SCALING AND MULTISCALING IN FINANCIAL INDEXES: A SIMPLE MODEL

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

Abstract. We propose a simple stochastic model for time series which is analytically tractable, easy to simulate and which captures some relevant stylized facts of financial indexes, including *scaling properties*. We show that the model fits the *Dow Jones Industrial Average* time series in the period 1935-2009 with a remarkable accuracy.

Despite its simplicity, the model has several interesting features. The volatility is not constant and displays high peaks. The empirical distribution of the *log-returns* (increments of the logarithm of the index) is non-Gaussian and may exhibit heavy tails. Log-returns corresponding to disjoint time intervals are uncorrelated but not independent: the correlation of their absolute values decays exponentially fast in the distance between the time intervals for large distances, while it has a slower decay for moderate distances. Finally, the distribution of the log-returns obeys *scaling relations* that are detected on real time series, but are not satisfied by most available models.

1. Introduction

1.1. Modeling financial indexes. Recent developments in stochastic modelling of time series have been strongly influenced by the analysis of financial indexes. The basic model, that has given rise to the celebrated Black & Scholes formula [11, 16], assumes that the logarithm X_t of the price of the underlying index, after subtracting the trend, is given by

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

where σ (the *volatility*) is a constant and $(W_t)_{t\geq 0}$ is a standard Brownian motion.

Despite its success, this model is not consistent with a number of *stylized facts* that are empirically detected in many real time series. Some of these facts are the following:

- the volatility is not constant: in particular, it may have high peaks, that may be interpreted as *shocks* in the market (see Figure 6(c) below);
- the empirical distribution of the increments $X_{t+h} - X_t$ of the logarithm of the price (the *log-returns*) has tails heavier than Gaussian (see Figure 4(b) below);
- log-returns corresponding to disjoint time-interval are uncorrelated, but not independent: in fact, the correlation between the absolute values $|X_{t+h} - X_t|$ and $|X_{s+h} - X_s|$ has a slow decay in $|t-s|$, up to moderate values for $|t-s|$. This phenomenon is known as *clustering of volatility* (see Figure 3 below).

In order to have a better fit with real data, many different models have been proposed to describe the volatility and the price process. In discrete-time, autoregressive models such as ARCH and GARCH [9, 6] have been widely used. In continuous time, the basic model (1.1) has been modified by letting $\sigma = \sigma_t$ be a stochastic process, often the solution of a

Date: June 1, 2010.

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

Key words and phrases. Financial Index, Scaling, Multiscaling, Brownian Motion, Stochastic Volatility.

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

stochastic differential equation driven by a general Lévy process. A systematic account of these *stochastic volatility models* can be found in [3]. More recent developments along these lines include the *generalized Ornstein-Uhlenbeck processes* and the COGARCH (GARCH in continuous time) [13, 14]. All these models involve several parameters, whose estimation on real data is itself a subject of research.

More recently (see [8, 5, 18]), other stylized facts of financial indexes have been pointed out, concerning the *scaling properties* of the empirical distribution of the log-returns. Consider the time series of an index $(s_i)_{1\le i\le T}$ over a period of $T \gg 1$ days and denote by p_h the *empirical distribution* of the (detrended) log-returns corresponding to an interval of h days:

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

where $\overline{d}_i$ is the local rate of linear growth of $\log(s_i)$ (see section 3 for details) and $\delta_x(\cdot)$ denotes the Dirac measure at $x \in \mathbb{R}$. The statistical analysis of various indexes, such as the *Dow Jones Industrial Average* (DJIA) or the *Nikkei 225*, shows that, for h within a suitable time scale, p_h obeys approximately a diffusive scaling relation (cf. Figure 1(A)):

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

where g is a probability density with tails heavier than Gaussian. If one considers the q-th empirical moment $m_q(h)$, defined by

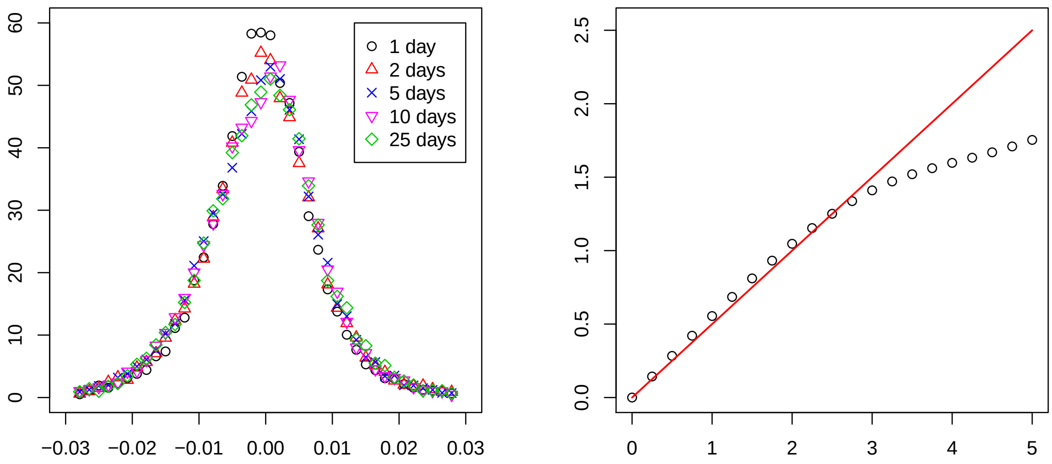
$$m_q(h) \,:=\, \frac{1}{T-h}\sum_{i=1}^{T-h} |x_{i+h} - x_i|^q \,=\, \int |r|^q\, p_h(\mathrm{d}r)\,, \tag{1.4}$$

from relation (1.3) it is natural to guess that $m_q(h)$ should scale as $h^{q/2}$. This is indeed what one observes for moments of small order $q \le \bar{q}$ (with $\bar{q} \simeq 3$ for the DJIA). However, for moments of higher order $q > \bar{q}$, the different scaling relation $h^{A(q)}$, with $A(q) < q/2$, takes place, cf. Figure 1(B) (see also [8]). This is the so-called *multiscaling of moments.*

Despite of the variety of models that can be found in the literature, it is quite nontrivial to identify a good one which agrees with *all* mentioned stylized facts. For example, the celebrated and widely used GARCH [2] exhibits non-constant volatility, clustering of volatility and non-Gaussian distribution of log-returns, but a closer analysis shows that multiscaling of moments is not present, at least for the range of values of the parameters that most often occur in practice. More refined version of GARCH (for instance FIGARCH, see [2, 7]) have been proposed. It should be stressed that although models with sufficiently many parameters can be adapted to data, the soundness of the procedure of statistical inference may be very weak. We shall make more comments on this issue in section 3.

1.2. **Baldovin and Stella's approach.** Motivated by renormalization group arguments from statistical physics, Baldovin and Stella [5, 18] have recently proposed a new model in which scaling considerations play a major role. Their construction may be summarized as follows. They first introduce a process $(Y_t)_{t\ge0}$ which satisfies the scaling relation (1.3) for a given function g, that is assumed to be even, so that its Fourier transform $\hat{g}(u) := \int_{\mathbb{R}} e^{iux} g(x)\mathrm{d}x$ is real (and even). The process $(Y_t)_{t\ge0}$ is defined by specifying its finite dimensional laws: for $t_1 < t_2 < \cdots < t_n$ the joint density of $Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}$ is given by

$$p(x_1,t_1;x_2,t_2;\ldots;x_n,t_n) = h\left(\frac{x_1}{\sqrt{t_1}}, \frac{x_2-x_1}{\sqrt{t_2-t_1}}, \ldots, \frac{x_n-x_{n-1}}{\sqrt{t_n-t_{n-1}}}\right), \tag{1.5}$$



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

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

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

$$\hat{h}(u_1, u_2, \dots, u_n) := \hat{g}\Big(\sqrt{u_1^2 + \dots + u_n^2}\Big). \tag{1.6}$$

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 (1.6) is not necessarily the Fourier transform of a probability on $\mathbb{R}^n$, so that some care is needed. We will see in section 2 for exactly what g's relations (1.5)-(1.6) define a consistent family of finite dimensional distributions, hence a stochastic process $(Y_t)_{t \geq 0}$, and it turns out that $(Y_t)_{t \geq 0}$ is always a *mixture* of Brownian motions. However, it is already clear from (1.5) that the increments of the process $(Y_t)_{t \geq 0}$ corresponding to time intervals of the same length (that is, for fixed $t_{i+1} - t_i$) have a *permutation invariant distribution* and therefore cannot exhibit any decay of correlations.

For this reason, Baldovin and Stella introduce what we believe is the most remarkable ingredient of their construction, namely a special form of time-inhomogeneity in the process $(Y_t)_{t \geq 0}$. Although they define it in terms of finite dimensional distributions, we find it simpler to give an equivalent pathwise construction. Given a sequence of times $0 < \tau_1 < \tau_2 < \cdots < \tau_n \uparrow +\infty$ and a fixed $0 < D \leq 1/2$, they introduce a new process $(X_t)_{t \geq 0}$ defined by

$$X_t := Y_{t^{2D}} \qquad \text{for } t \in [0, \tau_1), \tag{1.7}$$

and more generally

$$X_t := Y_{(t-\tau_n)^{2D} + \sum_{k=1}^{n} (\tau_k - \tau_{k-1})^{2D}} \qquad \text{for } t \in [\tau_n, \tau_{n+1})\,. \tag{1.8}$$

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

In their numerical study Baldovin and Stella, besides considering the periodic case $\tau_n := nt_0$ for t_0 fixed, do not really simulate the process $(X_t)_{t\geq 0}$ but rather an autoregressive version of it. More precisely, given a small time step δ and a natural number N, they first simulate x_δ, $x_{2\delta}$, $\ldots$, $x_{N\delta}$ according to the distribution of $(X_\delta, X_{2\delta}, \ldots, X_{N\delta})$. Then they compute the conditional distribution of $X_{(N+1)\delta}$ given $X_{2\delta} = x_{2\delta}, X_{3\delta} = x_{3\delta}, \ldots, X_{N\delta} = x_{N\delta}$ — thus neglecting x_δ — and sample $x_{(N+1)\delta}$ from this distribution. Similarly, $x_{(N+2)\delta}$ is sampled from the conditional distribution of $X_{(N+2)\delta}$ given $X_{3\delta} = x_{3\delta}$, $\ldots$, $X_{N\delta} = x_{N\delta}$, $X_{(N+1)\delta} = x_{(N+1)\delta}$, neglecting x_δ and $x_{2\delta}$, and so on.

After estimating the function g and the parameters t_0 and D on the DJIA time series, their simulated trajectories show a quite convincing agreement with both the basic scaling (1.3) and the multiscaling of moments, as well as with the clustering of volatility.

1.3. **Content of the paper.** The purpose of this paper is to define a simple continuous-time stochastic model $X = (X_t)_{t\geq 0}$ for a (detrended) log-index which, on the one hand, captures the essential features of the one proposed by Baldovin and Stella and, on the other hand, is easier to describe, to interpret and to simulate. For this model, all the stylized facts mentioned above (see §1.1) can be rigorously proved.

The computational simplicity of the model makes it easy, at least at a numerical level, to use it as an alternative to Black & Scholes and related models in applications to finance, such as pricing of options. Its performance on real markets of derivatives is the subject of current investigations.

The paper is organized as follows.

- In section 2 we give the precise definition of our model and we state its main properties.
- In section 3 we show the remarkable agreement of our model with the DJIA time series in the period 1935-2009.
- Sections 4, 5 and 6 contain the proofs of the main results stated in section 2, plus some additional material.
- Finally, some technical points have been deferred to the appendices A and B.

We believe that the numerical comparison with the DJIA, described in section 3, provides strong motivation for a further study of our model. We point out that the agreement with the Nikkei 225 index is very good as well. A systematic treatment of other time series, also beyond financial indexes, still has to be done; however, some preliminary analysis of the prices of single stocks shows that our model fits well some but not all of them. It would be very interesting to understand more precisely which of the properties we have mentioned are linked to *aggregation* of several stock prices, as in the DJIA.

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

2. The model and the main results

If a stochastic process $(Y_t)_{t\geq 0}$ is to satisfy relations (1.5)–(1.6), it must necessarily have *exchangeable increments*.[1] If we make the (very mild) assumption that $(Y_t)_{t\geq 0}$ has no fixed point of discontinuity, then a continuous-time version of the celebrated de Finetti's theorem ensures that $(Y_t)_{t\geq 0}$ is a mixture of Lévy processes, see Theorem 3 in [10] (cf. also [1]). Actually, more can be said: since by (1.3) the distribution of the increments of $(Y_t)_{t\geq 0}$ is *isotropic*, i.e., it has spherical symmetry in $\mathbb{R}^n$, by Theorem 4 in [10] the process $(Y_t)_{t\geq 0}$ is a mixture of Brownian motions. This means that we have the following representation:

$$Y_t = \sigma\, W_t, \tag{2.1}$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion and σ is a random variable (interpreted as a *random volatility*) taking values in $(0,\infty)$ and independent of $(W_t)_{t\geq 0}$. Viceversa, if a process $(Y_t)_{t\geq 0}$ satisfies (2.1), then, denoting by ν the law of σ, relations (1.5)–(1.6) hold with

$$g(x) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{x^2}{2\sigma^2}}\, \nu(\mathrm{d}\sigma)\,, \tag{2.2}$$

or, equivalently,

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

as one easily checks. This shows that the functions g for which (1.5)–(1.6) provide a consistent family of finite dimensional distributions are exactly those that may be expressed as in (2.2) for some probability ν on $(0,+\infty)$.

A sample path of (2.1) is obtained by sampling independently σ from ν and $(W_t)_{t\geq 0}$ from the Wiener measure. Note that this path *cannot* be distinguished from the path of a Brownian motion with *constant* volatility, and therefore correlations of increment cannot be detected empirically. The same observation applies to the time-inhomogeneous process $(X_t)_{t\geq 0}$ obtained by $(Y_t)_{t\geq 0}$ through (1.7)–(1.8). The reason why Baldovin and Stella measure nonzero correlations from their samples is that they do not simulate $(X_t)_{t\geq 0}$ but rather an autoregressive approximation of it.

