

SCALING AND MULTISCALING IN FINANCIAL SERIES: A SIMPLE MODEL

ALESSANDRO ANDREOLI,* *Università Politecnica delle Marche*

FRANCESCO CARAVENNA,** *Università degli Studi di Milano-Bicocca*

PAOLO DAI PRA,*** *Università degli Studi di Padova*

GUSTAVO POSTA,**** *Politecnico di Milano*

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 remarkable accuracy.

Keywords: Financial index; time series; scaling; multiscaling; Brownian motion; stochastic volatility; heavy tail; multifractal model

2010 Mathematics Subject Classification: Primary 60G44

Secondary 91B25; 91G70

1. Introduction

1.1. Modeling financial assets

Recent developments in stochastic modeling of time series have been strongly influenced by the analysis of financial assets, such as exchange rates, stocks, and market indexes. The basic model, which has given rise to the celebrated Black–Scholes formula [24], [25], assumes that the logarithm X_t of the price of the asset, after subtracting the trend, evolves through the simple equation

$$dX_t = \sigma dB_t, \quad (1)$$

where σ (the *volatility*) is a constant and $(B_t)_{t \geq 0}$ is a standard Brownian motion. It has been well known 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, which may be interpreted as *shocks* in the market;

Received 1 April 2011; revision received 16 April 2012.

* Postal address: Dipartimento di Management, Università Politecnica delle Marche, Piazzale Martelli 8, 60121 Ancona, Italy. Email address: alessandro.andreoli@univpm.it

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

*** Postal address: Dipartimento di Matematica, Università degli Studi di Padova, via Trieste 63, I-35121 Padova, Italy. Email address: daipra@math.unipd.it

**** Postal address: Dipartimento di Matematica, Politecnico di Milano, Piazzale Leonardo da Vinci 32, I-20133 Milano, Italy. Email address: gustavo.posta@polimi.it

- 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 a 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), allowing $\sigma = \sigma_t$ to vary with t and to be itself a stochastic process. This produces a wide class of processes, known as *stochastic volatility models*, determined by the process $(\sigma_t)_{t \geq 0}$, which are able to capture (at least some of) the abovementioned stylized facts; cf. [7], [29], and the 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:

$$p_h(\cdot) := \frac{1}{T-h} \sum_{i=1}^{T-h} \delta_{x_{i+h}-x_i}(\cdot), \quad x_i := \log(s_i) - \bar{d}_i. \tag{2}$$

Here \bar{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)):

$$p_h(dr) \simeq \frac{1}{\sqrt{h}} g\left(\frac{r}{\sqrt{h}}\right) dr. \tag{3}$$

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

$$m_q(h) := \frac{1}{T-h} \sum_{i=1}^{T-h} |x_{i+h} - x_i|^q = \int |r|^q p_h(dr), \tag{4}$$

from relation (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 $A(q) < q/2$, holds; cf. Figure 1(b). This is the so-called *multiscaling of moments*; cf. [18], [22], [23], and [32].

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*, such as the MMAR (multifractal model of asset returns; cf. [14], [15], and [20]) 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*, i.e. $X_t = W_{I(t)}$, where the time change $(I(t))_{t \geq 0}$ is a continuous and increasing process,

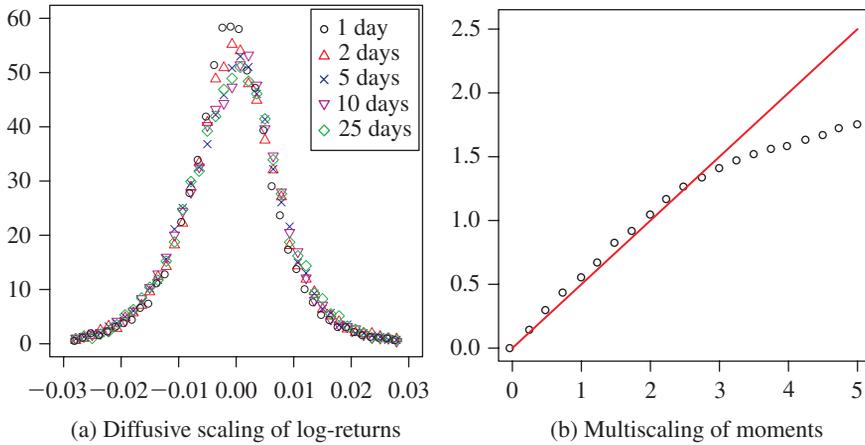


FIGURE 1: Scaling properties of the DJIA time series (opening prices 1935–2009). (a) The empirical densities of the log-returns over 1, 2, 5, 10, and 25 days show a remarkable overlap under diffusive scaling. (b) The scaling exponent $A(q)$ as a function of q , defined by the relation $m_q(h) \approx h^{A(q)}$ (cf. (4)), bends down from the Gaussian behavior $q/2$ (solid 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 $h \in \{1, \dots, 5\}$ (see Section 7 for more details).

sometimes called the *trading time*, which displays *multifractal features* and is usually taken to be independent of the Brownian motion $(W_t)_{t \geq 0}$ (see [12] for more details).

Modeling financial series through a random time change of Brownian motion is a classical topic, dating back to Clark [16], 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 almost surely (a.s.), any random time change of Brownian motion $X_t = W_{I(t)}$ can be written as a stochastic volatility model $dX_t = \sigma_t dB_t$, and vice versa. 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 $dX_t = \sigma_t dB_t$, where $(B_t)_{t \geq 0}$ is a Brownian motion and $(\sigma_t)_{t \geq 0}$ is an independent process with paths in $L^2_{\text{loc}}(\mathbb{R})$ a.s.—are the same class of processes; cf. [7] and [29]. The link between the two representations $dX_t = \sigma_t dB_t$ and $X_t - X_0 = W_{I(t)}$ is given by $\sigma_t = \sqrt{I'(t)}$ and $B_t = \int_0^t (I'(s))^{-1/2} dW_{I(s)} = \int_0^t (I'(I^{-1}(v)))^{-1/2} dW_u$. However, a key feature of multifractal models is precisely that their trading time $(I(t))_{t \geq 0}$ has *nonabsolutely continuous paths* a.s.; hence, they cannot be expressed as stochastic volatility models. This makes their analysis harder, as the standard tools available for Itô 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 abovementioned 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 the DJIA, S&P 500, FTSE 100, and 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 for by our model, such as the multiscale 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. [17]. As we discuss in Section 3, such features—which are relevant in the analysis of particular assets—can be incorporated into our model in a natural way. Generalizations in this sense are currently under investigation, as are the performances of our model in financial problems, such as the pricing of options (see [2]). In this paper we focus for simplicity on 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 generalizations; cf. [5], [8], [9], and [19]. More recently, continuous-time versions have been studied as well; cf. [26] and [27]. With no aim for completeness, let us mention that GARCH and FIGARCH do not display multiscaling of moments; cf. [12, Section 8.1.4]. We have also tested the model recently proposed in [10], which fits the statistics of the empirical volatility very accurately, and does exhibit multiscaling of moments. However, the model requires the calibration of more than 30 parameters.

We conclude by 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 in more detail the numerical comparison between our model and the DJIA time series.
- Finally, Appendix A contains the proofs of some technical results, while Appendix B is devoted to a brief discussion of the model introduced by Baldovin and Stella in [6] and [31], which has partially inspired the construction of our model.

1.3. Notation

Throughout the paper, the indexes s, t, u, x , and λ run over real numbers, while i, k, m , and n run over integers, so $t \geq 0$ means $t \in [0, \infty)$ while $n \geq 0$ means $n \in \{0, 1, 2, \dots\}$. The notation ‘ \sim ’ denotes asymptotic equivalence for positive sequences ($a_n \sim b_n$ if and only if $a_n/b_n \rightarrow 1$ as $n \rightarrow \infty$) and also equality in law for random variables, e.g. $W_1 \sim \mathcal{N}(0, 1)$.

Given two real functions $f(x)$ and $g(x)$, we write $f = O(g)$ as $x \rightarrow x_0$ if there exists $M > 0$ such that $|f(x)| \leq M|g(x)|$ for x in a neighborhood of x_0 , and we write $f = o(g)$ if $f(x)/g(x) \rightarrow 0$ as $x \rightarrow x_0$; in particular, $O(1)$ is a bounded quantity and $o(1)$ is a vanishing quantity. The standard exponential and Poisson laws are denoted by $\text{Exp}(\lambda)$ and $\text{Po}(\lambda)$ for $\lambda > 0$: $X \sim \text{Exp}(\lambda)$ means that $P(X \leq x) = (1 - e^{-\lambda x}) \mathbf{1}_{[0, \infty)}(x)$ for all $x \in \mathbb{R}$, while $Y \sim \text{Po}(\lambda)$ means that $P(Y = n) = e^{-\lambda} \lambda^n / n!$ for all $n \in \{0, 1, 2, \dots\}$. We sometimes write ‘(constant)’ 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 of, e.g. [4] and [16]. 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;
- $D \in (0, \frac{1}{2}]$ determines the sublinear 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 fit turns out to be obtained for a *nearly constant* σ . In any case, we stress 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; see Remark 5 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 , and ν , we need the following three sources of randomness:

- 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); for conciseness, we denote by σ a variable with the same law ν .

We assume that W , \mathcal{T} , and Σ are defined on some probability space (Ω, \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 , and $(\tau_{n+1} - \tau_n)_{n \geq 1}$ are independent and identically distributed with marginal laws $\text{Exp}(\lambda)$. 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 $2/\lambda$.

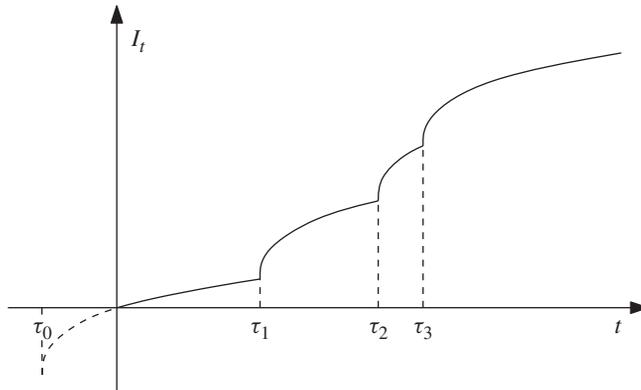


FIGURE 2: A sample trajectory of the process $(I_t)_{t \geq 0}$.

