{n=0}^{N} \sum_{w \in \mathcal{W}_n} a_w e_w, \quad a_w \in \mathbb{R}.$$

For $\mathbf{a} = \sum_{n=0}^{N} \sum_{w \in \mathcal{W}_n} a_w e_w \in T^N(\mathbb{R}^d)$ and $\mathbf{b} = \sum_{n=0}^{N} \sum_{w \in \mathcal{W}_n} b_w e_w \in T^N(\mathbb{R}^d)$ we set

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{n=0}^{N} \sum_{w \in \mathcal{W}_n} a_w b_w.$$

Observe in particular that $a_w = \langle e_w, \mathbf{a} \rangle$ for any word w.

Let X now be a continuous semimartingale. Then its truncated signature is recursively defined as the $T^N(\mathbb{R}^d)$-valued process $\mathbb{X}$ satisfying

$$\langle e_\emptyset, \mathbb{X}_t \rangle = 1, \quad \langle e_w, \mathbb{X}_t \rangle = \int_0^t \langle e_{w'_n}, \mathbb{X}_t \rangle \circ \mathrm{d}X_t^{i_n}, \quad w = (i_1, \dots, i_n) \in \mathcal{W}_n$$

for $n \in \{0, \dots, N\}$, where $\circ$ denotes Stratonovich integration, and w'_n is defined as $(i_1, \dots, i_{n-1})$ for the word $w = (i_1, \dots, i_n)$. Using the definition of the tensor product we see that the truncated signature of X can equivalently be defined as the unique solution to

$$\mathrm{d}\mathbb{X}_t = \mathbb{X}_t \otimes \circ \mathrm{d}X_t, \quad \mathbb{X}_0 = e_\emptyset. \tag{3.4}$$

We introduce the multi-factor signature volatility model from [1, Section 3.3]. Consider a d-dimensional Brownian motion $W = (W^1, \dots, W^d)$ on a filtered probability space satisfying the usual conditions. The signature volatility model is defined as

$$\begin{aligned} \frac{\mathrm{d}S_t}{S_t} &= \xi_t \mathrm{d}W_t^d, \\ \xi_t &= \langle \boldsymbol{\sigma}, \widehat{\mathbb{X}}_t \rangle, \end{aligned} \tag{3.5}$$

where $\widehat{\mathbb{X}}$ is the truncated signature of the time extended Brownian motion (t, W) and $\boldsymbol{\sigma} \in T^N(\mathbb{R}^{d+1})$. The question of integrability of the supremum (1.1) has been answered in [1, Theorem 3.2] for the case $m = 1$ via existence of higher moments. The existence of higher moments for the model (3.5) driven by multiple Brownian motions is not addressed in [1]. Hence the question of (1.1) is still open, and the next result directly characterizes (1.1) in this case.

Theorem 3.8. Consider the multi-factor signature volatility model (3.5) with $N \geq 2$ and define $\sigma' = \langle e_{d+1}^{\otimes N}, \boldsymbol{\sigma} \rangle$, which we assume to be non-zero. Then (1.1) holds if and only if N is odd and $\sigma' < 0$.

Proof. Under our assumptions, the process S is a martingale if and only if N is odd and $\sigma' < 0$, by [1, Theorem 3.5]. Hence it suffices to show (1.1) if N is odd and $\sigma' < 0$.

Denote $m = \dim T^N(\mathbb{R}^{d+1})$. Then $T^N(\mathbb{R}^{d+1})$ is isomorphic to $\mathbb{R}^m$ as a vector space. Define $\psi(t, \mathbf{y}) = \langle \boldsymbol{\sigma}, \mathbf{y}_t \rangle$ for $t \geq 0$ and $\mathbf{y} \in C([0, \infty), \mathbb{R}^m)$, and $K = 1$. Noting that $\widehat{\mathbb{X}}$ satisfies the linear SDE (3.4), with X replaced by (t, W), the signature volatility model fits into Definition 2.3.

The dynamics of $\widehat{\mathbb{X}}$, corresponding to (2.1), under the measure $\mathbb{P}'$ from Theorem 2.5 are now given by

$$\mathrm{d}\widehat{\mathbb{X}}_t = \big(\widehat{\mathbb{X}}_t \otimes e_1 + \langle \boldsymbol{\sigma}, \widehat{\mathbb{X}}_t \rangle \widehat{\mathbb{X}}_t \otimes e_{d+1}\big)\mathrm{d}t + \sum_{i=2}^{d+1} \widehat{\mathbb{X}}_t \otimes e_i \circ \mathrm{d}(W^{i-1})'_t. \tag{3.6}$$

Consider the signature SDE

$$\mathrm{d}Y_t = \langle \boldsymbol{\sigma}, \widehat{\mathbb{Y}}_t \rangle \mathrm{d}t + \mathrm{d}(W^d)'_t, \quad Y_0 = 0, \qquad t \in [0, T],$$

where $\widehat{\mathbb{Y}}_t$ is the signature of the process $(t, W'_{1:d-1}, Y)$. Note that the signature SDE (3.4) for $\widehat{\mathbb{Y}}_t$ agrees with (3.6). Since the vector fields therein are locally Lipschitz, the solution is unique up to a time of explosion τ_∞. In [2, Section 6] it is shown that $\mathbb{P}[\tau_\infty < T] = 0$ and thus, since $\mathbb{P}' \ll \mathbb{P}$ by construction, also $\mathbb{P}'[\tau_\infty < T] = 0$. Hence, we have the equality

$$\psi(t, \widehat{\mathbb{X}}_t) = \langle \boldsymbol{\sigma}, \widehat{\mathbb{Y}}_t \rangle, \quad t \in [0, T]. \tag{3.7}$$

By [1, Lemma 5.2 and Section 6] there exists a constant $C > 0$ such that

$$\mathbb{E}_{\mathbb{P}'}\Big[\int_0^T \langle \boldsymbol{\sigma}, \widehat{\mathbb{Y}}_t \rangle^2 \mathrm{d}t\Big] \leq C\Big(1 + \mathbb{E}_{\mathbb{P}'}\Big[\int_0^T Y_t^{2N} \mathrm{d}t\Big]\Big) \tag{3.8}$$

and

$$\mathbb{E}_{\mathbb{P}'}[Y_t^{2N}] \leq C, \quad t \in [0, T]. \tag{3.9}$$

Combining (3.7), (3.8), and (3.9) thus yields

$$\mathbb{E}_{\mathbb{P}'}\Big[\int_0^T \psi(t, \widehat{\mathbb{X}})^2 \mathrm{d}t\Big] \leq C\Big(1 + \int_0^T \mathbb{E}_{\mathbb{P}'}[Y_t^{2N}] \mathrm{d}t\Big) < \infty.$$

