SCALING AND MULTISCALING IN FINANCIAL SERIES: A SIMPLE MODEL

ALESSANDRO ANDREOLI, FRANCESCO CARAVENNA, PAOLO DAI PRA, AND GUSTAVO POSTA

ABSTRACT. We propose a simple stochastic volatility model which is analytically tractable, very easy to simulate and which captures some relevant stylized facts of financial assets, including *scaling properties*. In particular, the model displays a crossover in the log-return distribution from power-law tails (small time) to a Gaussian behavior (large time), slow decay in the volatility autocorrelation and multiscaling of moments. Despite its few parameters, the model is able to fit several key features of the time series of financial indexes, such as the *Dow Jones Industrial Average*, with a remarkable accuracy.

1. INTRODUCTION

1.1. **Modeling financial assets.** Recent developments in stochastic modelling of time series have been strongly influenced by the analysis of financial assets, such as exchange rates, stocks, and market indexes. The basic model, that has given rise to the celebrated Black & Scholes formula [22, 23], assumes that the logarithm X_t of the price of the asset, after subtracting the trend, evolves through the simple equation

$$\mathrm{d}X_t = \sigma \, \mathrm{d}B_t, \tag{1.1}$$

where σ (the *volatility*) is a constant and $(B_t)_{t\geq 0}$ is a standard Brownian motion. It has been well-know for a long time that, despite its success, this model is not consistent with a number of *stylized facts* that are empirically detected in many real time series, e.g.:

- the volatility is not constant and may exhibit high peaks, that may be interpreted as *shocks* in the market;
- the empirical distribution of the increments $X_{t+h} - X_t$ of the logarithm of the price — called *log-returns* — is non Gaussian, displaying *power-law tails* (see Figure 4(B) below), especially for small values of the time span h, while a Gaussian shape is approximately recovered for large values of h;
- log-returns corresponding to disjoint time-interval are uncorrelated, but not independent: in fact, the correlation between the absolute values $|X_{t+h} - X_t|$ and $|X_{s+h} - X_s|$ — called *volatility autocorrelation* — is positive (*clustering of volatility*) and has a slow decay in $|t - s|$ (*long memory*), at least up to moderate values for $|t - s|$ (cf. Figure 3(B)-(C) below).

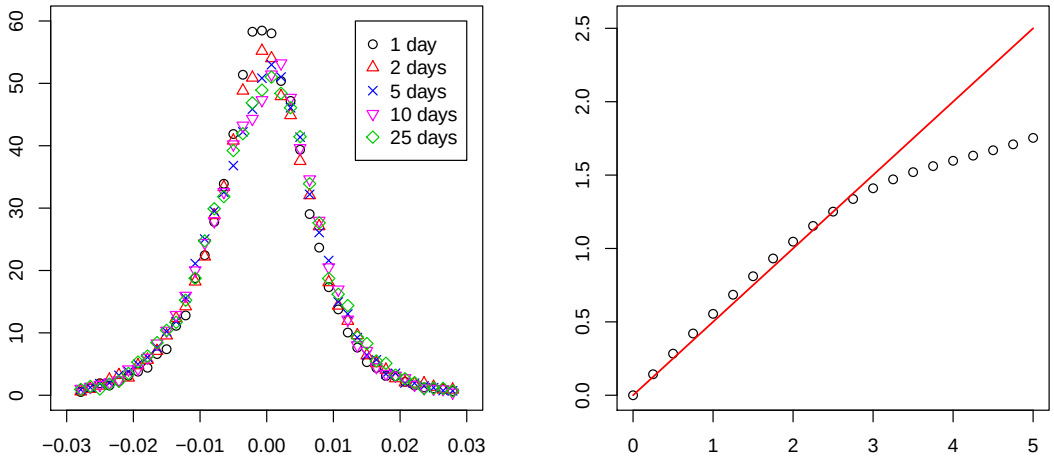
In order to account for these facts, a very popular choice in the literature of mathematical finance and financial economics has been to upgrade the basic model (1.1), allowing $\sigma = \sigma_t$ to vary with t and to be itself a stochastic process. This produces a wide class of processes,

Date: April 18, 2012.

2010 *Mathematics Subject Classification.* 60G44, 91B25, 91G70.

Key words and phrases. Financial Index, Time Series, Scaling, Multiscaling, Brownian Motion, Stochastic Volatility, Heavy Tails, Multifractal Models.

We gratefully acknowledge the support of the University of Padova under grant CPDA082105/08.



(A) *Diffusive scaling of log-returns.* (B) *Multiscaling of moments.*

FIGURE 1. *Scaling properties of the DJIA time series (opening prices 1935-2009).*
(A) The empirical densities of the log-returns over 1, 2, 5, 10, 25 days show a remarkable overlap under diffusive scaling.
(B) The scaling exponent $A(q)$ as a function of q, defined by the relation $m_q(h) \approx h^{A(q)}$ (cf. (1.4)), bends down from the Gaussian behavior $q/2$ (red line) for $q \geq \bar{q} \simeq 3$. The quantity $A(q)$ is evaluated empirically through a linear interpolation of $(\log m_q(h))$ versus $(\log h)$ for $h \in \{1, \dots, 5\}$ (cf. section 7 for more details).

known as *stochastic volatility models*, determined by the process $(\sigma_t)_{t \geq 0}$, which are able to capture (at least some of) the above-mentioned stylized facts, cf. [6, 29] and references therein.

More recently, other stylized facts have been pointed out concerning the *scaling properties* of the empirical distribution of the log-returns. Given a daily time series $(s_i)_{1 \leq i \leq T}$ over a period of $T \gg 1$ days, denote by p_h the *empirical distribution* of the (detrended) log-returns corresponding to an interval of h days:

$$p_h(\cdot) := \frac{1}{T-h} \sum_{i=1}^{T-h} \delta_{x_{i+h} - x_i}(\cdot)\,, \qquad x_i := \log(s_i) - \overline{d}_i\,, \tag{1.2}$$

where $\overline{d}_i$ is the local rate of linear growth of $\log(s_i)$ (see section 7 for details) and $\delta_x(\cdot)$ denotes the Dirac measure at $x \in \mathbb{R}$. The statistical analysis of various financial series, such as the *Dow Jones Industrial Average* (DJIA) or the *Nikkei 225*, shows that, for small values of h, p_h obeys approximately a *diffusive scaling relation* (cf. Figure 1(A)):

$$p_h(\mathrm{d}r) \simeq \frac{1}{\sqrt{h}}\, g\!\left(\frac{r}{\sqrt{h}}\right) \mathrm{d}r, \tag{1.3}$$

where g is a probability density with power-law tails. Considering the q-th empirical moment $m_q(h)$, defined by

$$m_q(h) \;:=\; \frac{1}{T-h} \sum_{i=1}^{T-h} |x_{i+h} - x_i|^q \;=\; \int |r|^q\, p_h(\mathrm{d}r)\,, \tag{1.4}$$

from relation (1.3) it is natural to guess that $m_q(h)$ should scale as $h^{q/2}$. This is indeed what one observes for moments of small order $q \leq \bar{q}$ (with $\bar{q} \simeq 3$ for the DJIA). However, for

moments of higher order $q > \bar{q}$, the different scaling relation $h^{A(q)}$, with $A(q) < q/2$, takes place, cf. Figure 1(B). This is the so-called *multiscaling of moments*, cf. [32, 21, 20, 17].

An interesting class of models that are able to reproduce the multiscaling of moments — as well as many other features, notably the persistence of volatility on different time scales — are the so-called *multifractal models*, like the MMAR (Multifractal Model of Asset Returns, cf. [14]) and the MSM (Markov-Switching Multifractal, cf. [13]). These models describe the evolution of the detrended log-price $(X_t)_{t\geq 0}$ as a *random time-change of Brownian motion*: $X_t = W_{I(t)}$, where the time-change $(I(t))_{t\geq 0}$ is a continuous and increasing process, sometimes called *trading time*, which displays *multifractal features* and is usually taken to be independent of the Brownian motion $(W_t)_{t\geq 0}$ (cf. [12] for more details).

Modeling financial series through a random time-change of Brownian motion is a classical topic, dating back to Clark [15], and reflects the natural idea that external information influences the speed at which exchanges take place in a market. It should be stressed that, under the mild regularity assumption that the time-change $(I(t))_{t\geq 0}$ has absolutely continuous paths a.s., any random time-change of Brownian motion $X_t = W_{I(t)}$ can be written as a stochastic volatility model $\mathrm{d}X_t = \sigma_t \,\mathrm{d}B_t$, and viceversa.[1] However, a key feature of multifractal models is precisely that their trading time $(I(t))_{t\geq 0}$ has *non absolutely continuous paths* a.s., hence they cannot be expressed as stochastic volatility models. This makes their analysis harder, as the standard tools available for Ito diffusions cannot be applied.

The purpose of this paper is to define a *simple* stochastic volatility model — or, equivalently, an independent random time change of Brownian motion, where the time-change process has absolutely continuous paths — which agrees with all the above mentioned stylized facts, displaying in particular a crossover in the log-return distribution from a power-law to a Gaussian behavior, slow decay in the volatility autocorrelation, diffusive scaling and multiscaling of moments. In its most basic version, the model contains only three real parameters and is defined as a simple, deterministic function of a Brownian motion and an independent Poisson process. This makes the process analytically tractable and very easy to simulate. Despite its few degrees of freedom, the model is able to fit remarkably well several key features of the time series of the main financial indexes, such as DJIA, S&P 500, FTSE 100, Nikkei 225. In this paper we present a detailed numerical analysis on the DJIA.

Let us mention that there are subtler stylized facts that are not properly accounted by our model, such as the multi-scale intermittency of the volatility profile, the possible skewness of the log-return distribution and the so-called *leverage effect* (negative correlation between log-returns and future volatilities), cf. [16]. As we discuss in section 3, such features — that are relevant in the analysis of particular assets — can be incorporated in our model in a natural way. Generalizations in this sense are currently under investigation, as are the performances of our model in financial problems, like the pricing of options (cf. A. Andreloli's Ph.D. Thesis [2]). In this article we stick for simplicity to the most basic formulation.

Finally, although we work in the framework of stochastic volatility models, we point out that an important alternative class of models in discrete time, widely used in the econometric literature, is given by autoregressive processes such as ARCH, GARCH, FIGARCH and their

[1] More precisely, "independent random time changes of Brownian motion with absolutely continuous time-change" — that is, processes $(X_t)_{t\geq 0}$ such that $X_t - X_0 = W_{I(t)}$, where $(W_t)_{t\geq 0}$ is a Brownian motion and $(I(t))_{t\geq 0}$ is an independent process with increasing and absolutely continuous paths a.s. — and "stochastic volatility models with independent volatility" — that is, processes $(X_t)_{t\geq 0}$ such that $\mathrm{d}X_t = \sigma_t \,\mathrm{d}B_t$, where $(B_t)_{t\geq 0}$ is a Brownian motion and $(\sigma_t)_{t\geq 0}$ is an independent process with paths in $L^2_{\mathrm{loc}}(\mathbb{R})$ a.s. — are the same class of processes, cf. [6, 29]. The link between the two representations $\mathrm{d}X_t = \sigma_t \,\mathrm{d}B_t$ and $X_t - X_0 = W_{I(t)}$ is given by $\sigma_t = \sqrt{I'(t)}$ and $B_t = \int_0^t (I'(s))^{-1/2} \mathrm{d}W_{I(s)} = \int_0^{I_t} (I'(I^{-1}(v)))^{-1/2} \mathrm{d}W_u$.

generalizations, cf. [18, 8, 5, 9]. More recently, continuous-time versions have been studied as well, cf. [24, 25]. With no aim for completeness, let us mention that GARCH and FIGARCH do not display multiscaling of moments, cf. [12, §8.1.4]. We have also tested the model recently proposed in [10], which is extremely accurate to fit the statistics of the empirical volatility, and it does exhibit multiscaling of moments. The price to pay is, however, that the model requires the calibration of more than 30 parameters.

We conclude noting that long memory effects in autoregressive models are obtained through a suitable dependence on the past in the equation for the volatility, while large price variations are usually controlled by specific features of the driving noise. In our model, we propose a *single* mechanism, modeling the reaction of the market to shocks, which is the source of all the mentioned stylized facts.

1.2. **Content of the paper.** The paper is organized as follows.

- In section 2 we give the definition of our model, we state its main properties and we discuss its ability to fit the DJIA time series in the period 1935-2009.
- In section 3 we discuss some key features and limitations of our model, point out possible generalizations, and compare it with other models.
- Sections 4, 5 and 6 contain the proofs of the main results, plus some additional material.
- In section 7 we discuss more in detail the numerical comparison between our model and the DJIA time series.
- Finally, appendix A contains the proof of some technical results, while appendix B is devoted to a brief discussion of the model introduced by F. Baldovin and A. Stella in [7, 31], which has partially inspired the construction of our model.

1.3. **Notation.** Throughout the paper, the indexes s, t, u, x, λ run over real numbers while i, k, m, n run over integers, so that $t \geq 0$ means $t \in [0, \infty)$ while $n \geq 0$ means $n \in \{0, 1, 2, \ldots\}$. The symbol "$\sim$" denotes asymptotic equivalence for positive sequences ($a_n \sim b_n$ if and only if $a_n/b_n \to 1$ as $n \to \infty$) and also equality in law for random variables, like $W_1 \sim \mathcal{N}(0, 1)$. Given two real functions $f(x)$ and $g(x)$, we write $f = O(g)$ as $x \to x_0$ if there exists $M > 0$ such that $|f(x)| \leq M|g(x)|$ for x in a neighborhood of x_0, while we write $f = o(g)$ if $f(x)/g(x) \to 0$ as $x \to x_0$; in particular, $O(1)$ (resp. $o(1)$) is a bounded (resp. a vanishing) quantity. The standard exponential and Poisson laws are denoted by $Exp(\lambda)$ and $Po(\lambda)$, for $\lambda > 0$: $X \sim Exp(\lambda)$ means that $P(X \leq x) = (1 - e^{-\lambda x})\mathbf{1}_{[0,\infty)}(x)$ for all $x \in \mathbb{R}$ while $Y \sim Po(\lambda)$ means that $P(Y = n) = e^{-\lambda}\lambda^n/n!$ for all $n \in \{0, 1, 2, \ldots\}$. We sometimes write (*const.*) to denote a positive constant, whose value may change from place to place.

2. The model and the main results

We introduce our model as an *independent random time change* of a Brownian motion, in the spirit e.g. of [15] and [4]. An alternative and equivalent definition, as a stochastic volatility model, is illustrated in section 2.2.

2.1. **Definition of the model.** In its basic version, our model contains only three real parameters:

- $\lambda \in (0, +\infty)$ is the inverse of the average waiting time between "shocks" in the market;
- $D \in (0, 1/2]$ determines the sub-linear time change $t \mapsto t^{2D}$, which expresses the "trading time" after shocks;

- $\sigma \in (0, +\infty)$ is proportional to the average volatility.

In order to have more flexibility, we actually let σ be a *random* parameter, i.e., a positive random variable, whose distribution ν becomes the relevant parameter:

- ν is a probability on $(0, \infty)$, *connected* to the volatility distribution.

Remark 1. When the model is calibrated to the main financial indexes (DJIA, S&P 500, FTSE 100, Nikkei 225), the best fitting turns out to be obtained for a *nearly constant* σ. In any case, we stess that the main properties of the model are only marginally dependent on the law ν of σ: in particular, the first two moments of ν, i.e. $E(\sigma)$ and $E(\sigma^2)$, are enough to determine the features of our model that are relevant for real-world times series, cf. Remark 10 below. Therefore, roughly speaking, we could say that in the general case of random σ our model has four "effective" real parameters.

Beyond the parameters λ, D, ν, we need the following three sources of alea:

- a standard Brownian motion $W = (W_t)_{t\geq 0}$;
- a Poisson point process $\mathcal{T} = (\tau_n)_{n\in\mathbb{Z}}$ on $\mathbb{R}$ with intensity λ;
- a sequence $\Sigma = (\sigma_n)_{n\geq 0}$ of independent and identically distributed positive random variables with law ν (so that $\sigma_n \sim \nu$ for all n); and for conciseness we denote by σ a variable with the same law ν.

We assume that $W, \mathcal{T}, \Sigma$ are defined on some probability space $(\Omega, \mathcal{F}, P)$ and that they are independent. By convention, we label the points of $\mathcal{T}$ so that $\tau_0 < 0 < \tau_1$. We will actually need only the points $(\tau_n)_{n\geq 0}$, and we recall that the random variables $(-\tau_0)$, τ_1, $(\tau_{n+1}-\tau_n)_{n\geq 1}$ are independent and identically distributed with marginal laws $Exp(\lambda)$. In particular, $1/\lambda$ is the mean distance between the points in $\mathcal{T}$, except for τ_0 and τ_1, whose average distance is $2/\lambda$. Although some of our results would hold for more general distributions of $\mathcal{T}$, we stick for simplicity to the (rather natural) choice of a Poisson process.