2.1. **Definition of the model.** We propose here a different mechanism, that replaces the autoregressive scheme by a random update of the volatility. Given two real numbers $D \in (0, 1/2]$, $\lambda \in (0,\infty)$ and a probability ν on $(0,\infty)$ (these may be viewed as our parameters), our model is defined upon the following three sources of alea:

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

We assume that $W, \mathcal{T}, \Sigma$ are defined on some probability space $(\Omega, \mathcal{F}, P)$ and that they are independent. By convention, we label the points of $\mathcal{T}$ so that $\tau_0 < 0 < \tau_1$. We will actually need only the points of $\mathcal{T} \cap [\tau_0, \infty)$, that is the variables $(\tau_n)_{n\geq 0}$. We recall that the random variables $-\tau_0$, τ_1, $(\tau_{n+1} - \tau_n)_{n\geq 1}$ are independent and identically distributed $Exp(\lambda)$, so that $1/\lambda$ is the mean distance between the points in $\mathcal{T}$. Although some of our results would hold

[1] By this we mean precisely the following: setting $\Delta Y_{(a,b)} := Y_b - Y_a$ for short, the distribution of the random vector $(\Delta Y_{I_1+y_1}, \ldots, \Delta Y_{I_n+y_n})$ — where the I_j's are intervals and y_j's real numbers — does not depend on $y_1, \ldots, y_n$, as long as the intervals $y_1 + I_1, \ldots, y_n + I_n$ are disjoint.

for more general distributions of $\mathcal{T}$, we stick for simplicity to the (rather natural) choice of a Poisson process.

We are now ready to define our model $X = (X_t)_{t\geq 0}$. For $t \in [0, \tau_1]$ we set

$$X_t := \sigma_0\big(W_{(t-\tau_0)^{2D}} - W_{(-\tau_0)^{2D}}\big)\,, \tag{2.3}$$

while for $t \in [\tau_n, \tau_{n+1}]$ (with $n \geq 1$) we set

$$X_t := X_{\tau_n} + \sigma_n \left(W_{(t-\tau_n)^{2D} + \sum_{k=1}^n (\tau_k - \tau_{k-1})^{2D}} - W_{\sum_{k=1}^n (\tau_k - \tau_{k-1})^{2D}}\right)\,. \tag{2.4}$$

In words: at the epochs τ_n the time inhomogeneity $t \mapsto t^{2D}$ is "refreshed" and the volatility is randomly updated: $\sigma_{n-1} \rightsquigarrow \sigma_n$. A possible financial interpretation of this mechanism is that jumps in the volatility correspond to shocks in the market. The reaction of the market is not homogeneous in time: if $D < 1/2$, the dynamics is fast immediately after the shock, and tends to slow down later, until a new jump occurs. For $D = 1/2$ our model reduces to a simple random volatility model $\mathrm{d}X_t = \sigma_t \,\mathrm{d}W_t$, where $\sigma_t := \sum_{k=0}^\infty \sigma_k \,\mathbf{1}_{[\tau_k, \tau_{k+1})}(t)$ is a (random) piecewise constant process.

Note that the pieces of the Brownian motion W used in each interval $[\tau_n, \tau_{n+1})$ (cf. (2.4)) are independent. Also observe that in (2.3) the instant $t = 0$ is not the beginning of the time inhomogeneity, as it was in (1.7): this is to ensure that the process X has stationary increments (see below).

Using the scale invariance of Brownian motion, we now give an alternative definition of our model X, that is equivalent in law with (2.4) but more convenient for the proofs in the following sections. For $t \geq 0$, define

$$i(t) \;:=\; \sup\{n \geq 0 :\, \tau_n \leq t\} \;=\; \#\{\mathcal{T} \cap (0, t]\}\,, \tag{2.5}$$

so that $\tau_{i(t)}$ is the location of the last point in $\mathcal{T}$ before t. Now we introduce the process $I = (I_t)_{t\geq 0}$ by

$$I_t \;:=\; \sigma_{i(t)}^2 \left(t - \tau_{i(t)}\right)^{2D} + \sum_{k=1}^{i(t)} \sigma_{k-1}^2 \left(\tau_k - \tau_{k-1}\right)^{2D} \;-\; \sigma_0^2 \left(-\tau_0\right)^{2D}, \tag{2.6}$$

with the agreement that the sum in the right hand side is zero if $i(t) = 0$. We can then redefine our basic process $X = (X_t)_{t\geq 0}$ by setting

$$X_t := W_{I_t}\,. \tag{2.7}$$

Note that I is a strictly increasing process with absolutely continuous paths, and it is independent of the Brownian motion W. Thus our model may be viewed as an independent random time change of a Brownian motion.

2.2. **Basic properties.** Some basic properties of our model can be easily established.

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

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

$$\mathrm{d}X_t = v_t \,\mathrm{d}B_t\,, \tag{2.8}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion. More precisely, denoting $I'(s) := \frac{\mathrm{d}}{\mathrm{d}s} I(s)$, the variables B_t and v_t are defined by

$$B_t := \int_0^{I_t} \frac{1}{\sqrt{I'(I^{-1}(u))}}\,\mathrm{d}W_u\,, \qquad v_t := \sqrt{I'(t)} = \sqrt{2D}\,\sigma_{i(t)} \left(t - \tau_{i(t)}\right)^{D - \frac{1}{2}}\,. \tag{2.9}$$

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

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

The proof of these properties is given in section 4, where we also show some relations between the distribution ν of the σ_k's and the law of the random variable X_t. Let us mention in particular that, for any $q > 0$,

$$E(|X_t|^q) < \infty \text{ for some (hence any) } t > 0 \quad \Longleftrightarrow \quad E(\sigma^q) < \infty \,. \tag{2.10}$$

As for most real financial indexes (like the DJIA, see section 3) there is empirical evidence that $E(|X_t|^q) < \infty$ for all $q > 0$, the main interest of our model is when σ *has finite moments of all orders*. In this case, a link between the exponential moments of σ and those of X_t is given in Proposition 14 in section 4 below.

Remark 1. If we look at the process X for a *fixed* realization of the variables $\mathcal{T}$ and Σ, averaging only on W — that is, if we work under the conditional probability $P(\,\cdot\,|\mathcal{T},\Sigma)$ — the increments of X are no longer stationary, as it is clear from (2.4). We stress however that the properties (B) and (C) above continue to hold also under $P(\,\cdot\,|\mathcal{T},\Sigma)$, as it will be clear in section 4. Of course, the condition $E(\sigma^2) < \infty$ in (C) is not needed under $P(\,\cdot\,|\mathcal{T},\Sigma)$.

Another important property of the process X is that the distribution of its increments is *ergodic*, as we show in section 4. This entails in particular that for every $\delta > 0$, $k \in \mathbb{N}$ and for every choice of the intervals $(a_1, b_1), \ldots, (a_k, b_k) \subseteq (0,\infty)$ and of the measurable function $F : \mathbb{R}^k \to \mathbb{R}$, we have almost surely

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} F(X_{n\delta+b_1} - X_{n\delta+a_1}, \ldots, X_{n\delta+b_k} - X_{n\delta+a_k}) \\ = E\big[F(X_{b_1} - X_{a_1}, \ldots, X_{b_k} - X_{a_k})\big] \,, \end{aligned} \tag{2.11}$$

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

Thanks to this property, our main results concerning the moments and the correlation of the increments of the process X, that we state in the next subsection, are of direct relevance for the comparison of our model with real data series. Some care is needed, however, because the accessible time length N in (2.11) may not be large enough to ensure that the empirical averages are close to their limit. We elaborate more on this issue in section 3, where we compare our model with the DJIA data from a numerical viewpoint.

2.3. **The main results.** We now state our main results concerning the process X. They correspond to the basic stylized facts that we have mentioned in the introduction: diffusive scaling of the distributions of log-returns (Theorem 2 below); multiscaling of moments (Theorem 3 and Corollary 4); clustering of volatility (Theorem 5 and Corollary 6).

The first result, proved in section 5, states that for small h the increments $(X_{t+h} - X_t)$ have an approximate diffusive scaling, in agreement with (1.3).

Theorem 2. *As $h \downarrow 0$ we have the convergence in distribution*

$$\frac{(X_{t+h} - X_t)}{\sqrt{h}} \xrightarrow[h\downarrow 0]{d} f(x)\, \mathrm{d}x \,, \tag{2.12}$$

where f is a mixture of centered Gaussian densities, namely

$$f(x) = \int_0^\infty \nu(\mathrm{d}\sigma) \int_0^\infty \mathrm{d}t\, \lambda e^{-\lambda t} \, \frac{t^{1/2-D}}{\sigma\sqrt{4D\pi}} \, \exp\left(-\frac{t^{1-2D}x^2}{4D\sigma^2}\right) . \tag{2.13}$$

We stress that the function f appearing in (2.12)–(2.13), which describes the asymptotic rescaled law of the increment $(X_{t+h}-X_t)$ in the limit of small h, has a different tail behavior from the density of $(X_{t+h}-X_t)$ for fixed h. For instance, when σ has finite moments of all orders, it follows by (2.10) that the same holds for $(X_{t+h}-X_t)$. However, from (2.13) and a simple change of variables we get

$$\int_{\mathbb{R}} |x|^q \, f(x) \, \mathrm{d}x = (2D)^{q/2} E(\sigma^q) \int_0^\infty \lambda e^{-\lambda s} \, s^{(D-1/2)q} \, \mathrm{d}s \,,$$

which is finite if and only if $q < q^* := (1/2 - D)^{-1}$. Therefore, independently of the law ν of σ, the density f *has always polynomial tails*: $\int_{\mathbb{R}} |x|^q f(x) \, \mathrm{d}x = \infty$ for $q \geq q^*$.

This feature of f has striking consequences on the scaling behavior of the moments of the increments our model. If we set for $q \in (0,\infty)$

$$m_q(h) := E(|X_{t+h} - X_t|^q) \,, \tag{2.14}$$

from the convergence result (2.12) it would be natural to guess that $m_q(h) \approx h^{q/2}$ as $h \downarrow 0$, in analogy with the Brownian motion case. However, this turns out to be true only for $q < q^*$. For $q \geq q^*$, 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 $q \geq q^*$. This transition in the scaling behavior of $m_q(h)$ goes under the name of *multiscaling of moments* and is discussed in detail, e.g., in [8]. Let us now state our result, that we prove in section 5.

Theorem 3 (Multiscaling of moments)**.** *Let $q > 0$, and assume $E\left(\sigma^q\right) < +\infty$. Then the quantity $m_q(h) := E(|X_{t+h} - X_t|^q) = E(|X_h|^q)$ is finite and has the following asymptotic behavior as $h \downarrow 0$:*

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

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

$$C_q := \begin{cases} E(|W_1|^q) \, E(\sigma^q) \, \lambda^{q/q^*} \, (2D)^{q/2} \, \Gamma(1 - q/q^*) & \text{if } q < q^* \\ E(|W_1|^q) \, E(\sigma^q) \, \lambda \, (2D)^{q/2} & \text{if } q = q^* \\ E(|W_1|^q) \, E(\sigma^q) \, \lambda \left[\int_0^\infty ((1+x)^{2D} - x^{2D})^{\frac{q}{2}} \, \mathrm{d}x \, + \, \frac{1}{Dq+1} \right] & \text{if } q > q^* \end{cases} , \tag{2.15}$$

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

Corollary 4. *The following relation holds true:*

$$A(q) := \lim_{h \downarrow 0} \frac{\log m_q(h)}{\log h} = \begin{cases} \frac{q}{2} & \text{if } q \leq q^* \\ Dq + 1 & \text{if } q \geq q^* \end{cases} . \tag{2.16}$$

The explicit form (2.15) of the multiplicative constant C_q will be used in section 3 for the estimation of the parameters of our model on the DJIA time series.

Our last theoretical result, proved in section 6, concerns the correlations of the absolute value of two increments, a quantity which is usually called *volatility autocorrelation.* We start determining the behavior of the covariance.

Theorem 5. *Assume that $E(\sigma^2) < \infty$. The following relation holds as $h \downarrow 0$, for all $s, t > 0$:*

$$Cov(|X_{s+h} - X_s|, |X_{t+h} - X_t|) = \frac{4D}{\pi} \, \lambda^{1-2D} \, e^{-\lambda|t-s|} \left(\phi(\lambda|t-s|) \, h \, + \, o(h) \right) , \tag{2.17}$$

where

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

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

We recall that $\rho(Y,Z) := Cov(Y,Z)/\sqrt{Var(Y)Var(Z)}$ is the correlation coefficient of two random variables Y, Z. As Theorem 3 yields

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

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

Corollary 6 (Volatility autocorrelation)**.** *Assume that $E(\sigma^2) < \infty$. The correlation of the increments of the process X has the following asymptotic behavior as $h \downarrow 0$:*

$$\lim_{h\downarrow 0} \rho(|X_{s+h} - X_s|, |X_{t+h} - X_t|) \;=\; \rho(t-s) \;:=\; \frac{2}{\pi\, Var(\sigma\,|W_1|\, S^{D-1/2})}\, e^{-\lambda|t-s|}\, \phi(\lambda|t-s|)\,, \tag{2.19}$$

where $\phi(\cdot)$ is defined in (2.18) *and $S \sim Exp(1)$ is independent of σ and W_1.*

This shows that the volatility autocorrelation of our process decays exponentially fast for time scales greater than the mean distance $1/\lambda$ between the epochs τ_k. For shorter time scales, a relevant contribution is given by the function $\phi(\cdot)$. Note that by (2.18) we can write

$$\phi(x) \;=\; Var(\sigma)\, E(S^{D-1/2}\,(S+x)^{D-1/2}) \;+\; E(\sigma)^2\, Cov(S^{D-1/2}, (S+x)^{D-1/2})\,, \tag{2.20}$$

where $S \sim Exp(1)$. As $x \to \infty$, the two terms in the right hand side decay as

$$E(S^{D-1/2}\,(S+x)^{D-1/2}) \;\approx\; x^{D-1/2}\,, \qquad Cov(S^{D-1/2}, (S+x)^{D-1/2}) \;\approx\; x^{D-3/2}\,,$$

while for $x = O(1)$ the decay of both terms is faster than polynomial but slower than exponential (see Figures 3(a) and 3(b) below).

3. Estimation and data analysis

We now consider some aspects of our model from a numerical viewpoint. We compare the theoretical predictions and the simulated data of our model with the time series of the Dow Jones Industrial Average (DJIA) over a period of 75 years: more precisely, we have considered the DJIA opening prices from 2 Jan 1935 to 31 Dec 2009, for a total of 18849 daily data. Of course, other indexes would be interesting as well, but we have decided to stick to the DJIA both for its importance and for its considerable time length.