Although some of our results would hold for more general distributions of \mathcal{T} , we focus for simplicity on the (rather natural) choice of a Poisson process.

For $t \geq 0$, we define

$$i(t) := \sup\{n \geq 0 : \tau_n \leq t\} = \#\{\mathcal{T} \cap (0, t]\}, \tag{5}$$

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

$$I_t = I(t) := \sigma_{i(t)}^2 (t - \tau_{i(t)})^{2D} + \sum_{k=1}^{i(t)} \sigma_{k-1}^2 (\tau_k - \tau_{k-1})^{2D} - \sigma_0^2 (-\tau_0)^{2D}, \tag{6}$$

with the agreement that the sum on the right-hand side is 0 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 $(dI_t/dt)_{t \geq 0}$ 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

$$X_t := W_{I_t}.$$

In words, our model is a 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 the Black–Scholes model with volatility σ . On the other hand, when $D < \frac{1}{2}$, the paths of I are singular (nondifferentiable) at the points in \mathcal{T} ; cf. Figure 2. This suggests a possible financial interpretation 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 the next subsection, where we give a stochastic volatility formulation of the model. We point out that the singularity is produced by the sublinear time change $t \mapsto t^{2D}$, which was first suggested by Baldovin and Stella in [6] and [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.

(P1) The process X has stationary increments.

(P2) The following relation between moments of X_t and σ holds: for any $q > 0$,

$$E(|X_t|^q) < \infty \text{ for some (hence any) } t > 0 \iff E(\sigma^q) < \infty. \tag{7}$$

(P3) The process X can be represented as a stochastic volatility model, i.e.

$$dX_t = v_t dB_t, \tag{8}$$

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

$$\begin{aligned} B_t &:= \int_0^t \frac{1}{\sqrt{I'(s)}} dW_{I(s)} = \int_0^{I_t} \frac{1}{\sqrt{I'(I^{-1}(u))}} dW_u, \\ v_t &:= \sqrt{I'(t)} = \sqrt{2D\sigma_{i(t)}(t - \tau_{i(t)})^{D-1/2}}. \end{aligned} \tag{9}$$

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

(P4) 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 $P(\cdot \mid \mathcal{T}, \Sigma)$ —the increments of X are no longer stationary, but properties (P3) and (P4) continue to hold (of course, the condition $E(\sigma^2) < \infty$ in (P4) is not required under $P(\cdot \mid \mathcal{T}, \Sigma)$).

Remark 3. It follows from (7) that if σ is chosen as a deterministic constant then X_t admits moments of all orders (actually, even exponential moments; cf. Proposition 1 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 taking σ with power-law tails. This, however, is *not* true, and is one of 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 in Remark 4, after having stated some results; see also Subsection 2.4 and Figure 4(b) below for a graphical comparison 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 every choice of the intervals $(a_1, b_1), \dots, (a_k, b_k) \subseteq (0, \infty)$ and of the measurable function $F: \mathbb{R}^k \rightarrow \mathbb{R}$, we have, a.s.,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} F(X_{n\delta+b_1} - X_{n\delta+a_1}, \dots, X_{n\delta+b_k} - X_{n\delta+a_k}) \\ = E(F(X_{b_1} - X_{a_1}, \dots, X_{b_k} - X_{a_k})), \end{aligned} \tag{10}$$

provided the expectation appearing on 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 , which 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 (10) 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 the main results for our model X , which correspond to the basic stylized facts mentioned in the introduction: diffusive scaling and crossover of the log-return distribution (Theorem 1), multiscaling of moments (Theorem 2 and Corollary 1), and clustering of the volatility (Theorem 3 and Corollary 2).

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

- *Small-time diffusive scaling:*

$$\frac{(X_{t+h} - X_t)}{\sqrt{h}} \xrightarrow{D}_{h \downarrow 0} f(x) dx := \text{law of } (\sqrt{2D}\lambda^{1/2-D})\sigma S^{D-1/2}W_1, \tag{11}$$

where $\sigma \sim \nu$, $S \sim \text{Exp}(1)$, 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 $q > 0$,

$$\int |x|^q f(x) dx < +\infty \iff q < q^* := \frac{1}{1/2 - D}. \tag{12}$$

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

$$\begin{aligned} \frac{(X_{t+h} - X_t)}{\sqrt{h}} &\xrightarrow{D}_{h \uparrow \infty} \frac{e^{-x^2/(2c^2)}}{\sqrt{2\pi c}} dx = \mathcal{N}(0, c^2), \\ c^2 &= \lambda^{1-2D} E(\sigma^2)\Gamma(2D + 1), \end{aligned} \tag{13}$$

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

Remark 4. 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. (7), so there are no heavy tails in the strict sense. However, for small h , the heavy-tailed density $f(x)$ is, by (11), an excellent approximation for the true distribution of $(X_{t+h} - X_t)/\sqrt{h}$ 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 (12) follows immediately from (11): in fact, when σ has finite moments of all orders,

$$\int |x|^q f(x) dx < +\infty \iff E(S^{(D-1/2)q}) = \int_0^\infty s^{(D-1/2)q} e^{-s} ds < +\infty,$$

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 0, 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 $q \in (0, \infty)$,

$$m_q(h) := E(|X_{t+h} - X_t|^q), \tag{14}$$

the natural scaling $m_q(h) \approx h^{q/2}$ as $h \downarrow 0$, which one would naively guess from (11), breaks down for $q > q^*$, for which 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, which we prove in Section 4.

Theorem 2. (Multiscaling of moments.) *Let $q > 0$, and assume that $E(\sigma^q) < +\infty$. The quantity $m_q(h)$ in (14) is finite and has the following asymptotic behavior as $h \downarrow 0$:*

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

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

$$C_q := \begin{cases} E(|W_1|^q) E(\sigma^q) \lambda^{q/q^*} (2D)^{q/2} \Gamma\left(1 - \frac{q}{q^*}\right) & \text{if } q < q^*, \\ E(|W_1|^q) E(\sigma^q) \lambda (2D)^{q/2} & \text{if } q = q^*, \\ E(|W_1|^q) E(\sigma^q) \lambda \left[\int_0^\infty ((1+x)^{2D} - x^{2D})^{q/2} dx + \frac{1}{Dq+1} \right] & \text{if } q > q^*, \end{cases} \tag{15}$$

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

Corollary 1. *The following relation holds:*

$$A(q) := \lim_{h \downarrow 0} \frac{\log m_q(h)}{\log h} = \begin{cases} \frac{q}{2} & \text{if } q \leq q^*, \\ Dq + 1 & \text{if } q \geq q^*, \end{cases} \quad \text{where } q^* := \frac{1}{1/2 - D}. \tag{16}$$

The explicit form of the multiplicative constant C_q in (15) 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 the *volatility autocorrelation*. We start by determining the behavior of the covariance.

Theorem 3. Assume that $E(\sigma^2) < \infty$. The following relation holds as $h \downarrow 0$ for all $s, t > 0$:

$$\text{cov}(|X_{s+h} - X_s|, |X_{t+h} - X_t|) = \frac{4D}{\pi} \lambda^{1-2D} e^{-\lambda|t-s|} (\phi(\lambda|t-s|)h + o(h)). \tag{17}$$

Here

$$\phi(x) := \text{cov}(\sigma S^{D-1/2}, \sigma(S+x)^{D-1/2}) \tag{18}$$

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

We recall that $\rho(Y, Z) := \text{cov}(Y, Z) / \sqrt{\text{var}(Y) \text{var}(Z)}$ is the correlation coefficient of two random variables Y and Z . As Theorem 2 yields

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

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

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

$$\begin{aligned} \lim_{h \downarrow 0} \rho(|X_{s+h} - X_s|, |X_{t+h} - X_t|) &= \rho(t-s) \\ &:= \frac{2}{\pi \text{var}(\sigma|W_1|S^{D-1/2})} e^{-\lambda|t-s|} \phi(\lambda|t-s|). \end{aligned} \tag{19}$$

Here $\phi(\cdot)$ is defined in (18) and $\sigma \sim \nu$, $S \sim \text{Exp}(1)$, and $W_1 \sim \mathcal{N}(0, 1)$ 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 (18) we can write

$$\phi(x) = \text{var}(\sigma) E(S^{D-1/2}(S+x)^{D-1/2}) + E(\sigma)^2 \text{cov}(S^{D-1/2}, (S+x)^{D-1/2}), \tag{20}$$

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

$$\begin{aligned} E(S^{D-1/2}(S+x)^{D-1/2}) &\sim c_1 x^{D-1/2}, \\ \text{cov}(S^{D-1/2}, (S+x)^{D-1/2}) &\sim c_2 x^{D-3/2}, \end{aligned} \tag{21}$$

where c_1 and c_2 are positive constants; hence, $\phi(x)$ has a power-law behavior as $x \rightarrow \infty$. 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 Figure 3(b) and (c)).

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 very good agreement. Here we describe in detail the case of the DJIA time series over a period of 75 years: we have considered the DJIA opening prices from 2 January 1935 to 31 December 2009 for a total of 18 849 daily data.