For $t \geq 0$, we define

$$i(t) \,:=\, \sup\{n \geq 0 :\, \tau_n \leq t\} \,=\, \#\{\mathcal{T} \cap (0, t]\}\,, \tag{2.1}$$

so that $\tau_{i(t)}$ is the location of the last point in $\mathcal{T}$ before t. Plainly, $i(t) \sim Po(\lambda t)$. Then we introduce the basic process $I = (I_t)_{t\geq 0}$ defined by

$$I_t \,=\, I(t) \,:=\, \sigma_{i(t)}^2 \left(t - \tau_{i(t)}\right)^{2D} \,+\, \sum_{k=1}^{i(t)} \sigma_{k-1}^2 \left(\tau_k - \tau_{k-1}\right)^{2D} \,-\, \sigma_0^2 \left(-\tau_0\right)^{2D}\,, \tag{2.2}$$

with the agreement that the sum in the right hand side is zero if $i(t) = 0$. More explicitly, $(I_t)_{t\geq 0}$ is a continuous process with $I_0 = 0$ and $I_{\tau_n+h} - I_{\tau_n} = (\sigma_n^2)\, h^{2D}$ for $0 \leq h \leq (\tau_{n+1}-\tau_n)$. We note that the derivative $(\frac{\mathrm{d}}{\mathrm{d}t} I_t)_{t\geq 0}$ is a *stationary regenerative process*, cf. [3]. See Figure 2 for a sample trajectory of $(I_t)_{t\geq 0}$ when $D < \frac{1}{2}$.

We then define our model $X = (X_t)_{t\geq 0}$ by setting

$$X_t := W_{I_t}\,. \tag{2.3}$$

In words: our model is an random time change of the Brownian motion $(W_t)_{t\geq 0}$ through the time-change process $(I_t)_{t\geq 0}$. Note that I is a strictly increasing process with absolutely continuous paths, and it is independent of W.

When $D = \frac{1}{2}$ and σ is constant, we have $I_t = \sigma^2 t$ and the model reduces to Black & Scholes with volatility σ. On the other hand, when $D < \frac{1}{2}$, the paths of I are singular (non differentiable) at the points in $\mathcal{T}$, cf. Figure 2. This suggests a possible financial interpretation

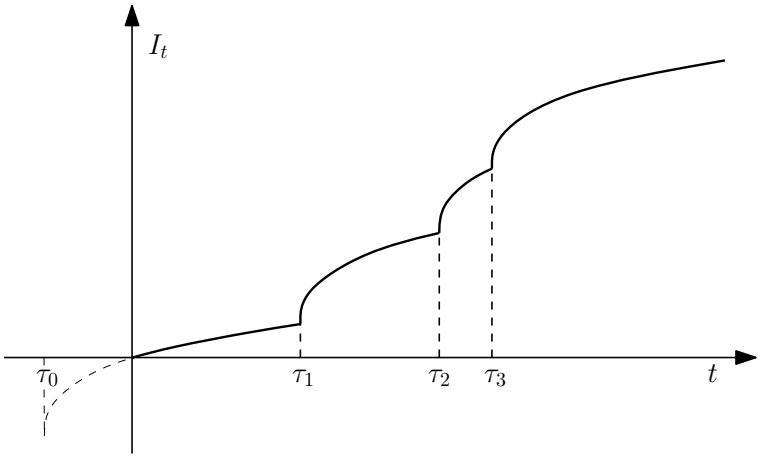


FIGURE 2. A sample trajectory of the process $(I_t)_{t\geq 0}$

of the instants in $\mathcal{T}$ as the epochs at which *big shocks* arrive in the market, making the volatility jump to infinity. This will be more apparent in next subsection, where we give a stochastic volatility formulation of the model. We point out that the singularity is produced by the sub-linear time change $t \mapsto t^{2D}$, that was first suggested by F. Baldovin and A. Stella in [7, 31] (their model is described in appendix B).

2.2. **Basic properties.** Let us state some basic properties of our model, that will be proved in section 6.

(A) *The process X has stationary increments.*

(B) *The following relation between moments of X_t and σ holds*: for any $q > 0$

$$E(|X_t|^q) < \infty \text{ for some (hence any) } t > 0 \quad \Longleftrightarrow \quad E(\sigma^q) < \infty\,. \tag{2.4}$$

(C) *The process X can be represented as a stochastic volatility model*:

$$\mathrm{d}X_t = v_t\,\mathrm{d}B_t\,, \tag{2.5}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion and $(v_t)_{t\geq 0}$ is an independent process, defined by (denoting $I'(s) := \frac{\mathrm{d}}{\mathrm{d}s} I(s)$)

$$\begin{aligned} B_t &:= \int_0^t \frac{1}{\sqrt{I'(s)}}\,\mathrm{d}W_{I(s)} = \int_0^{I_t} \frac{1}{\sqrt{I'(I^{-1}(u))}}\,\mathrm{d}W_u\,, \\ v_t &:= \sqrt{I'(t)} = \sqrt{2D}\,\sigma_{i(t)}\left(t - \tau_{i(t)}\right)^{D-\frac{1}{2}}\,. \end{aligned} \tag{2.6}$$

Note that, whenever $D < \frac{1}{2}$, the volatility v_t has singularities at the random times τ_n.

(D) *The process X is a zero-mean, square-integrable martingale* (provided $E(\sigma^2) < \infty$).

Remark 2. If we look at the process X for a *fixed* realization of the variables $\mathcal{T}$ and Σ, averaging only on W — that is, if we work under the conditional probability $P(\,\cdot\,|\mathcal{T}, \Sigma)$ — the increments of X are no longer stationary, but the properties (C) and (D) continue to hold (of course, condition $E(\sigma^2) < \infty$ in (D) is not required under $P(\,\cdot\,|\mathcal{T}, \Sigma)$).

Remark 3. It follows from (2.4) that if σ is chosen as a deterministic constant, then X_t admits moments of all order (actually, even exponential moments, cf. Proposition 11 in section 6). This seems to indicate that to *see* power-law tails in the distribution of $(X_{t+h} - X_t)$ — one of the basic stylized facts mentioned in the introduction — requires to take σ with

power-law tails. *This, however, is not true*, and is one on the surprising features of the simple model we propose: for typical choices of the parameters of our model, the distribution of $(X_{t+h} - X_t)$ displays a power-law tail behavior up to several standard deviations from the mean, irrespective of the law of σ. Thus, the eventually light tails are "invisible" for all practical purposes and real heavy-tailed distributions appear to be unnecessary to fit data. We discuss this issue below, cf. Remark 5, after having stated some results; see also subsection 2.4 and Figure 4(B) for a graphical comparisons with the DJIA time series.

Another important property of the process X is that its increments are *mixing*, as we show in section 6. This entails in particular that for every $\delta > 0$, $k \in \mathbb{N}$ and for every choice of the intervals $(a_1, b_1), \ldots, (a_k, b_k) \subseteq (0, \infty)$ and of the measurable function $F : \mathbb{R}^k \to \mathbb{R}$, we have almost surely

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} F(X_{n\delta+b_1} - X_{n\delta+a_1}, \ldots, X_{n\delta+b_k} - X_{n\delta+a_k}) = E\big[F(X_{b_1} - X_{a_1}, \ldots, X_{b_k} - X_{a_k})\big], \tag{2.7}$$

provided the expectation appearing in the right hand side is well defined. In words: the empirical average of a function of the increments of the process over a long time period is close to its expected value.

Thanks to this property, our main results concerning the distribution of the increments of the process X, that we state in the next subsection, are of direct relevance for the comparison of our model with real data series. Some care is needed, however, because the accessible time length N in (2.7) may not be large enough to ensure that the empirical averages are close to their limit. We elaborate more on this issue in section 7, where we compare our model with the DJIA data from a numerical viewpoint.

2.3. **The main results.** We now state our main results for our model X, that correspond to the basic stylized facts mentioned in the introduction: diffusive scaling and crossover of the log-return distribution (Theorem 4); multiscaling of moments (Theorem 6 and Corollary 7); clustering of volatility (Theorem 8 and Corollary 9).

Our first result, proved in section 4, shows that the increments $(X_{t+h} - X_t)$ have an approximate diffusive scaling both when $h \downarrow 0$, with a heavy-tailed limit distribution (in agreement with (1.3)), and when $h \uparrow \infty$, with a normal limit distribution. This is a precise mathematical formulation of a crossover phenomenon in the log-return distribution, from approximately heavy-tailed (for small time) to approximately Gaussian (for large time).

Theorem 4 (Diffusive scaling)**.** *The following convergences in distribution hold for any choice of the parameters D, λ and of the law ν of σ.*

- *Small-time diffusive scaling:*

$$\frac{(X_{t+h} - X_t)}{\sqrt{h}} \xrightarrow[h\downarrow 0]{d} f(x)\,\mathrm{d}x := \text{law of } \left(\sqrt{2D}\,\lambda^{\frac{1}{2}-D}\right)\sigma\, S^{D-\frac{1}{2}}\, W_1\,, \tag{2.8}$$

  where $\sigma \sim \nu$, $S \sim Exp(1)$ and $W_1 \sim \mathcal{N}(0,1)$ are independent random variables. The density f is thus a mixture of centered Gaussian densities and, when $D < \frac{1}{2}$, has power-law tails*: more precisely, if $E(\sigma^q) < \infty$ for all $q > 0$,*

$$\int |x|^q f(x)\,\mathrm{d}x < +\infty \iff q < q^* := \frac{1}{(\frac{1}{2} - D)}. \tag{2.9}$$

- *Large-time diffusive scaling: if* $E(\sigma^2) < \infty$

$$\frac{(X_{t+h} - X_t)}{\sqrt{h}} \xrightarrow[h\uparrow\infty]{d} \frac{e^{-x^2/(2c^2)}}{\sqrt{2\pi}c}\,\mathrm{d}x = \mathcal{N}(0, c^2)\,, \qquad c^2 = \lambda^{1-2D}\,E(\sigma^2)\,\Gamma(2D+1)\,, \quad (2.10)$$

where $\Gamma(\alpha) := \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x$ *denotes Euler's Gamma function.*

Remark 5. We have already observed that, when σ has finite moments of all orders, for $h > 0$ the increment $(X_{t+h} - X_t)$ has finite moments of all orders too, cf. (2.4), so there are no heavy tails in the strict sense. However, for h small, the heavy-tailed density $f(x)$ is by (2.8) an excellent approximation for the true distribution of $\frac{1}{\sqrt{h}}(X_{t+h} - X_t)$ up to a certain distance from the mean, which can be quite large. For instance, when the parameters of our model are calibrated to the DJIA time series, these "apparent power-law tails" are clearly visible for $h = 1$ (daily log-returns) up to a distance of about *six standard deviations from the mean*, cf. subsection 2.4 and Figure 4(B) below.

We also note that the moment condition (2.9) follows immediately from (2.8): in fact, when σ has finite moments of all orders,

$$\int |x|^q f(x)\,\mathrm{d}x < +\infty \iff E\big[S^{(D-1/2)q}\big] = \int_0^\infty s^{(D-1/2)q}\,e^{-s}\,\mathrm{d}s < +\infty\,, \qquad (2.11)$$

which clearly happens if and only if $q < q^* := (\frac{1}{2} - D)^{-1}$. This also shows that the heavy tails of f depend on the fact that the density of the random variable S, which represents (up to a constant) the distance between points in $\mathcal{T}$, is strictly positive around zero, and not on other details of the exponential distribution.

The power-law tails of f have striking consequences on the scaling behavior of the moments of the increments of our model. If we set for $q \in (0, \infty)$

$$m_q(h) := E(|X_{t+h} - X_t|^q)\,, \qquad (2.12)$$

the natural scaling $m_q(h) \approx h^{q/2}$ as $h \downarrow 0$, that one would naively guess from (2.8), breaks down for $q > q^*$, when the faster scaling $m_q(h) \approx h^{Dq+1}$ holds instead, the reason being precisely the fact that the q-moment of f is infinite for $q \geq q^*$. More precisely, we have the following result, that we prove in section 4.

Theorem 6 (Multiscaling of moments)**.** *Let* $q > 0$*, and assume* $E(\sigma^q) < +\infty$*. The quantity* $m_q(h)$ *in* (2.12) *is finite and has the following asymptotic behavior as* $h \downarrow 0$*:*

$$m_q(h) \sim \begin{cases} C_q\,h^{\frac{q}{2}} & \text{if } q < q^* \\ C_q\,h^{\frac{q}{2}}\log(\frac{1}{h}) & \text{if } q = q^* \\ C_q\,h^{Dq+1} & \text{if } q > q^* \end{cases}, \qquad \text{where } \; q^* := \frac{1}{(\frac{1}{2} - D)}\,.$$

The constant $C_q \in (0, \infty)$ *is given by*

$$C_q := \begin{cases} E(|W_1|^q)\,E(\sigma^q)\,\lambda^{q/q^*}\,(2D)^{q/2}\,\Gamma(1 - q/q^*) & \text{if } q < q^* \\ E(|W_1|^q)\,E(\sigma^q)\,\lambda\,(2D)^{q/2} & \text{if } q = q^* \\ E(|W_1|^q)\,E(\sigma^q)\,\lambda\,\big[\int_0^\infty ((1+x)^{2D} - x^{2D})^{\frac{q}{2}}\,\mathrm{d}x + \frac{1}{Dq+1}\big] & \text{if } q > q^* \end{cases}, \qquad (2.13)$$

where $\Gamma(\alpha) := \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x$ *denotes Euler's Gamma function.*

Corollary 7. *The following relation holds true:*

$$A(q) := \lim_{h\downarrow 0} \frac{\log m_q(h)}{\log h} = \begin{cases} \frac{q}{2} & \text{if } q \le q^* \\ Dq+1 & \text{if } q \ge q^* \end{cases}, \qquad \text{where } \; q^* := \frac{1}{(\frac{1}{2}-D)}\,. \tag{2.14}$$

The explicit form (2.13) of the multiplicative constant C_q will be used in section 7 for the estimation of the parameters of our model on the DJIA time series.

Our last theoretical result, proved in section 5, concerns the correlations of the absolute value of two increments, usually called *volatility autocorrelation*. We start determining the behavior of the covariance.

Theorem 8. *Assume that $E(\sigma^2) < \infty$. The following relation holds as $h \downarrow 0$, for all $s, t > 0$:*

$$Cov(|X_{s+h} - X_s|, |X_{t+h} - X_t|) = \frac{4D}{\pi}\,\lambda^{1-2D}\, e^{-\lambda|t-s|}\left(\phi(\lambda|t-s|)\,h \,+\, o(h)\right), \tag{2.15}$$

where

$$\phi(x) := Cov\big(\sigma\, S^{D-1/2}\,,\, \sigma\,\big(S+x\big)^{D-1/2}\big) \tag{2.16}$$

and $S \sim Exp(1)$ is independent of σ.

We recall that $\rho(Y,Z) := Cov(Y,Z)/\sqrt{Var(Y)Var(Z)}$ is the correlation coefficient of two random variables Y, Z. As Theorem 6 yields

$$\lim_{h\downarrow 0} \frac{1}{h}\, Var(|X_{t+h} - X_t|) \;=\; (2D)\,\lambda^{1-2D}\, Var(\sigma\,|W_1|\,S^{D-1/2})\,,$$

where $S \sim Exp(1)$ is independent of σ, W_1, we easily obtain the following result.