The data analysis, the simulations and the plots have been obtained with the software R [17]. The code we have used is available on the web page `http://www.math.unipd.it/~fcaraven/c.html`.

3.1. Preliminary considerations.

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

(a) The *multiscaling of moments*, cf. Corollary 4.

(b) The *volatility autocorrelation decay*, cf. Corollary 6.

(c) The *distribution* of X_t.

Roughly speaking, the idea is to compute *empirically* these quantities on the DJIA time series and then to compare the results with the *theoretical* predictions of our model. This is justified by the ergodicity of the increments of our process $(X_t)_{t\geq 0}$, cf. equation (2.11).

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

In order to estimate $(D, \lambda, E(\sigma), E(\sigma^2))$, we take into account four significant quantities that depend only on these parameters: the multiscaling coefficients C_1 and C_2 (see (2.15)), the multiscaling exponent $A(q)$ (see (2.16)) and the volatility autocorrelation function $\rho(t)$ (see (2.19)). We then consider a natural loss functional $\mathcal{L} = \mathcal{L}(D, \lambda, E(\sigma), E(\sigma^2))$, see (3.3) below, which describes some distance between these theoretical quantities and the corresponding empirical ones (evaluated on the DJIA time series), and we 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 $E(\sigma^2) \geq (E(\sigma))^2$.

Quite surprisingly, it turns out that the estimated values are such that $E(\sigma^2) \simeq (E(\sigma))^2$, that is σ *is nearly constant*. The constraint $E(\sigma^2) \geq (E(\sigma))^2$ is not playing a relevant role: the unconstrained minimum nearly coincides with the constrained one. Thus, the problem of determining the distribution of σ beyond its moments $E(\sigma)$ and $E(\sigma^2)$ does not appear in the case of the DJIA. If this was not the case, as it may be for other indexes, the law of σ should be estimated by requiring a good fit of the distribution of the log-return X_t with the empirical one. For the DJIA, all we can do is *to verify* that there is indeed a (remarkably) good fit between these distributions, with no further parameter to be estimated.

Before describing in detail the procedure we have just sketched, let us fix some notation: the DJIA time series will be denoted by $(s_i)_{0\leq i\leq N}$ (where $N = 18848$) and the corresponding detrended log-DJIA time series will be denoted by $(x_i)_{0\leq i\leq N}$:

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

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

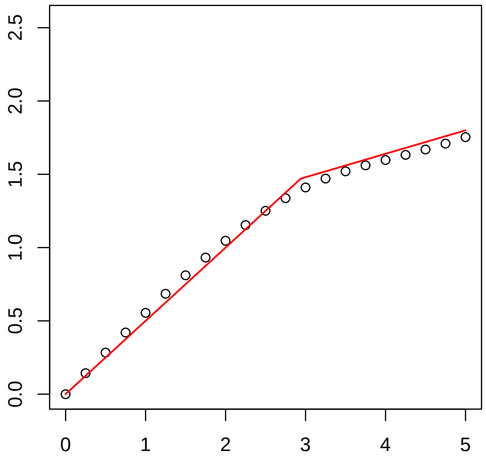
3.2. **Estimation of $D, \lambda, E(\sigma), E(\sigma^2)$.** The theoretical scaling exponent $A(q)$ is defined in (2.16) while the multiscaling constants C_1 and C_2 are given by (2.15) for $q = 1$ and $q = 2$. Since $q^* = (\frac{1}{2} - D)^{-1} > 2$ (we recall that $0 \leq D \leq \frac{1}{2}$), we can write more explicitly

$$C_1 = \frac{2\sqrt{D}\,\Gamma(\frac{1}{2}+D)\,E(\sigma)\,\lambda^{1/2-D}}{\sqrt{\pi}}\,, \qquad C_2 = 2D\,\Gamma(2D)\,E(\sigma^2)\,\lambda^{1-2D}\,. \tag{3.1}$$

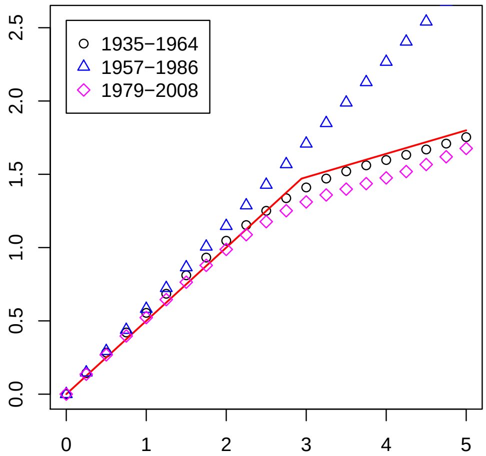
Defining the corresponding empirical quantities requires some care, because the DJIA data are in discrete-time and therefore no $h \downarrow 0$ limit is possible. We first evaluate the empirical q-moment $\widehat{m}_q(h)$ of the DJIA log-returns over h days, namely

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

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



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



(B) *Multiscaling in subperiods of 30 years.*

FIGURE 2. *Multiscaling of moments in the DJIA time series (1935-2009).*
(A) The empirical scaling exponent $\widehat{A}(q)$ in the period 1935-2009 (circles) and the theoretical scaling exponent $A(q)$ (line) as a function of q.
(B) The empirical scaling exponent $\widehat{A}(q)$ evaluated in subperiods of 30 years (see legend) and the theoretical scaling exponent $A(q)$ (line) as a function of q.
Note. The theoretical scaling exponent $A(q)$ is evaluated for $D = 0.16$, cf. (3.4).

For what concerns the theoretical volatility autocorrelation, Corollary 6 and the stationarity of the increments of our process $(X_t)_{t\geq 0}$ yield

$$\rho(t) := \lim_{h\downarrow 0} \rho(|X_h|, |X_{t+h} - X_t|) = \frac{2}{\pi\, Var(\sigma\,|W_1|\, S^{D-1/2})}\, e^{-\lambda t}\, \phi(\lambda t)\,, \tag{3.2}$$

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

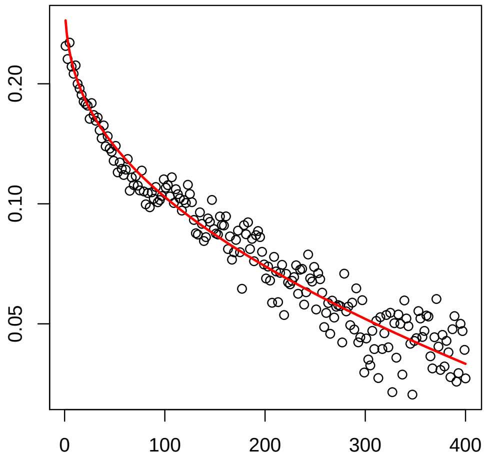
$$\phi(x) = Var(\sigma)\, E(S^{D-1/2}\, (S+x)^{D-1/2}) + E(\sigma)^2\, Cov(S^{D-1/2}, (S+x)^{D-1/2})\,,$$

cf. (2.20). Note that, although $\phi(\cdot)$ does not admit an explicit expression, it can be quite 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.

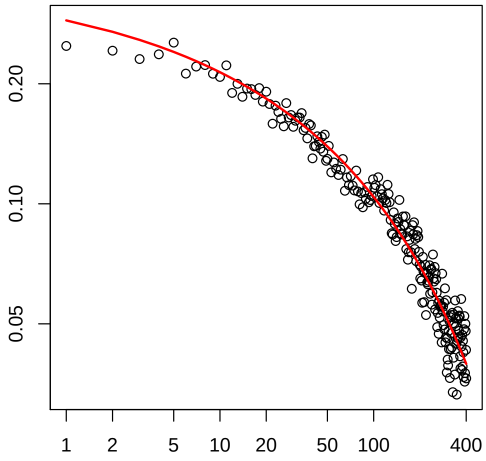
We can now define a *loss functional* $\mathcal{L}$ as follows:

$$\begin{aligned}\mathcal{L}(D, \lambda, E(\sigma), E(\sigma^2)) &= \frac{1}{2}\left\{\left(\frac{\widehat{C}_1}{C_1} - 1\right)^2 + \left(\frac{\widehat{C}_2}{C_2} - 1\right)^2\right\} + \frac{1}{20}\sum_{k=1}^{20}\left(\frac{\widehat{A}(k/4)}{A(k/4)} - 1\right)^2 \\ &\quad + \sum_{n=1}^{400} \frac{e^{-n/T}}{\left(\sum_{m=1}^{400} e^{-m/T}\right)}\left(\frac{\widehat{\rho}_1(n)}{\rho(n)} - 1\right)^2,\end{aligned} \tag{3.3}$$

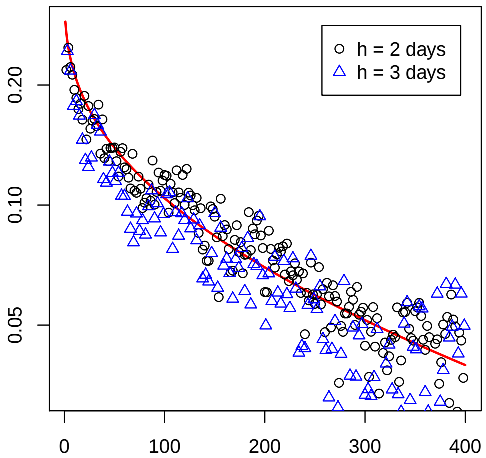
where the constant T controls a discount factor in long-range correlations. Of course, different weights for the four terms appearing in the functional could be assigned. We fix $T = 40$ (days) and we define the estimator $(\widehat{D}, \widehat{\lambda}, \widehat{E(\sigma)}, \widehat{E(\sigma^2)})$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



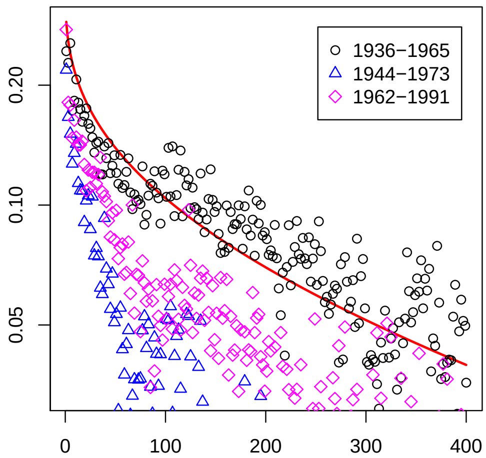
(A) *Volatility autocorrelation in the DJIA (1935-2009): log plot.*



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



(C) *Volatility autocorrelation over h days.*



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

FIGURE 3. *Volatility autocorrelation decay in the DJIA time series (1935-2009).*
(A) Log plot for the empirical 1-day volatility autocorrelation $\widehat{\rho}_1(t)$ in the period 1935-2009 (circles) and the theoretical prediction $\rho(t)$ (line), as functions of t (days). For clarity, only one data out of two is plotted.
(B) Same as (A), but log-log plot instead of log plot. For clarity, for $t \geq 50$ only one data out of two is plotted.
(C) The empirical h-day volatility autocorrelation $\widehat{\rho}_h(t)$ in the period 1935-2009 for $h = 2$ and $h = 3$ days (see legend) and the theoretical prediction $\rho(t)$ (line) as a function of t (days).
(D) The empirical 1-day volatility autocorrelation $\widehat{\rho}_1(t)$ evaluated in subperiods of 30 years (see legend) and the theoretical prediction $\rho(t)$ (line) as a function of t (days). For clarity, only one data out of two is plotted.
Note. The theoretical curve $\rho(t)$ is evaluated for the parameters values given in (3.4).

where the functional $\mathcal{L}$ attains its overall minimum, that is

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

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

We have then proceeded to the numerical study of the functional $\mathcal{L}$, which appears to be quite regular. With the help of the software Mathematica [15], we have obtained the following estimates for the parameters:

$$\widehat{D} \simeq 0.16, \quad \widehat{\lambda} \simeq 0.00097, \quad \widehat{E(\sigma)} \simeq 0.108, \quad \widehat{E(\sigma^2)} \simeq 0.117 \simeq \left(\widehat{E(\sigma)}\right)^2 \tag{3.4}$$

Using these values of the parameters, we find a very good agreement between the empirical and the theoretical quantities we have considered:

- the empirical values of the multiscaling constants are $C_1 \simeq 6.62\cdot 10^{-3}$, $C_2 \simeq 9.23\cdot 10^{-5}$, while the theoretical values are $\widehat{C}_1 \simeq 6.31 \cdot 10^{-3}$, $\widehat{C}_2 \simeq 9.38 \cdot 10^{-5}$;
- the theoretical and empirical multiscaling exponents $A(q)$ and $\widehat{A}(q)$ are remarkably close, see Figure 2(A) for a graphical comparison;
- the fit is very good also for the volatility autocorrelation functions $\rho(t)$ and $\widehat{\rho}_1(t)$, see Figures 3(A) (log plot) and 3(B) (log-log plot) for a graphical comparison, which also shows that the decay is faster than polynomial but slower than exponential. From Figure 3(C) one sees that the agreement between $\rho(t)$ and $\widehat{\rho}_h(t)$ is still good also for $h = 2$ and $h = 3$ days.

We point out that, evaluating the empirical multiscaling exponent $\widehat{A}(q)$ and the empirical volatility autocorrelation $\widehat{\rho}_1(t)$ over subperiods or 7500 data (roughly 30 years) of the DJIA time series, one finds a considerable amount of variability, cf. Figure 2(B) and Figure 3(D). This indicates that some uncertainty in the estimation of the parameters is unavoidable, at least for time series of the given length. We discuss this issue in subsection 3.4 below, where we show that *a comparable variability* is observed in data series simulated from our model.

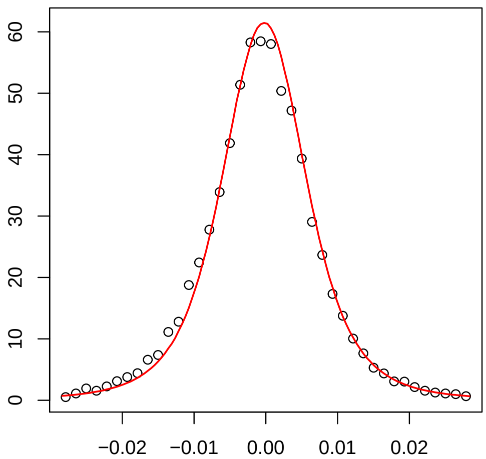
Remark 7. Baldovin and Stella [5, 18] find $D \simeq 0.24$ for their (different) model.