An application of Theorem 2.5 then concludes the proof. □

4 Martingales with non-integrable supremum

4.1 General remarks

By the Burkholder-Davis-Gundy inequalities, the set of local martingales with integrable supremum coincides with $\mathcal{H}^1$, the space of local martingales $(M_t)_{t \geq 0}$ for which

$$\|M\|_1 = \mathbb{E}\big[[M]_\infty^{1/2}\big]$$

is finite. By dominated convergence, "local" can be dropped from this statement. Clearly, every element of $\mathcal{H}^1$ is uniformly integrable. For details and further interesting properties of $\mathcal{H}^1$, such as its duality to the space of BMO martingales, we refer to Section IV.4 in [35].

A family of *discrete time* examples of martingales with *non*-integrable supremum is provided by Theorem 2 in [8]. In this section, we give three different approaches that yield examples in continuous time. Further examples can be found in [32]. First, recall that any continuous nonnegative martingale $(X_t)_{t\geq 0}$ with $X_0 = 1$ satisfies the reverse L^1 inequality

$$\mathbb{E}\Big[\sup_{t\geq 0} X_t\Big] \geq 1 + \mathbb{E}[X_\infty(\log(X_\infty))^+]. \tag{4.1}$$

For a proof, we refer to p. 149 in [13]; see also [22, 23] for earlier versions of similar inequalities. For the mere existence of martingales with non-integrable supremum, the following observation suffices: For $T \in (0,\infty]$, suppose that a probability space with filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ supports a non-negative $\mathcal{F}_T$-measurable random variable $X_T \in L^1$ which is not in the class $L\log L$, defined at the end of the introduction. If $X_t = \mathbb{E}[X_T|\mathcal{F}_t]$ is continuous, then it is a uniformly integrable martingale which is not in $\mathcal{H}^1$.

4.2 A construction by Dubins and Gilat

The second construction is due to [12] and appears also in [32, Example 1]. In our exposition, we add some details. In contrast to the above argument, the process will be explicitly defined. Unlike the rest of the paper, we drop the assumption of non-negativity, and aim at finding a real-valued martingale $(X_t)_{0\leq t\leq 1}$ with $\mathbb{E}[\sup_{t\in[0,1]} X_t] = \infty$. For convenience, the following assumption is in force throughout this subsection.

Assumption 4.1. Let F be the cumulative distribution function of some distribution on the real line with finite first moment, with full support and a continuous density.

The function

$$H_F(t) = \frac{1}{1-t}\int_t^1 F^{-1}(s)\mathrm{d}s, \quad 0 < t < 1, \tag{4.2}$$

is known as the Hardy–Littlewood maximal function of the distribution. For the significance of this function concerning analytic and martingale inequalities, we refer to [18] and the references therein. Note that, in financial terms, the Hardy–Littlewood maximal function is the tail value at risk, or expected

shortfall, of the distribution defined by F. In economics, the closely related function $t \mapsto \int_0^t F^{-1}(s)\mathrm{d}s / \int_0^1 F^{-1}(s)\mathrm{d}s$ is known as the Lorenz curve [25]. It has been shown in [8] that the distribution of H_F, viewed as a random variable on the probability space

$$\big((0,1), \mathcal{B}(0,1), \mathrm{Leb}\big), \tag{4.3}$$

is an upper bound for the set

$$\{\nu : \nu \text{ is the distribution of the supremum of a martingale closed by a random variable with cumulative distribution function } F\}$$

with respect to stochastic order. In [12] it was shown that the distribution of H_F is also a member of this set. This is achieved by defining a martingale X on the probability space (4.3) such that its closing element has cumulative distribution function F and the distribution of the supremum of X equals the distribution of H_F. If $H_F \notin L^1$, then X has the desired properties. See Proposition 4.4 for details.

Remark 4.2. (i) It is clear that only the upper tail of the supremum of a martingale can cause non-integrability, since

$$\Big|\sup_{t\in[0,1]} X_t\Big| \, \mathbf{1}_{\{\sup_{t\in[0,1]} X_t<0\}} \leq |X_0|.$$

(ii) Integrability of the Hardy–Littlewood maximal function (4.2) can only fail because of a blowup as t tends to 1. Indeed, $\lim_{t\downarrow 0} H_F(t) = \int_0^1 |F^{-1}(s)|\mathrm{d}s$ is finite, by existence of the first moment of F. Thus, for integrability criteria concerning H_F, it does not matter if $|F^{-1}(s)|$ is used in the definition (4.2) instead of $F^{-1}(s)$, as is the case in [38].

(iii) Concerning notation: The nondecreasing function f from [12] is our F^{-1}. On the other hand, f^* from [18] is assumed to be nonincreasing, and equals $F^{-1}(1-\cdot)$ in our notation. As a consequence, when reading [12, 18] it is important to note that the Hardy–Littlewood maximal functions H from [12] (which is our H_F) and F^* from [18] are not equal, but related by $F^*(1-t) = H(t)$, where $0 < t < 1$.

Lemma 4.3. If

$$\int_{\mathbb{R}} |x|(\log|x|)^+ F(\mathrm{d}x) = \infty, \tag{4.4}$$

then $H_F \notin L^1$.

Proof. This follows from a classical theorem due to Stein [38], which asserts that $H_F \notin L^1$ is equivalent to $\int_0^1 |F^{-1}(s)|(\log|F^{-1}(s)|)^+ \mathrm{d}s = \infty$, by substitution. □

Examples of cumulative distribution functions with finite first moment, but satisfying (4.4), can be easily found; e.g., assume that the density satisfies

$$F'(x) \leq \frac{c_1}{x^2(\log|x|)^\alpha}, \quad |x| \text{ large},$$

and

$$F'(x) \geq \frac{c_2}{(x\log x)^2}, \quad x > 0 \text{ large},$$

for some constants $\alpha \in (1,2]$ and $c_1, c_2 > 0$.