Corollary 9 (Volatility autocorrelation)**.** *Assume that $E(\sigma^2) < \infty$. The following relation holds as $h \downarrow 0$, for all $s, t > 0$:*

$$\begin{aligned} &\lim_{h\downarrow 0} \rho(|X_{s+h} - X_s|, |X_{t+h} - X_t|) \\ &\qquad = \rho(t-s) \;:=\; \frac{2}{\pi\, Var(\sigma\,|W_1|\,S^{D-1/2})}\, e^{-\lambda|t-s|}\,\phi(\lambda|t-s|)\,, \end{aligned} \tag{2.17}$$

where $\phi(\cdot)$ is defined in (2.16) *and $\sigma \sim \nu$, $S \sim Exp(1)$, $W_1 \sim \mathcal{N}(0,1)$ are independent random variables.*

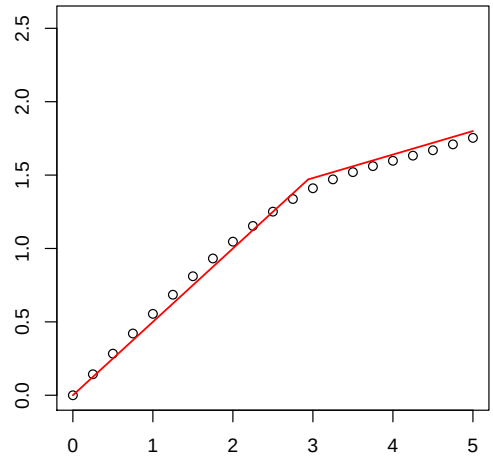
This shows that the volatility autocorrelation of our process decays exponentially fast for time scales larger than the mean distance $1/\lambda$ between the epochs τ_k. However, for shorter time scales the relevant contribution is given by the function $\phi(\cdot)$. By (2.16) we can write

$$\phi(x) \;=\; Var(\sigma)\, E(S^{D-1/2}\,(S+x)^{D-1/2}) \,+\, E(\sigma)^2\, Cov(S^{D-1/2}, (S+x)^{D-1/2})\,, \tag{2.18}$$

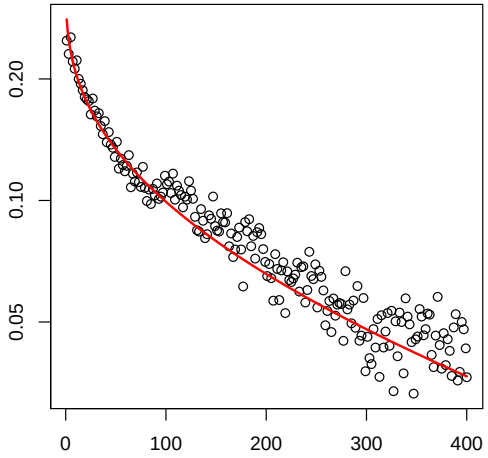
where $S \sim Exp(1)$. When $D < \frac{1}{2}$, as $x \to \infty$ the two terms in the right hand side decay as

$$E(S^{D-1/2}\,(S+x)^{D-1/2}) \;\sim\; c_1\, x^{D-1/2}\,, \quad Cov(S^{D-1/2}, (S+x)^{D-1/2}) \;\sim\; c_2\, x^{D-3/2}\,. \tag{2.19}$$

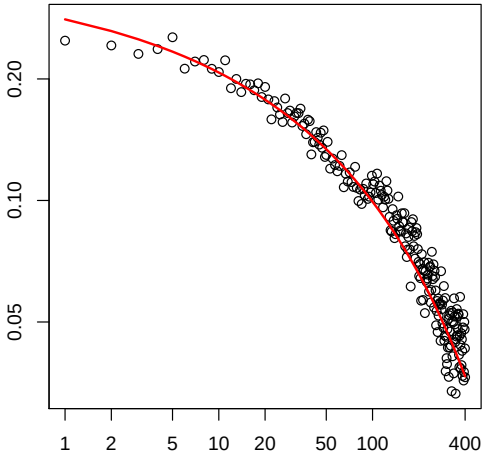
where c_1, c_2 are positive constants, hence $\phi(x)$ has a power-law behavior as $x \to \infty$. For $x = O(1)$, which is the relevant regime, the decay of $\phi(x)$ is, roughly speaking, slower than exponential but faster than polynomial (see Figures 3(B) and 3(C)).



(A) *Multiscaling in the DJIA (1935-2009).*



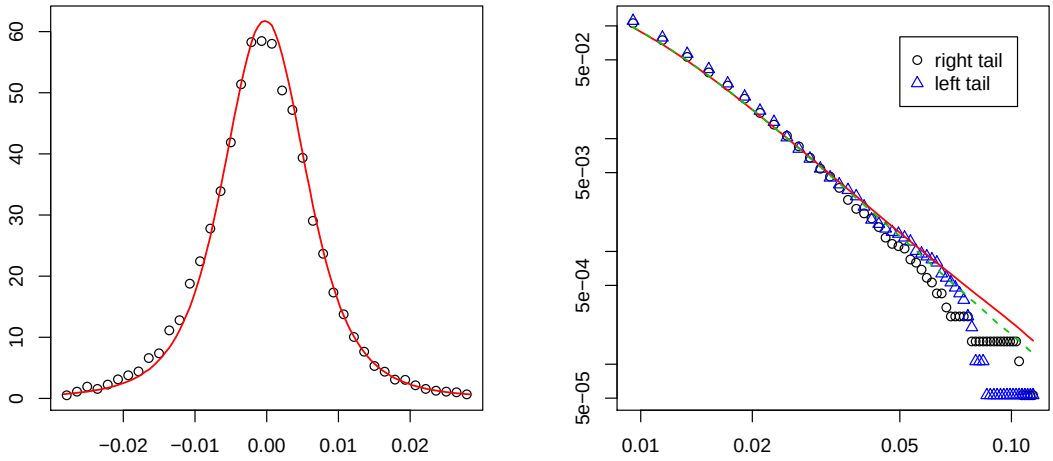
(B) *Volatility autocorrelation in the DJIA (1935-2009): log plot.*



(C) *Volatility autocorrelation in the DJIA (1935-2009): log-log plot.*

FIGURE 3. *Multiscaling of moments and volatility autocorrelation in the DJIA time series (1935-2009), compared with our model.*
(A) The DJIA empirical scaling exponent $\widehat{A}(q)$ (circles) and the theoretical scaling exponent $A(q)$ (line) as a function of q.
(B) Log plot for the DJIA empirical 1-day volatility autocorrelation $\widehat{\rho}_1(t)$ (circles) and the theoretical prediction $\rho(t)$ (line), as functions of t (days). For clarity, only one data out of two is plotted.
(C) Same as (B), but log-log plot instead of log plot. For clarity, for $t \geq 50$ only one data out of two is plotted.

2.4. **Fitting the DJIA time series.** We now consider some aspects of our model from a numerical viewpoint. More precisely, we have compared the theoretical predictions and the simulated data of our model with the time series of some of the main financial indexes (DJIA, S&P 500, FTSE 100 and Nikkei 225), finding a very good agreement. Here we describe in



(A) *Density of the empirical distribution of the DJIA log-returns (1935-2009).*

(B) *Tails of the empirical distribution of the DJIA log-returns (1935-2009).*

FIGURE 4. *Distribution of daily log-returns in the DJIA time series (1935-2009), compared with our model (cf. subsection 7.3 for details).*
(A) The density of the DJIA log-return empirical distribution $\hat{p}_1(\cdot)$ (circles) and the theoretical prediction $p_1(\cdot)$ (line). The plot ranges from zero to about three standard deviations ($\hat{s} \simeq 0.0095$) from the mean.
(B) Log-log plot of the right and left integrated tails of the DJIA log-return empirical distribution $\hat{p}_1(\cdot)$ (circles and triangles) and of the theoretical prediction $p_1(\cdot)$ (solid line). The plot ranges from one to about twelve standard deviations from the mean. Also plotted is the asymptotic density $f(\cdot)$ (dashed line) defined in equation (2.8).

detail the case of the DJIA time series over a period of 75 years: we have considered the DJIA opening prices from 2 Jan 1935 to 31 Dec 2009, for a total of 18849 daily data.

The four real parameters $D, \lambda, E(\sigma), E(\sigma^2)$ of our model have been chosen to optimize the fitting of the scaling function $A(q)$ of the moments (see Corollary 7), which only depends on D, and the curve $\rho(t)$ of the volatility autocorrelation (see Corollary 9), which depends on $D, \lambda, E(\sigma), E(\sigma^2)$ (more details on the parameter estimation are illustrated in section 7). We have obtained the following numerical estimates:

$$\widehat{D} \simeq 0.16, \quad \widehat{\lambda} \simeq 0.00097, \quad \widehat{E(\sigma)} \simeq 0.108, \quad \widehat{E(\sigma^2)} \simeq 0.0117 \simeq \left(\widehat{E(\sigma)}\right)^2. \tag{2.20}$$

Note that the estimated standard deviation of σ is negligible, so that σ is "nearly constant". We point out that the same is true for the other financial indexes that we have tested. In particular, in these cases there is no need to specify other details of the distribution ν of σ and our model is completely determined by the numerical values in (2.20).

As we show in Figure 3, there is an excellent fitting of the theoretical predictions to the observed data. We find remarkable that a rather simple mechanism of (relatively rare) volatilty shocks can account for the nontrivial profile of both the multiscaling exponent $A(q)$, cf. Figure 3(A), and the volatility autocorrelation $\rho(t)$, cf. Figure 3(B)-(C).

Last but not least, we have considered the distribution of daily log-returns: Figure 4 compares both the density and the integrated tails of the log-return empirical distribution, cf. (1.2), with the theoretical predictions of our model, i.e., the law of X_1. The agreement is remarkable, especially because the empirical distributions of log-returns *was not used for*

the calibration the model. This accuracy can therefore be regarded as structural property of the model.

In Figure 4(B) we have plotted the density of X_1, represented by the red solid line, and the asymptotic limiting density f appearing in equation (2.8) of Theorem 4, represented by the green dashed line. The two functions are practically indistinguishable up to six standard deviations from the mean, and still very close in the whole plotted range. We stress that f is a rather explicit function, cf. equation (2.8). Also note that the log-log plot in Figure 4(B) shows a clear power-law decay, as one would expect from f (the eventually light tails of X_1 are invisible).

Remark 10. We point out that, even if we had found $\widehat{Var}(\sigma) := \widehat{E(\sigma^2)} - (\widehat{E(\sigma)})^2 > 0$ (as could happen for different assets), detailed properties of the distribution of σ are not expected to be detectable from data — nor are they relevant. Indeed, the estimated mean distance between the successive epochs $(\tau_n)_{n\geq 0}$ of the Poisson process $\mathcal{T}$ is $1/\widehat{\lambda} \simeq 1031$ days, cf. (2.20). Therefore, in a time period of the length of the DJIA time series we are considering, only $18849/1031 \simeq 18$ variables σ_k are expected to be sampled, which is certainly not enough to allow more than a rough estimation of the distribution of σ. This should be viewed more as a robustness than a limitation of our model: even when σ is non-constant, its first two moments contain the information which is relevant for application to real data.

3. Discussion and further developments

Now that we have stated the main properties of the model, we can discuss more in depth its strength as well as its limitations, and consider possible generalizations.

3.1. **On the role of parameters.** A key feature of our model is its rigid structure. Let us focus for simplicity on the case in which σ *is a constant* (which, as we have discussed, is relevant for financial indexes). Not only is the model characterized by just three real parameters D, λ, σ: the role of λ and σ is reduced to *simple scale factors*. In fact, if we change the value of λ and σ in our process $(X_t)_{t\geq 0}$, keeping D fixed, the new process has the same distribution as $(aX_{bt})_{t\geq 0}$ for suitable a, b (depending of λ, σ), as it is clear from the definition (2.2) of $(I_t)_{t\geq 0}$. This means that D *is the only parameter that truly changes the shape* (beyond simple scale factors) of the relevant quantities of our model, as it is also clear from the explicit expressions we have derived for the small-time and large-time asymptotic distribution (Theorem 4), mulstiscaling of moments (Theorem 6) and volatility autocorrelation (Theorem 8).

More concretely, the structure of our model imposes strict relations between different quantities: for instance, the moment $q^* = (\frac{1}{2} - D)^{-1}$ beyond which one observes anomalous scaling, cf. (2.14), coincides with the power-law tail exponent of the (approximate) log-return distribution for small time, cf. (2.9), and is also linked (through D) to the slow decay of the volatility autocorrelation from short to moderate time of the, cf. (2.18) and (2.19). The fact that these quantitative constraints are indeed observed on the DJIA time series, cf. Figures 3 and 4, is not obvious a priori and is therefore particularly noteworthy.

3.2. **On the comparison with multifractal models.** As we observed in the introduction, the multiscaling of moments is a key feature of multifractal models. These are random time-changes of Brownian motion $X_t = W_{I_t}$, like our model, with the important difference that the time-change process $(I_t)_{t\geq 0}$ is rather singular, having non absolutely continuous paths. Since in our case the time-change process is quite regular and explicit, our model can be analyzed with more standard and elementary tools and is very easy to simulate.

A key property of multifractal models, which is at the origin of their multiscaling features, is that the law of X_t has power-law tails, for every $t > 0$. On the other hand, as we already discussed, the law of X_t in our model has finite moments of all orders — at least when $E(\sigma^q) < \infty$ for every $q > 0$, which is the typical case. In a sense, the source of multiscaling in our model is analogous, because (approximate) power-law tails appear in the distribution of X_t in the limit $t \downarrow 0$, but the point is that "true" power-law tails in the distribution of X_t are not necessary to have multiscaling properties.

We remark that the multiscaling exponent $A(q)$ of our model is piecewise linear with two different slopes, thus describing a *biscaling* phenomenon. Multifractal models are very flexible in this respect, allowing for a much wider class of behavior of $A(q)$. It appears however that a biscaling exponent is compatible with the time series of financial indexes (cf. also Remark 14 below).

We conclude with a semi-heuristic argument, which illustrates how heavy tails and multiscaling arise in our model. On the event $\{(-\tau_0) \le h, \tau_1 > h\}$ we can write, by (2.2), $I_h = \sigma_0^2\{(h-\tau_0)^{2D} - (-\tau_0)^{2D}\} \gtrsim h^{2D}$ and therefore $|X_h| = |W_{I_h}| \sim \sqrt{I_h}\,|W_1| \gtrsim h^D$. Consequently we get the bound

$$P(|X_h| \gtrsim h^D) \ge P((-\tau_0) \le h, \tau_1 > h) \gtrsim h\,, \tag{3.1}$$

which allows to draw a couple of interesting consequences.

- Relation (3.1) yields the lower bound $E(|X_h|^q) \gtrsim h^{Dq} P(|X_h| \gtrsim h^D) \gtrsim h^{Dq+1}$ on the moments of our process. Since $Dq + 1 < q/2$ for $q > q^* = (\frac{1}{2} - D)^{-1}$, this shows that the usual scaling $E(|X_h|^q) \simeq h^{q/2}$ cannot hold for $q > q^*$.
- Relation (3.1) can be rewritten as $P(\frac{1}{\sqrt{h}}|X_h| \gtrsim t) \gtrsim t^{-q^*}$, where $t = h^{-(\frac{1}{2}-D)}$ and $q^* = (\frac{1}{2} - D)^{-1}$. Since $t \to +\infty$ as $h \downarrow 0$, when $D < \frac{1}{2}$, this provides a glimpse of the appearance of power law tails in the distribution of X_h as $h \downarrow 0$, cf. (2.8) and (2.9), *with the correct tail exponent* q^*.

3.3. **On the stochastic volatility model representation.** We recall that our process $(X_t)_{t\ge 0}$ can be written as a stochastic volatility model $\mathrm{d}X_t = v_t\,\mathrm{d}B_t$, cf. (2.5). It is interesting to note that the squared volatility $(v_t)^2$ is the stationary solution of the following stochastic differential equation:

$$\mathrm{d}(v_t^2) = -\alpha_t \left(v_t^2\right)^\gamma \mathrm{d}t + \infty\,\mathrm{d}i(t), \tag{3.2}$$

where we recall that $(i(t))_{t\ge 0}$ is an ordinary Poisson process, while γ is a constant and α_t is a piecewise-constant function, defined by

$$\gamma := 2 + \frac{2D}{1-2D} > 2\,, \qquad \alpha_t := \frac{1-2D}{(2D)^{1/(1-2D)}} \frac{1}{\sigma_{i(t)}^{2/(1-2D)}} > 0\,.$$

We stress that $(v_t)^2$ is a *pathwise* solution of equation (3.2), i.e., it solves the equation for any fixed realization of the stochastic processes $i(t)$ and α_t. The infinite coefficient of the driving Poisson noise is no problem: in fact, thanks to the *superlinear* drift term $-\alpha_t \left(v_t^2\right)^\gamma$, the solution starting from infinity becomes instantaneously finite.[2]

The representation (3.2) of the volatility is also useful to understand the limitations of our model and to design possible generalizations. For instance, according to (3.2), the volatility has the rather unrealistic feature of being deterministic between jumps. This limitation could be weakened in various way, e.g. by replacing $i(t)$ in (3.2) with a more general Levy

[2]Note that the ordinary differential equation $\mathrm{d}x(t) = -\alpha\, x(t)^\gamma \mathrm{d}t$ with $x(0) = \infty$ has the explicit solution $x(t) = c\,t^{-1/(\gamma-1)}$ where $c = c(\alpha,\gamma) = (\alpha(\gamma-1))^{-1/(\gamma-1)}$.

subordinator, and/or adding to the volatility a continuous random component. Such addition should allow a more accurate description of the intermittent structure of the volatility profile, in the spirit of multifractal models.

In a sense, the model we have presented describes only the *relatively rare big jumps* of the volatility, ignoring the smaller random fluctuations that are present on smaller time scales. Besides obvious simplicity considerations, one of our aims is to point out that these big jumps, together with a nonlinear drift term as in (3.2), are sufficient to explain in a rather accurate way the several stylized facts we have discussed.

3.4. **On the skewness and leverage effect.** Our model predicts an even distribution for X_t, but it is known that several financial assets data exhibit a nonzero skewness. A reasonable way to introduce skewness is through the so-called *leverage effect*. This can be achieved, e.g., by modifying the stochastic volatility representation, given in equations (2.5) and (3.2), as follows:

$$\begin{aligned} \mathrm{d}X_t &= v_t \,\mathrm{d}B_t \,-\, \beta \,\mathrm{d}i(t) \\ \mathrm{d}v_t^2 &= -\alpha_t \left(v_t^2\right)^\gamma \mathrm{d}t \,+\, \infty \,\mathrm{d}i(t), \end{aligned}$$

where $\beta > 0$. In other words, when the volatility jumps (upward), the price jumps downward by an amount γ. The effect of this extension of the model is currently under investigation.