3.3. **The distribution of log-returns.** Having found that $\widehat{E(\sigma^2)} \simeq (\widehat{E(\sigma)})^2$ for the estimated parameters, cf. (3.4), the estimated variance of σ is equal to zero, that is σ is a constant. In particular, the model is now completely specified.

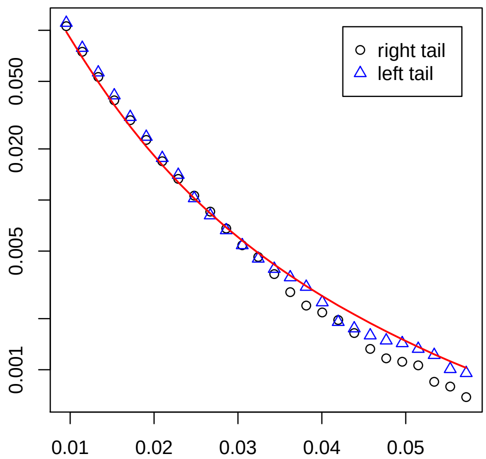
We have then compared the theoretical distribution $p_t(\cdot) := P(X_t \in \cdot) = P(X_t - X_0 \in \cdot)$ of our model for $t = 1$ (daily log-return) with the analogous quantity evaluated on the DJIA time series, i.e., the empirical distribution $\widehat{p}_t(\cdot)$ of the sequence $(x_{i+t} - x_i)_{0\leq i\leq N-t}$:

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

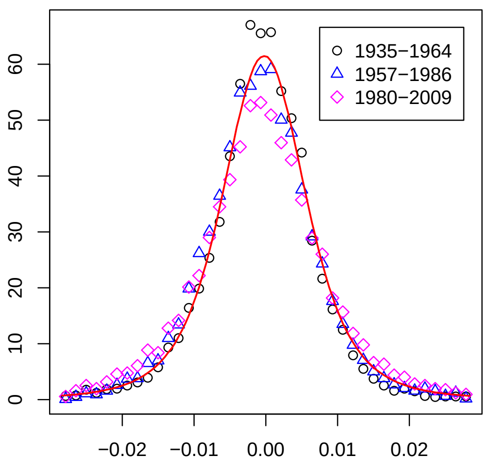
Although we do not have an analytic expression for the theoretical distribution $p_t(\cdot)$, it can be evaluated numerically via Monte Carlo simulations. In Figure 4(A) we have plotted the bulk of the distributions $p_1(\cdot)$ and $\widehat{p}_1(\cdot)$ or, more precisely, the corresponding densities in the range $[-3\hat{s}, +3\hat{s}]$, where $\hat{s} \simeq 0.0095$ is the standard deviation of $\widehat{p}_1(\cdot)$ (i.e., the empirical standard deviation of the daily log returns evaluated on the DJIA time series). In Figure 4(B) we have plotted the integrated tail of $p_1(\cdot)$, that is the function $z \mapsto P(X_1 > z) = P(X_1 < -z)$



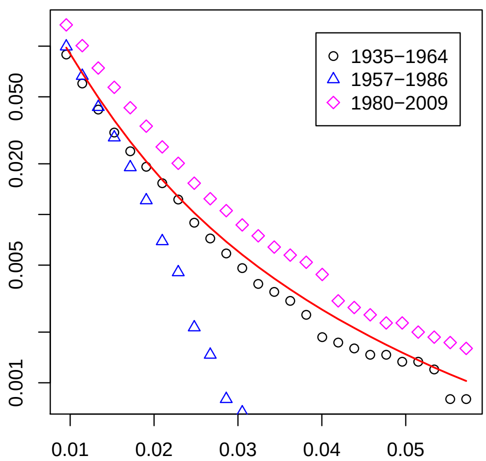
(A) *Density of the empirical distribution of the DJIA log-returns (1935-2009).*



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



(C) *Density in subperiods of 30 years.*



(D) *Left tail in subperiods of 30 years.*

FIGURE 4. *Distribution of 1-day log-returns in the DJIA time series (1935-2009).* (A) The density of the log-return empirical distribution $\hat{p}_1(\cdot)$ in the period 1935-2009 (circles) and the theoretical prediction $p_1(\cdot)$ (line).
(B) The integrated right and left tails of the log-return empirical distribution $\hat{p}_1(\cdot)$ in the period 1935-2009 (see legend) and the theoretical prediction $p_1(\cdot)$ (line).
(C) The density of the log-return empirical distribution $\hat{p}_1(\cdot)$ evaluated in subperiods of 30 years (see legend) and the theoretical prediction $p_1(\cdot)$ (line).
(D) The integrated left tail of the log-return empirical distribution $\hat{p}_1(\cdot)$ evaluated in subperiods of 30 years (see legend) and the theoretical prediction $p_1(\cdot)$ (line).
Note. The theoretical distribution $p_1(\cdot)$ is evaluated for the parameters values given in (3.4). The empirical standard deviation is $\hat{s} = \int x \, \hat{p}_1(\mathrm{d}x) \simeq 0.0095$; the range of plots (A) and (C) is $(-3\hat{s}, 3\hat{s})$ while the range of plots (B) and (D) is $(\hat{s}, 6\hat{s})$.

(note that $X_t \sim -X_t$ for our model) and the right and left integrated empirical tails $\widehat{R}(z)$ and $\widehat{L}(z)$ of $\widehat{p}_1(\cdot)$, defined for $z \geq 0$ by

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

in the range $z \in [\hat{s}, 6\hat{s}]$. As one can see, the agreement between the empirical and theoretical distributions is remarkably good for both figures, especially if one considers that *no parameter has been estimated using these curves.*

Figures 4(C) and 4(d) show that some amount of variability — less pronounced than for the multiscaling of moments and for the correlations — is present also in the empirical distribution $\widehat{p}_1(\cdot)$ evaluated over subperiods of 7500 data (roughly 30 years).

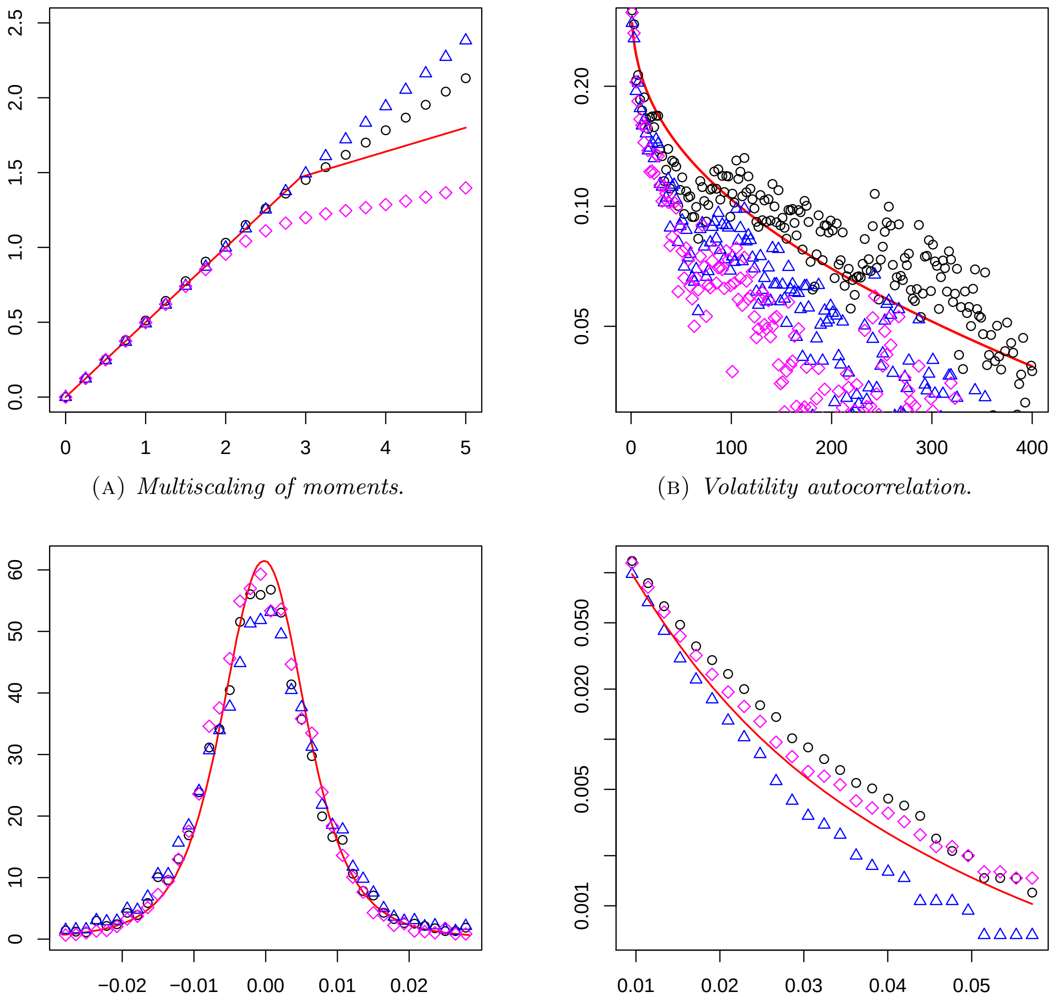
Remark 8. At this point, a fundamental difference between our model an that proposed by Baldovin and Stella is fully revealed. Baldovin and Stella's *starting point* is the non-Gaussian density $g(\cdot)$ of the log-return distribution: as a first step, they estimate $g(\cdot)$ on the DJIA time series and build a basic process having $g(\cdot)$ as marginal density, cf. (1.5) and (1.6); afterwards, they build a new process via nonlinear time change, in order to break exchangeability and produce decay of correlations.

On the other hand, when σ is constant (as for the DJIA), the basic process on which our model is built is just Brownian motion with constant volatility $(\sigma W_t)_{t \geq 0}$. Thus, the non-Gaussianity of the law of X_1 in our model is not imposed *a priori*, but is produced by the randomness of the time change, i.e. by the distribution of the shocks times $\mathcal{T}$. Having made the most natural choice for this distribution, namely that of a Poisson process, the fact that *a posteriori* we get an excellent fit with the empirical law of X_1 is quite remarkable.

Remark 9. It should be stressed that, even if we had found $\widehat{Var}(\sigma) := \widehat{E(\sigma^2)} - (\widehat{E(\sigma)})^2 > 0$ (as could happen for other indexes), detailed properties of the distribution of σ are not expected to be detectable from data. Indeed, the estimated mean distance between the successive epochs $(\tau_n)_{n \geq 0}$ of the Poisson process $\mathcal{T}$ is $1/\widehat{\lambda} \simeq 1031$ days, cf. (3.4). Therefore, in a time period of the length of the DJIA time series we are considering, only $18849/1031 \simeq 18$ variables σ_k are expected to be sampled, which is certainly not enough to allow more than a rough estimation of the distribution of σ. This should be viewed more as a robustness than a limitation of our model: even when $\widehat{Var}(\sigma) > 0$, the first two moments of σ contain the information which is relevant for application to real data.

Remark 10. Our model predicts an even distribution for X_t. On the other hand, it is known that DJIA data exhibit a small but nonzero skewness. At the discrete-time level, and therefore at the level of simulations, skewness could be easily introduced in the framework of our model. In fact, recalling that *conditionally* on $\mathcal{T}$ the increments ΔX of our model are centered Gaussian random variables, it would suffice to replace the distribution of the increments which follow closely the events of $\mathcal{T}$ — those that are most likely to be large — by a skew "deformation" of a Gaussian.

3.4. **Variability of estimators.** Generally speaking, it would be desirable to have "good" estimators, whose value does not change much if we compute them in sufficiently long but different time windows of the time series. On the contrary, we have already noticed that *considerable fluctuations* in different time windows of the DJIA time series are observed for the estimators we have used, namely the empirical scaling exponent $\widehat{A}(q)$, cf. Figure 2(A), the empirical volatility autocorrelation $\widehat{\rho}_1(t)$, cf. Figure 3(D), and — to a less extent — the empirical distribution $\widehat{p}_1(\cdot)$ of the 1-day log-return, cf. Figure 4(C) and 4(d). This is linked



(A) *Multiscaling of moments.*

(B) *Volatility autocorrelation.*

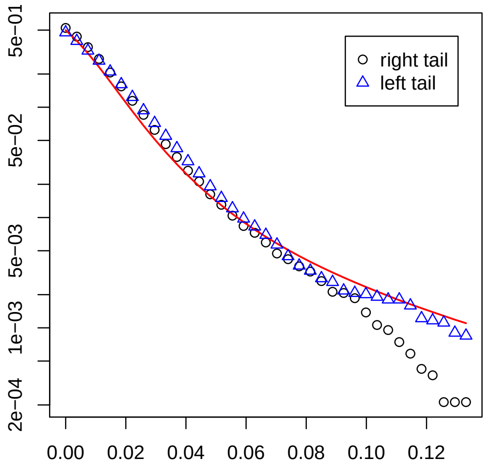
(C) *Density of the log-return empirical distribution.*

(D) *Tails of the log-return empirical distribution.*

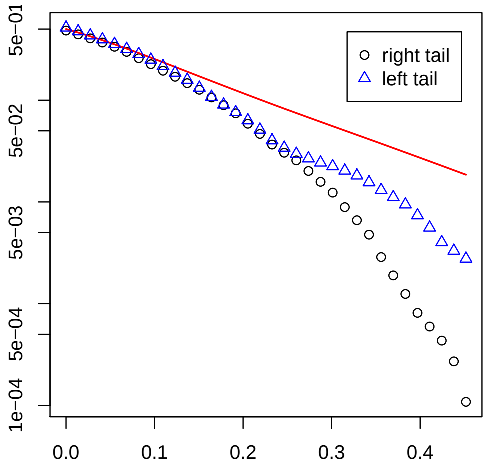
FIGURE 5. *Simulations over 75 years: variability in subperiods of 30 years.* Empirical evaluation of some observables in subperiods or 30 years for a 75-years-long time series sampled from our model $(X_t)_{t\geq 0}$ (with parameters fixed as in (3.4)): the scaling exponent $\widehat{A}(q)$ (A), the volatility autocorrelation $\widehat{\rho}_1(t)$ (B), the density (C) and the integrated tails (D) of the daily log-return distribution $\widehat{p}_1(\cdot)$.

to the fact that relation (2.11) is an equality only for $N \to \infty$, therefore there is a priori no guarantee that the DJIA time series over a few decades is *close to the ergodic limit*, i.e., long enough to guarantee a reasonably approximate equality.