Proposition 4.4. Recall Assumption 4.1, and suppose that F satisfies (4.4). Then the process $(X_t)_{0\leq t\leq 1}$ defined by

$$X_t(s) = \begin{cases} H_F(t) & \text{for} \quad t \leq s, \\ F^{-1}(s) & \text{for} \quad t > s \end{cases}$$

is a martingale on the probability space (4.3), with respect to its own filtration, and its supremum is not integrable.

Proof. According to the proof of [12, Lemma 2], the process $(X_t)_{0\leq t\leq 1}$ is a martingale. As this is not shown in [12], we give a proof for the reader's convenience. Let $U(s) = s$, so that U is a standard uniformly distributed random variable on the probability space (4.3). It is easy to see that $\mathcal{F}_t = \sigma(U\mathbf{1}_{\{U<t\}})$ is the filtration generated by X. Define $\tilde{X}_t = \mathbb{E}[F^{-1}(U)|\mathcal{F}_t]$ for $t \in [0,1]$. It is readily verified that we have $\tilde{X}_t = F^{-1}(U)$ on the event $\{U < t\} \in \mathcal{F}_t$, while on the event $\{U \geq t\} \in \mathcal{F}_t$ we have $\tilde{X}_t = H_F(t)$. Thus, $X = \tilde{X}$. Finally, we have

$$\sup_{0\leq t\leq 1} X_t(s) = H_F(s), \quad 0 < s < 1.$$

By Lemma 4.3, this is not integrable with respect to Lebesgue measure on $(0,1)$. □

4.3 A new approach

Our third construction starts with an arbitrary nonnegative local martingale M that is not a uniformly integrable martingale. In particular, M may be any nonnegative strict local martingale. See, e.g., Exercise 3.3.36 in [26]

for a standard example. We present the arguments for $T = \infty$, but they trivially apply to finite T as well. Observe that

$$\mathbb{E}\Big[\sup_{t\geq 0} M_t\Big] = \int_0^\infty \mathbb{P}\Big[\sup_{t\geq 0} M_t > u\Big]\mathrm{d}u$$

and

$$\sum_{n=1}^\infty \mathbb{P}\Big[\sup_{t\geq 0} M_t > n\Big] \leq \int_0^\infty \mathbb{P}\Big[\sup_{t\geq 0} M_t > u\Big]\mathrm{d}u \leq 1 + \sum_{n=1}^\infty \mathbb{P}\Big[\sup_{t\geq 0} M_t > n\Big].$$

As mentioned above, we have $M \in \mathcal{H}^1$ if and only if $\mathbb{E}[\sup_{t\geq 0} M_t] < \infty$, and so it follows that $M \in \mathcal{H}^1$ if and only if

$$\sum_{n=1}^\infty \mathbb{P}\Big[\sup_{t\geq 0} M_t > n\Big] < \infty. \tag{4.5}$$

This condition plays a key role in our construction.

Theorem 4.5. Let $(M_t)_{t\geq 0}$ be a nonnegative local martingale that is not a uniformly integrable martingale. Then there is an extended filtered probability space with a stopping time τ such that

(i) The lift of M to the extended probability space, again denoted by M, has the same properties.

(ii) The stopped process M^τ is a uniformly integrable martingale which is not in $\mathcal{H}^1$, i.e., $\mathbb{E}[\sup_{t\geq 0} M_t^\tau] = \infty$.

Proof. We assume that $T = \infty$; the construction clearly works for finite T, too. Define the nondecreasing sequence $(c_n)_{n\in\mathbb{N}}$, by setting

$$c_n = \log\left(e + \sum_{k=1}^n \mathbb{P}\Big[\sup_{t\geq 0} M_t > k\Big]\right) > 1, \tag{4.6}$$

for each $n \in \mathbb{N}$. Since $M \notin \mathcal{H}^1$, it follows that $\lim_{n\uparrow\infty} c_n = \infty$. Without loss of generality, the original probability space, with filtration $(\mathcal{F}_t)_{t\geq 0}$, accommodates an $\mathbb{N}$-valued and $\mathcal{F}_0$-measurable random variable Y that is independent of M, and whose distribution satisfies $\mathbb{P}[Y > n] = 1/c_n$, for each $n \in \mathbb{N}$. Let

$$\sigma = \inf\{t \geq 0 \,|\, M_t > Y\} \tag{4.7}$$

denote the first time M exceeds Y. It follows that

$$\mathbb{P}\Big[\sup_{t\geq 0} M_t^\sigma > n\Big] \geq \mathbb{P}\Big[\sup_{t\geq 0} M_t > n\Big]\mathbb{P}[Y > n] = \frac{1}{c_n}\mathbb{P}\Big[\sup_{t\geq 0} M_t > n\Big],$$

for each $n \in \mathbb{N}$. Consequently,

$$\sum_{n=1}^{\infty} \mathbb{P}\Big[\sup_{t\geq 0} M_t^{\sigma} > n\Big] \geq \lim_{m\uparrow\infty} \frac{1}{c_m} \sum_{n=1}^{m} \mathbb{P}\Big[\sup_{t\geq 0} M_t > n\Big] = \lim_{m\uparrow\infty} \frac{e^{c_m} - e}{c_m} = \infty,$$

since $(c_n)_{n\in\mathbb{N}}$ is non-decreasing and $\lim_{n\uparrow\infty} c_n = \infty$. This implies that $M^{\sigma} \notin \mathcal{H}^1$. On the other hand, the almost sure limit $M_{\infty}^{\sigma} = M_{\infty-}^{\sigma} \in [0, \infty)$ exists, since M^{σ} is a nonnegative local martingale, and hence also a nonnegative supermartingale. Moreover,

$$\begin{aligned}
\mathbb{E}[M_{\infty}^{\sigma}] &= \sum_{n=1}^{\infty} \mathbb{E}[M_{\infty}^{\sigma} \,|\, Y = n]\mathbb{P}[Y = n] \\
&= \sum_{n=1}^{\infty} \mathbb{E}[M_{\infty}^{\tau_n} \,|\, Y = n]\mathbb{P}[Y = n] \\
&= \sum_{n=1}^{\infty} \mathbb{E}[M_{\infty}^{\tau_n}]\mathbb{P}[Y = n] = \sum_{n=1}^{\infty} \mathbb{E}[M_0]\mathbb{P}[Y = n] = \mathbb{E}[M_0^{\sigma}],
\end{aligned}$$

where

$$\tau_n = \inf\{t \geq 0 \,|\, M_t > n\}, \quad n \in \mathbb{N}.$$

Here, the penultimate equality follows from the fact that M^{τ_n} is uniformly integrable, for each $n \in \mathbb{N}$. Consequently, M^{σ} is uniformly integrable. □