3.5. **On further developments.** A bivariate version $((X_t, Y_t))_{t\geq 0}$ of our model, where the two components are driven by possibly correlated Poisson point processes $\mathcal{T}^X, \mathcal{T}^Y$, has been investigated by P. Pigato in his Master's Thesis [27]. The model has been numerically calibrated on the joint time series of the DJIA and FTSE 100 indexes, finding in particular a very good agreement for the *volatility cross-correlation* between the two indexes: for this quantity, the model predicts *the same decay profile* as for the individual volatility autocorrelations, a fact which is not obvious a priori and is indeed observed on the real data.

We point out that an important ingredient in the numerical analysis on the bivariate model is a clever algorithm for finding the location of the relevant *big jumps* in the volatility (a concept which is of course not trivially defined). Such an algorithm has been devised by M. Bonino in his Master's Thesis [11], which deals with portfolio optimization problems in the framework of our model.

4. Scaling and multiscaling: proof of Theorems 4 and 6

We observe that for all fixed $t, h > 0$ we have the equality in law $X_{t+h} - X_t \sim \sqrt{I_h}\, W_1$, as it follows by the definition of our model $(X_t)_{t\geq 0} = (W_{I_t})_{t\geq 0}$. We also observe that $i(h) = \#\{\mathcal{T} \cap (0,h]\} \sim Po(\lambda h)$, as it follows from (2.1) and the properties of the Poisson process.

4.1. **Proof of Theorem 4.** Since $P(i(h) \geq 1) = 1 - e^{-\lambda h} \to 0$ as $h \downarrow 0$, we may focus on the event $\{i(h) = 0\} = \{\mathcal{T} \cap (0,h] = \emptyset\}$, on which we have $I_h = \sigma_0^2((h-\tau_0)^{2D} - (-\tau_0)^{2D})$, with $-\tau_0 \sim Exp(\lambda)$. In particular,

$$\lim_{h\downarrow 0} \frac{I_h}{h} \,=\, I'(0) \,=\, (2D)\,\sigma_0^2\,(-\tau_0)^{2D-1} \qquad \text{a.s.}\,.$$

Since $X_{t+h} - X_t \sim \sqrt{I_h}\, W_1$, the convergence in distribution (2.8) follows:

$$\frac{X_{t+h} - X_t}{\sqrt{h}} \xrightarrow{\;d\;} \sqrt{2D}\,\sigma_0\,(-\tau_0)^{D-1/2}\,W_1 \qquad \text{as } h \downarrow 0\,.$$

Next we focus on the case $h \uparrow \infty$. Under the assumption $E(\sigma^2) < \infty$, the random variables $\{\sigma_{k-1}^2(\tau_k - \tau_{k-1})^{2D}\}_{k\geq 1}$ are independent and identically distributed with finite mean, hence by the strong law of large numbers

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \sigma_{k-1}^2(\tau_k - \tau_{k-1})^{2D} \;=\; E(\sigma^2)E((\tau_1)^{2D}) \;=\; E(\sigma^2)\,\lambda^{-2D}\,\Gamma(2D+1) \qquad \text{a.s.}\,.$$

Plainly, $\lim_{h\to+\infty} i(h)/h = \lambda$ a.s., by the strong law of large numbers applied to the random variables $\{\tau_k\}_{k\geq 1}$. Recalling (2.2), it follows easily that

$$\lim_{h\uparrow\infty} \frac{I(h)}{h} \;=\; E(\sigma^2)\,\lambda^{1-2D}\,\Gamma(2D+1) \qquad \text{a.s.}\,.$$

Since $X_{t+h} - X_t \sim \sqrt{I_h}\,W_1$, we obtain the convergence in distribution

$$\frac{X_{t+h} - X_t}{\sqrt{h}} \xrightarrow{\;d\;} \sqrt{E(\sigma^2)\,\lambda^{1-2D}\,\Gamma(2D+1)}\,W_1 \qquad \text{as } h\uparrow\infty\,,$$

which coincides with (2.10). □

4.2. **Proof of Theorem 6.** Since $X_{t+h} - X_t \sim \sqrt{I_h}\,W_1$, we can write

$$E(|X_{t+h} - X_t|^q) = E(|I_h|^{q/2}|W_1|^q) = E(|W_1|^q)\,E(|I_h|^{q/2}) = c_q\,E(|I_h|^{q/2})\,, \tag{4.1}$$

where we set $c_q := E(|W_1|^q)$. We therefore focus on $E(|I_h|^{\frac{q}{2}})$, that we write as the sum of three terms, that will be analyzed separately:

$$E(|I_h|^{\frac{q}{2}}) = E(|I_h|^{\frac{q}{2}}\,\mathbf{1}_{\{i(h)=0\}}) + E(|I_h|^{\frac{q}{2}}\,\mathbf{1}_{\{i(h)=1\}}) + E(|I_h|^{\frac{q}{2}}\,\mathbf{1}_{\{i(h)\geq 2\}})\,. \tag{4.2}$$

For the first term in the right hand side of (4.2), we note that $P(i(h)=0) = e^{-\lambda h} \to 1$ as $h \downarrow 0$ and that $I_h = \sigma_0^2((h-\tau_0)^{2D} - (-\tau_0)^{2D})$ on the event $\{i(h)=0\}$. Setting $-\tau_0 =: \lambda^{-1}S$ with $S \sim Exp(1)$, we obtain as $h \downarrow 0$

$$E(|I_h|^{\frac{q}{2}}\,\mathbf{1}_{\{i(h)=0\}}) = E(\sigma^q)\,\lambda^{-Dq}\,E\big(((S+\lambda h)^{2D} - S^{2D})^{\frac{q}{2}}\big)\,\big(1+o(1)\big)\,. \tag{4.3}$$

Recalling that $q^* := (\frac{1}{2} - D)^{-1}$, we have

$$q \gtreqless q^* \quad\Longleftrightarrow\quad \frac{q}{2} \gtreqless Dq+1 \quad\Longleftrightarrow\quad -1 \gtreqless \left(D - \frac{1}{2}\right) q\,.$$

As $\delta \downarrow 0$ we have $\delta^{-1}((S+\delta)^{2D} - S^{2D}) \uparrow 2D\,S^{2D-1}$ and note that $E\big(S^{(D-\frac{1}{2})q}\big) = \Gamma(1-q/q^*)$ is finite if and only if $(D-\frac{1}{2})q > -1$, that is $q < q^*$. Therefore the monotone convergence theorem yields

$$\text{for } q < q^*: \quad \lim_{h\downarrow 0} \frac{E\big(((S+\lambda h)^{2D} - S^{2D})^{\frac{q}{2}}\big)}{\lambda^{\frac{q}{2}}\,h^{\frac{q}{2}}} \;=\; (2D)^{q/2}\,\Gamma(1-q/q^*) \;\in\; (0,\infty)\,. \tag{4.4}$$

Next observe that, by the change of variables $s = (\lambda h)x$, we can write

$$\begin{aligned} E\big(((S+\lambda h)^{2D} - S^{2D})^{\frac{q}{2}}\big) &= \int_0^\infty ((s+\lambda h)^{2D} - s^{2D})^{\frac{q}{2}}\,e^{-s}\,\mathrm{d}s \\ &= (\lambda h)^{Dq+1}\int_0^\infty ((1+x)^{2D} - x^{2D})^{\frac{q}{2}}\,e^{-\lambda h x}\,\mathrm{d}x\,. \end{aligned} \tag{4.5}$$

Note that $((1+x)^{2D} - x^{2D})^{\frac{q}{2}} \sim (2D)^{\frac{q}{2}}x^{(D-\frac{1}{2})q}$ as $x \to +\infty$ and that $(D-\frac{1}{2})q < -1$ if and only if $q > q^*$. Therefore, again by the monotone convergence theorem, we obtain

$$\text{for } q > q^*: \quad \lim_{h\downarrow 0} \frac{E\big(((S+\lambda h)^{2D} - S^{2D})^{\frac{q}{2}}\big)}{\lambda^{Dq+1}\,h^{Dq+1}} \;=\; \int_0^\infty ((1+x)^{2D} - x^{2D})^{\frac{q}{2}}\,\mathrm{d}x \;\in\; (0,\infty)\,. \tag{4.6}$$

Finally, in the case $q = q^*$ we have $((1+x)^{2D} - x^{2D})^{q^*/2} \sim (2D)^{q^*/2}\, x^{-1}$ as $x \to +\infty$ and we want to study the integral in the second line of (4.5). Fix an arbitrary (large) $M > 0$ and note that, integrating by parts and performing a change of variables, as $h \downarrow 0$ we have

$$\begin{aligned}\int_M^\infty \frac{e^{-\lambda h x}}{x}\,\mathrm{d}x \;&=\; -\log M e^{-\lambda h M} \;+\; \lambda h \int_M^\infty (\log x)\, e^{-\lambda h x}\,\mathrm{d}x \;=\; O(1) + \int_{\lambda h M}^\infty \log\Big(\frac{y}{\lambda h}\Big)\, e^{-y}\,\mathrm{d}y \\ &=\; O(1) + \int_{\lambda h M}^\infty \log\Big(\frac{y}{\lambda}\Big)\, e^{-y}\,\mathrm{d}y \;+\; \log\Big(\frac{1}{h}\Big) \int_{\lambda h M}^\infty e^{-y}\,\mathrm{d}y \;=\; \log\Big(\frac{1}{h}\Big)\,\big(1 + o(1)\big)\,.\end{aligned}$$

From this it is easy to see that as $h \downarrow 0$

$$\int_0^\infty ((1+x)^{2D} - x^{2D})^{\frac{q^*}{2}}\, e^{-\lambda h x}\,\mathrm{d}x \;\sim\; (2D)^{\frac{q^*}{2}} \log\Big(\frac{1}{h}\Big)\,.$$

Coming back to (4.5), noting that $Dq + 1 = \frac{q}{2}$ for $q = q^*$, it follows that

$$\lim_{h\downarrow 0} \frac{E\big(((S+h)^{2D} - S^{2D})^{\frac{q^*}{2}}\big)}{\lambda^{Dq^*+1}\, h^{\frac{q^*}{2}}\, \log(\frac{1}{h})} = (2D)^{\frac{q^*}{2}}\,. \tag{4.7}$$

Recalling (4.1) and (4.3), the relations (4.4), (4.6) and (4.7) show that the first term in the right hand side of (4.2) has the same asymptotic behavior as in the statement of the theorem, except for the regime $q > q^*$ where the constant does not match (the missing contribution will be obtained in a moment).

We now focus on the second term in the right hand side of (4.2). Note that, conditionally on the event $\{i(h) = 1\} = \{\tau_1 \le h, \tau_2 > h\}$, we have

$$I_h = \sigma_1^2 (h - \tau_1)^{2D} + \sigma_0^2 \big((\tau_1 - \tau_0)^{2D} - (-\tau_0)^{2D}\big) \sim \sigma_1^2 (h - hU)^{2D} + \sigma_0^2 \bigg(\Big(hU + \frac{S}{\lambda}\Big)^{2D} - \Big(\frac{S}{\lambda}\Big)^{2D}\bigg)\,,$$

where $S \sim Exp(1)$ and $U \sim U(0,1)$ (uniformly distributed on the interval $(0,1)$) are independent of σ_0 and σ_1. Since $P(i(h) = 1) = \lambda h + o(h)$ as $h \downarrow 0$, we obtain

$$E(|I_h|^{\frac{q}{2}}\, \mathbf{1}_{\{i(h)=1\}}) = \lambda h^{Dq+1}\, E\bigg[\bigg(\sigma_1^2 (1-U)^{2D} + \sigma_0^2 \bigg(\Big(U + \frac{S}{\lambda h}\Big)^{2D} - \Big(\frac{S}{\lambda h}\Big)^{2D}\bigg)^{\frac{q}{2}}\bigg)\bigg]\,. \tag{4.8}$$

Since $(u+x)^{2D} - x^{2D} \to 0$ as $x \to \infty$, for every $u \ge 0$, by the dominated convergence theorem we have (for every $q \in (0, \infty)$)

$$\lim_{h\downarrow 0} \frac{E(|I_h|^{\frac{q}{2}}\, \mathbf{1}_{\{i(h)=1\}})}{h^{Dq+1}} \;=\; \lambda E(\sigma_1^q)\, E\big((1-U)^{Dq}\big) \;=\; \lambda E(\sigma_1^q)\, \frac{1}{Dq+1}\,. \tag{4.9}$$

This shows that the second term in the right hand side of (4.2) gives a contribution of the order h^{Dq+1} as $h \downarrow 0$. This is relevant only for $q > q^*$, because for $q \le q^*$ the first term gives a much bigger contribution of the order $h^{q/2}$ (see (4.4) and (4.7)). Recalling (4.1), it follows from (4.9) and (4.6) that the contribution of the first and the second term in the right hand side of (4.2) matches the statement of the theorem (including the constant).

It only remains to show that the third term in the right hand side of (4.2) gives a negligible contribution. We begin by deriving a simple upper bound for I_h. Since $(a+b)^{2D} - b^{2D} \le a^{2D}$

for all $a, b \geq 0$ (we recall that $2D \leq 1$), when $i(h) \geq 1$, i.e. $\tau_1 \leq h$, we can write

$$\begin{aligned} I_h &= \sigma_{i(h)}^2 (h - \tau_{i(h)})^{2D} + \sum_{k=2}^{i(h)} \sigma_{k-1}^2 (\tau_k - \tau_{k-1})^{2D} + \sigma_0^2 \left[(\tau_1 - \tau_0)^{2D} - (-\tau_0)^{2D} \right] \\ &\leq \sigma_{i(h)}^2 (h - \tau_{i(h)})^{2D} + \sum_{k=2}^{i(h)} \sigma_{k-1}^2 (\tau_k - \tau_{k-1})^{2D} + \sigma_0^2 \tau_1^{2D} \,, \end{aligned} \tag{4.10}$$

where we agree that the sum over k is zero if $i(h) = 1$. Since $\tau_k \leq h$ for all $k \leq i(h)$, by the definition (2.1) of $i(h)$, relation (4.10) yields the bound $I_h \leq h^{2D} \sum_{k=0}^{i(h)} \sigma_k^2$, which holds clearly also when $i(h) = 0$. In conclusion, we have shown that for all $h, q > 0$

$$|I_h|^{q/2} \leq h^{Dq} \left(\sum_{k=0}^{i(h)} \sigma_k^2 \right)^{q/2} . \tag{4.11}$$

Consider first the case $q > 2$: by Jensen's inequality we have

$$\left(\sum_{k=0}^{i(h)} \sigma_k^2 \right)^{q/2} = (i(h)+1)^{q/2} \left(\frac{1}{i(h)+1} \sum_{k=0}^{i(h)} \sigma_k^2 \right)^{q/2} \leq (i(h)+1)^{q/2-1} \sum_{k=0}^{i(h)} \sigma_k^q \,. \tag{4.12}$$

By (4.11) and (4.12) we obtain

$$E\big(|I_h|^{q/2} \, \mathbf{1}_{\{i(h) \geq 2\}} \big) \;\leq\; h^{Dq} \, E(\sigma^q) \, E\big((i(h)+1)^{q/2} \, \mathbf{1}_{\{i(h) \geq 2\}} \big) \,. \tag{4.13}$$

A corresponding inequality for $q \leq 2$ is derived from (4.11) and the inequality $(\sum_{k=1}^{\infty} x_k)^{q/2} \leq \sum_{k=1}^{\infty} x_k^{q/2}$, which holds for every non-negative sequence $(x_n)_{n \in \mathbb{N}}$:

$$E\big(|I_h|^{q/2} \, \mathbf{1}_{\{i(h) \geq 2\}} \big) \;\leq\; h^{Dq} E\left(\sum_{k=0}^{i(h)} \sigma_k^q \, \mathbf{1}_{\{i(h) \geq 2\}} \right) \;\leq\; h^{Dq} \, E(\sigma^q) \, E\big((i(h)+1) \, \mathbf{1}_{\{i(h) \geq 2\}} \big) \,. \tag{4.14}$$

For any fixed $a > 0$, by the Hölder inequality with $p = 3$ and $p' = 3/2$ we can write for $h \leq 1$

$$\begin{aligned} E\big((i(h)+1)^a \, \mathbf{1}_{\{i(h) \geq 2\}} \big) \;&\leq\; E\big((i(h)+1)^{3a} \big)^{1/3} \, P(i(h) \geq 2)^{2/3} \\ &\leq\; E\big((i(1)+1)^{3a} \big)^{1/3} \, (1 - e^{-\lambda h} - e^{-\lambda h} \lambda h)^{2/3} \;\leq\; (const.) \, h^{4/3} \,, \end{aligned} \tag{4.15}$$

because $E\big((i(1)+1)^{3a} \big) < \infty$ (recall that $i(h) \sim Po(\lambda)$) and $(1 - e^{-\lambda h} - e^{-\lambda h} \lambda h) \sim \frac{1}{2} (\lambda h)^2$ as $h \downarrow 0$. Then it follows from (4.13) and (4.14) and (4.15) that

$$E\big(|I_h|^{q/2} \, \mathbf{1}_{\{i(h) \geq 2\}} \big) \;\leq\; (const.') \, h^{Dq+4/3} \,.$$

This shows that the contribution of the third term in the right hand side of (4.2) is always negligible with respect to the contribution of the second term (recall (4.9)). □

5. Decay of correlations: proof of Theorem 8

Given a Borel set $I \subseteq \mathbb{R}$, we let $\mathcal{G}_I$ denote the σ-algebra generated by the family of random variables $(\tau_k \mathbf{1}_{\{\tau_k \in I\}}, \sigma_k \mathbf{1}_{\{\tau_k \in I\}})_{k \geq 0}$. Informally, $\mathcal{G}_I$ may be viewed as the σ-algebra generated by the variables τ_k, σ_k for the values of k such that $\tau_k \in I$. From the basic property of the Poisson process and from the fact that the variables $(\sigma_k)_{k \geq 0}$ are independent, it follows that for disjoint Borel sets I, I' the σ-algebras $\mathcal{G}_I$, $\mathcal{G}_{I'}$ are independent. We set for short $\mathcal{G} := \mathcal{G}_{\mathbb{R}}$, which is by definition the σ-algebra generated by all the variables $(\tau_k)_{k \geq 0}$ and $(\sigma_k)_{k \geq 0}$, which coincides with the σ-algebra generated by the process $(I_t)_{t \geq 0}$.