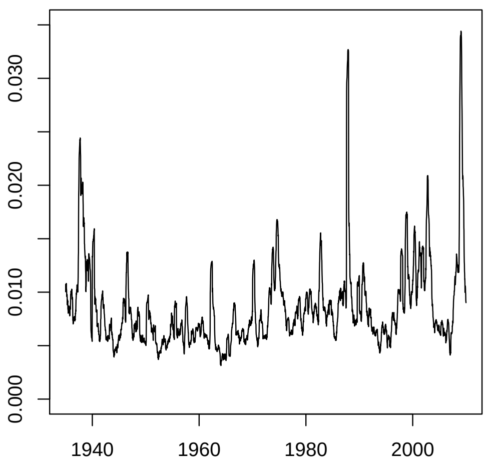
It is relevant to show that also this aspect is consistent with our model. In fact, it could happen that, unlike the DJIA, samples from our model over a few decades are long enough to damp out fluctuations. We have therefore simulated time series distributed according to our model, of the same length as the DJIA time series we are considering (18849 daily data), and we have evaluated the relevant quantities $\widehat{A}(q)$, $\widehat{\rho}_1(t)$ and $\widehat{p}_1(\cdot)$ over subperiods of 7500 days. Figure 5 shows that the variability displayed in such a sample is *qualitatively*



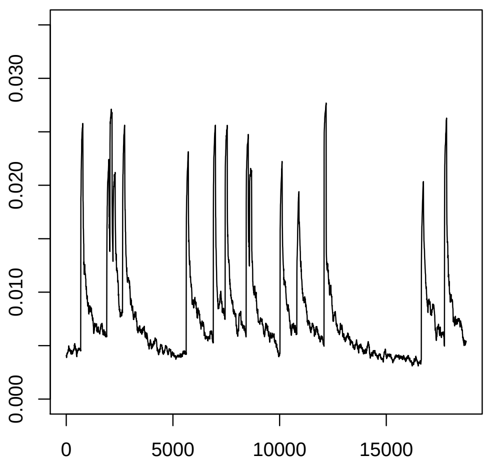
(A) *Tails of the empirical distribution of the DJIA log-returns over 5 days.*



(B) *Tails of the empirical distribution of the DJIA log-returns over 400 days.*



(C) *Empirical volatility of the DJIA (1935-2009).*



(D) *Empirical volatility of a simulation of our model.*

FIGURE 6. *Some features non completely caught by our model: tails and variance.*
• Integrated right and left tails (see legend) of the log-return empirical distribution $\hat{p}_h(\cdot)$ for the DJIA time series in the period 1935-2009 and the theoretical prediction $p_h(\cdot)$ (line) for $h = 5$ days (A) and $h = 400$ days (B).
• Empirical volatility for the daily log-returns of the DJIA time series in the period 1935-2009 (C) and for a simulation of our model over 75 years (D). The volatility at day i is evaluated through the formula $\left[\frac{1}{2R+1}\sum_{j=i-R}^{i+R}(x_j - x_{j-1})^2\right]^{1/2}$ with $R = 50$. The parameters of our model are fixed as in (3.4)

consistent with what is observed on the DJIA time series. We believe that this is a crucial point in showing the agreement of our model with the DJIA time series.

Remark 11. The multiscaling of empirical moments has been observed in several financial indexes in [8], where it is claimed that data provide solid arguments *against* model with linear or piecewise linear scaling exponents. Note that the theoretical scaling exponent $A(q)$

of our model is indeed piecewise linear, cf. (2.16). 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, cf. Figure 2(B). This shows that the objection against models with piecewise linear $A(q)$, raised in [8], cannot apply to the model we have proposed.

Remark 12. 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 $\widehat{A}(q)$.

3.5. **Final considerations.** The above statistical analysis shows that the model $(X_t)_{t\geq 0}$ that we propose is capable to explain several relevant features of the DJIA time series with a very good degree of accuracy, even though the fine estimation of the parameters is a delicate (and interesting) statistical problem. We close this section pointing out some empirical quantities that fit less precisely with our model.

- While for short range log-returns (over 1-5-10-50 days) the tails of the empirical distribution are close to the ones predicted by our model, cf. Figure 6(A), for longer ranges (such as 400 ore more days) some discrepancies emerge, cf. Figure 6(B).
- The graphic of the empirical volatility for the DJIA time series shows clearly peaks of various heights and of quite symmetric shape, cf. Figure 6(C). On the other hand, the peaks of the empirical volatility evaluated on time series simulated according to our model are all of about the same height and of sharply asymmetric shape, cf. Figure 6(D). While the asymmetric shape is a direct consequence of the definition of the process $(I_t)_{t\geq 0}$, cf. (2.6), the limited variability in the height of the peaks is mainly due to the fact that σ is constant.

  Numerical investigations lead us think that a higher variability in the heights of the peaks of the volatility of our model could be obtained through a slight modification of the model, using the following scaling function $f(t)$ instead of t^{2D}:

$$f(t) := \begin{cases} t^{2D} & \text{if } 0 < t < c \\ c^{2D} + 2\,D\,c^{2D-1}\,(t-c) & \text{if } t \geq c \end{cases},$$

  with c a constant parameter to be estimated. In fact, it turns out that for such a modified model the correlations decay faster than for the original model (with the same parameters). As a consequence, a non-constant σ is likely to be the right one to fit the DJIA time series, because as the variance of σ increases the theoretical volatility autocorrelation decay slows down. Of course, a non constant σ is then likely to produce more pronounced peaks.

4. Basic properties of the model

We start proving the properties (A)-(B)-(C) stated in section 2.2. We denote by $\mathcal{G}$ the σ-field generated by the whole process $(I_t)_{t\geq 0}$, which coincides with the σ-field generated by the sequences $\mathcal{T} = \{\tau_k\}_{k\geq 0}$ and $\Sigma = \{\sigma_k\}_{k\geq 0}$.

Proof of property (A). We first focus on the process $(I_t)_{t\geq 0}$, defined in (2.6). 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 that t, i.e., $i^h(t) = i(t+h)$. Recalling the definition (2.6), we can write

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

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

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

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

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

Proof of property (B). Observe first that $I'_s := \frac{\mathrm{d}}{\mathrm{d}s} I_s > 0$ a.s. and for Lebesgue–a.e. $s \geq 0$. By a change of variable, we can rewrite the process $(B_t)_{t\geq 0}$ defined in (2.9) as

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

which shows that relation (2.8) holds true. It remains to show that $(B_t)_{t\geq 0}$ is indeed a standard Brownian motion. Note that

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

Therefore, conditionally on $\mathcal{I}$, $(B_t)_{t\geq 0}$ is a centered Gaussian process (it is a Wiener integral), with conditional covariance given by

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

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

Proof of property (C). The assumption $E(\sigma^2) < \infty$ ensures that $E(|X_t|^2) < \infty$ for all $t \geq 0$, as we prove in Proposition 13 below. Let us now denote by $\mathcal{F}^X_t = \sigma(X_s, \, s \leq t)$ 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 $\mathcal{F}^X_t \vee \mathcal{G}$ the smallest σ-field containing $\mathcal{F}^X_t$ and $\mathcal{G}$. Since $E(W_{I_{t+h}} - W_{I_t} | \mathcal{F}^X_t \vee \mathcal{G}) = 0$ for all $h \geq 0$, by the basic properties of Brownian motion, recalling that $X_t = W_{I_t}$ we obtain

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

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

Next we show what was announced in (2.10).

Proposition 13. *For any $t > 0$*

$$E(|X_t|^q) < \infty \quad \Longleftrightarrow \quad E(\sigma^q) < \infty \, .$$

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

$$E(|I_t|^{q/2}) < \infty \quad \iff \quad E(\sigma^q) < \infty\,. \tag{4.1}$$

The implication "$\Rightarrow$" is easy: by the definition (2.6) of the process I we can write

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

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

Next we prove the implication "$\Leftarrow$". Note that $(a+b)^{2D} - b^{2D} \le a^{2D}$ for all $a, b \ge 0$ (we recall that $2D \le 1$). Then if $i(t) \ge 1$, i.e. $\tau_1 \le t$ we can write

$$\begin{aligned} I_t &= \sigma_{i(t)}^2 (t - \tau_{i(t)})^{2D} + \sum_{k=2}^{i(t)} \sigma_{k-1}^2 (\tau_k - \tau_{k-1})^{2D} + \sigma_0^2 \left[(\tau_1 - \tau_0)^{2D} - (-\tau_0)^{2D}\right] \\ &\le \sigma_{i(t)}^2 (t - \tau_{i(t)})^{2D} + \sum_{k=2}^{i(t)} \sigma_{k-1}^2 (\tau_k - \tau_{k-1})^{2D} + \sigma_0^2 \tau_1^{2D}\,, \end{aligned} \tag{4.2}$$

where we agree that the sum over k is zero if $i(t) = 1$. Since $\tau_k \le t$ for all $k \le i(t)$, by the definition (2.5) of $i(t)$, relation (4.2) yields the bound $I_t \le t^{2D} \sum_{k=0}^{i(t)} \sigma_k^2$, which holds clearly also when $i(t) = 0$. In conclusion, we have shown that for all $t, q > 0$

$$|I_t|^{q/2} \le t^{Dq} \left(\sum_{k=0}^{i(t)} \sigma_k^2 \right)^{q/2}\,. \tag{4.3}$$

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

$$\left(\sum_{k=0}^{i(t)} \sigma_k^2 \right)^{q/2} = (i(t)+1)^{q/2} \left(\frac{1}{i(t)+1} \sum_{k=0}^{i(t)} \sigma_k^2 \right)^{q/2} \le (i(t)+1)^{q/2-1} \sum_{k=0}^{i(t)} \sigma_k^q\,. \tag{4.4}$$

Therefore, if $E(\sigma^q) < \infty$, from (4.3) we can write

$$E(|I_t|^{q/2}) \le t^{Dq}\, E((i(t)+1)^{q/2})\, E(\sigma^q) < \infty\,, \tag{4.5}$$

because $i(t)$, being a Poisson random variable of mean λt, has finite moments of all orders and is independent of the variables $(\sigma_k)_{k \ge 0}$. Analogously, since $(\sum_{k=1}^\infty x_k)^\alpha \le \sum_{k=1}^\infty x_k^\alpha$ for $\alpha \in (0,1)$ and for every non-negative sequence $(x_n)_{n \in \mathbb{N}}$, for $q \le 2$ by (4.3) we can write

$$E(|I_t|^{q/2}) \le t^{Dq}\, E\left(\sum_{k=0}^{i(t)} \sigma_k^q \right) \le E((t-\tau_0)^{Dq})\, E(i(t)+1)\, E(\sigma^q) < \infty\,. \tag{4.6}$$

The proof of (4.1) is complete. ☐

We now relate the exponential moments of σ with those of X_t.

Proposition 14. *Regardless of the distribution of σ, for every $q > (1-D)^{-1}$ we have*

$$E\left[\exp\left(\gamma |X_t|^q\right)\right] = \infty\,, \qquad \forall t > 0\,,\ \forall \gamma > 0\,. \tag{4.7}$$

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

$$E\left[\exp\left(\gamma |X_t|^q\right)\right] < \infty \quad \forall \gamma > 0 \qquad \iff \qquad E\left[\exp\left(\alpha \sigma^{\frac{2q}{2-q}}\right)\right] < \infty \quad \forall \alpha > 0\,, \tag{4.8}$$

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

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

Remark 15. If ν has a bounded support, Proposition 14 implies that $E\left[\exp\left(\gamma|X_t|^q\right)\right] < +\infty$ for $q < (1-D)^{-1}$, and $E\left[\exp\left(\gamma|X_t|^q\right)\right] = +\infty$ for $q > (1-D)^{-1}$. This means that the tails of the distribution of X_t are heavier than Gaussian but lighter than exponential, *seemingly* in contrast with the red curve in Figure 4(B), that shows a decay *slower* than exponential for the distribution of the daily log-returns of our model (with parameters chosen to fit the DJIA time series, see (3.4)). As a matter of fact, simulations show that the predicted super-exponential decay of the tails can actually be seen:

- at a much larger distance from the mean than six standard deviations (which is the range of Figure 4(B)) for *daily* log-returns;
- at a more reasonable distance from the mean for log-returns over longer periods, such as hundreds of days (cf. the red curve in Figure 6(B)).

As both cases appear to have little interest in applications, the *relevant* tails are heavier than exponential, despite of Proposition 14.

We finally show that the distribution of its increments is *ergodic*. We actually show two stronger results; a 0–1 law for the tail σ-field of X, and the fact that the increments of X are exponentially mixing. In what follows, given a continuous or discrete-time stochastic process $\eta = (\eta_t)_{t\geq 0}$, we set $\Theta_t(\eta) := \sigma\{\eta_s : s \geq t\}$ and we define the so-called *tail* σ-field $\Theta(\eta) := \bigcap_{t\geq 0} \Theta_t(\eta)$. Also, for an interval $I \subseteq [0, +\infty)$, we let

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

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

Proposition 16. *The following properties hold:*

(a) *The tail σ-field $\Theta(X)$ is* trivial*, i.e., its events have probability either zero or one.*

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

$$|P(A \cap B) - P(A)P(B)| \leq 2e^{-\lambda(c-b)}. \tag{4.9}$$

As a consequence, equation (2.11) *holds true almost surely and in L^1, for every measurable function $F : \mathbb{R}^k \to \mathbb{R}$ such that $E[|F(X_{b_1} - X_{a_1}, \ldots, X_{b_k} - X_{a_k})|] < +\infty$.*

The proof of Proposition 16 is given in Appendix B.

5. Scaling and multiscaling: proof of Theorems 2 and 3

We observe that for all fixed $t, h > 0$ we have the equality in law $X_{t+h} - X_t \sim \sqrt{I_h}\, W_1$, as it follows by the definition of our model $(X_t)_{t\geq 0} = (W_{I_t})_{t\geq 0}$. We also observe that $i(h) = \#\{\mathcal{T} \cap (0,h]\} \sim Po(\lambda h)$, as it follows from (2.5) and the properties of the Poisson process.