The four real parameters D , λ , $E(\sigma)$, and $E(\sigma^2)$ of our model have been chosen to optimize the fitting of the scaling function $A(q)$ of the moments (see Corollary 1), which depends

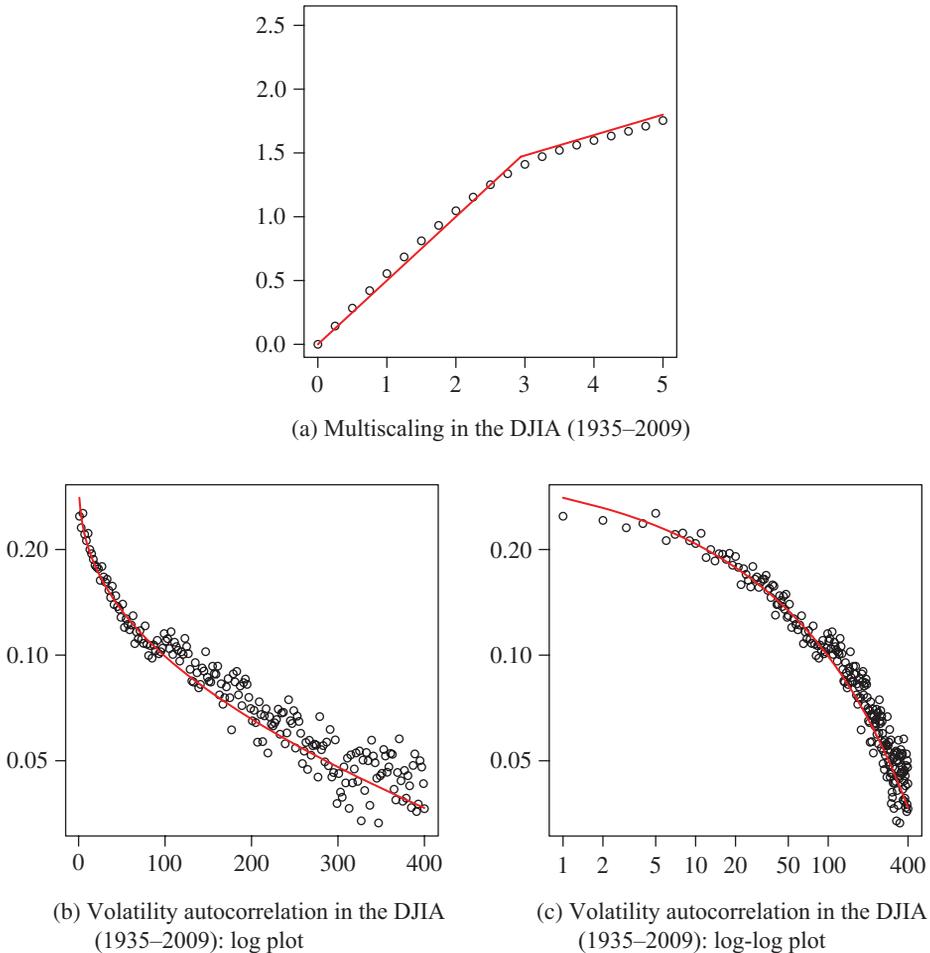


FIGURE 3: Multiscaling of moments and the volatility autocorrelation in the DJIA time series (1935–2009) compared with our model. (a) The DJIA empirical scaling exponent $\hat{A}(q)$ (circles) and the theoretical scaling exponent $A(q)$ (solid line) as a function of q . (b) Log plot for the DJIA empirical one-day volatility autocorrelation $\hat{\rho}_1(t)$ (circles) and the theoretical prediction $\rho(t)$ (solid line), as functions of t (days). For clarity, only the data for even values of t are plotted. (c) Same as (b), but a log-log plot instead of a log plot. For clarity, for $t \geq 50$, only the data for even values of t are plotted.

only on D and the curve $\rho(t)$ of the volatility autocorrelation (see Corollary 2), which depends on D , λ , $E(\sigma)$, and $E(\sigma^2)$ (more details on the parameter estimation are illustrated in Section 7). We have obtained the following numerical estimates:

$$\hat{D} \simeq 0.16, \quad \hat{\lambda} \simeq 0.00097, \quad \widehat{E(\sigma)} \simeq 0.108, \quad \widehat{E(\sigma^2)} \simeq 0.0117 \simeq (\widehat{E(\sigma)})^2. \quad (22)$$

Note that the estimated standard deviation of σ is negligible, so 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 (22).

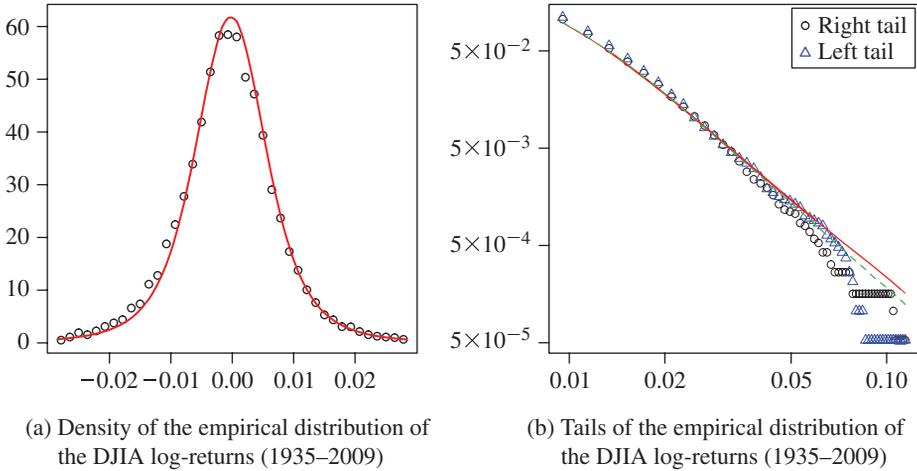


FIGURE 4: Distribution of daily log-returns in the DJIA time series (1935–2009) compared with our model (see Subsection 7.3 for details). (a) The density of the DJIA log-return empirical distribution $\hat{p}_1(\cdot)$ (circles) and the theoretical prediction $p_1(\cdot)$ (solid line). The plot ranges from 0 to about three standard deviations ($\hat{\sigma} \simeq 0.0095$) 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 1 to about twelve standard deviations from the mean. Also plotted is the asymptotic density $f(\cdot)$ (dashed line) defined in (11).

As we show in Figure 3, there is an excellent fitting of the theoretical predictions to the observed data. We find it remarkable that a rather simple mechanism of (relatively rare) volatility 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: in Figure 4 we compare both the density and the integrated tails of the log-return empirical distribution, cf. (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 a structural property of the model.

In Figure 4(b) we have plotted the density of X_1 , represented by the solid line, and the asymptotic limiting density f appearing in (11), represented by the 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; see (11). 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 5. We point out that, even if we had found $\widehat{\text{var}}(\sigma) := \widehat{E}(\sigma^2) - (\widehat{E}(\sigma))^2 > 0$ (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 $1/\hat{\lambda} \simeq 1031$ days; cf. (22). Therefore, in a time period of the length of the DJIA time series we are considering, only $18\,849/1031 \simeq 18$ 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 nonconstant, its first two moments contain the information that is relevant for application to real data.

3. Discussion and further developments

Having stated the main properties of the model, we can discuss in more 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 , and σ , the roles of λ and σ are reduced to *simple scale factors*. In fact, if we change the values 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 and b (depending on λ and σ), as is clear from the definition of $(I_t)_{t \geq 0}$ given in (6). 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 is also clear from the explicit expressions we have derived for the small-time and large-time asymptotic distributions (Theorem 1), multiscaling of moments (Theorem 2), and volatility autocorrelation (Theorem 3).

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 (see (16)) coincides with the power-law tail exponent of the (approximate) log-return distribution for small time (see (12)), and is also linked (through D) to the slow decay of the volatility autocorrelation from short to moderate time (see (20) and (21)). The fact that these quantitative constraints are indeed observed on the DJIA time series (see Figures 3 and 4) is not obvious a priori and is therefore particularly noteworthy.

3.2. On the comparison with multifractal models

As noted in the introduction, the multiscaling of moments is a key feature of multifractal models. These models are random time changes of Brownian motion, $X_t = W_{I_t}$, as does our model, with the important difference that the time-change process $(I_t)_{t \geq 0}$ is rather singular, having nonabsolutely 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 $t > 0$. However, as 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 what is important to note 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 (see also Remark 7 below).

We conclude with a semiheuristic argument, which illustrates how heavy tails and multiscaling arise in our model. On the event $\{(-\tau_0) \leq h, \tau_1 > h\}$ we can write, by (6), $I_h = \sigma_0^2 \{(h - \tau_0)^{2D} - (-\tau_0)^{2D}\} \gtrsim h^{2D}$ and, therefore, $|X_h| = |W_{I_h}| \sim \sqrt{I_h} |W_1| \gtrsim h^D$. Consequently, we obtain the bound

$$P(|X_h| \gtrsim h^D) \geq P((-\tau_0) \leq h, \tau_1 > h) \gtrsim h, \quad (23)$$

which allows us to draw a couple of interesting consequences.

- Relation (23) 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 (23) can be rewritten as $P(|X_h|/\sqrt{h} \gtrsim t) \gtrsim t^{-q^*}$, where $t = h^{-(1/2-D)}$ and $q^* = (\frac{1}{2} - D)^{-1}$. Since $t \rightarrow +\infty$ 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 $h \downarrow 0$ (see (11) and (12)) with the *correct* tail exponent q^* .

3.3. On the stochastic volatility model representation

We recall that our process $(X_t)_{t \geq 0}$ can be written as a stochastic volatility model $dX_t = v_t dB_t$; cf. (8). It is interesting to note that the squared volatility $(v_t)^2$ is the stationary solution of the stochastic differential equation

$$d(v_t^2) = -\alpha_t (v_t^2)^\gamma dt + \infty di(t), \tag{24}$$

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

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

We stress that $(v_t)^2$ is a *pathwise* solution of (24), 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 $-\alpha_t (v_t^2)^\gamma$, the solution starting from infinity becomes instantaneously finite (note that the ordinary differential equation $dx(t) = -\alpha x(t)^\gamma dt$ with $x(0) = \infty$ has the explicit solution $x(t) = ct^{-1/(\gamma-1)}$, where $c = c(\alpha, \gamma) = (\alpha(\gamma - 1))^{-1/(\gamma-1)}$).

The volatility representation (24) is also useful to understand the limitations of our model and to design possible generalizations. For instance, according to (24), the volatility has the rather unrealistic feature of being deterministic between jumps. This limitation could be weakened in various ways, e.g. by replacing $i(t)$ in (24) with a more general Lévy 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 (24), 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 (8) and (24), to be

$$dX_t = v_t dB_t - \beta di(t), \quad dv_t^2 = -\alpha_t(v_t^2)^\gamma dt + \infty di(t),$$

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 and \mathcal{T}^Y , has been investigated by Pigato [28]. The model has been numerically calibrated on the joint time series of the DJIA and FTSE 100 indexes, finding in particular 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 Bonino [11], which deals with portfolio optimization problems in the framework of our model.

4. Scaling and multiscaling: proofs of Theorems 1 and 2

We observe that, for all fixed $t, h > 0$, we have the equality in law $X_{t+h} - X_t \sim \sqrt{I_h}W_1$, which follows from the definition of our model, $(X_t)_{t \geq 0} = (W_t)_{t \geq 0}$. We also observe that $i(h) = \#\{\mathcal{T} \cap (0, h]\} \sim \text{Po}(\lambda h)$, which follows from (5) and the properties of the Poisson process.