We have to prove (2.15). Plainly, by translation invariance we can set $s = 0$ without loss of generality. We also assume that $h < t$. We start writing

$$\begin{aligned} &Cov(|X_h|, |X_{t+h} - X_t|) \\ &= Cov\big(E\big(|X_h|\big|\mathcal{G}\big)\,,\, E\big(|X_{t+h} - X_t|\big|\mathcal{G}\big)\big) \;+\; E\big(Cov\big(|X_h|, |X_{t+h} - X_t|\big|\mathcal{G}\big)\big)\,. \end{aligned} \tag{5.1}$$

We recall that $X_t = W_{I_t}$ and the process $(I_t)_{t\geq 0}$ is $\mathcal{G}$-measurable and independent of the process $(W_t)_{t\geq 0}$. It follows that, conditionally on $(I_t)_{t\geq 0}$, the process $(X_t)_{t\geq 0}$ has independent increments, hence the second term in the right hand side of (5.1) vanishes, because $Cov(|X_h|, |X_{t+h} - X_t||\mathcal{G}) = 0$ a.s.. For fixed h, from the equality in law $X_h = W_{I_h} \sim \sqrt{I_h}\, W_1$ it follows that $E(|X_h||\mathcal{G}) = c_1\sqrt{I_h}$, where $c_1 = E(|W_1|) = \sqrt{2/\pi}$. Analogously $E(|X_{t+h} - X_t||\mathcal{G}) = \sqrt{2/\pi}\sqrt{I_{t+h} - I_t}$ and (5.1) reduces to

$$Cov(|X_h|, |X_{t+h} - X_t|) \;=\; \frac{2}{\pi}\, Cov\big(\sqrt{I_h}, \sqrt{I_{t+h} - I_t}\big)\,. \tag{5.2}$$

Recall the definitions (2.1) and (2.2) of the variabes $i(t)$ and I_t. We now claim that we can replace $\sqrt{I_{t+h} - I_t}$ by $\sqrt{I_{t+h} - I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(h,t]=\emptyset\}}$ in (5.2). In fact from (2.2) we can write

$$I_{t+h} - I_t = \sigma^2_{i(t+h)}(t + h - \tau_{i(t+h)})^{2D} + \sum_{k=i(t)+1}^{i(t+h)} \sigma^2_{k-1}(\tau_k - \tau_{k-1})^{2D} - \sigma^2_{i(t)}(t - \tau_{i(t)})^{2D}\,,$$

where we agree that the sum in the right hand side is zero if $i(t + h) = i(t)$. This shows that $(I_{t+h} - I_t)$ is a function of the variables τ_k, σ_k with index $i(t) \leq k \leq i(t + h)$. Since $\{\mathcal{T} \cap (h, t] \neq \emptyset\} = \{\tau_{i(t)} > h\}$, this means that $\sqrt{I_{t+h} - I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(h,t]\neq\emptyset\}}$ is $\mathcal{G}_{(h,t+h]}$-measurable, hence independent of $\sqrt{I_h}$, which is clearly $\mathcal{G}_{(-\infty,h]}$-measurable. This shows that $Cov(\sqrt{I_h}, \sqrt{I_{t+h} - I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(h,t]\neq\emptyset\}}) = 0$, therefore from (5.2) we can write

$$Cov(|X_h|, |X_{t+h} - X_t|) \;=\; \frac{2}{\pi}\, Cov\big(\sqrt{I_h}, \sqrt{I_{t+h} - I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(h,t]=\emptyset\}}\big)\,. \tag{5.3}$$

Now we decompose this last covariance as follows:

$$\begin{aligned} &Cov\big(\sqrt{I_h}, \sqrt{I_{t+h} - I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(h,t]=\emptyset\}} \;=\; E\Big[\Big(\sqrt{I_h} - E\big(\sqrt{I_h}\big)\Big)\sqrt{I_{t+h} - I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(h,t]=\emptyset\}}\Big] \\ &= E\Big[\Big(\sqrt{I_h} - E\big(\sqrt{I_h}\big)\Big)\sqrt{I_{t+h} - I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(0,t+h]=\emptyset\}}\Big] \\ &\quad + E\Big[\Big(\sqrt{I_h} - E\big(\sqrt{I_h}\big)\Big)\sqrt{I_{t+h} - I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(h,t]=\emptyset\}}\mathbf{1}_{\{\mathcal{T}\cap([0,h]\cup(t,t+h])\neq\emptyset\}}\Big] \end{aligned} \tag{5.4}$$

We deal separately with the two terms in the r.h.s. of (5.4). The first gives the dominant contribution. To see this, observe that, on $\{\mathcal{T} \cap (0, t + h] = \emptyset\}$

$$I_h \;=\; \sigma_0^2\left[(h - \tau_0)^{2D} - (-\tau_0)^{2D}\right]$$

and

$$I_{t+h} - I_t \;=\; \sigma_0^2\left[(t + h - \tau_0)^{2D} - (t - \tau_0)^{2D}\right].$$

Since both $\sigma_0^2\left[(h-\tau_0)^{2D}-(-\tau_0)^{2D}\right]$ and $\sigma_0^2\left[(t+h-\tau_0)^{2D}-(t-\tau_0)^{2D}\right]$ are independent of $\{\mathcal{T}\cap(0,t+h]=\emptyset\}$, we have

$$\begin{aligned}
&E\left[\left(\sqrt{I_h}-E\big(\sqrt{I_h}\big)\right)\sqrt{I_{t+h}-I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(0,t+h]=\emptyset\}}\right]\\
&=\ E\Bigg[\left(\sigma_0\sqrt{(h-\tau_0)^{2D}-(-\tau_0)^{2D}}-E\big(\sqrt{I_h}\big)\right)\sigma_0\sqrt{(t+h-\tau_0)^{2D}-(t-\tau_0)^{2D}}\,\mathbf{1}_{\{\mathcal{T}\cap(0,t+h]=\emptyset\}}\Bigg]\\
&=\ e^{-\lambda(t+h)}E\left[\left(\sigma_0\sqrt{(h-\tau_0)^{2D}-(-\tau_0)^{2D}}-E\big(\sqrt{I_h}\big)\right)\sigma_0\sqrt{(t+h-\tau_0)^{2D}-(t-\tau_0)^{2D}}\right]\\
&=\ e^{-\lambda(t+h)}\left\{Cov\left(\sigma_0\sqrt{(h-\tau_0)^{2D}-(-\tau_0)^{2D}},\sigma_0\sqrt{(t+h-\tau_0)^{2D}-(t-\tau_0)^{2D}}\right)\right.\\
&\qquad+\left.\left[E\left(\sigma_0\sqrt{(h-\tau_0)^{2D}-(-\tau_0)^{2D}}\right)-E\big(\sqrt{I_h}\big)\right]E\left(\sigma_0\sqrt{(t+h-\tau_0)^{2D}-(t-\tau_0)^{2D}}\right)\right\}.
\end{aligned}\tag{5.5}$$

Since $\delta^{-1}((\delta+x)^{2D}-x^{2D})\uparrow 2Dx^{2D-1}$ as $\delta\downarrow 0$, by monotone convergence we obtain

$$\begin{aligned}
&\lim_{h\to0}\frac{1}{h}Cov\left(\sigma_0\sqrt{(h-\tau_0)^{2D}-(-\tau_0)^{2D}},\sigma_0\sqrt{(t+h-\tau_0)^{2D}-(t-\tau_0)^{2D}}\right)\\
&=\ 2DCov\left(\sigma_0(-\tau_0)^{D-1/2},\sigma_0(t-\tau_0)^{D-1/2}\right)\\
&=\ 2D\lambda^{1-2D}Cov\left(\sigma_0S^{D-1/2},\sigma_0(\lambda t+S)^{D-1/2}\right)\ =\ 2D\lambda^{1-2D}\phi(\lambda t),
\end{aligned}\tag{5.6}$$

with $S:=\lambda(-\tau_0)\sim Exp(1)$ and ϕ is defined in (2.16). Similarly

$$\lim_{h\to0}\frac{1}{\sqrt{h}}E\left(\sigma_0\sqrt{(t+h-\tau_0)^{2D}-(t-\tau_0)^{2D}}\right)\ =\ \sqrt{2D}E\left(\sigma_0(t-\tau_0)^{D-1/2}\right)<+\infty.\tag{5.7}$$

Therefore, if we show that

$$\lim_{h\to0}E\left(\sqrt{\frac{I_h}{h}}\right)\ =\ \sqrt{2D}E\left(\sigma_0(-\tau_0)^{D-1/2}\right)\tag{5.8}$$

using (5.5), (5.6), (5.7), we have

$$\lim_{h\to0}\frac{1}{h}E\left[\left(\sqrt{I_h}-E\sqrt{I_h}\right)\sqrt{I_{t+h}-I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(0,t+h]\,=\,\emptyset\}}\right]=2D\lambda^{1-2D}\phi(\lambda t)\tag{5.9}$$

To complete the proof of (5.9), we are left to show (5.8). But this is a nearly immediate consequence of Theorem 6: indeed, using (2.13) and the fact that $q^*>1$,

$$E\sqrt{I_h}\ =\ \frac{1}{E\sqrt{|W_1|}}E(|X_h|)\ =\ \frac{C_1}{E|W_1|}\sqrt{h}+o(\sqrt{h})\ =\ \sqrt{2D}E\left(\sigma_0(-\tau_0)^{D-1/2}\right)\sqrt{h}+o(\sqrt{h})\,.$$

The proof is now completed if we show that the second term in (5.4) is negligible, i.e. it is $o(h)$. By Cauchy-Schwarz inequality and the simple fact that $\left(\sqrt{I_h}-E\sqrt{I_h}\right)^2\le I_h+E(I_h)$

$$\begin{aligned}
&E\left[\left(\sqrt{I_h}-E\sqrt{I_h}\right)\sqrt{I_{t+h}-I_t}\,\mathbf{1}_{\{\mathcal{T}\cap(h,t]=\emptyset\}}\mathbf{1}_{\{\mathcal{T}\cap((0,h]\cup(t,t+h])\neq\emptyset\}}\right]\\
&\le\ \left(E\left[\left(\sqrt{I_h}-E\sqrt{I_h}\right)^2(I_{t+h}-I_t)\right]P(\mathcal{T}\cap((0,h]\cup(t,t+h])\neq\emptyset)\right)^{1/2}\\
&\le\ \left(E\left[(I_h+E(I_h))\,(I_{t+h}-I_t)\right]P(\mathcal{T}\cap((0,h]\cup(t,t+h])\neq\emptyset)\right)^{1/2}\\
&\le\ \left(2\,E\left[I_h^2\right]P(\mathcal{T}\cap((0,h]\cup(t,t+h])\neq\emptyset)\right)^{1/2}\ =\ \left(2\,E\left[I_h^2\right]\right)^{1/2}\sqrt{2\lambda h}\,.
\end{aligned}\tag{5.10}$$

By Theorem 6, $E\left[I_h^2\right]$ is of order h^2 if $4 < q^*$, and of order h^{4D+1} if $4 > q^*$, with a logarithmic correction for $q^* = 4$. In both cases $\left(E\left[I_h^2\right]\right)^{1/2} \sqrt{2\lambda h} = o(h)$, and the proof is completed. □

6. Basic properties of the model

In this Section we start by proving the properties (A)-(B)-(C)-(D) stated in section 2.2. Then we provide some connections between the tails of σ and those of X_t, also beyond the equivalence stated in (2.4). Finally, we establish a *mixing* property that yields relation (2.7). One of the proofs is postponed to the Appendix.

We denote by $\mathcal{G}$ the σ-field generated by the whole process $(I_t)_{t\geq 0}$, which coincides with the σ-field generated by the sequences $\mathcal{T} = \{\tau_k\}_{k\geq 0}$ and $\Sigma = \{\sigma_k\}_{k\geq 0}$.

Proof of property (A). We first focus on the process $(I_t)_{t\geq 0}$, defined in (2.2). For $h > 0$ let $\mathcal{T}^h := \mathcal{T} - h$ and denote the points in $\mathcal{T}^h$ by $\tau_k^h = \tau_k - h$. As before, let $\tau^h_{i^h(t)}$ be the largest point of $\mathcal{T}^h$ smaller that t, i.e., $i^h(t) = i(t+h)$. Recalling the definition (2.2), we can write

$$I_{t+h} - I_h = \sigma^2_{i^h(t)} \left(t - \tau^h_{i^h(t)}\right)^{2D} + \sum_{k=i^h(0)+1}^{i^h(t)} \sigma^2_{k-1} \left(\tau^h_k - \tau^h_{k-1}\right)^{2D} - \sigma^2_{i^h(0)} \left(-\tau^h_{i^h(0)}\right)^{2D},$$

where we agree that the sum in the right hand side is zero if $i^h(t) = i^h(0)$. This relation shows that $(I_{t+h} - I_h)_{t\geq 0}$ and $(I_t)_{t\geq 0}$ are *the same function* of the two random sets $\mathcal{T}^h$ and $\mathcal{T}$. Since $\mathcal{T}^h$ and $\mathcal{T}$ have the same distribution (both are Poisson point processes on $\mathbb{R}$ with intensity λ), the processes $(I_{t+h} - I_h)_{t\geq 0}$ and $(I_t)_{t\geq 0}$ have the same distribution too.

We recall that $\mathcal{G}$ is the σ-field generated by the whole process $(I_t)_{t\geq 0}$. From the definition $X_t = W_{I_t}$ and from the fact that Brownian motion has independent, stationary increments, it follows that for every Borel subset $A \subseteq \mathbb{R}^{[0,+\infty)}$

$$P\left(X_{h+\cdot} - X_h \in A\right) = E\left[P\left(W_{I_{h+\cdot}} - W_{I_h} \in A \,|\, \mathcal{G}\right)\right] = P\left(W_{I_\cdot} \in A\right) = P\left(X_\cdot \in A\right),$$

where we have used the stationarity property of the process I. Thus the processes $(X_t)_{t\geq 0}$ and $(X_{h+t} - X_h)_{t\geq 0}$ have the same distribution, which implies stationarity of the increments. □

Proof of property (B). Note that $E(|X_t|^q) = E(|W_{I_t}|^q) = E(|I_t|^{q/2})\, E(|W_1|^q)$, by the independence of W and I and the scaling properties of Brownian motion. We are therefore left with showing that

$$E(|I_t|^{q/2}) < \infty \quad \Longleftrightarrow \quad E(\sigma^q) < \infty\,. \tag{6.1}$$

The implication "⇒" is easy: by the definition (2.2) of the process I we can write

$$E(|I_t|^{q/2}) \geq E(|I_t|^{q/2}\, \mathbf{1}_{\{i(t)=0\}}) = E(\sigma_0^q)\, E(|(t-\tau_0)^{2D} - (-\tau_0)^{2D}|^{q/2}) P(i(t) = 0)\,,$$

therefore if $E(\sigma^q) = \infty$ then also $E(|I_t|^{q/2}) = \infty$.