5.1. **Proof of Theorem 2.** Since $P(i(h) \geq 1) = 1 - e^{-\lambda h} \to 0$ as $h \downarrow 0$, we may focus on the event $\{i(h) = 0\} = \{\mathcal{T} \cap (0,h] = \emptyset\}$, on which we have $I_h = \sigma_0^2((h-\tau_0)^{2D} - (-\tau_0)^{2D})$, with $-\tau_0 \sim Exp(\lambda)$. Since $X_{t+h} - X_t \sim \sqrt{I_h}\, W_1$, for every Borel set A we can write as $h \downarrow 0$

$$\begin{aligned} P(X_{t+h} - X_t \in A) &= P(\sqrt{I_h}\, W_1 \in A; [0,h] \cap \mathcal{T} = \emptyset) + o(1) \\ &= \int_0^\infty \nu(\mathrm{d}\sigma) \int_0^\infty \mathrm{d}s\, \lambda e^{-\lambda s}\, P\Big(\sigma\sqrt{(h+s)^{2D} - s^{2D}}\, W_1 \in A\Big) + o(1). \end{aligned}$$

Now observe that $h^{-1}((h+s)^{2D} - s^{2D}) \uparrow 2Ds^{2D-1}$ as $h \downarrow 0$, for every $s > 0$, hence, plainly,

$$\frac{1}{\sqrt{h}}\,\sigma\sqrt{(h+s)^{2D} - s^{2D}}\,W_1 \xrightarrow[h\downarrow 0]{d} \sigma\sqrt{2D}s^{D-1/2}\,W_1\,.$$

Therefore, if we choose $A = \sqrt{h}\,C$, where C is a Borel set whose boundary has zero Lebesgue measure, as $h \downarrow 0$ we have

$$P\left(\frac{X_{t+h} - X_t}{\sqrt{h}} \in C\right) = \int_0^\infty \nu(\mathrm{d}\sigma) \int_0^\infty \mathrm{d}s\,\lambda e^{-\lambda s}\, P\left(W_1 \in \frac{C}{\sigma\sqrt{2D}s^{D-1/2}}\right) + o(1).$$

Since $P(W_1 \in dx) = (2\pi)^{-1}\exp(-x^2/2)\,\mathrm{d}x$, the proof is completed. □

5.2. **Proof of Theorem 3.** Since $X_{t+h} - X_t \sim \sqrt{I_h}\,W_1$, we can write

$$E(|X_{t+h} - X_t|^q) = E(|I_h|^{q/2}|W_1|^q) = E(|W_1|^q)\,E(|I_h|^{q/2}) = c_q\,E(|I_h|^{q/2})\,, \tag{5.1}$$

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

$$E(|I_h|^{\frac{q}{2}}) = E(|I_h|^{\frac{q}{2}}\,\mathbf{1}_{\{i(h)=0\}}) + E(|I_h|^{\frac{q}{2}}\,\mathbf{1}_{\{i(h)=1\}}) + E(|I_h|^{\frac{q}{2}}\,\mathbf{1}_{\{i(h)\geq 2\}})\,. \tag{5.2}$$

For the first term in the right hand side of (5.2), we note that $P(i(h) = 0) = e^{-\lambda h} \to 1$ as $h \downarrow 0$ and that $I_h = \sigma_0^2((h-\tau_0)^{2D} - (-\tau_0)^{2D})$ on the event $\{i(h) = 0\}$. Setting $-\tau_0 =: \lambda^{-1}S$ with $S \sim Exp(1)$, we obtain as $h \downarrow 0$

$$E(|I_h|^{\frac{q}{2}}\,\mathbf{1}_{\{i(h)=0\}}) = E(\sigma^q)\,\lambda^{-Dq}\,E\big(((S+\lambda h)^{2D} - S^{2D})^{\frac{q}{2}}\big)\,\big(1+o(1)\big)\,. \tag{5.3}$$

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

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

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

$$\text{for } q < q^*: \quad \lim_{h\downarrow 0} \frac{E\big(((S+\lambda h)^{2D} - S^{2D})^{\frac{q}{2}}\big)}{\lambda^{\frac{q}{2}}\,h^{\frac{q}{2}}} \;=\; (2D)^{q/2}\,\Gamma(1 - q/q^*) \;\in\; (0,\infty)\,. \tag{5.4}$$

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

$$\begin{aligned} E\big(((S+\lambda h)^{2D} - S^{2D})^{\frac{q}{2}}\big) &= \int_0^\infty ((s+\lambda h)^{2D} - s^{2D})^{\frac{q}{2}}\,e^{-s}\,\mathrm{d}s \\ &= (\lambda h)^{Dq+1}\int_0^\infty ((1+x)^{2D} - x^{2D})^{\frac{q}{2}}\,e^{-\lambda h x}\,\mathrm{d}x\,. \end{aligned} \tag{5.5}$$

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

$$\text{for } q > q^*: \quad \lim_{h\downarrow 0} \frac{E\big(((S+\lambda h)^{2D} - S^{2D})^{\frac{q}{2}}\big)}{\lambda^{Dq+1}\,h^{Dq+1}} \;=\; \int_0^\infty ((1+x)^{2D} - x^{2D})^{\frac{q}{2}}\,\mathrm{d}x \;\in\; (0,\infty)\,. \tag{5.6}$$

Finally, in the case $q = q^*$ we have $((1+x)^{2D} - x^{2D})^{q^*/2} \sim (2D)^{q^*/2}\,x^{-1}$ as $x \to +\infty$ and we want to study the integral in the second line of (5.5). Fix an arbitrary (large) $M > 0$ and note that, integrating by parts and performing a change of variables, as $h \downarrow 0$ we have

$$\int_M^\infty \frac{e^{-\lambda hx}}{x}\,\mathrm{d}x \,=\, -\log M e^{-\lambda hM} \,+\, \lambda h \int_M^\infty (\log x)\, e^{-\lambda hx}\,\mathrm{d}x \,=\, O(1) + \int_{\lambda hM}^\infty \log\left(\frac{y}{\lambda h}\right) e^{-y}\,\mathrm{d}y$$
$$= O(1) + \int_{\lambda hM}^\infty \log\left(\frac{y}{\lambda}\right) e^{-y}\,\mathrm{d}y \,+\, \log\left(\frac{1}{h}\right) \int_{\lambda hM}^\infty e^{-y}\,\mathrm{d}y \,=\, \log\left(\frac{1}{h}\right)\,\big(1+o(1)\big)\,.$$

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

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

Coming back to (5.5), noting that $Dq+1 = \frac{q}{2}$ for $q = q^*$, it follows that

$$\lim_{h\downarrow 0} \frac{E\big(((S+h)^{2D} - S^{2D})^{\frac{q^*}{2}}\big)}{\lambda^{Dq^*+1}\, h^{\frac{q^*}{2}}\, \log(\frac{1}{h})} = (2D)^{\frac{q^*}{2}}\,. \tag{5.7}$$

Recalling (5.1) and (5.3), the relations (5.4), (5.6) and (5.7) show that the first term in the right hand side of (5.2) has the same asymptotic behavior as in the statement of the theorem, except for the regime $q > q^*$ where the constant does not match (the missing contribution will be obtained in a moment).

We now focus on the second term in the right hand side of (5.2). Note that, conditionally on the event $\{i(h) = 1\} = \{\tau_1 \le h, \tau_2 > h\}$, we have

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

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

$$E(|I_h|^{\frac{q}{2}}\, \mathbf{1}_{\{i(h)=1\}}) = \lambda h^{Dq+1}\, E\left[\left(\sigma_1^2(1-U)^{2D} + \sigma_0^2\left(\left(U + \frac{S}{\lambda h}\right)^{2D} - \left(\frac{S}{\lambda h}\right)^{2D}\right)^{\frac{q}{2}}\right)\right]. \tag{5.8}$$

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

$$\lim_{h\downarrow 0} \frac{E(|I_h|^{\frac{q}{2}}\, \mathbf{1}_{\{i(h)=1\}})}{h^{Dq+1}} \,=\, \lambda E(\sigma_1^q)\, E\big((1-U)^{Dq}\big) \,=\, \lambda E(\sigma_1^q)\, \frac{1}{Dq+1}\,. \tag{5.9}$$

This shows that the second term in the right hand side of (5.2) gives a contribution of the order h^{Dq+1} as $h \downarrow 0$. This is relevant only for $q > q^*$, because for $q \le q^*$ the first term gives a much bigger contribution of the order $h^{q/2}$ (see (5.4) and (5.7)). Recalling (5.1), it follows from (5.9) and (5.6) that the contribution of the first and the second term in the right hand side of (5.2) matches the statement of the theorem (including the constant).

It only remains to show that the third term in the right hand side of (5.2) gives a negligible contribution. Consider first the case $q > 2$: from the basic bound (4.3) and Jensen's inequality, in analogy with (4.4) and (4.5), we obtain

$$E\big(|I_h|^{q/2}\, \mathbf{1}_{\{i(h)\ge 2\}}\big) \,\le\, t^{Dq}\, E(\sigma^q)\, E\big((i(h)+1)^{q/2}\, \mathbf{1}_{\{i(h)\ge 2\}}\big)\,. \tag{5.10}$$

On the other hand, since for $\alpha \in (0,1)$ we have $(\sum_{k=1}^\infty x_k)^\alpha \le \sum_{k=1}^\infty x_k^\alpha$, when $q \le 2$ from (4.3) we can write, in analogy with (4.6),

$$E\big(|I_h|^{q/2}\, \mathbf{1}_{\{i(h)\ge 2\}}\big) \,\le\, t^{Dq}\, E(\sigma^q)\, E\big((i(h)+1)\, \mathbf{1}_{\{i(h)\ge 2\}}\big)\,. \tag{5.11}$$

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

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

because $E\big((i(1)+1)^{3a}\big) < \infty$ (recall that $i(h) \sim Po(\lambda)$) and $(1 - e^{-\lambda h} - e^{-\lambda h}\lambda h) \sim \frac{1}{2}(\lambda h)^2$ as $h \downarrow 0$. Then it follows from (5.10) and (5.11) that

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

This shows that the contribution of the third term in the right hand side of (5.2) is always negligible with respect to the contribution of the second term (recall (5.9)). $\square$

6. Decay of correlations: proof of Theorem 5

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

We have to prove (2.17). Plainly, by translation invariance we can set $s = 0$ without loss of generality. We also assume that $h < t$. We start writing

$$\begin{aligned} &Cov(|X_h|, |X_{t+h} - X_t|) \\ &\quad = Cov\big(E\big(|X_h|\big|\mathcal{G}\big)\,,\, E\big(|X_{t+h} - X_t|\big|\mathcal{G}\big)\big) \;+\; E\big(Cov\big(|X_h|, |X_{t+h} - X_t|\big|\mathcal{G}\big)\big)\,. \end{aligned} \tag{6.1}$$

We recall that $X_t = W_{I_t}$ and the process $(I_t)_{t\ge 0}$ is $\mathcal{G}$-measurable and independent of the process $(W_t)_{t\ge 0}$. It follows that, conditionally on $(I_t)_{t \ge 0}$, the process $(X_t)_{t\ge 0}$ has independent increments, hence the second term in the right hand side of (6.1) vanishes, because $Cov(|X_h|, |X_{t+h} - X_t||\mathcal{G}) = 0$ a.s.. For fixed h, from the equality in law $X_h = W_{I_h} \sim \sqrt{I_h}\, W_1$ it follows that $E(|X_h||\mathcal{G}) = c_1 \sqrt{I_h}$, where $c_1 = E(|W_1|) = \sqrt{2/\pi}$. Analogously $E(|X_{t+h} - X_t||\mathcal{G}) = \sqrt{I_{t+h} - I_t}$ and (6.1) reduces to

$$Cov(|X_h|, |X_{t+h} - X_t|) \;=\; \frac{2}{\pi}\, Cov\big(\sqrt{I_h}, \sqrt{I_{t+h} - I_t}\big)\,. \tag{6.2}$$

Recall the definitions (2.5) and (2.6) of the variabes $i(t)$ and I_t. We now claim that we can replace $\sqrt{I_{t+h} - I_t}$ by $\sqrt{I_{t+h} - I_t}\, \mathbf{1}_{\{\mathcal{T} \cap (h,t] = \emptyset\}}$ in (6.2). In fact from (2.6) we can write

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

where we agree that the sum in the right hand side is zero if $i(t+h) = i(t)$. This shows that $(I_{t+h} - I_t)$ is a function of the variables τ_k, σ_k with index $i(t) \le k \le i(t+h)$. Since $\{\mathcal{T} \cap (h,t] \ne \emptyset\} = \{\tau_{i(t)} > h\}$, this means that $\sqrt{I_{t+h} - I_t}\, \mathbf{1}_{\{\mathcal{T} \cap (h,t] \ne \emptyset\}}$ is $\mathcal{G}_{(h,t+h]}$-measurable, hence independent of $\sqrt{I_h}$, which is clearly $\mathcal{G}_{(-\infty,h]}$-measurable. This shows that $Cov(\sqrt{I_h}, \sqrt{I_{t+h} - I_t}\, \mathbf{1}_{\{\mathcal{T} \cap (h,t] \ne \emptyset\}}) = 0$, therefore from (6.2) we can write

$$Cov(|X_h|, |X_{t+h} - X_t|) \;=\; \frac{2}{\pi}\, Cov\big(\sqrt{I_h}, \sqrt{I_{t+h} - I_t}\, \mathbf{1}_{\{\mathcal{T} \cap (h,t] = \emptyset\}}\big)\,. \tag{6.3}$$

It is convenient to define

$$j(t) := \inf\{k \in \mathbb{Z} : \tau_k > t\} = i(t) + 1\,,$$

and note that the variables $\tau_{j(t)}$, $\tau_{j(t)+1}$, ... are $\mathcal{G}_{(t,\infty)}$-measurable. We now introduce a variable $A_{t,h}$, which is essentially $I_{t+h} - I_t$ with $\tau_{i(t)}$ replaced by $\tau_{i(h)}$. More precisely, we set

$$A_{t,h} := \sigma^2_{i(h)} \big[(t + h - \tau_{i(h)})^{2D} - (t - \tau_{i(h)})^{2D}\big] \qquad \text{if } \mathcal{T} \cap (t, t + h] = \emptyset\,, \tag{6.4}$$

and

$$\begin{aligned} A_{t,h} &:= \sigma^2_{i(t+h)}(t + h - \tau_{i(t+h)})^{2D} + \sum_{k=j(t)+1}^{i(t+h)} \sigma^2_{k-1}(\tau_k - \tau_{k-1})^{2D} \\ &\quad + \sigma^2_{i(h)}\big[(\tau_{j(t)} - \tau_{i(h)})^{2D} - (t - \tau_{i(h)})^{2D}\big] \qquad \text{if } \mathcal{T} \cap (t, t + h] \neq \emptyset\,, \end{aligned} \tag{6.5}$$

where we agree that the sum in the right hand side of (6.5) is zero unless $i(t + h) > j(t)$, that is $\#(\mathcal{T} \cap (t, t + h]) \geq 2$. We stress that $A_{t,h}$ is $\mathcal{G}_{(-\infty,h]\cup(t,t+h]}$-measurable: in fact, $A_{t,h}$ is nothing but $I_{t+h} - I_t$ when we erase from $\mathcal{T}$ the points τ_k (if any) falling in $(h, t]$. In particular, on the event $\{\mathcal{T} \cap (h, t] = \emptyset\}$ we can replace $I_{t+h} - I_t$ by $A_{t,h}$, rewriting (6.3) as

$$\begin{aligned} Cov(|X_h|, |X_{t+h} - X_t|) &= \frac{2}{\pi}\, Cov\big(\sqrt{I_h}, \sqrt{A_{t,h}}\, \mathbf{1}_{\{\mathcal{T}\cap(h,t]=\emptyset\}}\big) \\ &= \frac{2}{\pi}\, E\big(\big(\sqrt{I_h} - E(\sqrt{I_h})\big) \cdot \sqrt{A_{t,h}} \cdot \mathbf{1}_{\{\mathcal{T}\cap(h,t]=\emptyset\}}\big) \\ &= \frac{2}{\pi}\, e^{-\lambda(t-h)}\, E\big(\big(\sqrt{I_h} - E(\sqrt{I_h})\big) \cdot \sqrt{A_{t,h}}\,\big)\,, \end{aligned} \tag{6.6}$$

where in the last equality we have used the independence of the σ-algebras $\mathcal{G}_{(-\infty,h]\cup(t,\infty)}$ and $\mathcal{G}_{(h,t]}$ and the fact that $P(\mathcal{T} \cap (h, t] = \emptyset) = P(i(t) - i(h) = 0) = e^{-\lambda(t-h)}$.