4.1. Proof of Theorem 1

Since $P(i(h) \geq 1) = 1 - e^{-\lambda h} \rightarrow 0$ 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 $-\tau_0 \sim \text{Exp}(\lambda)$. In particular,

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

Since $X_{t+h} - X_t \sim \sqrt{I_h}W_1$, the convergence in distribution (11) follows:

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

Next we focus on the case $h \uparrow \infty$. Under the assumption that $E(\sigma^2) < \infty$, 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,

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

Clearly, $\lim_{h \rightarrow +\infty} i(h)/h = \lambda$ a.s., by the strong law of large numbers applied to the random variables $\{\tau_k\}_{k \geq 1}$. Recalling (6), it easily follows that

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

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

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

which coincides with (13).

4.2. Proof of Theorem 2

Since $X_{t+h} - X_t \sim \sqrt{I_h}W_1$, we can write

$$E(|X_{t+h} - X_t|^q) = E(|I_h|^{q/2}|W_1|^q) = E(|W_1|^q)E(|I_h|^{q/2}) = c_q E(|I_h|^{q/2}), \tag{25}$$

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

$$E(|I_h|^{q/2}) = E(|I_h|^{q/2} \mathbf{1}_{\{i(h)=0\}}) + E(|I_h|^{q/2} \mathbf{1}_{\{i(h)=1\}}) + E(|I_h|^{q/2} \mathbf{1}_{\{i(h)\geq 2\}}). \tag{26}$$

For the first term on the right-hand side of (26), we note that $P(i(h) = 0) = e^{-\lambda h} \rightarrow 1$ 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 $S \sim \text{Exp}(1)$, we obtain, as $h \downarrow 0$,

$$E(|I_h|^{q/2} \mathbf{1}_{\{i(h)=0\}}) = E(\sigma^q)\lambda^{-Dq} E(((S + \lambda h)^{2D} - S^{2D})^{q/2})(1 + o(1)). \tag{27}$$

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

$$q \gtrless q^* \iff \frac{q}{2} \gtrless Dq + 1 \iff -1 \gtrless \left(D - \frac{1}{2}\right)q.$$

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

$$\lim_{h \downarrow 0} \frac{E(((S + \lambda h)^{2D} - S^{2D})^{q/2})}{\lambda^{q/2}h^{q/2}} = (2D)^{q/2}\Gamma\left(1 - \frac{q}{q^*}\right) \in (0, \infty). \tag{28}$$

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

$$\begin{aligned} E(((S + \lambda h)^{2D} - S^{2D})^{q/2}) &= \int_0^\infty ((s + \lambda h)^{2D} - s^{2D})^{q/2} e^{-s} ds \\ &= (\lambda h)^{Dq+1} \int_0^\infty ((1 + x)^{2D} - x^{2D})^{q/2} e^{-\lambda hx} dx. \end{aligned} \tag{29}$$

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

$$\lim_{h \downarrow 0} \frac{E(((S + \lambda h)^{2D} - S^{2D})^{q/2})}{\lambda^{Dq+1}h^{Dq+1}} = \int_0^\infty ((1 + x)^{2D} - x^{2D})^{q/2} dx \in (0, \infty). \tag{30}$$

Finally, in the case $q = q^*$ we have $((1 + x)^{2D} - x^{2D})^{q^*/2} \sim (2D)^{q^*/2}x^{-1}$ as $x \rightarrow +\infty$ and we want to study the integral in the second line of (29). Fix an arbitrary (large) $M > 0$, and

note that, integrating by parts and performing a change of variable, as $h \downarrow 0$, we have

$$\begin{aligned} \int_M^\infty \frac{e^{-\lambda hx}}{x} dx &= -\log M e^{-\lambda hM} + \lambda h \int_M^\infty (\log x) e^{-\lambda hx} dx \\ &= O(1) + \int_{\lambda hM}^\infty \log\left(\frac{y}{\lambda h}\right) e^{-y} dy \\ &= O(1) + \int_{\lambda hM}^\infty \log\left(\frac{y}{\lambda}\right) e^{-y} dy + \log\left(\frac{1}{h}\right) \int_{\lambda hM}^\infty e^{-y} dy \\ &= \log\left(\frac{1}{h}\right)(1 + o(1)). \end{aligned}$$

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

$$\int_0^\infty ((1+x)^{2D} - x^{2D})^{q^*/2} e^{-\lambda hx} dx \sim (2D)^{q^*/2} \log\left(\frac{1}{h}\right).$$

From (29), noting that $Dq + 1 = q/2$ for $q = q^*$, it follows that

$$\lim_{h \downarrow 0} \frac{E(((S+h)^{2D} - S^{2D})^{q^*/2})}{\lambda^{Dq^*+1} h^{q^*/2} \log(1/h)} = (2D)^{q^*/2}. \tag{31}$$

Recalling (25) and (27), relations (28), (30), and (31) show that the first term on the right-hand side of (26) 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 on the right-hand side of (26). Note that, conditionally on the event $\{i(h) = 1\} = \{\tau_1 \leq h, \tau_2 > h\}$, we have

$$\begin{aligned} I_h &= \sigma_1^2 (h - \tau_1)^{2D} + \sigma_0^2 ((\tau_1 - \tau_0)^{2D} - (-\tau_0)^{2D}) \\ &\sim \sigma_1^2 (h - hU)^{2D} + \sigma_0^2 \left(\left(hU + \frac{S}{\lambda} \right)^{2D} - \left(\frac{S}{\lambda} \right)^{2D} \right), \end{aligned}$$

where $S \sim \text{Exp}(1)$ 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

$$E(|I_h|^{q/2} \mathbf{1}_{\{i(h)=1\}}) = \lambda h^{Dq+1} E\left(\left(\sigma_1^2(1-U)^{2D} + \sigma_0^2\left(\left(U + \frac{S}{\lambda h}\right)^{2D} - \left(\frac{S}{\lambda h}\right)^{2D}\right)^{q/2}\right)\right).$$

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

$$\lim_{h \downarrow 0} \frac{E(|I_h|^{q/2} \mathbf{1}_{\{i(h)=1\}})}{h^{Dq+1}} = \lambda E(\sigma_1^q) E((1-U)^{Dq}) = \lambda E(\sigma_1^q) \frac{1}{Dq+1}. \tag{32}$$

This shows that the second term on the right-hand side of (26) gives a contribution of the order h^{Dq+1} as $h \downarrow 0$. This is relevant only for $q > q^*$ because, for $q \leq q^*$, the first term gives a much bigger contribution of the order $h^{q/2}$ (see (28) and (31)). Recalling (25), it follows from (32) and (30) that the contribution of the first and the second terms on the right-hand side of (26) matches the statement of the theorem (including the constant).

It only remains to show that the third term on the right-hand side of (26) gives a negligible contribution. We begin by deriving a simple upper bound for I_h . Since $(a + b)^{2D} - b^{2D} \leq a^{2D}$ 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

$$\begin{aligned}
 I_h &= \sigma_{i(h)}^2 (h - \tau_{i(h)})^{2D} + \sum_{k=2}^{i(h)} \sigma_{k-1}^2 (\tau_k - \tau_{k-1})^{2D} + \sigma_0^2 [(\tau_1 - \tau_0)^{2D} - (-\tau_0)^{2D}] \\
 &\leq \sigma_{i(h)}^2 (h - \tau_{i(h)})^{2D} + \sum_{k=2}^{i(h)} \sigma_{k-1}^2 (\tau_k - \tau_{k-1})^{2D} + \sigma_0^2 \tau_1^{2D}, \tag{33}
 \end{aligned}$$

with the convention that the sum over k is 0 if $i(h) = 1$. Since $\tau_k \leq h$ for all $k \leq i(h)$, by the definition of $i(h)$ given in (5), relation (33) yields the bound $I_h \leq h^{2D} \sum_{k=0}^{i(h)} \sigma_k^2$, which clearly also holds when $i(h) = 0$. In conclusion, we have shown that, for all $h, q > 0$,

$$|I_h|^{q/2} \leq h^{Dq} \left(\sum_{k=0}^{i(h)} \sigma_k^2 \right)^{q/2}. \tag{34}$$

Consider first the case $q > 2$. By Jensen’s inequality we have

$$\left(\sum_{k=0}^{i(h)} \sigma_k^2 \right)^{q/2} = (i(h) + 1)^{q/2} \left(\frac{1}{i(h) + 1} \sum_{k=0}^{i(h)} \sigma_k^2 \right)^{q/2} \leq (i(h) + 1)^{q/2-1} \sum_{k=0}^{i(h)} \sigma_k^q. \tag{35}$$

By (34) and (35), we obtain

$$E(|I_h|^{q/2} \mathbf{1}_{\{i(h) \geq 2\}}) \leq h^{Dq} E(\sigma^q) E((i(h) + 1)^{q/2} \mathbf{1}_{\{i(h) \geq 2\}}). \tag{36}$$

A corresponding inequality for $q \leq 2$ is derived from (34): since $(\sum_{k=1}^\infty x_k)^{q/2} \leq \sum_{k=1}^\infty x_k^{q/2}$ for every nonnegative sequence $(x_n)_{n \in \mathbb{N}}$, we obtain

$$E(|I_h|^{q/2} \mathbf{1}_{\{i(h) \geq 2\}}) \leq h^{Dq} E\left(\sum_{k=0}^{i(h)} \sigma_k^q \mathbf{1}_{\{i(h) \geq 2\}} \right) \leq h^{Dq} E(\sigma^q) E((i(h) + 1) \mathbf{1}_{\{i(h) \geq 2\}}). \tag{37}$$

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

$$\begin{aligned}
 E((i(h) + 1)^a \mathbf{1}_{\{i(h) \geq 2\}}) &\leq E((i(h) + 1)^{3a})^{1/3} P(i(h) \geq 2)^{2/3} \\
 &\leq E((i(1) + 1)^{3a})^{1/3} (1 - e^{-\lambda h} - e^{-\lambda h} \lambda h)^{2/3} \\
 &\leq (\text{constant}) h^{4/3}, \tag{38}
 \end{aligned}$$

because $E((i(1) + 1)^{3a}) < \infty$ (recall that $i(h) \sim \text{Po}(\lambda)$) 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 (36), (37), and (38) that

$$E(|I_h|^{q/2} \mathbf{1}_{\{i(h) \geq 2\}}) \leq (\text{constant}^1) h^{Dq+4/3}.$$

This shows that the contribution of the third term on the right-hand side of (26) is always negligible with respect to the contribution of the second term (recall (32)).