The implication "⇐" follows immediately from the bounds (4.13) and (4.14), which hold also without the indicator $\mathbf{1}_{\{i(h)\geq 2\}}$. □

Proof of property (C). Observe first that $I'_s := \frac{\mathrm{d}}{\mathrm{d}s} I_s > 0$ a.s. and for Lebesgue–a.e. $s \geq 0$. By a change of variable, we can rewrite the process $(B_t)_{t\geq 0}$ defined in (2.6) as

$$B_t = \int_0^{I_t} \frac{1}{\sqrt{I'(I^{-1}(u))}}\, \mathrm{d}W_u = \int_0^t \frac{1}{\sqrt{I'_s}}\, \mathrm{d}W_{I_s} = \int_0^t \frac{1}{\sqrt{I'_s}}\, \mathrm{d}X_s\,,$$

which shows that relation (2.5) holds true. It remains to show that $(B_t)_{t\geq 0}$ is indeed a standard Brownian motion. Note that

$$B_t = \int_0^{I_t} \sqrt{(I^{-1})'(u)}\, \mathrm{d}W_u\,.$$

Therefore, conditionally on $\mathcal{G}$ (the σ-field generated by $(I_t)_{t\geq 0}$), $(B_t)_{t\geq 0}$ is a centered Gaussian process — it is a Wiener integral — with conditional covariance given by

$$Cov(B_s, B_t \,|\, \mathcal{G}) = \int_0^{\min\{I_s, I_t\}} (I^{-1})'(u)\, \mathrm{d}u = \min\{s,t\}\,.$$

This shows that, conditionally on $\mathcal{G}$, $(B_t)_{t\geq 0}$ is a Brownian motion. Therefore, it is *a fortiori* a Brownian motion without conditioning. □

Proof of property (D). The assumption $E(\sigma^2) < \infty$ ensures that $E(|X_t|^2) < \infty$ for all $t \geq 0$, as we have already shown. Let us now denote by $\mathcal{F}_t^X = \sigma(X_s,\, s \leq t)$ the natural filtration of the process X. We recall that $\mathcal{G}$ denotes the σ-field generated by the whole process $(I_t)_{t\geq 0}$ and we denote by $\mathcal{F}_t^X \vee \mathcal{G}$ the smallest σ-field containing $\mathcal{F}_t^X$ and $\mathcal{G}$. Since $E(W_{I_{t+h}} - W_{I_t} | \mathcal{F}_t^X \vee \mathcal{G}) = 0$ for all $h \geq 0$, by the basic properties of Brownian motion, recalling that $X_t = W_{I_t}$ we obtain

$$E(X_{t+h} | \mathcal{F}_t^X \vee \mathcal{G}) = X_t + E(W_{I_{t+h}} - W_{I_t} | \mathcal{F}_t^X \vee \mathcal{G}) = X_t.$$

Taking the conditional expectation with respect to $\mathcal{F}_t^X$ on both sides, we obtain the martingale property for $(X_t)_{t\geq 0}$. □

Let us state a proposition, proved in Appendix A, that relates the exponential moments of σ to those of X_t. We recall that, when our model is calibrated to real time series, like the DJIA, the "observable tails" of X_t are quite insensitive to the details of the distribution of σ, cf. Remarks 3 and 5.

Proposition 11. *Regardless of the distribution of σ, for every $q > (1-D)^{-1}$ we have*

$$E\left[\exp\left(\gamma |X_t|^q\right)\right] = \infty\,, \qquad \forall t > 0\,,\ \forall \gamma > 0\,. \tag{6.2}$$

On the other hand, for all $q < (1-D)^{-1}$ and $t > 0$ we have

$$E\left[\exp\left(\gamma |X_t|^q\right)\right] < \infty \quad \forall \gamma > 0 \qquad \Longleftrightarrow \qquad E\left[\exp\left(\alpha \sigma^{\frac{2q}{2-q}}\right)\right] < \infty \quad \forall \alpha > 0\,, \tag{6.3}$$

and the same relation holds for $q = (1-D)^{-1}$ provided $D < \frac{1}{2}$.

Note that $(1-D)^{-1} \in (1,2]$, because $D \in (0, \frac{1}{2}]$, so that for $D < \frac{1}{2}$ the distribution of X_t has always tails heavier than Gaussian.

We finally show a mixing property for the increments of our process. In what follows, for an interval $I \subseteq [0, +\infty)$, we let

$$\mathcal{F}_I^{\mathcal{D}} := \sigma\left(X_t - X_s : s, t \in I\right)$$

to denote the σ-field generated by the increments in I of the process X.

Proposition 12. *Let $I = [a,b)$, $J = [c,d)$, with $0 \leq a < b \leq c < d$. Then, for every $A \in \mathcal{F}_I^{\mathcal{D}}$ and $B \in \mathcal{F}_J^{\mathcal{D}}$*

$$|P(A \cap B) - P(A)P(B)| \leq e^{-\lambda(c-b)}\,. \tag{6.4}$$

As a consequence, equation (2.7) *holds true almost surely and in L^1, for every measurable function $F : \mathbb{R}^k \to \mathbb{R}$ such that $E[|F(X_{b_1} - X_{a_1}, \ldots, X_{b_k} - X_{a_k})|] < +\infty$.*

Proof. We recall that $\mathcal{T}$ denotes the set $\{\tau_k : k \in \mathbb{Z}\}$ and, for $I \subseteq \mathbb{R}$, $\mathcal{G}_I$ denotes the σ-algebra generated by the family of random variables $(\tau_k \mathbf{1}_{\{\tau_k \in I\}}, \sigma_k \mathbf{1}_{\{\tau_k \in I\}})_{k \geq 0}$, where $(\sigma_k)_{k \geq 0}$ is the sequence of volatilities. We introduce the $\mathcal{G}_{[b,c)}$-measurable event

$$\Gamma := \{\mathcal{T} \cap [b, c) \neq \emptyset\}.$$

(We recall that the σ-field $\mathcal{G}_I$ was defined at the beginning of section 5.) We claim that, for $A \in \mathcal{F}_I^{\mathcal{D}}$, $B \in \mathcal{F}_J^{\mathcal{D}}$, we have

$$P(A \cap B \cap \Gamma) = P(A)P(B \cap \Gamma). \tag{6.5}$$

To see this, the key is in the following two remarks.

- $\mathcal{F}_I^{\mathcal{D}}$ and $\mathcal{F}_J^{\mathcal{D}}$ are independent *conditionally on* $\mathcal{G} = \mathcal{G}_{\mathbb{R}}$. This follows immediately from the independence of W and (I_t). As a consequence, $P(A \cap B|\mathcal{G}) = P(A|\mathcal{G})P(B|\mathcal{G})$ a.s..
- Conditionally to $\mathcal{G}$, the family of random variables $(X_t - X_s)_{s,t \in [c,d)}$ is a Gaussian process whose covariances are measurable with respect to the σ-field generated by the random variables $\{I_t - I_c : t \in (c, d)\}$. In particular, $P(B|\mathcal{G})$ is measurable with respect to this σ-field. Similarly for $[a, b)$ in place of $[c, d)$. Note also that the increment $I_t - I_c$ is a measurable function of the random variables

  $$\{(\tau_k \mathbf{1}_{\{\tau_k \in I\}}, \sigma_k \mathbf{1}_{\{\tau_k \in I\}}) : k \geq 0\} \cup \{(\sigma_{i(c)}, \tau_{i(c)})\}.$$

  It follows that the random variable $(P(B|\mathcal{G})\mathbf{1}_\Gamma)$ is $\mathcal{G}_{(b,d)}$ measurable, and it is therefore independent of $P(A|\mathcal{G})$, which is $\mathcal{G}_{(-\infty,b]}$ measurable.

Thus we have

$$P(A \cap B \cap \Gamma) = E(P(A \cap B|\mathcal{G})\mathbf{1}_\Gamma) = E(P(A|\mathcal{G})P(B|\mathcal{G})\mathbf{1}_\Gamma) = P(A)P(B \cap \Gamma)$$

where the two remarks above have been used. Thus (6.5) is established. Finally

$$\begin{aligned}
&|P(A \cap B) - P(A)P(B)| \\
&= |P(A \cap B \cap \Gamma) + P(A \cap B \cap \Gamma^c) - P(A)P(B \cap \Gamma) - P(A)P(B \cap \Gamma^c)| \\
&= |P(A \cap B \cap \Gamma^c) - P(A)P(B \cap \Gamma^c)| = |P(A \cap B|\Gamma^c) - P(A|\Gamma^c)P(B|\Gamma^c)|P(\Gamma^c) \\
&\leq P(\Gamma^c) = e^{-\lambda(c-b)}\,.
\end{aligned}$$

We finally show that equation (2.7) holds true almost surely and in L^1, for every measurable function $F : \mathbb{R}^k \to \mathbb{R}$ such that $E[|F(X_{b_1} - X_{a_1}, \ldots, X_{b_k} - X_{a_k})|] < +\infty$. Consider the $\mathbb{R}^k$-valued stochastic process $\xi = (\xi_n)_{n \in \mathbb{N}}$ defined by

$$\xi_n := (X_{n\delta + b_1} - X_{n\delta + a_1}, \ldots, X_{n\delta + b_k} - X_{n\delta + a_k})\,,$$

for fixed $\delta > 0$, $k \in \mathbb{N}$ and $(a_1, b_1), \ldots, (a_k, b_k) \subseteq (0, \infty)$. The process ξ is stationary, because we have proven in section 6 that X has stationary increments. Moreover, inequality (6.4) implies that ξ is *mixing*, and therefore ergodic (see e.g. [30], Ch. 5, §2, Definition 4 and Theorem 2). The existence of the limit in (2.7), both a.s. and in L^1, is then a consequence of the classical Ergodic Theorem, (see e.g. [30], Ch. 5, §3, Theorems 1 and 2). □

7. Estimation and data analysis

In this Section we present the main steps that led to the calibration of the model to the DJIA over a period of 75 years; the essential results have been sketched in Section 2.4. We point out that the agreement with the S&P 500, FTSE 100 and Nikkei 225 indexes is very good as well. A systematic treatment of other time series, beyond financial indexes, still has to be done, but some preliminary analysis of single stocks shows that our model fits well

some but not all of them. It would be interesting to understand which of the properties we have mentioned are linked to *aggregation* of several stock prices, as in the DJIA.

The data analysis, the simulations and the plots have been obtained with the software R [28]. The code we have used is publicly available on the web page `http://www.matapp.unimib.it/~fcaraven/c.html`.

7.1. **Overview.** For the numerical comparison of our process $(X_t)_{t\geq 0}$ with the DJIA time series, we have decided to focus on the following quantities:

(a) The *multiscaling of moments*, cf. Corollary 7.

(b) The *volatility autocorrelation decay*, cf. Corollary 9.

Roughly speaking, the idea is to compute *empirically* these quantities on the DJIA time series and then to compare the results with the *theoretical* predictions of our model. This is justified by the ergodic properties of the increments of our process $(X_t)_{t\geq 0}$, cf. equation (2.7).

The first problem that one faces is the *estimation of the parameters* of our model: the two scalars $\lambda \in (0,\infty)$, $D \in (0, \frac{1}{2}]$ and the *distribution* ν *of* σ. This in principle belongs to an infinite dimensional space, but in a first time we focus on the moments $E(\sigma)$ and $E(\sigma^2)$. In order to estimate $(D, \lambda, E(\sigma), E(\sigma^2))$, we take into account four significant quantities that depend only on these parameters:

- the multiscaling coefficients C_1 and C_2 (see (2.13));
- the multiscaling exponent $A(q)$ (see (2.14));
- the volatility autocorrelation function $\rho(t)$ (see (2.17)).

We consider a natural loss functional $\mathcal{L} = \mathcal{L}(D, \lambda, E(\sigma), E(\sigma^2))$ which measures the distance between these theoretical quantities and the corresponding empirical ones, evaluated on the DJIA time series, see (7.3) below. We then define the estimator for $(D, \lambda, E(\sigma), E(\sigma^2))$ as the point at which $\mathcal{L}$ attains its overall minimum, subject to the constraint $E(\sigma^2) \geq (E(\sigma))^2$.

It turns out that the estimated values are such that $E(\sigma^2) \simeq (E(\sigma))^2$, that is σ *is nearly constant* and the estimated parameters completely specify the model. (The constraint $E(\sigma^2) \geq (E(\sigma))^2$ is not playing a relevant role: the unconstrained minimum nearly coincides with the constrained one.) Thus, the problem of determining the distribution of σ beyond its moments $E(\sigma)$ and $E(\sigma^2)$ does not appear in the case of the DJIA. More generally, even we had found $\widehat{Var}(\sigma) := \widehat{E(\sigma^2)} - (\widehat{E(\sigma)})^2 > 0$ and hence σ is not constant, fine details of its distribution ν beyond the first two moments give a negligible contribution to the properties that are relevant for application to real data series, as we observed in Remark 10.

7.2. **Estimation of the parameters** $D, \lambda, E(\sigma), E(\sigma^2)$**.** Let us fix some notation: the DJIA time series will be denoted by $(s_i)_{0\leq i\leq N}$ (where $N = 18848$) and the corresponding detrended log-DJIA time series will be denoted by $(x_i)_{0\leq i\leq N}$:

$$x_i := \log(s_i) - \overline{d}(i)\,,$$

where $\overline{d}(i) := \frac{1}{250}\sum_{k=i-250}^{i-1} \log(s_i)$ is the mean log-DJIA price on the previous 250 days. (Other reasonable choices for $\overline{d}(i)$ affect the analysis only in a minor way.)

The theoretical scaling exponent $A(q)$ is defined in (2.14) while the multiscaling constants C_1 and C_2 are given by (2.13) for $q = 1$ and $q = 2$. Since $q^* = (\frac{1}{2} - D)^{-1} > 2$ (we recall that $0 \leq D \leq \frac{1}{2}$), we can write more explicitly

$$C_1 = \frac{2\sqrt{D}\,\Gamma(\frac{1}{2} + D)\,E(\sigma)\,\lambda^{1/2-D}}{\sqrt{\pi}}\,, \qquad C_2 = 2D\,\Gamma(2D)\,E(\sigma^2)\,\lambda^{1-2D}\,. \tag{7.1}$$

Defining the corresponding empirical quantities requires some care, because the DJIA data are in discrete-time and therefore no $h \downarrow 0$ limit is possible. We first evaluate the empirical q-moment $\widehat{m}_q(h)$ of the DJIA log-returns over h days, namely

$$\widehat{m}_q(h) := \frac{1}{N+1-h} \sum_{i=0}^{N-h} |x_{i+h} - x_i|^q \,.$$

By Theorem 6, the relation $\log \widehat{m}_q(h) \sim A(q)(\log h) + \log(C_q)$ should hold for h small. By plotting $(\log \widehat{m}_q(h))$ versus $(\log h)$ one finds indeed an approximate linear behavior, for moderate values of h and when q is not too large ($q \lesssim 5$). By a standard linear regression of $(\log \widehat{m}_q(h))$ versus $(\log h)$ for $h = 1, 2, 3, 4, 5$ days we therefore determine the empirical values of $A(q)$ and C_q on the DJIA time series, that we call $\widehat{A}(q)$ and $\widehat{C}_q$.

For what concerns the theoretical volatility autocorrelation, Corollary 9 and the stationarity of the increments of our process $(X_t)_{t\geq 0}$ yield

$$\rho(t) := \lim_{h\downarrow 0} \rho(|X_h|, |X_{t+h} - X_t|) = \frac{2}{\pi \, Var(\sigma \, |W_1| \, S^{D-1/2})} \, e^{-\lambda t} \, \phi(\lambda t) \,, \tag{7.2}$$

where $S \sim Exp(1)$ is independent of σ and W_1 and where the function $\phi(\cdot)$ is given by

$$\phi(x) = Var(\sigma) \, E(S^{D-1/2} \, (S+x)^{D-1/2}) + E(\sigma)^2 \, Cov(S^{D-1/2}, (S+x)^{D-1/2}) \,,$$

cf. (2.18). Note that, although $\phi(\cdot)$ does not admit an explicit expression, it can be easily evaluated numerically. For the analogous empirical quantity, we define the empirical DJIA volatility autocorrelation $\widehat{\rho}_h(t)$ over h-days as the *sample correlation coefficient* of the two sequences $(|x_{i+h} - x_i|)_{0\leq i\leq N-h-t}$ and $(|x_{i+h+t} - x_{i+t}|)_{0\leq i\leq N-h-t}$. Since no $h \downarrow 0$ limit can be taken on discrete data, we are going to compare $\rho(t)$ with $\widehat{\rho}_h(t)$ for $h = 1$ day.

We can then define a *loss functional* $\mathcal{L}$ as follows:

$$\begin{aligned} \mathcal{L}(D, \lambda, E(\sigma), E(\sigma^2)) \, = \, & \frac{1}{2} \left\{ \left(\frac{\widehat{C}_1}{C_1} - 1 \right)^2 + \left(\frac{\widehat{C}_2}{C_2} - 1 \right)^2 \right\} + \frac{1}{20} \sum_{k=1}^{20} \left(\frac{\widehat{A}(k/4)}{A(k/4)} - 1 \right)^2 \\ & + \sum_{n=1}^{400} \frac{e^{-n/T}}{\left(\sum_{m=1}^{400} e^{-m/T} \right)} \left(\frac{\widehat{\rho}_1(n)}{\rho(n)} - 1 \right)^2 , \end{aligned} \tag{7.3}$$

where the constant T controls a discount factor in long-range correlations. Of course, different weights for the four terms appearing in the functional could be assigned. We fix $T = 40$ (days) and we define the estimator $(\widehat{D}, \widehat{\lambda}, \widehat{E(\sigma)}, \widehat{E(\sigma^2)})$ of the parameters of our model as the point where the functional $\mathcal{L}$ attains its overall minimum, that is

$$(\widehat{D}, \widehat{\lambda}, \widehat{E(\sigma)}, \widehat{E(\sigma^2)}) := \underset{\substack{D\in(0,\frac{1}{2}],\ \lambda, E(\sigma), E(\sigma^2)\in(0,\infty) \\ \text{such that } E(\sigma^2)\geq (E(\sigma))^2}}{\arg\min} \left\{ \mathcal{L}(D, \lambda, E(\sigma), E(\sigma^2)) \right\} ,$$

where the constraint $E(\sigma^2) \geq (E(\sigma))^2$ is due to $Var(\sigma) = E(\sigma^2) - (E(\sigma))^2 \geq 0$. We expect that such an estimator has good properties, such as asymptotic consistency and normality (we omit a proof of these properties, as it goes beyond the spirit of this paper).