We now focus on

$$E\big(\big(\sqrt{I_h} - E(\sqrt{I_h})\big) \cdot \sqrt{A_{t,h}}\,\big) = Cov\big(\sqrt{I_h}\,,\, \sqrt{A_{t,h}}\,\big)\,, \tag{6.7}$$

that we can write as the sum of four terms:

$$\begin{aligned} Cov\big(\sqrt{I_h}\,,\, \sqrt{A_{t,h}}\,\big) &= Cov\big(\sqrt{I_h}\, \mathbf{1}_{\{\mathcal{T}\cap[0,h]=\emptyset\}}\,,\, \sqrt{A_{t,h}}\, \mathbf{1}_{\{\mathcal{T}\cap[t,t+h]=\emptyset\}}\big) \\ &\quad + Cov\big(\sqrt{I_h}\, \mathbf{1}_{\{\mathcal{T}\cap[0,h]\neq\emptyset\}}\,,\, \sqrt{A_{t,h}}\, \mathbf{1}_{\{\mathcal{T}\cap[t,t+h]=\emptyset\}}\big) \\ &\quad + Cov\big(\sqrt{I_h}\, \mathbf{1}_{\{\mathcal{T}\cap[0,h]=\emptyset\}}\,,\, \sqrt{A_{t,h}}\, \mathbf{1}_{\{\mathcal{T}\cap[t,t+h]\neq\emptyset\}}\big) \\ &\quad + Cov\big(\sqrt{I_h}\, \mathbf{1}_{\{\mathcal{T}\cap[0,h]\neq\emptyset\}}\,,\, \sqrt{A_{t,h}}\, \mathbf{1}_{\{\mathcal{T}\cap[t,t+h]\neq\emptyset\}}\big)\,. \end{aligned} \tag{6.8}$$

The first term in the right hand side gives the leading contribution:

$$\begin{aligned} &Cov\big(\sqrt{I_h}\, \mathbf{1}_{\{\mathcal{T}\cap[0,h]=\emptyset\}}\,,\, \sqrt{A_{t,h}}\, \mathbf{1}_{\{\mathcal{T}\cap[t,t+h]=\emptyset\}}\big) \\ &\ = Cov\big(\sigma_0\sqrt{(h - \tau_0)^{2D} - (-\tau_0)^{2D}}\, \mathbf{1}_{\{\mathcal{T}\cap[0,h]=\emptyset\}}\,,\, \sigma_0\sqrt{(t + h - \tau_0)^{2D} - (t - \tau_0)^{2D}}\, \mathbf{1}_{\{\mathcal{T}\cap[t,t+h]=\emptyset\}}\big) \\ &\ = \lambda^{-2D}\, e^{-2\lambda h}\, Cov\big(\sigma_0\sqrt{(\lambda h + S)^{2D} - S^{2D}}\,,\, \sigma_0\sqrt{(\lambda t + \lambda h + S)^{2D} - (\lambda t + S)^{2D}}\big)\,, \end{aligned}$$

where $S := \lambda(-\tau_0) \sim Exp(1)$ and we have used the fact that the variables τ_0, $\mathbf{1}_{\{\mathcal{T}\cap[t,t+h]=\emptyset\}}$ and $\mathbf{1}_{\{\mathcal{T}\cap[0,h]=\emptyset\}}$ are independent. Since $\delta^{-1}((\delta + x)^{2D} - x^{2D}) \uparrow 2Dx^{2D-1}$ as $\delta \downarrow 0$, by monotone convergence we obtain

$$\frac{1}{h}\, Cov\big(\sqrt{I_h}\, \mathbf{1}_{\{\mathcal{T}\cap[0,h]=\emptyset\}}\,,\, \sqrt{A_{t,h}}\, \mathbf{1}_{\{\mathcal{T}\cap[t,t+h]=\emptyset\}}\big) \xrightarrow{h\downarrow 0} 2D\, \lambda^{1-2D}\, Cov\big(\sigma S^{2D-1}, \sigma(S{+}\lambda t)^{2D-1}\big)\,,$$

in agreement with (2.17) and (2.18).

It remains to show that the second, the third and the fourth term in the right hand side of (6.8) give a negligible contribution. Using the relation $|Cov(X,Y)| \le \sqrt{E(X^2)E(Y^2)}$, valid for any couple of random variables X, Y, by the Cauchy-Schwarz inequality we can write

$$\big|Cov\big(\sqrt{I_h}\,\mathbf{1}_C\,,\,\sqrt{A_{t,h}}\,\mathbf{1}_D\big)\big| \;\le\; \sqrt{E(I_h\,\mathbf{1}_C)\,E(A_{t,h}\,\mathbf{1}_D)} \;\le\; \big(E(I_h^2)\,E(A_{t,h}^2)\,P(C)\,P(D)\big)^{1/4},$$

for any choice of the events C, D. We now claim that $E(A_{t,h}^2) \le E(I_h^2)$. In fact, from (6.4) and (6.5) we see that $A_{t,h}$ coincides with $I_{t+h} - I_t$ once we replace $\tau_{i(h)}$ by $\tau_{i(t)}$. Note that $\tau_{i(h)}$ appears in $A_{t,h}$ in a term of the form $(b-\tau_{i(h)})^{2D} - (a-\tau_{i(h)})^{2D}$, for suitable $a < b$, hence replacing $\tau_{i(h)}$ by $\tau_{i(t)}$ we get something bigger, because $\tau_{i(h)} \le \tau_{i(t)}$ (we assume that $h < t$). We have thus shown that $A_{t,h} \le I_{t+h} - I_t \sim I_h$, whence $E(A_{t,h}^2) \le E(I_h^2)$. By Theorem 3 we have $E(I_h^2) \le (const.)\,h^2$ (observe that $q^* > 2$), therefore

$$Cov\big(\sqrt{I_h}\,\mathbf{1}_C\,,\,\sqrt{A_{t,h}}\,\mathbf{1}_D\big) \;\le\; (const.)\,h \cdot \big(P(C)\,P(D)\big)^{1/4}. \tag{6.9}$$

To deal with the second, third and fourth term in the right hand side of (6.8), we must take either $C = \{\mathcal{T} \cap [0,h] \ne \emptyset\}$ or $D = \{\mathcal{T} \cap [t, t+h] \ne \emptyset\}$. In any case $P(C)P(D) \le \lambda h$ and therefore the right hand side of (6.9) is $o(h)$, uniformly in $t \in \mathbb{R}$ (it does not depend on t). This completes the proof. □

Appendix A. Proof of Proposition 14

We first need two simple technical lemmas.

Lemma 17. *For $0 < q < 2$, consider the function $\varphi_q : [0, +\infty) \to [0, +\infty)$ define by*

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

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

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

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

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

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

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

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

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

Lemma 18. *Let $X_1, X_2, \ldots, X_n$ be independent random variables uniformly distributed in $[0,1]$, and $U_1 < U_2 < \ldots < U_n$ be the associated order statistics. For $n \geq 2$ and $k = 2, \ldots, n$, set $\xi_k := U_k - U_{k-1}$. Then, for every $\epsilon > 0$*

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

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

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

see [19] for a proof. $\square$

Proof of Proposition 14. Since $X_t = W_{I_t}$ and $\sqrt{I_t}\, W_1$ have the same law, we can write

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

We begin with the proof of (4.8), hence we work in the regime $q < (1-D)^{-1}$, or $q = (1-D)^{-1}$ and $D < \frac{1}{2}$; in any case, $q < 2$. We start with the "$\Leftarrow$" implication. Since I_t and W_1 are independent, it follows by Lemma 17 that

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

for some $C, \delta > 0$. For the moment we work on the event $\{i(t) \geq 1\}$. It follows by the basic bound (4.2) that

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

where we set

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

Note that $\sum_{k=0}^{i(t)} \xi_k = 1$. It is easily shown by Lagrange multipliers that the function $(x_0, x_1, \ldots, x_n) \mapsto \sum_{k=0}^n x_k^{2D}\sigma_k^2$ subject to the constraint $\sum_{k=0}^n x_k = 1$ attains its maximum at $(x_0^*, x_1^*, \ldots, x_n^*)$, with

$$x_k^* = \frac{\sigma_k^{\frac{2}{1-2D}}}{\sum_{h=0}^n \sigma_h^{\frac{2}{1-2D}}}.$$

Then it follows from (A.3) that

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

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

$$I_1^{\frac{q}{2-q}} \ \leq \ \left(\sum_{k=0}^{i(t)} \sigma_k^{\frac{2}{1-2D}}\right)^{(1-2D)\frac{q}{2-q}} \ \leq \ \sum_{k=0}^{i(t)} \sigma_k^{\frac{2q}{2-q}}.$$

Now observe that if $i(t)=0$ we have $I_t=(t-\tau_0)^{2D}-(-\tau_0)^{2D}\le t^{2D}$. Therefore, by (A.2)

$$\begin{aligned} E\left[e^{\gamma|X_t|^q}\right] &\le C\left(E\left[\exp\left(\delta\sum_{k=0}^{i(t)}\sigma_k^{\frac{2q}{2-q}}\right)\mathbf{1}_{\{i(t)\ge 1\}}\right]+E\left[\exp\left(\delta\, t^{\frac{2Dq}{2-q}}\right)\mathbf{1}_{\{i(t)=0\}}\right]\right) \\ &\le C\left(E\left[\rho^{i(t)+1}\right]+\exp\left(\delta\, t^{\frac{2Dq}{2-q}}\right)\right), \end{aligned} \tag{A.4}$$

where we have set

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

Therefore, if $\rho<\infty$, the right hand side of (A.4) is finite, because $i(t)\sim Po(\lambda t)$ has finite exponential moments of all order. This proves the "$\Leftarrow$" implication in (4.8).

The "$\Rightarrow$" implication in (4.8) is simpler. By the lower bound in Lemma 17 we have

$$E\left[e^{\gamma|X_t|^q}\right]=E\left[\exp\left(\gamma I_t^{q/2}|W_1|^q\right)\right]\ge CE\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\right], \tag{A.5}$$

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

$$\begin{aligned} E\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\right] &\ge E\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\mathbf{1}_{\{i(t)=0\}}\right] \\ &= E\left[\exp\left(\delta\left[(t-\tau_0)^{2D}-(-\tau_0)^{2D}\right]\sigma_0^{\frac{2q}{2-q}}\right)\right]P(i(t)=0). \end{aligned} \tag{A.6}$$

Under the condition

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

the last expectation in (A.6) is infinite, since $\left[(t-\tau_0)^{2D}-(-\tau_0)^{2D}\right]>0$ almost surely and is independent of σ_0. Looking back at (A.5), we have proved the "$\Rightarrow$" implication in (4.8).

Next we prove (4.7), hence we assume that $q>(1-D)^{-1}$. Consider first the case $q<2$ (which may happen only for $D<\frac{1}{2}$). By (A.5)

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

We note that, by the definition (2.5) of I_t, we can write

$$I_t\ge\sum_{k=2}^{i(t)}\sigma_{k-1}^2(\tau_k-\tau_{k-1})^{2D}, \tag{A.7}$$

where we agree that the sum is zero if $i(t)<2$. For $n\ge 0$, we let P_n to denote the conditional probability $P(\,\cdot\,|i(t)=n)$ and E_n the corresponding expectation. Note that, under P_n, the random variables $(\tau_k-\tau_{k-1})_{k=2}^n$ have the same law of the random variables $(\xi_k)_{k=2}^n$ in Lemma 18, for $n\ge 2$. Consider the following events:

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

where $a>0$ is such that $\nu([a,+\infty))=:\rho>0$ and $\epsilon>0$ will be chosen later. Note that $P_n(A_n)=\rho^{n-1}$ while $P_n(B_n)\to 1$ as $n\to+\infty$, by Lemma 18. In particular, there is $c>0$

such that $P_n(B_n) \ge c$ for every n. Plainly, A_n and B_n are independent under P_n. We have

$$\begin{aligned}(\mathrm{A.8})\quad \psi(n) := E_n\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\right] &\ge E_n\left[\exp\left(\delta I_t^{\frac{q}{2-q}}\right)\mathbf{1}_{A_n\cap B_n}\right] \\ &\ge c\,\rho^{n-1}\exp\left[\delta a^{\frac{q}{2-q}}\left(\frac{1}{n^{1+\epsilon}}\right)^{2D\frac{q}{2-q}} n^{(1-\epsilon)\frac{q}{2-q}}\right] \\ &= c\,\rho^{n-1}\exp\left[\delta a^{\frac{q}{2-q}} n^{(1-2D-\epsilon(1+2D))\frac{q}{2-q}}\right]\end{aligned}$$

Note that $q > \frac{1}{1-D}$ is equivalent to $(1-2D)\frac{q}{2-q} > 1$, therefore ϵ can be chosen small enough so that $b := (1-2D-\epsilon(1+2D))\frac{q}{2-q} > 1$. It then follows by (A.8) that $\psi(n) \ge d\,\exp(d\,n^b)$ for every $n \in \mathbb{N}$, for a suitable $d > 0$. Therefore

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

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

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

$$(\mathrm{A.9})\qquad E\left[e^{\gamma|X_t|^q}\right] = E\left[\exp\left(\gamma I_t^{q/2}|W_1|^q\right)\right],$$

hence if $q > 2$ we have $E\left[e^{\gamma|X_t|^q}\right] = \infty$, because $E[\exp(c|W_1|^q)] = \infty$ for every $c > 0$, $I_t > 0$ almost surely and I_t is independent of W_1. On the other hand, if $q = 2$ we must have $D < \frac{1}{2}$ (recall that we are in the regime $q > (1-D)^{-1}$) and the steps leading to (A.8) have shown that in this case I_t is unbounded. It then follows again from (A.9) that $E[e^{\gamma|X_t|^2}] = \infty$. □

Appendix B. Proof of Proposition 16

We first need the following technical Lemma.