5. Decay of correlations: proof of Theorem 3

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 and 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 and I' , the σ -algebras \mathcal{G}_I and $\mathcal{G}_{I'}$ are independent. We write $\mathcal{G} := \mathcal{G}_{\mathbb{R}}$ for short, 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 (17). Clearly, by translation invariance we can set $s = 0$ without loss of generality. We also assume that $h < t$. We start by writing

$$\begin{aligned} & \text{cov}(|X_h|, |X_{t+h} - X_t|) \\ &= \text{cov}(\mathbb{E}(|X_h| \mid \mathcal{G}), \mathbb{E}(|X_{t+h} - X_t| \mid \mathcal{G})) + \mathbb{E}(\text{cov}(|X_h|, |X_{t+h} - X_t| \mid \mathcal{G})). \end{aligned} \tag{39}$$

We recall that $X_t = W_{I_t}$, and that 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 on the right-hand side of (39) vanishes, because $\text{cov}(|X_h|, |X_{t+h} - X_t| \mid \mathcal{G}) = 0$ a.s. For fixed h , from the equality in law $X_h = W_{I_h} \sim \sqrt{I_h}W_1$, it follows that $\mathbb{E}(|X_h| \mid \mathcal{G}) = c_1 \sqrt{I_h}$, where $c_1 = \mathbb{E}(|W_1|) = \sqrt{2/\pi}$. Analogously $\mathbb{E}(|X_{t+h} - X_t| \mid \mathcal{G}) = \sqrt{2/\pi} \sqrt{I_{t+h} - I_t}$ and (39) reduces to

$$\text{cov}(|X_h|, |X_{t+h} - X_t|) = \frac{2}{\pi} \text{cov}(\sqrt{I_h}, \sqrt{I_{t+h} - I_t}). \tag{40}$$

Recall the definitions of the variables $i(t)$ and I_t given in (5) and (6). 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 (40). In fact, from (6) we can write

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

with the convention that the sum on the right-hand side is 0 if $i(t+h) = i(t)$. This shows that $(I_{t+h} - I_t)$ is a function of the variables τ_k and σ_k with index $i(t) \leq k \leq i(t+h)$. Since $\{\mathcal{T} \cap (h,t] \neq \emptyset\} = \{\tau_{i(t)} > h\}$, then $\sqrt{I_{t+h} - I_t} \mathbf{1}_{\{\mathcal{T} \cap (h,t] \neq \emptyset\}}$ is $\mathcal{G}_{(h,t+h]}$ -measurable and, hence, independent of $\sqrt{I_h}$, which is clearly $\mathcal{G}_{(-\infty,h]}$ -measurable. This shows that $\text{cov}(\sqrt{I_h}, \sqrt{I_{t+h} - I_t} \mathbf{1}_{\{\mathcal{T} \cap (h,t] \neq \emptyset\}}) = 0$; therefore, from (40) we can write

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

Now we decompose this last covariance as follows:

$$\begin{aligned} & \text{cov}(\sqrt{I_h}, \sqrt{I_{t+h} - I_t} \mathbf{1}_{\{\mathcal{T} \cap (h,t] = \emptyset\}}) \\ &= \mathbb{E}((\sqrt{I_h} - \mathbb{E}(\sqrt{I_h})) \sqrt{I_{t+h} - I_t} \mathbf{1}_{\{\mathcal{T} \cap (h,t] = \emptyset\}}) \\ &= \mathbb{E}((\sqrt{I_h} - \mathbb{E}(\sqrt{I_h})) \sqrt{I_{t+h} - I_t} \mathbf{1}_{\{\mathcal{T} \cap (0,t+h] = \emptyset\}}) \\ & \quad + \mathbb{E}((\sqrt{I_h} - \mathbb{E}(\sqrt{I_h})) \sqrt{I_{t+h} - I_t} \mathbf{1}_{\{\mathcal{T} \cap (h,t] = \emptyset\}} \mathbf{1}_{\{\mathcal{T} \cap ([0,h] \cup (t,t+h]) \neq \emptyset\}}). \end{aligned} \tag{41}$$

We deal with the two terms on the right-hand side of (41) separately. The first gives the dominant contribution. To see this, observe that, on $\{\mathcal{T} \cap (0, t+h] = \emptyset\}$,

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

and

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

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

$$\begin{aligned} & E[(\sqrt{I_h} - E(\sqrt{I_h}))\sqrt{I_{t+h} - I_t} \mathbf{1}_{\{\mathcal{T} \cap (0, t+h) = \emptyset\}}] \\ &= E((\sigma_0\sqrt{(h - \tau_0)^{2D} - (-\tau_0)^{2D}} - E(\sqrt{I_h})) \\ &\quad \times \sigma_0\sqrt{(t + h - \tau_0)^{2D} - (t - \tau_0)^{2D}} \mathbf{1}_{\{\mathcal{T} \cap (0, t+h) = \emptyset\}}) \\ &= e^{-\lambda(t+h)} E((\sigma_0\sqrt{(h - \tau_0)^{2D} - (-\tau_0)^{2D}} - E(\sqrt{I_h})) \\ &\quad \times \sigma_0\sqrt{(t + h - \tau_0)^{2D} - (t - \tau_0)^{2D}}) \\ &= e^{-\lambda(t+h)} \{ \text{cov}(\sigma_0\sqrt{(h - \tau_0)^{2D} - (-\tau_0)^{2D}}, \sigma_0\sqrt{(t + h - \tau_0)^{2D} - (t - \tau_0)^{2D}}) \\ &\quad + [E(\sigma_0\sqrt{(h - \tau_0)^{2D} - (-\tau_0)^{2D}}) - E(\sqrt{I_h})] \\ &\quad \times E(\sigma_0\sqrt{(t + h - \tau_0)^{2D} - (t - \tau_0)^{2D}}) \}. \end{aligned} \tag{42}$$

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

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{1}{h} \text{cov}(\sigma_0\sqrt{(h - \tau_0)^{2D} - (-\tau_0)^{2D}}, \sigma_0\sqrt{(t + h - \tau_0)^{2D} - (t - \tau_0)^{2D}}) \\ &= 2D \text{cov}(\sigma_0(-\tau_0)^{D-1/2}, \sigma_0(t - \tau_0)^{D-1/2}) \\ &= 2D\lambda^{1-2D} \text{cov}(\sigma_0 S^{D-1/2}, \sigma_0(\lambda t + S)^{D-1/2}) \\ &= 2D\lambda^{1-2D} \phi(\lambda t), \end{aligned} \tag{43}$$

with $S := \lambda(-\tau_0) \sim \text{Exp}(1)$ and ϕ defined as in (18). Similarly,

$$\lim_{h \rightarrow 0} \frac{1}{\sqrt{h}} E(\sigma_0\sqrt{(t + h - \tau_0)^{2D} - (t - \tau_0)^{2D}}) = \sqrt{2D} E(\sigma_0(t - \tau_0)^{D-1/2}) < +\infty. \tag{44}$$

Therefore, if we show that

$$\lim_{h \rightarrow 0} E\left(\sqrt{\frac{I_h}{h}}\right) = \sqrt{2D} E(\sigma_0(-\tau_0)^{D-1/2}) \tag{45}$$

using (42), (43), and (44), we have

$$\lim_{h \rightarrow 0} \frac{1}{h} E[(\sqrt{I_h} - E(\sqrt{I_h}))\sqrt{I_{t+h} - I_t} \mathbf{1}_{\{\mathcal{T} \cap (0, t+h) = \emptyset\}}] = 2D\lambda^{1-2D} \phi(\lambda t). \tag{46}$$

To complete the proof of (46), it remains to show (45). But, this is a nearly immediate consequence of Theorem 2: indeed, using (15) and the fact that $q^* > 1$,

$$\begin{aligned} E(\sqrt{I_h}) &= \frac{1}{E\sqrt{|W_1|}} E(|X_h|) \\ &= \frac{C_1}{E|W_1|} \sqrt{h} + o(\sqrt{h}) \\ &= \sqrt{2D} E(\sigma_0(-\tau_0)^{D-1/2}) \sqrt{h} + o(\sqrt{h}). \end{aligned}$$

The proof is now completed by showing that the second term in (41) is negligible, i.e. $o(h)$. By the Cauchy–Schwarz inequality and the simple fact that $(\sqrt{I_h} - E(\sqrt{I_h}))^2 \leq I_h + E(I_h)$,

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

By Theorem 2, $E(I_h^2)$ 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 $(E(I_h^2))^{1/2} \sqrt{2\lambda h} = o(h)$, and the proof is completed.

6. Basic properties of the model

In this section we start by proving properties (P1)–(P4) stated in Section 2.2. We then provide some connections between the tails of σ and those of X_t , also beyond the equivalence stated in (7). Finally, we establish a *mixing* property that yields relation (10). One of the proofs is postponed to Appendix A.

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 (P1). We first focus on the process $(I_t)_{t \geq 0}$ defined in (6). 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 $\tau_{i^h(t)}^h$ be the largest point of \mathcal{T}^h smaller than t , i.e. $i^h(t) = i(t+h)$. Recalling definition (6), we can write

$$I_{t+h} - I_h = \sigma_{i^h(t)}^2 (t - \tau_{i^h(t)}^h)^{2D} + \sum_{k=i^h(0)+1}^{i^h(t)} \sigma_{k-1}^2 (\tau_k^h - \tau_{k-1}^h)^{2D} - \sigma_{i^h(0)}^2 (-\tau_{i^h(0)}^h)^{2D},$$

with the convention that the sum on the right-hand side is 0 if $i^h(t) = i^h(0)$. This relation shows that $(I_{t+h} - I_h)_{t \geq 0}$ and $(I_t)_{t \geq 0}$ can be expressed as the *same* function applied to the two random sets \mathcal{T}^h and \mathcal{T} , respectively. 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 $X_t = W_t$ and the fact that Brownian motion has independent, stationary increments, it follows that, for every Borel subset $A \subseteq \mathbb{R}^{[0, +\infty)}$,

$$\mathbb{P}(X_{h+} - X_h \in A) = E(\mathbb{P}(W_{I_{h+}} - W_{I_h} \in A \mid \mathcal{G})) = \mathbb{P}(W_I \in A) = \mathbb{P}(X \in A),$$

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 (P2). Note that $E(|X_t|^q) = E(|W_t|^q) = E(|I_t|^{q/2}) E(|W_1|^q)$, by the independence of W and I and the scaling properties of Brownian motion. It therefore remains to show that

$$E(|I_t|^{q/2}) < \infty \iff E(\sigma^q) < \infty.$$

The implication ‘ \Rightarrow ’ is easy: by the definition of the process I given in (6) 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 $E(|I_t|^{q/2}) = \infty$ also.