Recall from the beginning of this section that the existence of martingales not in $\mathcal{H}^1$ is an immediate consequence of (4.1). If M is continuous, then the process M^{σ} from the proof of Theorem 4.5 is of the kind mentioned there, i.e., $M_{\infty}^{\sigma} \notin L \log L$. Besides accommodating càdlàg processes M, we make the following pedagogical remark: Theorem 4.5 and its proof require only material from an introductory course on stochastic calculus, and not the reverse L^1 inequality. Moreover, the natural way to prove $M^{\sigma} \notin \mathcal{H}^1$ is to show that (4.5) fails, and not to verify $M_{\infty}^{\sigma} \notin L \log L$.

Example 4.6. Consider a nonnegative local martingale M that belongs to Class $(\mathcal{C}_0)$, according to the terminology of [33], and suppose that $M_0 = 1$. In that case, M is a strictly positive local martingale without any positive jumps, for which $M_{\infty} = M_{\infty-} = 0$. The construction from Theorem 4.5 is then applicable, since $\mathbb{E}[M_{\infty}] = 0 < 1 = \mathbb{E}[M_0]$ implies that M is not uniformly integrable. Moreover, an application of Doob's maximal identity [33, Lemma 2.1] provides the following concrete representation for the non-decreasing sequence $(c_n)_{n\in\mathbb{N}}$, defined by (4.6):

$$c_n = \log\bigg(e + \sum_{k=1}^{n} \frac{1}{k}\bigg), \quad n \in \mathbb{N}.$$

As the harmonic numbers tend to infinity, we have $\lim_{n\uparrow\infty} c_n = \infty$, which is the crucial ingredient for showing that $M^\sigma \notin \mathcal{H}^1$, where the stopping time σ is given by (4.7).

A Auxiliary results for SVEs

This appendix provides auxiliary results needed in Appendix B for the proof of Theorem 3.2. Theorem A.3 establishes a comparison and uniqueness result for deterministic Volterra equations and Theorem A.4 shows existence and uniqueness for stochastic Volterra equations. We first recall the following definition from [19].

Definition A.1 (Volterra kernel of continuous type)**.** A measurable function $\kappa : \Delta \to [0, \infty)$ is a Volterra kernel of continuous type[1] if for each $t \in [0, \infty)$ we have $\kappa(t, .) \in L^1([0, \infty))$ and the map $[0, \infty) \ni t \mapsto \kappa(t, .)$ is continuous.

To apply the results of [39] in the proof of the upcoming existence result for stochastic Volterra equations (SVEs), we establish two facts for Volterra kernels of continuous type.

Lemma A.2. Let κ be a Volterra kernel of continuous type. Then for each bounded measurable function $g : \mathbb{R} \to \mathbb{R}$, the map

$$t \mapsto \int_0^t \kappa(t, s)g(s)\mathrm{d}s$$

is continuous and for each $T > 0$ we have

$$\limsup_{\varepsilon\downarrow 0} \Big\| \int_\cdot^{\cdot+\varepsilon} \kappa(\cdot + \varepsilon, s)\mathrm{d}s \Big\|_{L^\infty([0,T])} = 0.$$

Proof. The first assertion is a special case of [19, Theorem 9.5.3], but we still prove it, as the following steps are required to prove the second assertion. Let $g : \mathbb{R} \to \mathbb{R}$ be bounded and measurable and let $h \geq 0$. Fix $t, u \geq 0$ and first assume $u > t$. Then

$$\begin{aligned}\Big| \int_0^u \kappa(u + h, s)g(s)\mathrm{d}s - \int_0^t \kappa(t + h, s)g(s)\mathrm{d}s \Big| \\ \leq \|g\|_\infty \Big(\int_t^u \kappa(u + h, s)\mathrm{d}s + \int_0^t |\kappa(u + h, s) - \kappa(t + h, s)|\mathrm{d}s \Big).\end{aligned}$$

[1] By abuse of notation, for $t \in [0, \infty)$ we extend $\kappa(t, .)$ from $[0, t)$ to $[0, \infty)$ and set $\kappa(t, s) = 0$ for $s \geq t$.

For $u \downarrow t$, the first integral on the right hand side vanishes via an application of dominated convergence and the second due to the assumed L^1-continuity. The case $u < t$ is similar. Hence, taking $h = 0$, the first claim is shown.

For the second claim, fix $\varepsilon > 0$ and note that

$$\int_t^{t+\varepsilon} \kappa(t+\varepsilon, s)\mathrm{d}s = \int_0^{t+\varepsilon} \kappa(t+\varepsilon, s)\mathrm{d}s - \int_0^t \kappa(t+\varepsilon, s)\mathrm{d}s$$

is a continuous function of t, since it is the difference of two continuous functions by the first part of the proof by taking $g = 1$ and $h = 0$ and $h = \varepsilon$, respectively. Hence, the supremum over the compact interval $[0, T]$ is attained by some maximizing $\hat{t}_\varepsilon \in [0, T]$. Hence, we have

$$\Big\| \int_\cdot^{\cdot+\varepsilon} \kappa(\cdot+\varepsilon, s)\mathrm{d}s \Big\|_{L^\infty([0,T])} = \int_{\hat{t}_\varepsilon}^{\hat{t}_\varepsilon+\varepsilon} \kappa(\hat{t}_\varepsilon+\varepsilon, s)\mathrm{d}s.$$