We have then proceeded to the numerical study of the functional $\mathcal{L}$, which appears to be quite regular. With the help of the software Mathematica [26], we have obtained the estimates for the parameters, given already in (2.20):

$$\widehat{D} \simeq 0.16, \quad \widehat{\lambda} \simeq 0.00097, \quad \widehat{E(\sigma)} \simeq 0.108, \quad \widehat{E(\sigma^2)} \simeq 0.0117 \simeq \left(\widehat{E(\sigma)} \right)^2 \tag{7.4}$$

7.3. **Graphical comparison.** Having found that $\widehat{E(\sigma^2)} \simeq (\widehat{E(\sigma)})^2$, the estimated variance of σ is equal to zero, that is σ is a constant. In particular, the model is completely specified and we can compare some quantities, as predicted by our model, with the corresponding numerical ones evaluated on the DJIA time series. The graphical results have been already described in section 2.4 and show a very good agreement, cf. Figure 3 for the multiscaling of moments and the volatility autocorrelation and Figure 4 for the log-return distribution.

Let us give some details about Figure 4. The theoretical distribution $p_t(\cdot) := P(X_t \in \cdot) = P(X_t - X_0 \in \cdot)$ of our model, for which we do not have an analytic expression, can be easily evaluated numerically via Monte Carlo simulations. The analogous quantity evaluated for the DJIA time series is the empirical distribution $\widehat{p}_t(\cdot)$ of the sequence $(x_{i+t} - x_i)_{0 \le i \le N-t}$:

$$\widehat{p}_t(\cdot) := \frac{1}{N+1-t} \sum_{i=0}^{N-t} \delta_{x_{i+t}-x_i}(\cdot)\,. \tag{7.5}$$

In Figure 4(A) we have plotted the bulk of the distributions $p_t(\cdot)$ and $\widehat{p}_t(\cdot)$ for $t = 1$ (daily log-returns) or, more precisely, the corresponding densities, in the range $[-3\hat{s}, +3\hat{s}]$, where $\hat{s} \simeq 0.0095$ is the standard deviation of $\widehat{p}_1(\cdot)$ (i.e., the empirical standard deviation of the daily log returns evaluated on the DJIA time series). In Figure 4(B) we have plotted the tail of $p_1(\cdot)$, that is the function $z \mapsto P(X_1 > z) = P(X_1 < -z)$ (note that $X_t \sim -X_t$ for our model) and the right and left empirical tails $\widehat{R}(z)$ and $\widehat{L}(z)$ of $\widehat{p}_1(\cdot)$, defined for $z \ge 0$ by

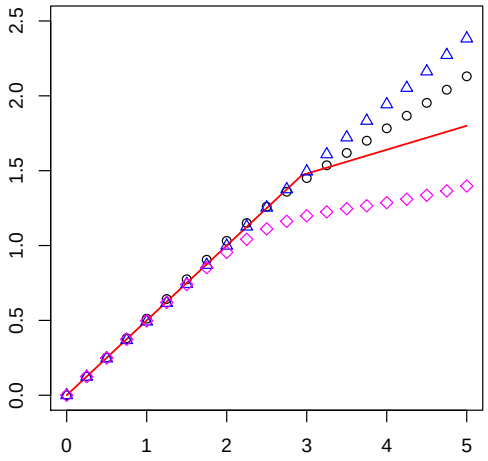
$$\widehat{L}(z) := \frac{\#\{1 \le i \le N :\ x_i - x_{i-1} < -z\}}{N}\,, \quad \widehat{R}(z) := \frac{\#\{1 \le i \le N :\ x_i - x_{i-1} > z\}}{N}\,,$$

in the range $z \in [\hat{s}, 12\hat{s}]$.

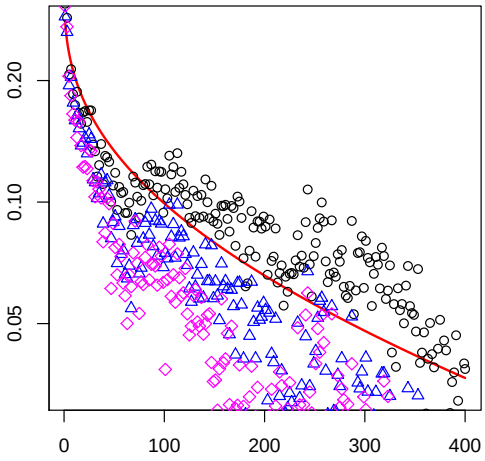
7.4. **Variability of estimators.** In this paper we have identified relatively rare but dramatic shocks in the volatility as the main common source of various stylized facts such as multiscaling, autocorrelations and heavy tails. As observed in Remark 10, the expected number of shocks in a period of 75 years is about 18, which is a rather low number; this means that empirical averages may be not very close to their ergodic limit or, in different words, estimators should have non-negligible variance. A way to detect this is to simulate data from our model for 75 years, and then compute estimators using data in different *subperiods*, that we have chosen of 30 years. Figure 5(A) and 5(B) show indeed a considerable variability of the values of the estimators for the multiscaling exponent and the volatility autocorrelations, when computed in different subperiods. We have then repeated the same computations on the DJIA time series, see Figure 5(C) and 5(D), and we have observed a similar variability. We regard this as a significant test for this model.

Remark 13. We point out that, among the different quantities that we have considered, the scaling exponent $\widehat{A}(q)$ appears to be the most sensitive. For instance, if instead of the opening prices one took the closing prices of the DJIA time series (over the same time period 1935-2009), one would obtain a different (though qualitatively similar) graph of $\widehat{A}(q)$.

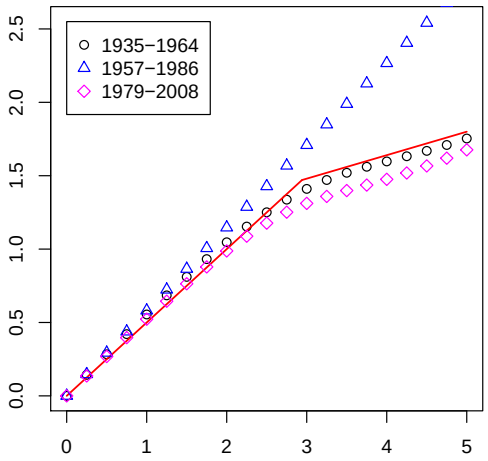
Remark 14. The multiscaling of empirical moments has been observed in several financial indexes in [17], where it is claimed that data provide solid arguments *against* model with linear or piecewise linear scaling exponents. Note that the theoretical scaling exponent $A(q)$ of our model is indeed piecewise linear, cf. (2.14). However, Figure 5(A) shows that the empirical scaling exponent $\widehat{A}(q)$ evaluated on data simulated from our model "smooths out" the change of slope, yielding graphs that are analogous to those obtained for the DJIA time series, cf. Figure 5(C). This shows that the objection against models with piecewise linear $A(q)$, raised in [17], cannot apply to the model we have proposed.



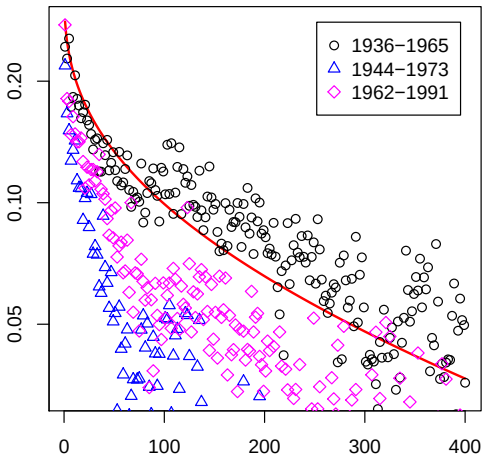
(A) *Multiscaling exponent in subperiods of 30 years: simulated data.*



(B) *Volatility autocorrelation in subperiods of 30 years: sumilated data.*



(C) *Multiscaling exponent in subperiods of 30 years: DJIA data.*



(D) *Volatility autocorrelation in subperiods of 30 years: DJIA data.*

FIGURE 5. *Variability of estimators in subperiods of 30 years.*
Empirical evaluation of the observables $\widehat{A}(q)$ and $\widehat{\rho}_1(t)$ in subperiods or 30 years for a 75-years-long time series, sampled from our model $(X_t)_{t\geq 0}$ ((A) and (B)) and from the DJIA time series ((C) and (D))

APPENDIX A. PROOF OF PROPOSITION 11

We first need two simple technical lemmas.

Lemma 15. *For $0 < q < 2$, consider the function $\varphi_q : [0, +\infty) \to [0, +\infty)$ define by*

$$\varphi_q(\beta) := \int_{-\infty}^{+\infty} e^{\beta|x|^q - \frac{1}{2}x^2} \, \mathrm{d}x.$$

Then there are constants $C_1, C_2 > 0$, that depend on q, such that for all $\beta > 0$

$$C_1 e^{C_1 \beta^{\frac{2}{2-q}}} \leq \varphi_q(\beta) \leq C_2 e^{C_2 \beta^{\frac{2}{2-q}}} . \tag{A.1}$$

Proof. We begin by observing that it is enough to establish the bounds in (A.1) for β large enough. Consider the function of positive real variable $f(r) := e^{\beta r^q - \frac{1}{2} r^2}$. It is easily checked that f is increasing for $0 \leq r \leq (\beta q)^{\frac{1}{2-q}}$. Thus

$$\varphi_q(\beta) \geq \int_{\frac{1}{2}(\beta q)^{\frac{1}{2-q}}}^{(\beta q)^{\frac{1}{2-q}}} f(r)\, \mathrm{d}r \geq \frac{1}{2}(\beta q)^{\frac{1}{2-q}} f\left(\frac{1}{2}(\beta q)^{\frac{1}{2-q}}\right) = \frac{1}{2}(\beta q)^{\frac{1}{2-q}} \exp\left[c(q)\beta^{\frac{2}{2-q}}\right],$$

with $c(q) := \frac{1}{2^q} q^{\frac{q}{2-q}} - \frac{1}{8} q^{\frac{2}{2-q}} > 0$. The lower bound in (A.1) easily follows for β large.

For the upper bound, by direct computation one observes that $f(r) \leq e^{-\frac{1}{4} r^2}$ for $r > (4\beta)^{\frac{1}{2-q}}$. We have:

$$\varphi_q(\beta) \leq \int_{|x| \leq (4\beta)^{\frac{1}{2-q}}} f(|x|)\, \mathrm{d}x + \int_{|x| > (4\beta)^{\frac{1}{2-q}}} e^{-\frac{1}{4}x^2} \mathrm{d}x \leq 2(4\beta)^{\frac{1}{2-q}} \|f\|_\infty + \int_{-\infty}^{+\infty} e^{-\frac{1}{4}x^2} \mathrm{d}x\,.$$

Since $\|f\|_\infty = f((\beta q)^{\frac{1}{2-q}}) = \exp\left[C(q)\beta^{\frac{2}{2-q}}\right]$ for a suitable $C(q)$, also the upper bound follows, for β large. □

Lemma 16. *Let $X_1, X_2, \ldots, X_n$ be independent random variables uniformly distributed in $[0,1]$, and $U_1 < U_2 < \ldots < U_n$ be the associated order statistics. For $n \geq 2$ and $k = 2, \ldots, n$, set $\xi_k := U_k - U_{k-1}$. Then, for every $\epsilon > 0$*

$$\lim_{n \to +\infty} P\left(\left| \left\{ k \in \{2, \ldots, n\} : \ \xi_k > \frac{1}{n^{1+\epsilon}} \right\} \right| \geq n^{1-\epsilon} \right) = 1.$$

Proof. This is a consequence of the following stronger result: for every $x > 0$, as $n \to \infty$ we have the convergence in probability

$$\frac{1}{n} \left| \left\{ k \in \{2, \ldots, n\} : \ \xi_k > \frac{x}{n} \right\} \right| \longrightarrow e^{-x}\,,$$

see [33] for a proof. □

Proof of Proposition 11. Since $X_t = W_{I_t}$ and $\sqrt{I_t}\, W_1$ have the same law, we can write

$$E\left[e^{\gamma |X_t|^q}\right] = E\left[\exp\left(\gamma I_t^{q/2} |W_1|^q\right)\right].$$

We begin with the proof of (6.3), hence we work in the regime $q < (1-D)^{-1}$, or $q = (1-D)^{-1}$ and $D < \frac{1}{2}$; in any case, $q < 2$. We start with the "$\Leftarrow$" implication. Since I_t and W_1 are independent, it follows by Lemma 15 that

$$E\left[\exp\left(\gamma I_t^{q/2} |W_1|^q\right)\right] \leq C E\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\right], \tag{A.2}$$

for some $C, \delta > 0$. For the moment we work on the event $\{i(t) \geq 1\}$. It follows by the basic bound (4.10) that

$$I_t \leq \sum_{k=0}^{i(t)} \xi_k^{2D}\, \sigma_k^2\,, \tag{A.3}$$

where we set

$$\xi_k := \begin{cases} \tau_1 & \text{for } k = 0 \\ \tau_{k+1} - \tau_k & \text{for } 1 \le k \le i(t) - 1 \\ t - \tau_{i(t)} & \text{for } k = i(t) \end{cases}$$

Note that $\sum_{k=0}^{i(t)} \xi_k = t$. By applying Hölder inequality to (A.3) with exponents $p = \frac{1}{2D}$, $p' = \frac{1}{1-2D}$, we obtain:

$$I_t \,\le\, t^{2D} \left(\sum_{k=0}^{i(t)} \sigma_k^{\frac{2}{1-2D}} \right)^{1-2D} .$$

By assumption $q \le \frac{1}{1-D}$, which is the same as $(1-2D)\frac{q}{2-q} \le 1$. Thus

$$I_t^{\frac{q}{2-q}} \,\le\, t^{\frac{2Dq}{2-q}} \left(\sum_{k=0}^{i(t)} \sigma_k^{\frac{2}{1-2D}} \right)^{(1-2D)\frac{q}{2-q}} \,\le\, t^{\frac{2Dq}{2-q}} \sum_{k=0}^{i(t)} \sigma_k^{\frac{2q}{2-q}} . \tag{A.4}$$

Now observe that if $i(t) = 0$ we have $I_t = \sigma_0^2[(t-\tau_0)^{2D} - (-\tau_0)^{2D}] \le \sigma_0^2\, t^{2D}$, hence (A.4) holds also when $i(t) = 0$. Therefore, by (A.2)

$$E\left[e^{\gamma |X_t|^q}\right] \,\le\, CE\left[\exp\left(\delta\, t^{\frac{2Dq}{2-q}} \sum_{k=0}^{i(t)} \sigma_k^{\frac{2q}{2-q}}\right)\right] \,=\, C\,E\left[\rho^{i(t)+1}\right] , \tag{A.5}$$

where we have set

$$\rho = \rho_t := E\left[\exp\left(\delta\, t^{\frac{2Dq}{2-q}}\, \sigma_0^{\frac{2q}{2-q}}\right)\right] .$$

Therefore, if $\rho < \infty$, the right hand side of (A.5) is finite, because $i(t) \sim Po(\lambda t)$ has finite exponential moments of all order. This proves the "$\Leftarrow$" implication in (6.3).

The "$\Rightarrow$" implication in (6.3) is simpler. By the lower bound in Lemma 15 we have

$$E\left[e^{\gamma |X_t|^q}\right] = E\left[\exp\left(\gamma I_t^{q/2} |W_1|^q\right)\right] \ge CE\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\right], \tag{A.6}$$

for suitable $C, \delta > 0$. We note that

$$\begin{aligned} E\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\right] &\ge E\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right) \mathbf{1}_{\{i(t)=0\}}\right] \\ &= E\left[\exp\left(\delta \left[(t-\tau_0)^{2D} - (-\tau_0)^{2D}\right]^{\frac{q}{2-q}}\, \sigma_0^{\frac{2q}{2-q}}\right)\right] P(i(t)=0)\,. \end{aligned} \tag{A.7}$$

Under the condition

$$E\left[\exp\left(\alpha \sigma^{\frac{2q}{2-q}}\right)\right] = +\infty \;\; \forall \alpha > 0,$$

the last expectation in (A.7) is infinite, since $\left[(t-\tau_0)^{2D} - (-\tau_0)^{2D}\right] > 0$ almost surely and is independent of σ_0. Looking back at (A.6), we have proved the "$\Rightarrow$" implication in (6.3).