The implication ‘ \Leftarrow ’ follows immediately from bounds (36) and (37), which also hold without the indicator $\mathbf{1}_{\{i(h) \geq 2\}}$.

Proof of property (P3). Observe first that $I'_s := dI_s/ds > 0$ a.s. and for Lebesgue almost everywhere $s \geq 0$. By a change of variable, we can rewrite the process $(B_t)_{t \geq 0}$ defined in (9) as

$$B_t = \int_0^{I_t} \frac{1}{\sqrt{I'(I^{-1}(u))}} dW_u = \int_0^t \frac{1}{\sqrt{I'_s}} dW_{I_s} = \int_0^t \frac{1}{\sqrt{I'_s}} dX_s,$$

which shows that relation (8) holds. It remains to show that $(B_t)_{t \geq 0}$ is indeed a standard Brownian motion. Note that

$$B_t = \int_0^{I_t} \sqrt{(I^{-1})'(u)} dW_u.$$

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

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

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 (P4). The assumption that $E(\sigma^2) < \infty$ 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, such as the DJIA, the ‘observable tails’ of X_t are quite insensitive to the details of the distribution of σ ; see Remarks 3 and 4.

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

$$E(\exp(\gamma|X_t|^q)) = \infty \text{ for all } t > 0 \text{ and } \gamma > 0. \tag{47}$$

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

$$E(\exp(\gamma|X_t|^q)) < \infty \text{ for all } \gamma > 0 \iff E(\exp(\alpha\sigma^{2q/(2-q)})) < \infty \text{ for all } \alpha > 0, \tag{48}$$

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, for $D < \frac{1}{2}$, the distribution of X_t always has 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

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

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

Proposition 2. *Let $I = [a, b)$ and $J = [c, d)$ with $0 \leq a < b \leq c < d$. Then, for every $A \in \mathcal{F}_I^{\mathcal{D}}$ and $B \in \mathcal{F}_J^{\mathcal{D}}$,*

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq e^{-\lambda(c-b)}. \tag{49}$$

As a consequence, (10) holds a.s. and in L^1 for every measurable function $F : \mathbb{R}^k \rightarrow \mathbb{R}$ such that $\mathbb{E}(|F(X_{b_1} - X_{a_1}, \dots, X_{b_k} - X_{a_k})|) < +\infty$.

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 $A \in \mathcal{F}_I^{\mathcal{D}}$ and $B \in \mathcal{F}_J^{\mathcal{D}}$, we have

$$\mathbb{P}(A \cap B \cap \Gamma) = \mathbb{P}(A)\mathbb{P}(B \cap \Gamma). \tag{50}$$

The key to seeing this 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,

$$\mathbb{P}(A \cap B \mid \mathcal{G}) = \mathbb{P}(A \mid \mathcal{G})\mathbb{P}(B \mid \mathcal{G}) \quad \text{a.s.}$$

- Conditionally on \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, $\mathbb{P}(B \mid \mathcal{G})$ is measurable with respect to this σ -field. The same holds 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

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

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

Thus, we have

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

where the two remarks above have been used. Thus, (50) is established. Finally,

$$\begin{aligned} &|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \\ &= |\mathbb{P}(A \cap B \cap \Gamma) + \mathbb{P}(A \cap B \cap \Gamma^c) - \mathbb{P}(A)\mathbb{P}(B \cap \Gamma) - \mathbb{P}(A)\mathbb{P}(B \cap \Gamma^c)| \\ &= |\mathbb{P}(A \cap B \cap \Gamma^c) - \mathbb{P}(A)\mathbb{P}(B \cap \Gamma^c)| \end{aligned}$$

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

We finally show that (10) holds a.s. and in L^1 for every measurable function $F: \mathbb{R}^k \rightarrow \mathbb{R}$ such that $\mathbb{E}(|F(X_{b_1} - X_{a_1}, \dots, X_{b_k} - X_{a_k})|) < +\infty$. Consider the \mathbb{R}^k -valued stochastic process $\xi = (\xi_n)_{n \in \mathbb{N}}$ defined by

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

for fixed $\delta > 0$, $k \in \mathbb{N}$, and $(a_1, b_1), \dots, (a_k, b_k) \subseteq (0, \infty)$. The process ξ is stationary, because we have proven in Section 6 that X has stationary increments. Moreover, inequality (49) implies that ξ is *mixing*, and, therefore, ergodic (see, e.g. [30, Chapter 5, Section 2, Definition 4 and Theorem 2]). The existence of the limit in (10), both a.s. and in L^1 , is then a consequence of the classical ergodic theorem (see, e.g. [30, Chapter 5, Section 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 (see <http://www.R-project.org>). The code we have used is publicly available online at <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*; see Corollary 1.
- (b) The *volatility autocorrelation decay*; see Corollary 2.

Roughly speaking, the idea is to *empirically* compute 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. (10).

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

- the multiscaling coefficients C_1 and C_2 (see (15));
- the multiscaling exponent $A(q)$ (see (16));
- the volatility autocorrelation function $\rho(t)$ (see (19)).

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 (51) 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 $E(\sigma^2) \geq (E(\sigma))^2$ does not play 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 found $\widehat{\text{var}}(\sigma) := E(\widehat{\sigma^2}) - (\widehat{E(\sigma)})^2 > 0$ 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 5.

7.2. Estimation of the parameters $D, \lambda, E(\sigma)$, and $E(\sigma^2)$

Let us fix some notation. The DJIA time series will be denoted by $(s_i)_{0 \leq i \leq N}$ (where $N = 18\,848$) and the corresponding detrended log-DJIA time series will be denoted by $(x_i)_{0 \leq i \leq N}$, where

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

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

The theoretical scaling exponent $A(q)$ is defined in (16), while the multiscaling constants C_1 and C_2 are given in (15) 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

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

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 2, the relation $\log \widehat{m}_q(h) \sim A(q)(\log h) + \log(C_q)$ should hold for small h . By plotting $\log \widehat{m}_q(h)$ versus $\log h$ we indeed find 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, which we call $\widehat{A}(q)$ and \widehat{C}_q .

As for the theoretical volatility autocorrelation, Corollary 2 and the stationarity of the increments of our process $(X_t)_{t \geq 0}$ yield

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

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

$$\phi(x) = \text{var}(\sigma) E(S^{D-1/2} (S + x)^{D-1/2}) + E(\sigma)^2 \text{cov}(S^{D-1/2}, (S + x)^{D-1/2});$$

cf. (20). 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

$$\begin{aligned} \mathcal{L}(D, \lambda, E(\sigma), E(\sigma^2)) = & \frac{1}{2} \left\{ \left(\frac{\widehat{C}_1}{C_1} - 1 \right)^2 + \left(\frac{\widehat{C}_2}{C_2} - 1 \right)^2 \right\} + \frac{1}{20} \sum_{k=1}^{20} \left(\frac{\widehat{A}(k/4)}{A(k/4)} - 1 \right)^2 \\ & + \sum_{n=1}^{400} \frac{e^{-n/T}}{\left(\sum_{m=1}^{400} e^{-m/T} \right)} \left(\frac{\widehat{\rho}_1(n)}{\rho(n)} - 1 \right)^2, \end{aligned} \tag{51}$$

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,

$$\begin{aligned} (\widehat{D}, \widehat{\lambda}, \widehat{E}(\sigma), \widehat{E}(\sigma^2)) := & \arg \min \{ \mathcal{L}(D, \lambda, E(\sigma), E(\sigma^2)) \}, \\ & D \in (0, 1/2], \lambda, E(\sigma), E(\sigma^2) \in (0, \infty) \\ & \text{such that } E(\sigma^2) \geq (E(\sigma))^2 \end{aligned}$$

where the constraint $E(\sigma^2) \geq (E(\sigma))^2$ is due to $\text{var}(\sigma) = E(\sigma^2) - (E(\sigma))^2 \geq 0$. 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).

The numerical study of the functional \mathcal{L} , which appears to be quite regular, was carried out using MATHEMATICA[®], yielding the estimates for the parameters given in (22), i.e.

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

7.3. Graphical comparison

Having found that $\widehat{E}(\sigma^2) \simeq (\widehat{E}(\sigma))^2$, the estimated variance of σ is equal to 0, 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 for the DJIA time series. The graphical results have already been described in Section 2.4 and show very good agreement; see 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 $(x_{i+t} - x_i)_{0 \leq i \leq N-t}$:

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

In Figure 4(a) we 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 $[-3\hat{\delta}, +3\hat{\delta}]$, where $\hat{\delta} \simeq 0.0095$ is the standard deviation of $\widehat{p}_1(\cdot)$ (i.e. the empirical standard deviation of the daily log-returns evaluated for the DJIA time series). In Figure 4(b) we 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

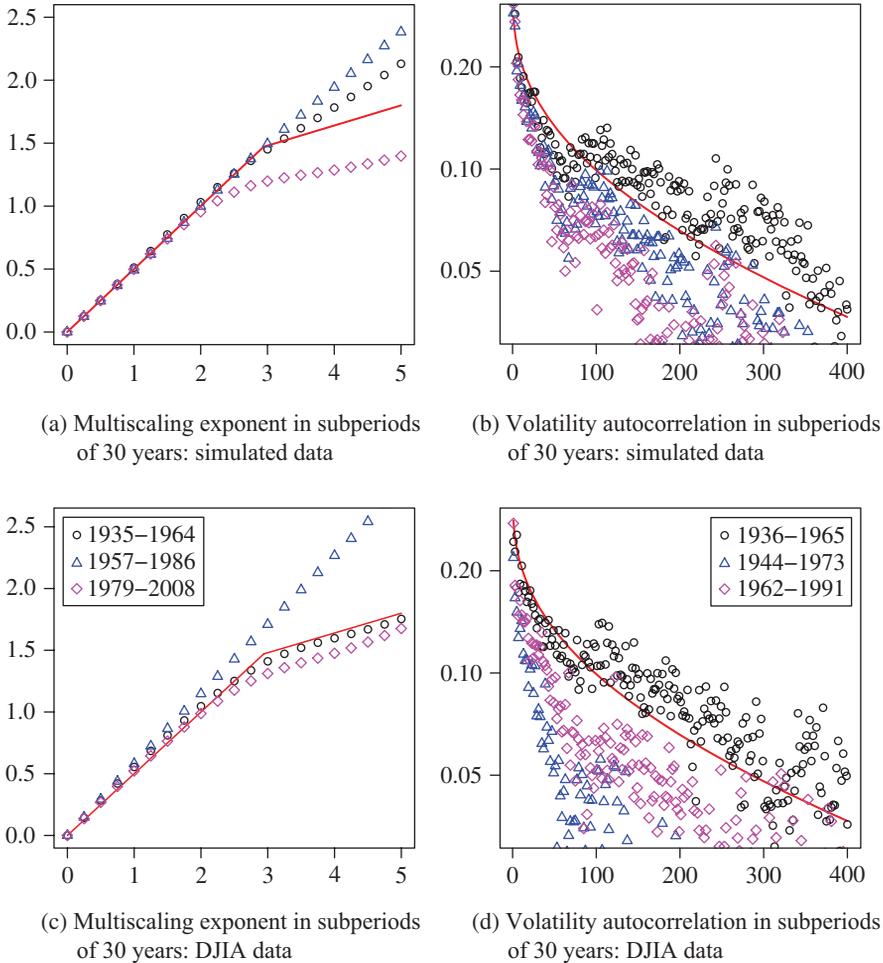


FIGURE 5: Variability of estimators in subperiods of 30 years. Empirical evaluation of the observables $\hat{A}(q)$ and $\hat{\rho}_1(t)$ in subperiods of 30 years for a 75-year-long time series, sampled from our model $(X_t)_{t \geq 0}$ (a) and (b) and from the DJIA time series ((c) and (d)).