Since the continuous maps

$$t \mapsto \int_0^t \kappa(t, s)\mathrm{d}s \in \mathbb{R} \quad \text{and} \quad t \mapsto \kappa(t, \cdot) \in L^1([0, \infty))$$

are uniformly continuous on the compact set $[0, T+1]$, there exists for every $\eta > 0$ some $\delta > 0$ such that $|u - v| < \delta$ implies

$$\Big| \int_0^u \kappa(u, s)\mathrm{d}s - \int_0^v \kappa(v, s)\mathrm{d}s \Big| < \eta \quad \text{and} \quad \int_0^\infty \big|\kappa(u, s) - \kappa(v, s)\big|\mathrm{d}s < \eta.$$

Let $\eta > 0$ be arbitrary, choose δ accordingly, and fix an arbitrary $\varepsilon < \delta$. Then we have

$$\begin{aligned}
\Big| \int_{\hat{t}_\varepsilon}^{\hat{t}_\varepsilon+\varepsilon} \kappa(\hat{t}_\varepsilon+\varepsilon, s)\mathrm{d}s \Big| &= \Big| \int_0^{\hat{t}_\varepsilon+\varepsilon} \kappa(\hat{t}_\varepsilon+\varepsilon, s)\mathrm{d}s - \int_0^{\hat{t}_\varepsilon} \kappa(\hat{t}_\varepsilon+\varepsilon, s)\mathrm{d}s \Big| \\
&\leq \Big| \int_0^{\hat{t}_\varepsilon+\varepsilon} \kappa(\hat{t}_\varepsilon+\varepsilon, s)\mathrm{d}s - \int_0^{\hat{t}_\varepsilon} \kappa(\hat{t}_\varepsilon, s)\mathrm{d}s \Big| \\
&\qquad + \Big| \int_0^{\hat{t}_\varepsilon} \kappa(\hat{t}_\varepsilon, s)\mathrm{d}s - \int_0^{\hat{t}_\varepsilon} \kappa(\hat{t}_\varepsilon+\varepsilon, s)\mathrm{d}s \Big| \\
&< \eta + \int_0^\infty \big|\kappa(\hat{t}_\varepsilon+\varepsilon, s) - \kappa(\hat{t}_\varepsilon, s)\big|\mathrm{d}s < 2\eta.
\end{aligned}$$

Since η was arbitrary, this completes the proof. □

Next, we state our main result for deterministic Volterra equations. Using [19, Theorem 13.4.7] we state a comparison result and using [19, Corollary 9.5.6] we argue that a given solution must be unique if the vector field is locally Lipschitz continuous. Note that this deterministic result can be applied pathwise to the SVEs we are interested in.

Theorem A.3. Let κ be a Volterra kernel of continuous type, let $f \in C([0,\infty),\mathbb{R}^m)$ and $h \in C([0,\infty)\times\mathbb{R}^m,\mathbb{R}^m)$ be locally Lipschitz uniformly in time, i.e. for every $n\in\mathbb{N}$ there is some constant $L(n)>0$ with

$$\sup_{t\in[0,n]} |h(t,x)-h(t,y)| \le L(n)|x-y|, \quad |x|,|y|\le n.$$

Suppose the Volterra equation

$$y(t) = f(t) + \int_0^t \kappa(t,s)h(s,y(s))\mathrm{d}s, \quad t\ge 0, \tag{A.1}$$

admits a solution y up to a time of explosion T_∞. Then the solution y is unique on $[0,T_\infty)$. If h is additionally (component-wise) increasing in the second argument, for a continuous function $u\in C([0,T_\infty),\mathbb{R}^m)$ with

$$u(t) \overset{(\le)}{\ge} f(t) + \int_0^t k(t,s)h(s,u(s))\mathrm{d}s, \quad t\in[0,T_\infty)$$

we have $u(t) \overset{(\le)}{\ge} y(t)$ on $[0,T_\infty)$. (Here the inequalities are to be understood component-wise.)

Proof. We argue first that the comparison result holds if the solution y is unique. It suffices to show that [19, Theorem 13.4.7] is applicable, i.e., the assumptions of [19, Theorem 12.2.7] need to be checked. By continuity of f and h, assumptions (i) and (ii) are satisfied. For (iii) we note that for any $T>0$ and $n\in\mathbb{N}$ there exists $C>0$ such that

$$|\kappa(T,s)h(s,y)| \le \kappa(T,s) \sup_{s\in[0,T],|y|\le n} |h(s,y)| \le C\kappa(T,s)$$

for each $s\in[0,T]$ and $|y|\le n$, and, by definition of a Volterra kernel of continuous type, $\kappa(T,\cdot)$ is integrable on $[0,T]$. To verify (iv) we check that the set

$$\left\{ [0,T]\ni t\mapsto \int_0^t \kappa(t,s)h(s,\psi(s))\mathrm{d}s \,\middle|\, \psi\in C_{f(0),n}([0,T],\mathbb{R}^m)\right\}$$

is equi-continuous, where $C_{f(0),n}([0,T],\mathbb{R}^m)$ is the set of $\mathbb{R}^m$-valued continuous functions starting in $f(0)$ and modulus bounded by some $n\in\mathbb{N}$. Consider $u,t\in[0,T]$ with $u<t$ and $\psi\in C_{f(0),n}([0,T],\mathbb{R}^m)$. We then have

$$\begin{aligned}
&\Big|\int_0^t \kappa(t,s)h(s,\psi(s))\mathrm{d}s - \int_0^u \kappa(u,s)h(s,\psi(s))\mathrm{d}s\Big| \\
&\qquad \le \Big|\int_0^u (k(t,s)-k(u,s))h(s,\psi(s))\mathrm{d}s + \int_u^t k(t,s)h(s,\psi(s))\mathrm{d}s\Big| \\
&\qquad \le C\left(\int_0^\infty |k(t,s)-k(u,s)|\mathrm{d}s + \int_u^t |k(t,s)|\mathrm{d}s\right),
\end{aligned}$$

for some $C > 0$. The last two terms converge to zero as u tends to t, since κ is of continuous type. A similar argument applies if $u > t$, hence we obtain equi-continuity. Thus the comparison result holds.