Lemma 19. *Let η and θ be two independent stochastic processes, such that both $\Theta(\eta)$ and $\Theta(\theta)$ are trivial. Then the σ-field*

$$\mathcal{G} := \bigcap_{s,t\ge 0}\,[\Theta_t(\eta)\vee\Theta_s(\theta)]$$

is trivial, where $\Theta_t(\eta)\vee\Theta_s(\theta)$ denotes the smallest σ-field containing $\Theta_t(\eta)\cup\Theta_s(\theta)$

Proof. We introduce the function space $E := \mathbb{R}^{[0,+\infty)}$, equipped with the σ-field generated by cylinder sets. If Z is a bounded $\mathcal{G}$-measurable random variable, then for every $s,t \ge 0$ we can write

$$(\mathrm{B.1})\qquad Z = f_{s,t}\left(\eta_{s+\cdot},\theta_{t+\cdot}\right),$$

where $\eta_{s+\cdot}$ denotes the process $(\eta_{s+u})_{u\ge 0}$, and $f_{s,t} : E\times E \to \mathbb{R}$ is bounded and measurable. We set $f := f_{0,0}$ and we denote by μ_η, μ_θ respectively the laws of η, θ on E.

By (B.1), for μ_η-a.e. $x \in E$ we have $f(x,\theta) = f_{0,t}(x,\theta_{t+\cdot})$, for every $t \ge 0$. Thus $f(x,\theta)$ is $\Theta(\theta)$-measurable and hence a.s. constant: $f(x,\theta) = f_1(x)$ a.s., where $f_1(x) := E(f(x,\theta))$. Therefore, for every measurable $B \subseteq E$ we can write

$$(\mathrm{B.2})\qquad \int_E f(x,y)\,\mathbf{1}_B(y)\,\mu_\theta(\mathrm{d}y) = f_1(x)\cdot\mu_\theta(B)\,,\qquad \text{for } \mu_\eta\text{-a.e. } x \in E\,.$$

Observe now that, again by (B.1), we have a.s.

$$f_1(\eta) = E(f(\eta,\theta)|\eta) = E(f_{t,0}(\eta_{t+\cdot},\theta)|\eta) = E(f_{t,0}(x,\theta))|_{x=\eta_{t+\cdot}}\,.$$

This shows that $f_1(\eta)$ is $\Theta(\eta)$-measurable and hence a.s. constant: $f_1(\eta) = c$ a.s., where $c := E(f_1(\eta)) = E(f(\eta, \theta))$. Therefore, for every measurable $A \subseteq E$ we can write

$$\int_E f_1(x)\, \mathbf{1}_A(x)\, \mu_\eta(\mathrm{d}x) = c \cdot \mu_\eta(A)\,. \tag{B.3}$$

Putting together (B.2), (B.3) and Fubini's theorem, it follows that for all measurable $A, B \subseteq E$

$$\int_{E\times E} f(x,y)\, \mathbf{1}_{A\times B}(x,y)\, (\mu_\eta \otimes \mu_\theta)(\mathrm{d}x, \mathrm{d}y) = c\, (\mu_\eta \otimes \mu_\theta)(A \times B)\,, \tag{B.4}$$

where $\mu_\eta \otimes \mu_\theta$ denotes of course the product measure of μ_η and μ_θ on $E \times E$.

Since the sets of the form $A \times B$ are a basis of the product σ-field of $E \times E$, relation (B.4) still holds true with $A \times B$ replaced by any measurable subset $C \subseteq E \times E$. Choosing $C_n^+ = \{(x,y) : f(x,y) > c + \frac{1}{n}\}$ and $C_n^- = \{(x,y) : f(x,y) < c - \frac{1}{n}\}$, it follows easily that $P(C_n^+) = P(C_n^-) = 0$ for every $n \in \mathbb{N}$, that is $f(x,y) = c$ for $(\mu_\eta \otimes \mu_\theta)$-a.e. $(x,y) \in E \times E$. Therefore $Z = f(\eta, \theta) = c$ a.s.. We have proved that every bounded $\mathcal{G}$-measurable random variable is almost surely constant, which is equivalent to the triviality of $\mathcal{G}$. □

Proof of Proposition 16. We begin with the proof of statement (a), the $0-1$ law for $\Theta(X)$. If Y is a bounded $\Theta(X)$-measurable random variable, by definition for every $t \geq 0$ it can be written in the form

$$Y = \varphi_t\left(X_{t+\cdot}\right) = \varphi_t\left(W_{I_{t+\cdot}}\right),$$

where $\varphi_t : \mathcal{C}[0, +\infty) \to \mathbb{R}$ is bounded and measurable. Now, for $s \geq 0$, set

$$Y_{t,s} = \varphi_t\left(W_{I_{t+\cdot}}\right) \mathbf{1}_{\{I_t > s\}}.$$

Note that $Y_{t,s}$ is $\Theta_s(W) \vee \Theta_t(I)$-measurable. Plainly, $I_n \uparrow +\infty$ almost surely as $n \uparrow +\infty$, therefore for every $s \geq 0$

$$\lim_{n\to+\infty} Y_{n,s} = \lim_{n\to+\infty} \varphi_n\left(W_{I_{n+\cdot}}\right) = Y.$$

It follows that Y is $\bigcap_{s,t\geq 0} \left[\Theta_s(W) \vee \Theta_t(I)\right]$-measurable. In other words

$$\Theta(X) \subseteq \bigcap_{s,t\geq 0} \left[\Theta_s(W) \vee \Theta_t(I)\right]\,,$$

and by Lemma 19 all we need to show is that both $\Theta(W)$ and $\Theta(I)$ are trivial.

The triviality of $\Theta(W)$ is well known: it follows, e.g., from Blumenthal 0-1 law (cf. Theorem 2.7.17 in [12]) applied to the Brownian motion $(sW_{1/s})_{s\geq 0}$. For the triviality of $\Theta(I)$ we proceed as follows. If $A \in \Theta(I)$, for each $n \in \mathbb{N}$

$$\mathbf{1}_A = \varphi_n(I_{n+\cdot})\,,$$

for a suitable measurable function φ_n. Recalling the definition (2.5) of the variable $i(t)$, it is clear that $i(t) \uparrow +\infty$ as $t \uparrow +\infty$, hence for every $m \in \mathbb{N}$

$$\mathbf{1}_A = \lim_{n\to+\infty} \varphi_n(I_{n+\cdot}) \mathbf{1}_{\{i(n) > m\}}\,.$$

Now observe that the function $\varphi_n(I_{n+\cdot})\mathbf{1}_{\{i(n)>m\}}$ is measurable with respect to the σ-field $\sigma(\eta_k : k \geq 0)$, where $\eta_k := (\sigma_k, \tau_{k+1} - \tau_k)$, and is invariant under any permutation of $(\eta_0, \ldots, \eta_{m-1})$. This shows that A is measurable for $\sigma(\eta_k : k \geq 0)$ and invariant by finite permutations of the η_k's. By the Hewitt-Savage 0-1 law, A has probability either zero or one. The proof of (a) is now completed.

We now prove part (b). 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 are considering $I = [a, b)$, $J = [c, d)$, with $0 \leq a < b \leq c < d$. We introduce the $\mathcal{G}_{[b,c)}$-measurable event

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

We claim that, for $A \in \mathcal{F}_I^{\mathcal{D}}$, $B \in \mathcal{F}_J^{\mathcal{D}}$, we have

$$P(A \cap B|\Gamma) = P(A)P(B). \tag{B.5}$$

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

- $\mathcal{F}_I^{\mathcal{D}}$ and $\mathcal{F}_J^{\mathcal{D}}$ are independent *conditionally* to $\mathcal{G} = \mathcal{G}_{\mathbb{R}}$. This follows immediately from the independence of W and (I_t).
- Conditionally to $\mathcal{G}$, the process $(X_t)_{t\in I}$ is a Gaussian process whose covariances are all $\mathcal{G}_I$-measurable. It follows that $P(A|\mathcal{G})$ is $\mathcal{G}_I$-measurable. Similarly, $P(B|\mathcal{G})$ is $\mathcal{G}_J$-measurable.

Thus

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

where the first of the remarks above has been used. Now, by the second remark, $P(A|\mathcal{G})\mathbf{1}_\Gamma$ is $\mathcal{G}_{[a,c)}$-measurable, and $P(B|\mathcal{G})$ is $\mathcal{G}_{[c,d)}$-measurable. Therefore they are independent, and we obtain

$$P(A \cap B|\Gamma) = P(A|\Gamma)P(B) = P(A)P(B)$$

where, in the last equality, we have used the fact that, since $\mathcal{G}_I$ and Γ are independent, $P(A|\Gamma) = E(P(A|\mathcal{G})|\Gamma) = E[P(A|\mathcal{G})] = P(A)$. To complete the proof of part (b) we observe that, by what just shown,

$$|P(A \cap B) - P(A)P(B)| = |P(A \cap B) - P(A \cap B|\Gamma)|.$$

But, letting $C := A \cap B$

$$|P(C) - P(C|\Gamma)| = \frac{|E\left[(\mathbf{1}_\Gamma - P(\Gamma))\mathbf{1}_C\right]|}{P(\Gamma)} \leq \frac{E\left[|(\mathbf{1}_\Gamma - P(\Gamma))|\right]}{P(\Gamma)} = 2(1 - P(\Gamma)).$$

Since $1 - P(\Gamma) = e^{-\lambda(c-b)}$, we obtain

$$|P(A \cap B) - P(A)P(B)| \leq 2(1 - P(\Gamma)) = 2e^{-\lambda(c-b)}$$

as desired.

We finally show that equation (2.11) holds true almost surely and in L^1, for every measurable function $F : \mathbb{R}^k \to \mathbb{R}$ such that $E[|F(X_{b_1} - X_{a_1}, \ldots, X_{b_k} - X_{a_k})|] < +\infty$. Consider the $\mathbb{R}^k$-valued stochastic process $\xi = (\xi_n)_{n\in\mathbb{N}}$ defined by

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

for fixed $\delta > 0$, $k \in \mathbb{N}$ and (a_1, b_1), $\ldots$, $(a_k, b_k) \subseteq (0, \infty)$. The process ξ is stationary, because we have proven in section 4 that X has stationary increments. Furthermore, the tail triviality of X implies that ξ is *ergodic*, because every invariant event of ξ is clearly contained in the tail σ-field of X. The existence of the limit in (2.11), both a.s. and in L^1, is then a consequence of the classical Ergodic Theorem, cf. Theorem 24.1 in [4]. Note that the ergodicity of ξ also follows form the fact that ξ is *mixing*, as shown in part (b). □

Acknowledgements

We thank Fulvio Baldovin, Massimiliano Caporin, Wolfgang Runggaldier and Attilio Stella for fruitful discussions.

References

[1] L. Accardi, Y. G. Lu, *A continuous version of de Finetti's theorem*, Ann. Probab. **21** (1993), 1478–1493.
[2] R. T. Baillie, *Long memory processes and fractional integration in econometrics*, J. Econometrics **73** (1996), 5–59.
[3] O. E. Barndorff-Nielsen, N. Shephard, *Non Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics*, J. R. Statist. Soc. B **63** (2001), 167–241.
[4] P. Billingsley, *Probability and Measure*, Third Edition, John Wiley and Sons (1995).
[5] F. Baldovin, A. Stella, *Scaling and efficiency determine the irreversible evolution of a market*, PNAS **104**, n. 50 (2007), 19741–19744.
[6] T. Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity*, J. Econometrics **31** (1986), 307–327.
[7] T. Bollerslev, H.O. Mikkelsen, *Modeling and pricing long memory in stock market volatility*, J. Econometrics **31** (1996), 151–184.
[8] T. Di Matteo, T. Aste, M. M. Dacorogna, *Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development*, J. Banking Finance **29** (2005), 827–851.
[9] R. F. Engle, *Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation*, Econometrica **50** (1982), 987–1008.
[10] D. A. Freedman, *Invariants Under Mixing Which Generalize de Finetti's Theorem: Continuous Time Parameter*, Ann. Math. Statist. **34** (1963), 1194–1216.
[11] J. C. Hull, *Options, Futures and Other Derivatives*, Pearson/Prentice Hall (2009).
[12] I. Karatzas, S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer (1988).
[13] C. Kluppelberg, A. Lindner, R. A. Maller, *A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour*, J. Appl. Probab. **41** (2004) 601–622.
[14] C. Kluppelberg, A. Lindner, R. A. Maller, *Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models*, in *From Stochastic Calculus to Mathematical Finance*, Yu. Kabanov, R. Lipster and J. Stoyanov (Eds.), Springer (2007).
[15] Wolfram Research, Inc., *Mathematica*, Version 7.0, Champaign, IL (2008).
[16] B. Øksendal, Stochastic Differential Equations, Springer-Verlag (2003).
[17] R Development Core Team (2009), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. URL: `http://www.R-project.org`.
[18] A. L. Stella, F. Baldovin, *Role of scaling in the statistical modeling of finance*, Pramana **71** (2008), 341–352.
[19] L. Weiss, *The Stochastic Convergence of a Function of Sample Successive Differences*, Ann. Math. Statist. **26** (1955), 532–536.

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

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

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

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