and left empirical tails $\hat{R}(z)$ and $\hat{L}(z)$ of $\hat{p}_1(\cdot)$, defined for $z \geq 0$ by

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

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 5, 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 not be very close to their ergodic limit or, in other words, estimators should have nonnegligible 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*, which we have chosen to be 30 years. Figure 5(a) and (b) indeed

show a considerable variability of the values of the estimators for the multiscaling exponent and the volatility autocorrelations, when computed in different subperiods. We repeated the same computations on the DJIA time series (see Figure 5(c) and (d)), observing a similar variability. We regard this as a significant test for this model.

Remark 6. 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 7. The multiscaling of empirical moments has been observed in several financial indexes in [18], where it is claimed that data provide solid arguments *against* models with linear or piecewise-linear scaling exponents. Note that the theoretical scaling exponent $A(q)$ of our model is indeed piecewise linear; see (16). 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; see Figure 5(c). This shows that the objection against models with piecewise linear $A(q)$, raised in [18], cannot apply to the model we have proposed.

Appendix A. Proof of Proposition 1

We first need two simple technical lemmas.

Lemma 1. For $0 < q < 2$, consider the function $\varphi_q : [0, +\infty) \rightarrow [0, +\infty)$ defined by

$$\varphi_q(\beta) := \int_{-\infty}^{+\infty} \exp\left[\beta|x|^q - \frac{x^2}{2}\right] dx.$$

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

$$C_1 \exp[C_1\beta^{2/(2-q)}] \leq \varphi_q(\beta) \leq C_2 \exp[C_2\beta^{2/(2-q)}]. \tag{52}$$

Proof. We begin by observing that it is enough to establish the bounds in (52) for large enough β . Consider the function of positive real variable $f(r) := e^{\beta r^q - r^2/2}$. It is easily checked that f is increasing for $0 \leq r \leq (\beta q)^{1/(2-q)}$. Thus,

$$\begin{aligned} \varphi_q(\beta) &\geq \int_{(\beta q)^{1/(2-q)/2}}^{(\beta q)^{1/(2-q)}} f(r) dr \\ &\geq \frac{1}{2}(\beta q)^{1/(2-q)} f\left(\frac{1}{2}(\beta q)^{1/(2-q)}\right) \\ &= \frac{1}{2}(\beta q)^{1/(2-q)} \exp[c(q)\beta^{2/(2-q)}], \end{aligned}$$

with

$$c(q) := \frac{1}{2q}q^{q/(2-q)} - \frac{1}{8}q^{2/(2-q)} > 0.$$

The lower bound in (52) easily follows for large β .

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

$$\begin{aligned} \varphi_q(\beta) &\leq \int_{|x| \leq (4\beta)^{1/(2-q)}} f(|x|) dx + \int_{|x| > (4\beta)^{1/(2-q)}} e^{-x^2/4} dx \\ &\leq 2(4\beta)^{1/(2-q)} \|f\|_\infty + \int_{-\infty}^{+\infty} e^{-x^2/4} dx. \end{aligned}$$

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

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

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

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

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

see [33] for a proof.

Proof of Proposition 1. Since $X_t = W_t$ and $\sqrt{I_t} W_1$ have the same law, we can write

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

We begin with the proof of (48); 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 from Lemma 1 that

$$\mathbb{E}(\exp[\gamma I_t^{q/2} |W_1|^q]) \leq C \mathbb{E}(\exp[\delta I_t^{q/(2-q)}]) \tag{53}$$

for some $C, \delta > 0$. For the moment, we work on the event $\{i(t) \geq 1\}$. It follows from the basic bound (33) that

$$I_t \leq \sum_{k=0}^{i(t)} \xi_k^{2D} \sigma_k^2, \tag{54}$$

where we set

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

Note that $\sum_{k=0}^{i(t)} \xi_k = t$. By applying Hölder’s inequality to (54) with exponents $p = 1/2D$ and $p' = 1/(1 - 2D)$, we obtain

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

By assumption, $q \leq 1/(1 - D)$, which is the same as $(1 - 2D)q/(2 - q) \leq 1$. Thus,

$$I_t^{q/(2-q)} \leq t^{2Dq/(2-q)} \left(\sum_{k=0}^{i(t)} \sigma_k^{2/(1-2D)} \right)^{(1-2D)q/(2-q)} \leq t^{2Dq/(2-q)} \sum_{k=0}^{i(t)} \sigma_k^{2q/(2-q)}. \tag{55}$$

Now observe that if $i(t) = 0$ we have $I_t = \sigma_0^2[(t - \tau_0)^{2D} - (-\tau_0)^{2D}] \leq \sigma_0^2 t^{2D}$; hence, (55) also holds when $i(t) = 0$. Therefore, by (53),

$$E(e^{\gamma |X_t|^q}) \leq C E\left(\exp\left[\delta t^{2Dq/(2-q)} \sum_{k=0}^{i(t)} \sigma_k^{2q/(2-q)}\right]\right) = C E(\rho^{i(t)+1}), \tag{56}$$

where we have set

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

Therefore, if $\rho < \infty$, the right-hand side of (56) is finite, because $i(t) \sim \text{Po}(\lambda t)$ has finite exponential moments of all orders. This proves the ‘ \Leftarrow ’ implication in (48).

The ‘ \Rightarrow ’ implication in (48) is simpler. By the lower bound in Lemma 1 we have

$$E(e^{\gamma |X_t|^q}) = E(\exp[\gamma I_t^{q/2} |W_1|^q]) \geq C E(\exp[\delta I_t^{q/(2-q)}]) \tag{57}$$

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

$$\begin{aligned} E(\exp[\delta I_t^{q/(2-q)}]) &\geq E(\exp[\delta I_t^{q/(2-q)}] \mathbf{1}_{\{i(t)=0\}}) \\ &= E(\exp[\delta[(t - \tau_0)^{2D} - (-\tau_0)^{2D}]^{q/(2-q)} \sigma_0^{2q/(2-q)}]) P(i(t) = 0). \end{aligned} \tag{58}$$

Under the condition

$$E[\exp(\alpha \sigma^{2q/(2-q)})] = +\infty \quad \text{for all } \alpha > 0,$$

the last expectation in (58) is infinite, since $[(t - \tau_0)^{2D} - (-\tau_0)^{2D}] > 0$ a.s., which is independent of σ_0 . Looking back at (57), we have proved the ‘ \Rightarrow ’ implication in (48).

Next we prove (47); hence, we assume that $q > (1 - D)^{-1}$. Consider first the case $q < 2$ (which may happen only for $D < \frac{1}{2}$). By (57),

$$E(e^{\gamma |X_t|^q}) \geq C E(\exp[\delta I_t^{q/(2-q)}]).$$

We note that, by the definition of I_t given in (5), we can write

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

with the convention that the sum is 0 if $i(t) < 2$. For $n \geq 0$, we let P_n denote the conditional probability $P(\cdot \mid i(t) = n)$ and we let E_n be the corresponding expectation. Note that, under P_n , the random variables $(\tau_k - \tau_{k-1})_{k=2}^n$ have the same law as the random variables $(\xi_k)_{k=2}^n$ in Lemma 2 for $n \geq 2$. Consider the events

$$A_n := \{\sigma_k^2 \geq a \text{ for all } k = 2, \dots, n\}, \quad B_n := \left\{ \left\{ k = 2, \dots, n : \xi_k > \frac{1}{n^{1+\varepsilon}} \right\} \geq n^{1-\varepsilon} \right\},$$

where $a > 0$ is such that $\nu([a, +\infty)) =: \rho > 0$ and $\varepsilon > 0$ will be chosen later. Note that $P_n(A_n) = \rho^{n-1}$ while $P_n(B_n) \rightarrow 1$ as $n \rightarrow +\infty$ by Lemma 2. In particular, there exists a

$c > 0$ such that $P_n(B_n) \geq c$ for every n . Clearly, A_n and B_n are independent under P_n . We have

$$\begin{aligned} \psi(n) &:= E_n(\exp[\delta I_t^{q/(2-q)}]) \\ &\geq E_n(\exp[\delta I_t^{q/(2-q)}] \mathbf{1}_{A_n \cap B_n}) \\ &\geq c \rho^{n-1} \exp\left[\delta a^{q/(2-q)} \left(\frac{1}{n^{1+\varepsilon}}\right)^{2Dq/(2-q)} n^{(1-\varepsilon)q/(2-q)}\right] \\ &= c \rho^{n-1} \exp[\delta a^{q/(2-q)} n^{(1-2D-\varepsilon(1+2D))q/(2-q)}]. \end{aligned} \tag{59}$$

Note that $q > 1/(1 - D)$ is equivalent to $(1 - 2D)q/(2 - q) > 1$; therefore, ε can be chosen small enough so that $b := (1 - 2D - \varepsilon(1 + 2D))q/(2 - q) > 1$. It then follows from (59) that $\psi(n) \geq d \exp[dn^b]$ for every $n \in \mathbb{N}$ and a suitable $d > 0$. Therefore,

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

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

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

$$E(e^{\gamma |X_t|^q}) = E(\exp[\gamma I_t^{q/2} |W_1|^q]); \tag{60}$$

hence, if $q > 2$, we have $E(e^{\gamma |X_t|^q}) = \infty$ because $E(\exp[c|W_1|^q]) = \infty$ for every $c > 0$, $I_t > 0$ a.s., 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 (59) have shown that in this case I_t is unbounded. It then follows again from (60) that $E(e^{\gamma |X_t|^2}) = \infty$.