We next argue that the solution is unique on $[0, T]$ for $T \in (0, T_\infty)$. Let y' be another solution defined on $[0, T]$. Note that the function h restricted to the compact set $[0, T] \times [-(\|y\|_\infty + \|y'\|_\infty), \|y\|_\infty + \|y'\|_\infty]^m$ is Lipschitz continuous with constant $L > 0$. Thus for $z = |y - y'|$ it holds

$$z(t) \leq L \int_0^t k(t, s) z(s) \mathrm{d}s, \quad t \in [0, T].$$

Knowing that the linear Volterra equation $z(t) = L \int_0^t k(t, s) z(s) \mathrm{d}s$, $t \in [0, T]$, has zero as its unique solution, see [19, Corollary 9.5.6], we can apply the previously established comparison. Thus $z \equiv 0$ and we have shown uniqueness, which completes the proof. □

In our SVEs of interest we allow for forcing terms given by continuous stochastic processes Z. We show that for locally Lipschitz vector fields, solutions always exist up to a time of explosion. A path-wise application of the previous result immediately yields path-wise uniqueness for these SVEs. Using [30] we have uniqueness in law as well.

Theorem A.4. Let κ be a Volterra kernel of continuous type and consider the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0}, \mathbb{P})$. Let Z be a continuous and $\mathbb{R}^m$-valued adapted process starting in zero. Further, let $h \in C([0, \infty) \times \mathbb{R}^m, \mathbb{R}^m)$ be locally Lipschitz uniformly in time, see Theorem A.3. Then the stochastic Volterra integral equation

$$Y_t = Z_t + \int_0^t \kappa(t, s) h(s, Y_s) \mathrm{d}s, \quad t \geq 0,$$

admits a path-wise unique solution Y, which is adapted and continuous and exists up to a time of explosion $\tau_\infty = \lim_{n \uparrow \infty} \tau_n$, where $\tau_n = \inf\{t \geq 0 : |Y_t| > n\}$. Furthermore, this solution is unique in law.

Proof. Fix $n \in \mathbb{N}$ for the moment. To truncate the function h define

$$h_n(s, y) = \begin{cases} h(s \wedge n, y), & \text{for} \quad |y| \leq n, \\ h\left(s \wedge n, \frac{y}{|y|} n\right), & \text{for} \quad |y| > n, \end{cases}$$

for $s \geq 0$. The resulting function h_n is continuous and bounded by the continuity of h. Further, we have, for each $t \geq 0$ and all $x, y \in \mathbb{R}^m$,

$$|h_n(t, x) - h_n(t, y)| \leq L(n)|x - y|.$$

Define next the stopping time $\sigma_n = \inf\{t \geq 0 : |Z_t| \geq n\}$. By Lemma A.2 we can apply [39, Theorem 3.1], which yields that there exists a unique adapted processes Y^n such that for Lebesgue almost all $t \geq 0$ we have

$$Y_t^n = Z_t^{\sigma_n} + \int_0^t \kappa(t,s) h_n(s, Y_s^n) \mathrm{d}s.$$

By the boundedness of h_n, the right-hand side of the above equation is continuous according to Lemma A.2. Hence, we may assume that Y^n is continuous and satisfies the above equation for all $t \geq 0$ and up to indistinguishability. To remove the dependence on n, first define the stopping times

$$\rho_0 = 0, \quad \rho_n = n \wedge \sigma_n \wedge \inf\{t \geq 0 : |Y_t^n| \geq n\}, \quad n \geq 1.$$

We now establish two facts:

(1) For each $n \in \mathbb{N}$ and $t \in [0, \rho_{n-1}]$ we have $Y_t^{n-1} = Y_t^n$. To see this, note that for $t \in [0, \rho_{n-1}]$ we have $|Y_t^{n-1}| \leq n - 1 < n$ and hence

$$\begin{aligned} Y_t^{n-1} &= Z_t^{\sigma_{n-1}} + \int_0^t \kappa(t,s) h_{n-1}(s, Y_s^{n-1}) \mathrm{d}s \\ &= Z_t^{\sigma_n} + \int_0^t \kappa(t,s) h_n(s, Y_s^{n-1}) \mathrm{d}s. \end{aligned}$$

Uniqueness and continuity therefore yield $Y^{n-1} = Y^n$ on $[0, \rho_{n-1}]$.

(2) We have $\rho_{n-1} \leq \rho_n$ for each $n \in \mathbb{N}$, which follows immediately from (1): On $[0, \rho_{n-1}]$ the process Y^n agrees with Y^{n-1}, which implies $|Y_t^n| \leq n-1$ for each $t \in [0, \rho_{n-1}]$. Hence, for $t \in [0, \rho_{n-1}]$ we have $|Y_t^n| < n$.

Now, we define the process Y by

$$Y_t = \sum_{n=1}^{\infty} Y_t^n 1_{[\rho_{n-1}, \rho_n)}(t), \qquad t \geq 0.$$

Then Y is adapted because Y^n is and $\{\rho_{n-1} \leq t < \rho_n\} \in \mathcal{F}_t$ for each $n \in \mathbb{N}$ and $t \geq 0$. Clearly, Y is continuous on $[0, \lim_{n\uparrow\infty} \rho_n)$ by (1) and (2), and jumps to zero afterwards. To see that Y is a solution of the Volterra equation on $[0, \lim_{n\uparrow\infty} \rho_n)$, we note for $t \in [\rho_{n-1}, \rho_n)$, using $Z_t^{\sigma_n} = Z_t$, that

$$\begin{aligned} Y_t = Y_t^n &= Z_t + \int_0^t \kappa(t,s) h_n(s, Y_s^n) \mathrm{d}s \\ &= Z_t + \sum_{l=1}^{n} \int_{\rho_{l-1}\wedge t}^{\rho_l \wedge t} \kappa(t,s) h_n(s, Y_s^n) \mathrm{d}s \\ &= Z_t + \sum_{l=1}^{n} \int_{\rho_{l-1}\wedge t}^{\rho_l \wedge t} \kappa(t,s) h(s, Y_s^l) \mathrm{d}s = Z_t + \int_0^t \kappa(t,s) h(s, Y_s) \mathrm{d}s. \end{aligned}$$

Hence we have established existence, and path-wise uniqueness immediately follows from Theorem A.3. Uniqueness in law follows from [30, Proposition 2.10]. This completes the proof. □

B Proof of the reverse direction of Theorem 3.2

The proof strategy relies on reducing the multi-factor model to a a one-dimensional Volterra equation, which is studied in [15, Theorem 1.1(2)]. We shall make use of Theorems A.3 and A.4 from Appendix A.