Next we prove (6.2), hence we assume that $q > (1-D)^{-1}$. Consider first the case $q < 2$ (which may happen only for $D < \frac{1}{2}$). By (A.6)

$$E\left[e^{\gamma |X_t|^q}\right] \ge CE\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\right] .$$

We note that, by the definition (2.1) of I_t, we can write

$$I_t \geq \sum_{k=2}^{i(t)} \sigma_{k-1}^2 (\tau_k - \tau_{k-1})^{2D}\,, \tag{A.8}$$

where we agree that the sum is zero if $i(t) < 2$. For $n \geq 0$, we let P_n to denote the conditional probability $P(\,\cdot\,|i(t) = n)$ and E_n the corresponding expectation. Note that, under P_n, the random variables $(\tau_k - \tau_{k-1})_{k=2}^n$ have the same law of the random variables $(\xi_k)_{k=2}^n$ in Lemma 16, for $n \geq 2$. Consider the following events:

$$A_n := \{\sigma_k^2 \geq a,\ \forall k = 2, \dots, n\}\,, \qquad B_n := \left\{ \left| \left\{ k = 2, \dots, n :\ \xi_k > \frac{1}{n^{1+\epsilon}} \right\} \right| \geq n^{1-\epsilon} \right\},$$

where $a > 0$ is such that $\nu([a, +\infty)) =: \rho > 0$ and $\epsilon > 0$ will be chosen later. Note that $P_n(A_n) = \rho^{n-1}$ while $P_n(B_n) \to 1$ as $n \to +\infty$, by Lemma 16. In particular, there is $c > 0$ such that $P_n(B_n) \geq c$ for every n. Plainly, A_n and B_n are independent under P_n. We have

$$\begin{aligned}
\psi(n) := E_n\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\right] &\geq E_n\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\mathbf{1}_{A_n \cap B_n}\right] \\
&\geq c\,\rho^{n-1} \exp\left[\delta a^{\frac{q}{2-q}} \left(\frac{1}{n^{1+\epsilon}}\right)^{2D\frac{q}{2-q}} n^{(1-\epsilon)\frac{q}{2-q}}\right] \\
&= c\,\rho^{n-1} \exp\left[\delta a^{\frac{q}{2-q}} n^{(1-2D-\epsilon(1+2D))\frac{q}{2-q}}\right]
\end{aligned} \tag{A.9}$$

Note that $q > \frac{1}{1-D}$ is equivalent to $(1-2D)\frac{q}{2-q} > 1$, therefore ϵ can be chosen small enough so that $b := (1 - 2D - \epsilon(1+2D))\frac{q}{2-q} > 1$. It then follows by (A.9) that $\psi(n) \geq d\,\exp(d\,n^b)$ for every $n \in \mathbb{N}$, for a suitable $d > 0$. Therefore

$$E\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\right] = E[\psi(i(t))] = +\infty\,,$$

because $i(t) \sim Po(\lambda t)$ and hence $E[\exp(d\,i(t)^b)] = \infty$ for all $d > 0$ and $b > 1$.

Next we consider the case $q \geq 2$. Note that

$$E\left[e^{\gamma |X_t|^q}\right] = E\left[\exp\left(\gamma I_t^{q/2} |W_1|^q\right)\right]\,, \tag{A.10}$$

hence if $q > 2$ we have $E\left[e^{\gamma |X_t|^q}\right] = \infty$, because $E[\exp(c|W_1|^q)] = \infty$ for every $c > 0$, $I_t > 0$ almost surely and I_t is independent of W_1. On the other hand, if $q = 2$ we must have $D < \frac{1}{2}$ (recall that we are in the regime $q > (1-D)^{-1}$) and the steps leading to (A.9) have shown that in this case I_t is unbounded. It then follows again from (A.10) that $E[e^{\gamma |X_t|^2}] = \infty$. □

Appendix B. The model of Baldovin and Stella

Let us briefly discuss the model proposed by F. Baldovin and A. Stella [7, 31], motivated by renormalization group arguments from statistical physics. They first introduce a process $(Y_t)_{t \geq 0}$ which satisfies the scaling relation (1.3) for a given function g, that is assumed to be even, so that its Fourier transform $\hat{g}(u) := \int_{\mathbb{R}} e^{iux} g(x) \mathrm{d}x$ is real (and even). The process $(Y_t)_{t \geq 0}$ is defined by specifying its finite dimensional laws: for $t_1 < t_2 < \cdots < t_n$ the joint density of $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$ is given by

$$p(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = h\left(\frac{x_1}{\sqrt{t_1}}, \frac{x_2 - x_1}{\sqrt{t_2 - t_1}}, \dots, \frac{x_n - x_{n-1}}{\sqrt{t_n - t_{n-1}}}\right), \tag{B.1}$$

where h is the function whose Fourier transform $\hat{h}$ is given by

$$\hat{h}(u_1, u_2, \dots, u_n) := \hat{g}\Big(\sqrt{u_1^2 + \ldots + u_n^2}\Big). \tag{B.2}$$

Note that if g is the standard Gaussian density, then $(Y_t)_{t\geq 0}$ is the ordinary Brownian motion. For a non Gaussian g, the expression in (B.2) is not necessarily the Fourier transform of a probability on $\mathbb{R}^n$, so that some care is needed (we come back to this point in a moment). However, it is clear from (B.1) that the increments of the process $(Y_t)_{t\geq 0}$ corresponding to time intervals of the same length (that is, for fixed $t_{i+1} - t_i$) have a *permutation invariant distribution* and therefore cannot exhibit any decay of correlations.

For this reason, Baldovin and Stella introduce what is probably the most interesting ingredient of their construction, namely a special form of time-inhomogeneity. They define it in terms of finite dimensional distributions, bur it is simpler to give a pathwise construction: given a sequence of (possibly random) times $0 < \tau_1 < \tau_2 < \cdots < \tau_n \uparrow +\infty$ and a fixed $0 < D \leq 1/2$, they introduce a new process $(X_t)_{t\geq 0}$ defined by

$$X_t := Y_{t^{2D}} \qquad \text{for } t \in [0, \tau_1), \tag{B.3}$$

and more generally

$$X_t := Y_{(t-\tau_n)^{2D} + \sum_{k=1}^{n} (\tau_k - \tau_{k-1})^{2D}} \qquad \text{for } t \in [\tau_n, \tau_{n+1})\,. \tag{B.4}$$

For $D = 1/2$ we have clearly $X_t \equiv Y_t$, while for $D < 1/2$ the process $(X_t)_{t\geq 0}$ is obtained from $(Y_t)_{t\geq 0}$ by a nonlinear time-change, that is "refreshed" at each time τ_n. This transformation has the effect of amplifying the increments of the process for t immediately after the times $(\tau_n)_{n\geq 1}$, while the increments tend to become small for larger t.

Let us shed some light into the implicit relations (B.1)–(B.2). If a stochastic process $(Y_t)_{t\geq 0}$ is to satisfy these relations, it must necessarily have *exchangeable increments*: by this we mean (cf. [19, p.1210]) that, setting $\Delta Y_{(a,b)} := Y_b - Y_a$ for short, the distribution of the random vector $(\Delta Y_{I_1+y_1}, \dots, \Delta Y_{I_n+y_n})$ — where the I_j's are intervals and y_j's real numbers — does not depend on $y_1, \dots, y_n$, as long as the intervals $y_1 + I_1, \dots, y_n + I_n$ are disjoint. If we make the (very mild) assumption that $(Y_t)_{t\geq 0}$ has no fixed point of discontinuity, then a continuous-time version of the celebrated de Finetti's theorem ensures that $(Y_t)_{t\geq 0}$ is a mixture of Lévy processes, cf. Theorem 3 in [19] (cf. also [1]). Actually, more can be said: since by (1.3) the distribution of the increments of $(Y_t)_{t\geq 0}$ is *isotropic*, i.e., it has spherical symmetry in $\mathbb{R}^n$, by Theorem 4 in [19] the process $(Y_t)_{t\geq 0}$ is necessarily *a mixture of Brownian motions*. This means that we have the following representation:

$$Y_t = \sigma\, W_t, \tag{B.5}$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion and σ is an independent real random variable (a random, but time-independent, volatility). Viceversa, if a process $(Y_t)_{t\geq 0}$ satisfies (B.5), then, denoting by ν the law of σ, it is easy to check that relations (B.1)–(B.2) hold with

$$g(x) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma}}\, e^{-\frac{x^2}{2\sigma^2}}\, \nu(\mathrm{d}\sigma)\,, \tag{B.6}$$

or, equivalently,

$$\hat{g}(u) = \int_{\mathbb{R}} e^{-\frac{\sigma^2 u^2}{2}}\, \nu(\mathrm{d}\sigma)\,.$$

This shows that the functions g for which (B.1)–(B.2) provide a consistent family of finite dimensional distributions are exactly those that may be expressed as in (B.6) for some probability ν on $(0, +\infty)$.

Note that a path of (B.5) is obtained by sampling independently σ from ν and $(W_t)_{t\geq 0}$ from the Wiener measure, hence this path *cannot* be distinguished from the path of a Brownian motion with *constant* volatility. In particular, the (possible) correlation of the increments of the process $(Y_t)_{t\geq 0}$ cannot be detected empirically, and the same observation applies to the time-inhomogeneous process $(X_t)_{t\geq 0}$ obtained by $(Y_t)_{t\geq 0}$ through (B.3)–(B.4). In other words, the processes obtained through this construction have non ergodic increments.

Nevertheless, Baldovin and Stella claim to measure nonzero correlations from their samples: after estimating the function g and the parameters t_0 and D on the DJIA time series, their simulated trajectories show a good agreement with the clustering of volatility, as well as with the basic scaling (1.3) and the multiscaling of moments. The explanation of this apparent contradiction is that Baldovin and Stella do not simulate the process $(X_t)_{t\geq 0}$ defined through the above construction, but rather an autoregressive approximation of it. In fact, besides making a periodic choice of the times $\tau_n := nt_0$, they fix a small time step δ and a natural number N and they first simulate x_δ, $x_{2\delta}$, …, $x_{N\delta}$ according to the true distribution of $(X_\delta, X_{2\delta}, \ldots, X_{N\delta})$. Then they compute the conditional distribution of $X_{(N+1)\delta}$ given $X_{2\delta} = x_{2\delta}$, $X_{3\delta} = x_{3\delta}$, …, $X_{N\delta} = x_{N\delta}$ — thus *neglecting* x_δ — and sample $x_{(N+1)\delta}$ from this distribution. Similarly, $x_{(N+2)\delta}$ is sampled from the conditional distribution of $X_{(N+2)\delta}$ given $X_{3\delta} = x_{3\delta}$, …, $X_{N\delta} = x_{N\delta}$, $X_{(N+1)\delta} = x_{(N+1)\delta}$, neglecting both x_δ and $x_{2\delta}$, and so on. It is plausible that such an autoregressive procedure may produce an ergodic process.

Acknowledgements

We thank Fulvio Baldovin, Massimiliano Caporin, Wolfgang Runggaldier and Attilio Stella for fruitful discussions. We are very grateful to the anonymous referee for several important remarks and suggestions.

References

[1] L. Accardi, Y. G. Lu, *A continuous version of de Finetti's theorem*, Ann. Probab. **21** (1993), 1478–1493.
[2] A. Andreoli, *Scaling and Multiscaling in Financial Indexes: a Simple Model*, Ph.D. Thesis, University of Padova (2011). Available at the address `http://www.matapp.unimib.it/~fcaraven/c.html`.
[3] S. Asmussen, *Applied probability and queues*, Second Edition, Application of Mathematics **51**, Springer–Verlag, New York (2003).
[4] T. Ané and H. Geman, *Order Flow, Transaction Clock, and Normality of Asset Returns*, The Journal of Finance, **55**, No. 5 (2000), 2259-2284
[5] R. T. Baillie, *Long memory processes and fractional integration in econometrics*, J. Econometrics **73** (1996), 5–59.
[6] O. E. Barndorff-Nielsen, N. Shephard, *Non Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics*, J. R. Statist. Soc. B **63** (2001), 167–241.
[7] F. Baldovin, A. Stella, *Scaling and efficiency determine the irreversible evolution of a market*, PNAS **104**, n. 50 (2007), 19741–19744.
[8] T. Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity*, J. Econometrics **31** (1986), 307–327.
[9] T. Bollerslev, H.O. Mikkelsen, *Modeling and pricing long memory in stock market volatility*, J. Econometrics **31** (1996), 151–184.
[10] T. Bollerslev, U. Kretschmer, C. Pigorsch, G. Tauchen, *A discrete-time model for daily S & P500 returns and realized variations: Jumps and leverage effects*, J. Econometrics **150** (2009), 151–166.
[11] M. Bonino, *Portfolio allocation and monitoring under volatility shocks*, Master's Thesis, University of Padova (2011). Available at the address `http://www.matapp.unimib.it/~fcaraven/c.html`.
[12] L.E. Calvet, A.J. Fisher, *Multifractal Volatility*, Academic Press (2008).
[13] L.E. Calvet, A.J. Fisher, *Forecasting multifractal volatility*, Journal of Econometrics **105** (2001), 27-58.
[14] L.E. Calvet, A. J. Fisher, B. B. Mandelbrot, Cowles Foundation Discussion Papers No. 1164-1166, Yale University (1997). Papers available from `http://cowles.econ.yale.eduorhttp://www.ssrn.com`.

[15] P.K. Clark, *A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices*, Econometrica **41** (1973), 135-155
[16] R. Cont, *Empirical properties of asset returns: stylized facts and statistical issues*, Quantitative Finance **1** (2001), 223-236.
[17] T. Di Matteo, T. Aste, M. M. Dacorogna, *Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development*, J. Banking Finance **29** (2005), 827–851.
[18] R. F. Engle, *Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation*, Econometrica **50** (1982), 987–1008.
[19] D. A. Freedman, *Invariants Under Mixing Which Generalize de Finetti's Theorem: Continuous Time Parameter*, Ann. Math. Statist. **34** (1963), 1194–1216.
[20] S. Galluccio, G. Caldarelli, M. Marsili, Y. C. Zhang, *Scaling in currency exchange*, Physica A **245** (1997), 423-36.
[21] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, Y. Dodge, *Turbulent cascades in foreign exchange markets*, Nature **381** (1996), 767-70.
[22] J. C. Hull, *Options, Futures and Other Derivatives*, Pearson/Prentice Hall (2009).
[23] I. Karatzas, S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer (1988).
[24] C. Kluppelberg, A. Lindner, R. A. Maller, *A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour*, J. Appl. Probab. **41** (2004) 601–622.
[25] C. Kluppelberg, A. Lindner, R. A. Maller, *Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models*, in *From Stochastic Calculus to Mathematical Finance*, Yu. Kabanov, R. Lipster and J. Stoyanov (Eds.), Springer (2007).
[26] Wolfram Research, Inc., *Mathematica*, Version 7.0, Champaign, IL (2008).
[27] P. Pigato, *A multivariate model for financial indexes subject to volatility shocks*, Master's Thesis, University of Padova (2011). Available at the address `http://www.matapp.unimib.it/~fcaraven/c.html`.
[28] R Development Core Team (2009), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. URL: `http://www.R-project.org`.
[29] N. Shephard, T.G. Andersen, *Stochastic Volatility: Origins and Overview*, in *Handbook of Financial Time Series*, Springer (2009), 233-254.
[30] A.N. Shiryaev, *Probability*, Graduate Texts in Mathematics, 2^{nd} edition, Springer, 1995.
[31] A. L. Stella, F. Baldovin, *Role of scaling in the statistical modeling of finance*, Pramana **71** (2008), 341–352.
[32] J.C. Vassilicos, A. Demos, F. Tata, *No evidence of chaos but some evidence of multifractals in the foreign exchange and the stock market*, in *Applications of Fractals and Chaos*, eds. A.J. Crilly, R.A. Earnshaw, H. Jones, pp. 249-65, Springer (1993).
[33] L. Weiss, *The Stochastic Convergence of a Function of Sample Successive Differences*, Ann. Math. Statist. **26** (1955), 532–536.

Dipartimento di Management, Università Politecnica delle Marche, Piazzale Martelli 8, 60121 Ancona Italy.
E-mail address: `alessandro.andreoli@univpm.it`

Dipartimento di Matematica e Applicazioni, Università degli Studi di Milano-Bicocca, via Cozzi 53, I-20125 Milano, Italy
E-mail address: `francesco.caravenna@unimib.it`

Dipartimento di Matematica Pura ed Applicata, Università degli Studi di Padova, via Trieste 63, I-35121 Padova, Italy
E-mail address: `daipra@math.unipd.it`

Dipartimento di Matematica, Politecnico di Milano, Piazzale Leonardo da Vinci 32, I-20133 Milano, Italy
E-mail address: `gustavo.posta@polimi.it`