Appendix B. The model of Baldovin and Stella

Let us briefly discuss the model proposed by Baldovin and Stella [6] and [31], motivated by renormalization group arguments from statistical physics. They first introduced a process $(Y_t)_{t \geq 0}$ which satisfies the scaling relation (3) for a given function g , which is assumed to be even, so that its Fourier transform $\hat{g}(u) := \int_{\mathbb{R}} e^{iux} g(x) dx$ is real (and even). The process $(Y_t)_{t \geq 0}$ is defined by specifying its finite-dimensional laws: for $t_1 < t_2 < \dots < t_n$, the joint density of $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$ is given by

$$p(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = h\left(\frac{x_1}{\sqrt{t_1}}, \frac{x_2 - x_1}{\sqrt{t_2 - t_1}}, \dots, \frac{x_n - x_{n-1}}{\sqrt{t_n - t_{n-1}}}\right), \tag{61}$$

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

$$\hat{h}(u_1, u_2, \dots, u_n) := \hat{g}(\sqrt{u_1^2 + \dots + u_n^2}). \tag{62}$$

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 (62) is not necessarily the Fourier transform of a probability on \mathbb{R}^n , so some care is needed (we come back to this point in a moment). However, it is clear from (61) 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 introduced what is probably the most interesting ingredient of their construction, namely, a special form of time inhomogeneity. They defined

it in terms of finite-dimensional distributions, but it is simpler to give a pathwise construction: given a sequence of (possibly random) times $0 < \tau_1 < \tau_2 < \dots < \tau_n \uparrow +\infty$ and a fixed $0 < D \leq \frac{1}{2}$, they introduced a new process $(X_t)_{t \geq 0}$ defined by

$$X_t := Y_{t^{2D}} \quad \text{for } t \in [0, \tau_1), \tag{63}$$

and, more generally,

$$X_t := Y_{(t - \tau_n)^{2D} + \sum_{k=1}^n (\tau_k - \tau_{k-1})^{2D}} \quad \text{for } t \in [\tau_n, \tau_{n+1}). \tag{64}$$

For $D = \frac{1}{2}$, we clearly have $X_t \equiv Y_t$, while, for $D < \frac{1}{2}$, the process $(X_t)_{t \geq 0}$ is obtained from $(Y_t)_{t \geq 0}$ by a nonlinear time change, which 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 on the implicit relations (61)–(62). If a stochastic process $(Y_t)_{t \geq 0}$ is to satisfy these relations, it must necessarily have *exchangeable increments*: by this we mean (see [21, 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 the y_j s are 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 theorem ensures that $(Y_t)_{t \geq 0}$ is a mixture of Lévy processes; see Theorem 3 of [21] (see also [1]). Actually, more can be said: since, by (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 of [21], the process $(Y_t)_{t \geq 0}$ is necessarily a *mixture of Brownian motions*. This means that we have the representation

$$Y_t = \sigma W_t, \tag{65}$$

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). If a process $(Y_t)_{t \geq 0}$ satisfies (65) then, denoting by ν the law of σ , it is easy to check that relations (61)–(62) hold with

$$g(x) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \nu(d\sigma), \tag{66}$$

or, equivalently,

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

This shows that the functions g for which (61)–(62) provide a consistent family of finite-dimensional distributions are exactly those that may be expressed as in (66) for some probability ν on $(0, +\infty)$.

Note that a path of (65) 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 (63)–(64). In other words, the processes obtained through this construction have nonergodic increments.

Nevertheless, Baldovin and Stella claimed 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 showed good agreement with the clustering of volatility, as well as with the basic scaling (3) and the multiscaling of moments. The explanation of this apparent contradiction is that Baldovin and Stella did 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 fixed a small time step δ and a natural number N and they first simulated $x_\delta, x_{2\delta}, \dots, x_{N\delta}$ according to the true distribution of $(X_\delta, X_{2\delta}, \dots, X_{N\delta})$. Then they computed the conditional distribution of $X_{(N+1)\delta}$ given $X_{2\delta} = x_{2\delta}, X_{3\delta} = x_{3\delta}, \dots, X_{N\delta} = x_{N\delta}$ —thus, neglecting x_δ —and sampled $x_{(N+1)\delta}$ from this distribution. Similarly, $x_{(N+2)\delta}$ is sampled from the conditional distribution of $X_{(N+2)\delta}$ given $X_{3\delta} = x_{3\delta}, \dots, X_{N\delta} = x_{N\delta}, X_{(N+1)\delta} = x_{(N+1)\delta}$, 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. We gratefully acknowledge the support of the University of Padova under grant CPDA082105/08.

References

- [1] ACCARDI, L. AND LU, Y. G. (1993). A continuous version of de Finetti's theorem. *Ann. Prob.* **21**, 1478–1493.
- [2] ANDREOLI, A. (2011). Scaling and multiscaling in financial indexes: a simple model. Doctoral Thesis, University of Padova. Available at <http://www.matapp.unimib.it/~fcaraven/c.html>.
- [3] ASMUSSEN, S. (2003). *Applied Probability and Queues* (Appl. Math. **51**), 2nd edn. Springer, New York.
- [4] ANÉ, T. AND GEMAN, H. (2000). Order flow, transaction clock, and normality of asset returns. *J. Finance* **55**, 2259–2284.
- [5] BAILLIE, R. T. (1996). Long memory processes and fractional integration in econometrics. *J. Econometrics* **73**, 5–59.
- [6] BALDOVIN, F. AND STELLA, A. (2007). Scaling and efficiency determine the irreversible evolution of a market. *Proc. Nat. Acad. Sci. USA* **104**, 19741–19744.
- [7] BARNENDORFF-NIELSEN, O. E. AND SHEPHARD, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Statist. Soc. B* **63**, 167–241.
- [8] BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307–327.
- [9] BOLLERSLEV, T. AND MIKKELSEN, H. O. (1996). Modeling and pricing long memory in stock market volatility. *J. Econometrics* **31**, 151–184.
- [10] BOLLERSLEV, T., KRETSCHMER, U., PIGORSCH, C. AND TAUCHEN, G. (2009). A discrete-time model for daily S&P500 returns and realized variations: jumps and leverage effects. *J. Econometrics* **150**, 151–166.
- [11] BONINO, M. (2011). Portfolio allocation and monitoring under volatility shocks. Master's Thesis, University of Padova. Available at <http://www.matapp.unimib.it/~fcaraven/c.html>.
- [12] CALVET, L. E. AND FISHER, A. J. (2008). *Multifractal Volatility*. Academic Press.
- [13] CALVET, L. AND FISHER, A. (2001). Forecasting multifractal volatility. *J. Econometrics* **105**, 27–58.
- [14] CALVET, L., FISHER, A. AND MANDELBROT, B. (1997). A multifractal model of asset returns. Discussion paper 1164, Yale University. Available at <http://cowles.econ.yale.edu>.
- [15] CALVET, L., FISHER, A. AND MANDELBROT, B. (1997). Large deviations and the distribution of price changes. Discussion paper 1165, Yale University. Available at <http://cowles.econ.yale.edu>.
- [16] CLARK, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* **41**, 135–155.
- [17] CONT, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quant. Finance* **1**, 223–236.
- [18] DI MATTEO, T., ASTE, T. AND DACOROGNA, M. M. (2005). Long-term memories of developed and emerging markets: using the scaling analysis to characterize their stage of development. *J. Banking Finance* **29**, 827–851.
- [19] ENGLE, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica* **50**, 987–1007.

- [20] FISHER, L., CALVET, A. AND MANDELBROT, B. (1997). Multifractality of Deutschemark/US dollar exchange rates. Discussion paper 1166, Yale University. Available at <http://cowles.econ.yale.edu>.
- [21] FREEDMAN, D. A. (1963). Invariants under mixing which generalize de Finetti's theorem: continuous time parameter. *Ann. Math. Statist.* **34**, 1194–1216.
- [22] GALLUCCIO, S., CALDARELLI, G., MARSILI, M. AND ZHANG, Y.-C. (1997). Scaling in currency exchange. *Physica A* **245**, 423–436.
- [23] GHASHGHAIE, S. *et al.* (1996). Turbulent cascades in foreign exchange markets. *Nature* **381**, 767–770.
- [24] HULL, J. C. (2009). *Options, Futures and Other Derivatives*. Pearson/Prentice Hall.
- [25] KARATZAS, I. AND SHREVE, S. E. (1988). *Brownian Motion and Stochastic Calculus*. Springer, New York.
- [26] KLÜPPELBERG, C., LINDNER, A. AND MALLER, R. (2004). A continuous-time GARCH process driven by a Lévy process: stationarity and second-order behaviour. *J. Appl. Prob.* **41**, 601–622.
- [27] KLÜPPELBERG, C., LINDNER, A. AND MALLER, R. (2006). Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models. In *From Stochastic Calculus to Mathematical Finance*, eds Y. Kabanov, R. Lipster and J. Stoyanov, Springer, Berlin, pp. 393–419.
- [28] PIGATO P. (2011). A multivariate model for financial indexes subject to volatility shocks. Masters Thesis, University of Padova. Available at <http://www.matapp.unimib.it/~fcaraven/c.html>.
- [29] SHEPHARD, N. AND ANDERSEN, T. G. (2009). Stochastic volatility: origins and overview. In *Handbook of Financial Time Series*, Springer, Berlin, pp. 233–254.
- [30] SHIRYAEV, A. N. (1995). *Probability* (Graduate Texts Math. **95**), 2nd edn. Springer, New York.
- [31] STELLA, A. L. AND BALDOVIN, F. (2008). Role of scaling in the statistical modeling of finance. *Pramana* **71**, 341–352.
- [32] VASSILICOS, J. C., DEMOS, A. AND TATA, F. (1993). 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 and H. Jones, Springer, Berlin, pp. 249–265.
- [33] WEISS, L. (1955). The stochastic convergence of a function of sample successive differences. *Ann. Math. Statist.* **26**, 532–536.