Proof of Theorem 3.2 (Reverse direction). We prove the reverse direction of Theorem 3.2 by contraposition. For convenience we define $\rho = (\sigma^S_{(1:m)})^\top$. By assumption and upon relabelling, we can assume w.l.o.g. that we have $\rho^\top L_1^\top \nu_1 > 0$. Recalling $W_{1:m} = (W^1, \dots, W^m)^\top$, by Lévy's characterization of Brownian motion, the process

$$B^i = \tfrac{1}{|\nu_i^\top L_i|} \nu_i^\top L_i W_{1:m}, \quad i = 1, \dots, N,$$

is a correlated $\mathbb{R}^N$-valued Brownian motion. Let the matrix $\Gamma \in \mathbb{R}^{N\times m}$ be given such that $B = \Gamma W_{1:m}$. For $i = 1, \dots, N$, we define the constants $\beta_i = |\nu_i^\top L_i| > 0$ and the continuous functions

$$g_i : [0,\infty) \ni t \longmapsto \gamma_i \xi_0(t) \exp\big(-\tfrac{\beta_i^2}{2} t^{2\alpha-1}\big) \in (0,\infty).$$

Define the process $\overline{Y}_t = \int_0^t K_{\alpha,1}(t,s)\mathrm{d}B_s$ and the function

$$h(t,y) = \Big(\sum_{i=1}^N g_i(t)\exp(\beta_i y_i)\Big)^{\frac{1}{2}}, \quad t \geq 0, \quad y \in \mathbb{R}^N,$$

such that the variance process $\psi(t,Y)^2$ is equivalently given as $h(t,\overline{Y}_t)^2$. By Novikov's criterion, the sequence of stopping times defined by $\overline{\tau}_n = \inf\{t \geq 0 : \|\overline{Y}_t\|_\infty > n\}$ for $n \in \mathbb{N}$ is a localizing sequence for S. We define the measures $\mathrm{d}\mathbb{P}'_n = S_T^{\overline{\tau}_n}/S_0 \mathrm{d}\mathbb{P}$ for fixed $T > 0$ and $n \in \mathbb{N}$. By Girsanov's theorem, the $\mathbb{R}^N$-valued process

$$B^n = B - \Gamma\rho \int_0^{\cdot\wedge\tau_n} h(s,\overline{Y}_s)\mathrm{d}s$$

is a correlated $\mathbb{P}'_n$-Brownian motion. Using this, we see that $\overline{Y}$ solves the stochastic Volterra equation

$$\overline{Y}_t = \int_0^t K_{\alpha,1}(t,s)\mathrm{d}B^n_s + \Gamma\rho\int_0^{t\wedge\tau_n} K_{\alpha,1}(t,s)h(s,\overline{Y}_s)\mathrm{d}s, \quad t \geq 0.$$

Analogously, the stochastic Volterra equation

$$\widetilde{Y}_t = \int_0^t K_{\alpha,1}(t,s)\mathrm{d}B_s + \Gamma\rho \int_0^t K_{\alpha,1}(t,s)h(s,\widetilde{Y}_s)\mathrm{d}s, \quad t \geq 0, \tag{B.1}$$

admits a unique solution up to a time of explosion $\widetilde{\tau}_\infty$ by Theorem A.4 with announcing sequence $\widetilde{\tau}_n = \inf\{t \geq 0 : \|\widetilde{Y}_t\|_\infty > n\}$ for $n \in \mathbb{N}$, satisfying $\lim_{n\uparrow\infty} \widetilde{\tau}_n = \widetilde{\tau}_\infty$. By uniqueness of solutions in law, the law of $\overline{Y}^{\overline{\tau}_n}$ under $\mathbb{P}'_n$ agrees with the law of $\widetilde{Y}^{\widetilde{\tau}_n}$ under $\mathbb{P}$. Combining this with the classical argument used by [37], we arrive at

$$S_0 - \mathbb{E}\left[S_T\right] = S_0 \lim_{n\uparrow\infty} \mathbb{P}'_n[\overline{\tau}_n \leq T] = S_0\mathbb{P}[\widetilde{\tau}_\infty \leq T], \tag{B.2}$$

and it suffices to show that $\mathbb{P}[\widetilde{\tau}_\infty \leq T] > 0$.

We next reduce the multi-variate setting to the one-dimensional situation. To this end, consider the one-dimensional Volterra equation

$$X_t = \overline{Y}^1_t + (\Gamma\rho)_1 \int_0^t K_{\alpha,1}(t,s)\sqrt{g_1(s)}\exp(\tfrac{1}{2}\beta_1 X_s)\mathrm{d}s, \quad t \geq 0.$$

We claim that the explosion time of the process X is greater than or equal to $\widetilde{\tau}_\infty$. Note that for $t < \widetilde{\tau}_\infty$ we have by (B.1)

$$\widetilde{Y}^1_t \geq \overline{Y}^1_t + (\Gamma\rho)_1 \int_0^t K_{\alpha,1}(t,s)\sqrt{g_1(s)}\exp(\tfrac{1}{2}\beta_1 \widetilde{Y}^1_s)\mathrm{d}s, \quad t \geq 0.$$

By the comparison result Theorem A.3 we must have $\widetilde{Y}^1 \geq X$ as long as both processes exist. Since $X \geq \overline{Y}^1$ and $\overline{Y}$ is finite, the process X cannot explode before $\widetilde{Y}^1$ and the claim holds. Thus it suffices to show that the probability of X exploding before time T is strictly positive under $\mathbb{P}$. This follows from the arguments in [15, Proof of Theorem 1.1(2)]. Thus $\mathbb{P}[\widetilde{\tau}_\infty \leq T] > 0$ and S is a strict local martingale by (B.2), concluding the proof. □

Acknowledgements

We are very grateful to two anonymous referees and an associate editor for their constructive comments. We thank Hardy Hulley for helpful suggestions.

References

[1] E. Abi Jaber, P. Gassiat, and D. Sotnikov, *Martingale property and moment explosions in signature volatility models.* arXiv:2503.17103v2, 2025.

[2] E. Abi Jaber and L.-A. Gérard, *Signature volatility models: Pricing and hedging with Fourier*, SIAM Journal on Financial Mathematics, to appear.

[3] E. Abi Jaber, M. Larsson, and S. Pulido, *Affine Volterra processes*, Ann. Appl. Probab., 29 (2019), pp. 3155–3200.

[4] B. Acciaio, M. Beiglböck, F. Penkner, and W. Schachermayer, *A model-free version of the fundamental theorem of asset pricing and the super-replication theorem*, Math. Finance, 26 (2016), pp. 233–251.

[5] B. Acciaio, M. Beiglböck, F. Penkner, W. Schachermayer, and J. Temme, *A trajectorial interpretation of Doob's martingale inequalities*, Ann. Appl. Probab., 23 (2013), pp. 1494–1505.

[6] G. Alobaidi and R. Mallier, *Laplace transforms and the American straddle*, J. Appl. Math., 2 (2002), pp. 121–129.

[7] C. Bayer, P. Friz, and J. Gatheral, *Pricing under rough volatility*, Quant. Finance, 16 (2016), pp. 887–904.

[8] D. Blackwell and L. E. Dubins, *A converse to the dominated convergence theorem*, Illinois J. Math., 7 (1963), pp. 508–514.

[9] O. Bonesini, A. Jacquier, and A. Muguruza, *Risk premium and rough volatility*, Front. Math. Finance, 7 (2025), pp. 78–94.

[10] C. Cuchiero, G. Gazzani, J. Möller, and S. Svaluto-Ferro, *Joint calibration to SPX and VIX options with signature-based models*, Mathematical Finance, 35 (2025), pp. 161–213.

[11] C. Cuchiero, G. Gazzani, and S. Svaluto-Ferro, *Signature-based models: Theory and calibration*, SIAM Journal on Financial Mathematics, 14 (2023), pp. 910–957.

[12] L. E. Dubins and D. Gilat, *On the distribution of maxima of martingales*, Proc. Amer. Math. Soc., 68 (1978), pp. 337–338.

[13] R. Durrett, *Brownian Motion and Martingales in Analysis*, Wadsworth International Group, Belmont, CA, 1984.

[14] H. Föllmer and A. Schied, *Stochastic Finance*, De Gruyter, Berlin, 4th ed., 2016.

[15] P. Gassiat, *On the martingale property in the rough Bergomi model*, Electron. Commun. Probab., 24 (2019), pp. Paper No. 33, 9 pages.

[16] S. Gerhold, C. Gerstenecker, and A. Pinter, *Moment explosions in the rough Heston model*, Decis. Econ. Finance, 42 (2019), pp. 575–608.

[17] S. Gerhold and A. Pinter, *Difference equation theory meets mathematical finance*, in Algorithmic Combinatorics: Enumerative Combinatorics, Special Functions and Computer Algebra—In Honour of Peter Paule on his 60th Birthday, Springer, Cham, 2020, pp. 197–213.

[18] D. Gilat, *The best bound in the $L \log L$ inequality of Hardy and Littlewood and its martingale counterpart*, Proc. Amer. Math. Soc., 97 (1986), pp. 429–436.

[19] G. Gripenberg, S.-O. Londen, and O. Staffans, *Volterra Integral and Functional Equations*, vol. 34 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1990.

[20] A. Gulisashvili, *Analytically Tractable Stochastic Stock Price Models*, Springer Finance, Springer, Heidelberg, 2012.

[21] ——, *Gaussian stochastic volatility models: scaling regimes, large deviations, and moment explosions*, Stochastic Process. Appl., 130 (2020), pp. 3648–3686.

[22] R. F. Gundy, *On the class $L \log L$, martingales, and singular integrals*, Studia Math., 33 (1969), pp. 109–118.

[23] R. F. Gundy, *Inégalités pour martingales à un et deux indices: l'espace H^p*, in Eighth Saint Flour Probability Summer School–1978, vol. 774 of Lecture Notes in Math., Springer, Berlin, 1980, pp. 251–334.

[24] J. Kallsen and C. Kühn, *Convertible bonds: financial derivatives of game type*, in Exotic Option Pricing and Advanced Lévy Models, Wiley, Chichester, 2005, pp. 277–291.

[25] T. Kämpke and F. J. Radermacher, *Income Modeling and Balancing*, Springer, 2015.

[26] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, vol. 113 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 1991.

[27] ——, *Methods of Mathematical Finance*, Springer-Verlag, New York, 1998.

[28] M. Keller-Ressel, *Moment explosions and long-term behavior of affine stochastic volatility models*, Math. Finance, 21 (2011), pp. 73–98.

[29] M. Keller-Ressel and A. Majid, *A comparison principle between rough and non-rough Heston models—with applications to the volatility surface*, Quant. Finance, 20 (2020), pp. 919–933.

[30] T. Kurtz, *The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities*, Electronic Journal of Probability, 12 (2007), pp. 951 – 965.

[31] C. Lacombe, A. Muguruza, and H. Stone, *Asymptotics for volatility derivatives in multi-factor rough volatility models*, Math. Financ. Econ., 15 (2021), pp. 545–577.

[32] G. Lowther, *Martingales with non-integrable maximum*, 2016. `https://almostsuremath.com/2016/09/12/martingales-with-non-integrable-maximum/` [Accessed: (2026-03-11)].

[33] A. Nikeghbali and M. Yor, *Doob's maximal identity, multiplicative decompositions and enlargements of filtrations*, Illinois J. Math., 50 (2006), pp. 791–814.

[34] G. Peškir, *On the exponential Orlicz norms of stopped Brownian motion*, Studia Math., 117 (1996), pp. 253–273.

[35] P. E. Protter, *Stochastic Integration and Differential Equations*, vol. 21 of Applications of Mathematics (New York), Springer-Verlag, Berlin, second ed., 2004.

[36] J. Ruf, *The martingale property in the context of stochastic differential equations*, Electron. Commun. Probab., 20 (2015), pp. No. 34, 10.

[37] C. A. Sin, *Complications with stochastic volatility models*, Adv. in Appl. Probab., 30 (1998), pp. 256–268.

[38] E. M. Stein, *Note on the class $L \log L$*, Studia Math., 32 (1969), pp. 305–310.

[39] X. Zhang, *Stochastic Volterra equations in Banach spaces and stochastic partial differential equation*, J. Funct. Anal., 258 (2010), pp. 1361–1425.